Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 436 (2025) 113268

Journal of Differential Equations

www.elsevier.com/locate/jde

Polynomial normal forms for ODEs near a center-saddle equilibrium point

Amadeu Delshams [a,1], Piotr Zgliczyński [b,*,2]

[a] *Lab of Geometry and Dynamical Systems and IMTech, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain*
[b] *Jagiellonian University, Institute of Computer Science and Computational Mathematics, Łojasiewicza 6, 30–348 Kraków, Poland*

Received 7 November 2024; revised 1 March 2025; accepted 25 March 2025
Available online 2 April 2025

Abstract

In this work we consider a saddle-center equilibrium for general vector fields as well as Hamiltonian systems, and we transform it locally into a polynomial normal form in the saddle variables by a change of coordinates. This problem was first solved by Bronstein and Kopanskii in 1994 [3], as well as by Banyaga, de la Llave and Wayne in 1996 [5] in the saddle case. The proof used relies on the deformation method used in [5], which in particular implies the preservation of the symplectic form for a Hamiltonian system, although our proof is different and, we believe, simpler. We also show that if the system has sign-symmetry, then the transformation can be chosen so that it also has sign-symmetry. This issue is important in our study of shadowing non-transverse heteroclinic chains (Delshams and Zgliczynski 2018 and 2024) for the toy model systems (TMS) of the cubic defocusing nonlinear Schrödinger equation (NLSE) on $2D$-torus or similar Hamiltonian PDE, which are used to prove energy transfer in these PDE.

MSC: 34C20; 34C45

Keywords: Polynomial normal form; Center manifold; Normally hyperbolic manifold

* Corresponding author.
E-mail addresses: Amadeu.Delshams@upc.edu (A. Delshams), umzglicz@cyf-kr.edu.pl (P. Zgliczyński).
1 Partially supported by the Spanish MINECO/FEDER Grant PID2021-123968NB-I00.
2 Research has been supported by polish NCN grant 2016/22/A/ST1/00077.

https://doi.org/10.1016/j.jde.2025.113268

1. Introduction

The main question addressed in this work is: Can a vector field near a center-saddle equilibrium be transformed into a polynomial normal form in the saddle variables by a change of coordinates? (finite smoothness is allowed). This problem has been solved by Bronstein and Kopanskii [3,4], as well as by Banyaga, de la Llave and Wayne [5] (they give a complete proof for maps, but the result also applies to ODEs) for a saddle equilibrium point, where it is also shown that the coordinate transformation can be chosen preserving the symplectic form, volume or contact if the original vector field preserves it.

In this work we consider a saddle-center equilibrium for general vector fields as well as Hamiltonian systems, and we apply the deformation method as in [5]. These systems arose in our work on shadowing non-transversal heteroclinic chains [10,11]. Although the technique of the proof is inspired by the approach taken in [5], our proof is different and, we believe, simpler. We also show that if the system has sign-symmetry, then the transformation can be chosen so that it also has sign-symmetry. This issue is important in our study of shadowing non-transverse heteroclinic chains [10,11] for the toy model systems (TMS) of the nonlinear cubic defocusing Schrödinger equation on 2-dimensional torus (NLSE) or similar Hamiltonian PDE, which are used to prove energy transfer in these PDE.

The plan of the paper is as follows. We first prove the result on general vector fields and Hamiltonian systems, which is stated as Theorem 3 in Section 2. This proof is concluded in section 6. Later on we introduce some additional invariance requirements (the sign-symmetry) and we show in Theorem 30 that the transformation constructed in the general case also has these properties.

1.1. Notation

We will denote $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ the natural, integer, real and complex numbers, respectively. We will assume that $0 \in \mathbb{N}$. For $x \in \mathbb{R}$, we denote $[x]$ its integer part, $\Re x$ its real part and $\Im x$ its imaginary part. For a matrix A we will denote $A^\top$ its transposed matrix. Finally we will denote $B_n(R)$ the ball of radius R centered at the origin of $\mathbb{R}^n$.

If $u = h(z)$ is a (local) diffeomorphism in $\mathbb{R}^n$, expressing new coordinates u in terms of the old ones denoted by z, and $\mathcal{Z}(z)$ is a vector field, then it is transformed by $u = h(z)$ to the pushforward vector field $h_*\mathcal{Z}$ defined as

$$h_*\mathcal{Z}(u) := Dh(h^{-1}(u))\mathcal{Z}(h^{-1}(u)).$$

Conversely, given $z = g(u)$ (for example $g = h^{-1}$) expressing the old coordinates z in terms of the new coordinates u, $\mathcal{Z}(z)$ is transformed by $z = g(u)$ to the pullback vector field $g^*\mathcal{Z}$ defined as

$$g^*\mathcal{Z}(u) = Dg(u)^{-1}\mathcal{Z}(g(u)). \tag{1}$$

The action on functions is as follows

$$h_*Z(u) = Z(h^{-1}(u)),$$
$$g^*Z(u) = Z(g(u)).$$

The commutator of two vector fields X, Y is given by

$$[X,Y]_i = \sum_j \left(\frac{\partial Y_i}{\partial z_j} X_j - \frac{\partial X_i}{\partial z_j} Y_j \right),$$
$$[X,Y](z) = DY(z)X(z) - DX(z)Y(z). \tag{2}$$

For a function $g(\varepsilon, x)$ depending on ε, which is treated as a parameter, we will use the following notations

$$g_\varepsilon(x) = g(\varepsilon, x),$$
$$Dg_\varepsilon(x) = \frac{\partial g}{\partial x}(\varepsilon, x).$$

For $g(z) \in C^r$ we will say that $g(z) = O_k(|z|^q)$ (for $z \to 0$—this might not be written explicitly) for some $k \leq r$ if

$$D^j g(z) = O(|z|^{q-j}), \quad j = 0, \ldots, \min(k,q).$$

If $z = (x, y)$ and $g(z) \in C^r$, we will say $g(z) = O_k(|x|^q)$ for $z \in W$ (bounded set) and $x \to 0$ for $k \leq r$ if

$$D^j g(z) = O(|x|^{q-j}), \quad j = 0, \ldots, \min(k,q),$$

where D is the differentiation with respect to $z = (x, y)$, i.e., all partial derivatives are involved.

Given a vector space V with the basis $\{e_i\}_{i=1}^m$ and a multiindex $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{N}^m$, for a vector $z = \sum_i z_i e_i$ we define z^α by

$$z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_m^{\alpha_m}.$$

By $|\alpha|$ we will denote the degree of multiindex, defined by $|\alpha| = \sum_i \alpha_i$.

Often, when discussing normal forms we will use complex coordinates $c_k = x_k + i y_k$, and then a part of the phase space will be described by $\mathbb{C}^r$ and the polynomials on $\mathbb{R}^m \times \mathbb{C}^r$ will be of the following form

$$z^\alpha c^\beta \overline{c}^{\beta^*} = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_m^{\alpha_m} \cdot c_1^{\beta_1} \cdots c_r^{\beta_r} \cdot \overline{c}_1^{\beta_1^*} \cdots \overline{c}_r^{\beta_r^*}. \tag{3}$$

In such situation the degree of the monomial $z^\alpha c^\beta \overline{c}^{\beta^*}$ is given by $|\alpha| + |\beta| + |\beta^*|$.

1.1.1. Symplectic forms and Hamiltonian equations

By Ω we will denote a symplectic form (a closed non-degenerate differential 2-form) on some phase space V. A map is called symplectic (with respect to Ω) iff it preserves the form Ω.

The standard (real) symplectic form is given by

$$\Omega = \sum_j dx_j \wedge dy_j.$$

Occasionally we will use more general symplectic form, for example $\Omega' = g^*\Omega$.

Given a real function $H(t,z)$ for real t and $z=(x,y)$, and a real symplectic form Ω the Hamiltonian vector field $X_H(t,z)$ is defined by [1]

$$\Omega(X_H(t,z),\eta) = D_z H(t,z)\eta, \quad \forall \eta \in T_z V. \tag{4}$$

Therefore the Hamiltonian equations can be written as

$$\dot{z} = J(z)\nabla_z H(t,z) \tag{5}$$

where $\nabla_z H(t,z) = (D_z H(t,z))^\top$ and $J(z)$ is the regular antisymmetric matrix defined by (4).

In the case of the standard symplectic form the Hamiltonian equations of motion are just

$$\dot{x}_j = \frac{\partial H}{\partial y_j}, \quad \dot{y}_j = -\frac{\partial H}{\partial x_j}. \tag{6}$$

Since we are going to consider systems $\dot{z} = Az + O_2(z)$ in a neighborhood of an equilibrium point, it will be very convenient to consider complex variables $c_l = x_l + iy_l$ for some $l \in L$ associated to pure imaginary eigenvalues $\nu = \pm i\omega$, $\omega \neq 0$, of the matrix A. Then the standard symplectic form Ω becomes

$$\Omega = \sum_{k\in K} dx_k \wedge dy_k + \frac{i}{2}\sum_{l\in L} dc_l \wedge d\bar{c}_l.$$

In such case the Hamiltonian equations of motion (6) become

$$\begin{aligned} \dot{x}_k &= \frac{\partial H}{\partial y_k}, \quad \dot{y}_k = -\frac{\partial H}{\partial x_k}, \quad k \in K, \\ \dot{c}_l &= -2i\frac{\partial H}{\partial \bar{c}_l}, \quad l \in L. \end{aligned}$$

One can get a more compact notation introducing the complex variables $c_k = x_k + iy_k$ for the indexes $k \in K$ associated to real eigenvalues to get a canonical complex Hamiltonian system

$$\dot{c}_j = -2i\frac{\partial H}{\partial \bar{c}_j}$$

associated to the standard complex symplectic form $\Omega = \frac{i}{2}\sum_j dc_j \wedge d\bar{c}_j$.

In addition, in the presence of saddle variables associated to complex eigenvalues $\nu = \pm a \pm ib$, $ab \neq 0$, of the matrix A, complex pairs of variables are handy and are of the form $c_{lm} = x_l + ix_m$ and $g_{lm} = y_l + iy_m$. Then in the standard symplectic form the contribution of $dx_l \wedge dy_l + dx_m \wedge dy_m$ becomes $\frac{1}{2}dc_{lm} \wedge d\bar{g}_{lm} + \frac{1}{2}d\bar{c}_{lm} \wedge dg_{lm}$ so that Ω becomes

$$\Omega = \sum_{k\in K} dx_k \wedge dy_k + \frac{1}{2}\sum_{(l,m)\in L'\times L''} (dc_{lm} \wedge d\bar{g}_{lm} + d\bar{c}_{lm} \wedge dg_{lm}) + \frac{i}{2}\sum_{l\in L} dc_l \wedge d\bar{c}_l,$$

and the equations of motion for the c_{lm} and g_{lm} variables are

$$\dot{c}_{lm} = 2\frac{\partial H}{\partial \bar{g}_{lm}}, \quad \dot{g}_{lm} = -2\frac{\partial H}{\partial \bar{c}_{lm}}.$$

For a real system associated to a real Hamiltonian, the equations of motion of $\overline{c}_l$, $\overline{c}_{lm}$ and $\overline{g}_{lm}$ are redundant and there is no need to write them, since they are conjugate to those of c_l, c_{lm}, g_{lm}, respectively. Indeed

$$\dot{\overline{c}}_l = 2i\frac{\partial H}{\partial c_l}, \quad \dot{\overline{c}}_{lm} = -2\frac{\partial H}{\partial g_{lm}}, \quad \dot{\overline{g}}_{lm} = -2\frac{\partial H}{\partial c_{lm}}.$$

1.2. Resonant monomials

Assume that we have a coordinate system $(z_1, \ldots, z_n)$ which brings matrix $A \in \mathbb{R}^{n\times n}$ to its complex Jordan normal form. Let $\nu_1, \ldots, \nu_n$ be the corresponding eigenvalues. In case some of them are not real, they appear within pairs of conjugate complex numbers, and the Jordan normal form for A takes complex values.

Let us denote by e_j, $j = 1, \ldots, n$ the n-dimensional canonical basis, and consider a polynomial vector field

$$\dot{z} = Az + \sum_{2\leq r\leq N} \sum_{|\alpha|=r} \left(p^1_\alpha z^\alpha, \ldots, p^n_\alpha z^\alpha\right) \tag{7}$$

The following definition is taken from [3].

Definition 1. [3, Def. 2.4] Let $j \in \{1, \ldots, n\}$ and $\alpha \in \mathbb{Z}^n_+$. The pair (j, α) as well as the monomial $z^\alpha e_j$ of system (7) is called *resonant* (in the sense of Poincaré) if the following equality holds:

$$\nu_j = \langle \alpha, \nu\rangle := \sum_{k=1}^{n} \alpha_k \nu_k.$$

The integer $r = |\alpha|$ is called the *order of resonance*.
We will denote by $\mathcal{R}\mathrm{es}(A, k_1, k_2)$ the set of resonant monomials of equation (7) of order greater than or equal to k_1 and of order less than or equal to k_2.

We will occasionally say that z^α is resonant for the j-th variable if $z^\alpha e_j$ is resonant.

In the case that system (7) is a Hamiltonian of $m = n/2$ degrees of freedom

$$H = \sum_{2\leq|\alpha|+|\beta|+|\beta^*|\leq N+1} a_{\alpha\beta\beta^*} z^\alpha c^\beta \overline{c}^{\beta^*}, \tag{8}$$

where $z = (x, y) \in \mathbb{R}^{2m-2r}$, $c \in \mathbb{C}^r$ and $a_{\alpha\beta\beta^*} = \overline{a_{\alpha\beta^*\beta}}$ for a real Hamiltonian, the eigenvalues $(\nu_1, \ldots, \nu_{m-r}, \nu_{m-r+1}, \ldots, \nu_m, \nu_{m+1}, \ldots, \nu_{2m-r}, \nu_{2m-r+1}, \ldots, \nu_{2m}) := (\nu^+_{\mathrm{R}}, \nu_{\mathrm{C}}, \nu^-_{\mathrm{R}}, \nu^*_{\mathrm{C}})$ of the Hamiltonian matrix A appear in pairs with both signs and can be ordered, for example, so that the $2(m-r)$ complex ones, $0 \leq r \leq m$, satisfy $\nu_i = \bar{\nu}_{m+i}$ for $i = 1, \ldots, r$. Then the resonant monomials of Hamiltonian (8) are those that satisfy $\langle \nu_{\mathrm{R}}, \alpha\rangle + \langle \beta, \nu_{\mathrm{C}}\rangle + \langle \beta^*, \nu^*_{\mathrm{C}}\rangle = 0$, where

$\nu_\mathrm{R} = \left(\nu_\mathrm{R}^+, \nu_\mathrm{R}^-\right)$ (see, for example, [18, Sec. 10.4]) We will denote by $\mathrm{Res}(A, k_1, k_2)$ the set of resonant monomials of Hamiltonian (8) of order greater than or equal to k_1 and of order less than or equal to k_2.

2. The polynomial normal form theorem for general vector fields and Hamiltonian system

2.1. Normal form around an equilibrium point, the Poincaré-Dulac theorem

Assume that we have a coordinate system $(z_1, \ldots, z_n)$ which brings a matrix $A \in \mathbb{R}^{n\times n}$ to its complex Jordan normal form, and consider the system

$$\dot{z} = Az + O_2(|z|^2). \tag{9}$$

One would like to bring (9) to the simplest possible form. The first well-known result in this direction is the Poincaré-Dulac theorem [1].

Theorem 1. *If the vector field (9) is C^{Q+1} with $Q \geq 2$, then for any $J \subset \mathbb{Z}_+^n$ such that $\mathcal{R}\mathrm{es}(A, 2, Q) \subset J$ and $2 \leq |\alpha| \leq Q$ for $\alpha \in J$, system (9) after finite number of analytic changes of variables, can be written as*

$$\dot{z} = Az + \mathcal{N}_Q(z) + \mathcal{R}_{Q+1}(z),$$

where $\mathcal{N}_Q(z) = \sum_{\alpha\in J} \mathcal{N}_{Q,\alpha} z^\alpha$ and the remainder vector field satisfies $\mathcal{R}_{Q+1}(z) = O_{Q+1}(|z|^{Q+1})$. Moreover, if system (9) is Hamiltonian, the changes of variables can be chosen symplectic, so that the transformed system is also Hamiltonian.

Regarding the notation z^α, it should be noted that the n-dimensional variable $z = (x, c)$ can contain both real variables x associated with real eigenvalues of the matrix A, as well as conjugate variables $c, \overline{c}$ associated with complex eigenvalues of the matrix A, so any monomial z^α is of the already mentioned form (3), i.e., $z^\alpha = x^{\hat{\alpha}} c^\beta \overline{c}^{\beta^*}$.

In [3] the authors show that also terms which are not weakly resonant can be removed, but then the transformation is of finite smoothness.

2.2. Normal form for an equilibrium point with the local center manifold

We will now consider a system around a saddle-center equilibrium, thus splitting the $n \times n$ matrix A in complex Jordan normal form of (9) into two matrices, a $n_A \times n_A$ hyperbolic (just called *saddle* from now on) matrix A, whose eigenvalues have nonzero real part, and a $n_B \times n_B$ *center* matrix B whose eigenvalues have zero real part, with $n = n_A + n_B$. Since we are working with real systems, all eigenvalues of B appear in pairs and the Jordan normal form variables for B are complex variables $c \in \mathbb{C}^r$, with $2r = n_B$, whereas the Jordan normal form variables for a matrix A with k_A real eigenvalues and $2l_A$ non-real eigenvalues are of the type $z = (x, \zeta) \in \mathbb{R}^{k_A} \times \mathbb{C}^{l_A}$, with $k_A + 2l_A = n_A$, so that the phase space of our system is described by $(z, c) = (x, \zeta, c)$:

$$\begin{aligned} \dot{z} &= Az + N_z(z, c), \qquad z \in \mathbb{R}^{k_A} \times \mathbb{C}^{l_A}, \\ \dot{c} &= Bc + N_c(z, c), \qquad c \in \mathbb{C}^r, \end{aligned} \tag{10}$$

with $N(z,c) = (N_z(z,c), N_c(z,c)) = O_2(|z|^2 + |c|^2)$.

From now on we will denote our phase space as

$$\mathcal{P} = \left\{ z \in \mathbb{R}^{k_A} \times \mathbb{C}^{l_A}, c \in \mathbb{C}^r \right\}. \tag{11}$$

We now introduce some notation for the *spectral gap* of A.

Definition 2. Given a saddle matrix A (i.e., $\Re\lambda \neq 0$ for $\lambda \in \mathrm{Sp}(A)$), there exist four positive real numbers delimiting two minimal intervals $[-\lambda_{\max}, -\lambda_{\min}]$ and $[\mu_{\min}, \mu_{\max}]$ such that $\Re\mathrm{Sp}(A) \subset [-\lambda_{\max}, -\lambda_{\min}] \cup [\mu_{\min}, \mu_{\max}]$. For the sake of brevity, we will just call $\mathfrak{S} = \mathfrak{S}(A)$ to the vector formed by these numbers characterizing the *spectral gap* of A:

$$\mathfrak{S} = (\lambda_{\min}, \lambda_{\max}, \mu_{\min}, \mu_{\max}) \implies \Re\mathrm{Sp}(A) \subset [-\lambda_{\max}, -\lambda_{\min}] \cup [\mu_{\min}, \mu_{\max}].$$

Remark 2. This definition works well only for a genuine saddle matrix A possessing both eigenvalues with positive real part and eigenvalues with negative real part. In particular, if A is a Hamiltonian matrix, then $\lambda_{\min} = \mu_{\min}$ and $\lambda_{\max} = \mu_{\max}$. If $\Re\mathrm{Sp}(A) < 0$ then the interval $[\mu_{\min}, \mu_{\max}]$ is empty and $\mathfrak{S}(A) = (\lambda_{\min}, \lambda_{\max})$. Analogously, if $\Re\mathrm{Sp}(A) > 0$, the interval $[-\lambda_{\max}, -\lambda_{\min}]$ is empty and $\mathfrak{S}(A) = (\mu_{\min}, \mu_{\max})$.

We can now state the first main result of this paper.

Theorem 3. *Consider system (10), such that $(z,c) = 0$ is a non-degenerate equilibrium point and the subspace $\{z = 0\}$ is invariant, and let $\mathfrak{S}$ be the spectral gap of A as introduced in Definition 2.*

Then for any $k \geq 1$, there exists $Q_0 = Q_0(k, \mathfrak{S})$, such that for any $Q \geq Q_0$ there exists $q_0 = q_0(Q, k, \mathfrak{S})$, so that the following assertion of the theorem is valid.

If system (10) is C^q, with $q \geq q_0$ and if $q - 1 \geq P \geq Q$ and $q \geq Q + 3$, then there exists a C^k change of variables in a neighborhood of the origin, transforming it to the system

$$\begin{aligned} \dot{z} &= Az + \mathcal{N}(z,c) + \mathcal{R}(z,c), \\ \dot{c} &= Bc + O_2(|z|^2 + |c|^2), \end{aligned} \tag{12}$$

with

$$\mathcal{N}(z,c) = \sum_{\substack{j=1,\dots,k_A+l_A, \\ (j,(\alpha,\alpha^*,\beta,\beta^*)) \in \mathcal{R}\mathrm{es}((A,B),2,P)}} p^j_{\alpha,\alpha^*,\beta,\beta^*} z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} e_j,$$

$$\begin{aligned} \mathcal{R}(z,c) &= \sum_{\substack{j=1,\dots,k_A+l_A, \\ |\alpha|+|\alpha^*|=1, |\beta|+|\beta^*|=P+1-Q}} g^j_{\alpha,\alpha^*,\beta,\beta^*}(z,c) z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} e_j \\ &= O_{P+1-Q}\left(|c|^{P+1-Q}\right) O_1(z), \end{aligned}$$

where $p^j_{\alpha,\alpha^*,\beta,\beta^*}$ are constants and $g^j_{\alpha,\alpha^*,\beta,\beta^*}(z,c)$ are continuous functions.

Moreover, if system (10) is Hamiltonian, the change of variables can be chosen symplectic, and therefore the transformed system is also Hamiltonian.

Our second main result, Theorem 30, is that if system (10) has *sign-symmetry* (see Section 7), that is, it is $\mathcal{S}_s$-symmetric for any symmetry $\mathcal{S}_s(z_+, z_-, c_1, \ldots, c_m) = (s_+z_+, s_-z_-, s_1c_1, \ldots, s_mc_m)$ and any choice of signs $\{s = (s_+, s_-, s_1, \ldots, s_m)\}$, then the change of variables can be chosen such that the transformed system also has sign-symmetry.

Remark 4. The sign symmetry is a property introduced in [8], which guarantees that the variational equations of system (10) along ingoing and outgoing orbits have a block diagonal structure, and are an essential property to prove that heteroclinic connections between different systems (10) can be shadowed, even when these heteroclinic connections are not transversal, see [11]. Theorem 30 is adapted in Lemma 31 to the toy model system derived in [8].

Remark 5. The assumption that the subspace $\{z = 0\}$ is invariant could be easily satisfied. Namely, for a saddle-center fixed after a coordinate change where a local center manifold is set as $\{z = 0\}$.

Remark 6. The first results about polynomial normal forms in the saddle variables for saddle-center equilibria are due to Bronstein and Kopanskii [3]. In the pure saddle case the assertions of Theorem 3 are valid for $P = Q$ and $\mathcal{R}(z) = 0$ and were proved using the deformation method by Banyaga, de la Llave and Wayne [5] (in this second paper the authors give a complete proof for maps, but the result also applies to ODEs), showing the preservation of geometric (Hamiltonian, contact, volume preserving) structures. Bronstein and Kopanskii [4] applied later on these methods to show also the preservation of geometric structures.

Remark 7. Bringing the remainder term to the form stated in the theorem (or removing it in the pure saddle case) does not require that $\mathcal{N}$ contains only the resonant terms, it fact it can contain any polynomial terms of order up to P.

Remark 8. In the transformed system (12), the term $\mathcal{N}(z, c)$ is the polynomial normal form up to degree P in both variables z and c provided by the Poincaré-Dulac Theorem 1, whereas the term $\mathcal{R}(z, c)$ contains the linear terms in the variable z with coefficients of order $Q + P - 1$ in the variable c of the remainder of Theorem 1. It is important to notice that, in general, the linear part of the z-equation in (10) cannot be reduced to a constant matrix (see, for instance the review [19] for sufficient conditions), since the frequencies in the matrix B are not assumed to be even incommensurable.

Remark 9. For a Hamiltonian system (10), the Hamiltonian associated to the transformed Hamiltonian system (12) provided by Theorem 3 takes the form $H = H_1(c) + H_2(z) + N(z, c) + R(z, c)$, where $H_1(c)$ has nondegenerate quadratic part, $H_2(z)$ is quadratic (in a Hamiltonian Jordan form) and

$$\begin{aligned} N(z, \bar{z}, c, \bar{c}) &= \sum_{(\alpha,\alpha^*,\beta,\beta^*)\in \mathrm{Res}((A,B),3,P+1)} P_{\alpha,\alpha^*,\beta,\beta^*}\, z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} \\ R(z, \bar{z}, c, \bar{c}) &= \sum_{|\alpha|+|\alpha^*|=2,\, |\beta|+|\beta^*|=P+1-Q} G_{\alpha,\alpha^*,\beta,\beta^*}(z, \bar{z}, c, \bar{c})\, z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} \\ &= O_{P+1-Q}\left(|c|^{P+1-Q}\right) O_2(z^2), \end{aligned}$$

where $P_{\alpha,\alpha^*,\beta,\beta^*}$ are constants and $G_{\alpha,\alpha^*,\beta,\beta^*}(z,c)$ are continuous functions.

Remark 10. The values of the constants Q, Q_0, q_0, k from Theorem 3 for the general vector field as follows

$$Q_0(k) = \max\left(k + \left[k\frac{\lambda_{\max}}{\mu_{\min}} + \frac{\mu_{\max}}{\mu_{\min}}\right], \left[(k+1)\frac{\mu_{\max}}{\mu_{\min}}\right]\right) + \max\left(k + \left[k\frac{\mu_{\max}}{\lambda_{\min}} + \frac{\lambda_{\max}}{\lambda_{\min}}\right], \left[(k+1)\frac{\lambda_{\max}}{\lambda_{\min}}\right]\right), \tag{13}$$

and for the Hamiltonian case

$$Q_0(k) = 2k + 1 + 2\left[(k+1)\frac{\mu_{\max}}{\mu_{\min}}\right]. \tag{14}$$

In both cases

$$q_0(k, Q) = Q + 3,$$

and in the pure saddle case $q_0(k, Q) = Q + 2$.

The comparison of Q_0 and q_0 with the numbers obtained in previous works [3–5] in the case of a saddle is discussed in Appendix C.

The proof of the above theorem has two parts. In the first part we will use the Poincaré-Dulac theorem to remove some non-resonant terms and in the second part we will remove the remainder using the deformation method following the idea from [5]. The proof of this theorem will start in the next section.

If the matrix A has only real eigenvalues, we have the following corollary.

Corollary 11. *Under the same assumptions as in Theorem 3, assume further that all eigenvalues of A are real.*

Then with the same k, Q, P as in Theorem 3, system (10) can be brought locally to the following form

$$\begin{aligned} \dot{z} &= Az + \sum_{j=1,\dots,n_A,(j,\alpha)\in\mathcal{R}\mathrm{es}(A,2,P)} g^j_\alpha(c) z^\alpha e_j \\ &\quad + \sum_{j=1,\dots,n_A,|\alpha|=1,|\beta|+|\beta^*|=P+1-Q} g^j_{\alpha,\beta,\beta^*}(z,c) z^\alpha c^\beta \bar{c}^{\beta^*} e_j, \\ \dot{c} &= Bc + O_2(|z|^2 + |c|^2), \end{aligned}$$

where $g^j_\alpha(c)$ are polynomials and $g^j_{\alpha,\beta,\beta^*}(z,c)$ are continuous functions.

Proof. First note that since A has only real eigenvalues, the variables $\bar{z}$ do not appear.

Let us take a closer look at the sum of monomials in the first equation of (12). Since all eigenvalues of B are purely imaginary and those of A are real, then for any $(j, (\alpha, \beta)) \in \mathcal{R}\mathrm{es}((A, B), 2, P)$ we have that $(j, \alpha) \in \mathcal{R}\mathrm{es}(A, 2, P)$. Therefore we have

$$\sum_{\substack{j=1,\dots,n_A,\\ (j,(\alpha,(\beta,\beta^*)))\in\mathcal{R}\mathrm{es}((A,B),2,P)}} p^j_{\alpha,\beta,\beta^*} z^\alpha c^\beta \bar{c}^{\beta^*} e_j = \sum_{(j,\alpha)\in\mathcal{R}\mathrm{es}(A,2,P)} g^j_\alpha(c) z^\alpha e_j$$

for the polynomials $g^j_\alpha(c) = g^j_\alpha(c,\bar{c}) = \sum_{2-|\alpha|\le|\beta|+|\beta^*|\le P-|\alpha|} p^j_{\alpha,\beta,\beta^*} c^\beta \bar{c}^{\beta^*}$. □

Remark 12. For a Hamiltonian system (10), when all the eigenvalues of A are real, the Hamiltonian associated to the transformed Hamiltonian system (12) takes the form $H(z,c) = H_1(c) + H_2(z) + N(z,c) + R(z,c)$, with

$$N(z,c,\bar{c}) = \sum_{\alpha\in\mathrm{Res}(A,3,P+1)} P_\alpha(c,\bar{c}) z^\alpha,$$

$$R(z,c,\bar{c}) = \sum_{|\alpha|=2,|\beta|+|\beta^*|=P+1-Q} G_{\alpha,\beta,\beta^*}(z,c)\, z^\alpha c^\beta \bar{c}^{\beta^*},$$

where $P_\alpha(c,\bar{c})$ are polynomials and $G_{\alpha,\beta,\beta^*}(z,c)$ are continuous functions.

3. The first part of the proof of Theorem 3, removal of non-resonant terms up to some order

Let us fix $P \ge Q \ge 2$. Let us write a Taylor formula for the z-component of our vector field

$$\dot{z} = Az + \sum_{\substack{j=1,\dots,k_A+l_A,\\ 2\le|\alpha|+|\alpha^*|+|\beta|+|\beta^*|\le P}} p^j_{\alpha,\alpha^*,\beta,\beta^*} z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} e_j + \mathcal{R}(z,c), \tag{15}$$

$$\dot{c} = Bc + O_2(|z|^2 + |c|^2), \tag{16}$$

where

$$\mathcal{R}(z,c) = O_{P+1}\left((|z|+|c|)^{P+1}\right). \tag{17}$$

Using Theorem 1 we can remove the resonant terms of the z-component in system (15)–(16) up to order P to obtain (with a different remainder term which will denote again by $\mathcal{R}$, and which satisfies (17))

$$\dot{z} = Az + \mathcal{N}_1(z,c) + \mathcal{R}(z,c),$$
$$\dot{c} = Bc + O_2(|z|^2 + |c|^2),$$

with

$$\mathcal{N}_1(z,c) = \sum_{\substack{j=1,\dots,k_A+l_A,\\ \left(j,\left(\alpha,\alpha^*,\beta,\beta^*\right)\right)\in\mathcal{R}\mathrm{es}((A,B),2,P)}} p^j_{\alpha,\alpha^*,\beta,\beta^*} z^\alpha \bar{z}^{\alpha^*} c^\beta \bar{c}^{\beta^*} e_j.$$

The remainder can be written as sum

$$\mathcal{R}(c,z) = \sum_{|\alpha|=P+1} \left(\int_0^1 \frac{(1-t)^P}{P!} D_\alpha \mathcal{R}(t(c,z))\,\mathrm{d}t \right) (c,z)^\alpha .$$

Now we collect in $\mathcal{R}_2(z,c)$ the terms of $\mathcal{R}(z,c)$ of order greater than or equal to $Q+1$ *in the variable* z, and name $\mathcal{R}_1(z,c) := \mathcal{R}(z,c) - \mathcal{R}_2(z,c)$. In other words, split the remainder term $\mathcal{R}(z,c)$ as

$$\begin{aligned} \mathcal{R}(z,c) &= \mathcal{R}_1(z,c) + \mathcal{R}_2(z,c) \\ \mathcal{R}_1(z,c) &= O_{P+1-Q}\left(c^{P+1-Q}\right) O_1(z), \quad \mathcal{R}_2(z,c) = O_{Q+1}(z^{Q+1}). \end{aligned} \tag{18}$$

Notice that for a general remainder $\mathcal{R}(z,c)$ it would simply follow that $\mathcal{R}_1(z,c) = O_{P+1-Q}\left(c^{P+1-Q}\right)$. The special expression (18) comes from the fact that $\mathcal{R}(z,c)$ vanishes for $z=0$.

Remark 13. By assumption (17), $\mathcal{R}(z,c)$ is of order $P+1$ in the variables (z,c), so we could even write

$$\begin{aligned} \mathcal{R}_1(z,c) &= O_{P+1-Q}\left(c^{P+1-Q}\right) O_1(z) O_{Q-1}\left((|z|+|c|)^{Q-1}\right), \\ \mathcal{R}_2(z,c) &= O_{Q+1}(z^{Q+1}) O_{P-Q}\left((|z|+|c|)^{P-Q}\right), \end{aligned}$$

but as we are going to bound the factors containing O_{Q-1} and O_{P-Q} by constants later on, the estimates (18) will be enough for us. On the other hand, since $\mathcal{R}(0,c)$ vanishes for any c, the same happens to the $O(|c|^{P+1})$ term in $\mathcal{R}(z,c)$

$$\left(\int_0^1 \frac{(1-t)^P}{P!} D_c^{(P+1)} \mathcal{R}(tc,tz)\,\mathrm{d}t \right) c^{P+1}.$$

As a result of these transformations we obtain the following system

$$\dot{z} = Az + \mathcal{N}(z,c) + \mathcal{R}(z,c), \tag{19}$$

$$\dot{c} = Bc + O_2(|z|^2 + |c|^2), \tag{20}$$

where

$$\begin{aligned} \mathcal{N}(z,c) &= \mathcal{N}_1(z,c) + \mathcal{R}_1(z,c), \\ \mathcal{R}(z,c) &= \mathcal{R}_2(z,c) = O_{Q+1}(|z|^{Q+1}). \end{aligned} \tag{21}$$

Moreover, due to the saddle character of the matrix A, it splits as

$$A = \begin{pmatrix} A_u & 0 \\ 0 & A_s \end{pmatrix},$$

and the matrices A_u, A_s, and B are written in an appropriate Jordan form with

$$m_l(A_u) > 0, \qquad \mu_{log}(A_s) < 0, \qquad \mu_{log}(B) < \delta, \qquad \mu_{log}(-B) < \delta \tag{22}$$

for any $\delta > 0$ small enough fixed in advance, where m_l and μ_{log} are the logarithmic norms recalled in Appendix A.

The logarithmic norm $\mu_{log}(B)$ might be not equal to zero in the presence of non-trivial Jordan blocks, however in such situation it can be made arbitrarily close to 0, by choosing a linear coordinate system in the center direction so that the off-diagonal terms are very small.

In the Hamiltonian case, all these linear transformations to Jordan normal form can be chosen to be symplectic, see [20,2]. In order to have the vector field to be in the form (19)–(21) for Hamiltonian case we proceed as follows. For Hamiltonian $H(z, c) - H(c, 0)$ we consider the Taylor formula with remainder of order $P + 2$ (notice that we consider order greater than in the case of the general vector field, because the vector fields are obtained from the derivatives of Hamiltonian) and remove all or some non-resonant terms up to order $P + 1$ and then we split the reminder gathering all terms with order $Q + 2$ or higher in z. As a result our Hamiltonian takes the following form

$$H(z, c) = H_{c,2}(c) + N_c(c) + H_2(z) + N_z(z, c) + R_1(z, c) + R_2(z, c), \tag{23}$$

with $H_2(z)$ and $H_{c,2}(c)$ being quadratic Hamiltonians in suitable Jordan forms and

$$\begin{aligned}
H(c, 0) &= H_{c,2}(c) + N_c(c) \\
N_c(c) &= O_{q+1}(|c|^3), \\
N_z(z, \overline{z}, c, \overline{c}) &= \sum_{(\alpha,\alpha^*,\beta,\beta^*)\in \mathrm{Res}((A,B),3,P+1)} P_{\alpha,\alpha^*,\beta,\beta^*}\, z^\alpha \overline{z}^{\alpha^*} c^\beta \overline{c}^{\beta^*} \\
R_1(z, \overline{z}, c, \overline{c}) &= \sum_{|\alpha|+|\alpha^*|=2,\, |\beta|+|\beta^*|=P+1-Q} G_{\alpha,\alpha^*,\beta,\beta^*}(z, \overline{z}, c, \overline{c})\, z^\alpha \overline{z}^{\alpha^*} c^\beta \overline{c}^{\beta^*} \\
&= O_{P+1-Q}\left(|c|^{P+1-Q}\right) O_2(z^2), \\
R_2(z, c) &= O_{Q+2}(|z|^{Q+2}),
\end{aligned} \tag{24}$$

where $P_{\alpha,\alpha^*,\beta,\beta^*}$ are constants and $G_{\alpha,\alpha^*,\beta,\beta^*}(z, c)$ are continuous functions.

Now our goal is to apply the deformation method to remove the remainder term $\mathcal{R}_2$ (or R_2 in Hamiltonian setting). This starts in the next section.

4. Derivation of the cohomological equation

The process of removing the remainder term $\mathcal{R}$ in (19)–(20) through the deformation method, involves several stages and is concluded at the end of section 6. The first step is to derive a *cohomological equation*.

Consider a vector field dependent on a parameter ε

$$\mathcal{Z}_\varepsilon = \mathcal{Z}_0 + \varepsilon \mathcal{R}(z). \tag{25}$$

Following the idea of the deformation method taken from [5], to remove the term $\varepsilon\mathcal{R}(z)$ from equation (25) we would like to find a family of diffeomorphisms g_ε such that

$$g_\varepsilon^* \mathcal{Z}_\varepsilon = \mathcal{Z}_0. \tag{26}$$

Since g_ε is a smooth family of diffeomorphisms, we can find a family of vector fields $\mathcal{G}_\varepsilon$ such that

$$\frac{d}{d\varepsilon} g_\varepsilon = \mathcal{G}_\varepsilon \circ g_\varepsilon. \tag{27}$$

This more carefully written means that

$$\frac{\partial g}{\partial \varepsilon}(\varepsilon, x) = \mathcal{G}(\varepsilon, g(\varepsilon, x)),$$

so that

$$\mathcal{G}(\varepsilon, x) = \frac{\partial g}{\partial \varepsilon}(\varepsilon, g_\varepsilon^{-1}(x)).$$

Note that (27) is a non-autonomous o.d.e.

In [5] the main focus was on diffeomorphisms and for vector fields the authors claim (without a proof) that for the pushforward $(g_\varepsilon)_*$ it holds that

$$\frac{d}{d\varepsilon}(g_\varepsilon)_* \mathcal{Z}_\varepsilon = (g_\varepsilon)_* \left([\mathcal{Z}_\varepsilon, \mathcal{G}_\varepsilon] + \frac{d}{d\varepsilon} \mathcal{Z}_\varepsilon \right), \tag{28}$$

which is not true, since the right formula is

$$\frac{d}{d\varepsilon}(g_{\varepsilon *})\mathcal{Z}_\varepsilon = (g_{\varepsilon *}) \left([\mathcal{Z}_\varepsilon, \mathcal{G}_\varepsilon] + \frac{d\mathcal{Z}_\varepsilon}{d\varepsilon} \right) + [(g_{\varepsilon *})\mathcal{Z}_\varepsilon, \mathcal{G}_\varepsilon - (g_{\varepsilon *})\mathcal{G}_\varepsilon].$$

However it turns out that equation (28) works for the pullback g^*. Namely, we have

Lemma 14. *Assume that $g \in C^2$ and $\mathcal{G} \in C^1$ are such that $g_0 = \mathrm{Id}$ and*

$$\frac{d}{d\varepsilon} g_\varepsilon = \mathcal{G}_\varepsilon \circ g_\varepsilon. \tag{29}$$

Assume that $\mathcal{Z}(\varepsilon, z)$ is C^1. Then

$$\frac{d(g_\varepsilon^* \mathcal{Z}_\varepsilon)}{d\varepsilon} = g_\varepsilon^* \left([\mathcal{Z}_\varepsilon, \mathcal{G}_\varepsilon] + \frac{d\mathcal{Z}_\varepsilon}{d\varepsilon} \right). \tag{30}$$

Proof. From (29) we obtain

$$\frac{\partial}{\partial \varepsilon} \frac{\partial}{\partial x} g(\varepsilon, x) = \frac{\partial \mathcal{G}}{\partial x}(\varepsilon, g(\varepsilon, x)) \frac{\partial g}{\partial x}(\varepsilon, x) = D\mathcal{G}_\varepsilon(g(\varepsilon, x)) Dg_\varepsilon(x). \tag{31}$$

Let us denote

$$Y(\varepsilon, u) = \left(g_\varepsilon^* \mathcal{Z}_\varepsilon\right)(u).$$

Then by (1)

$$Dg_\varepsilon(u)Y(\varepsilon, u) = \mathcal{Z}_\varepsilon(g_\varepsilon(u)).$$

Observe that $Y(\varepsilon, u)$ is C^1.

We differentiate the above equality with respect to ε. For the left hand side we obtain (we use (31))

$$\begin{aligned}\frac{d}{d\varepsilon}\left(Dg_\varepsilon(u)Y(\varepsilon,u)\right) &= \left(\frac{\partial}{\partial\varepsilon}Dg_\varepsilon(u)\right)Y(\varepsilon,u) + Dg_\varepsilon(u)\frac{\partial Y}{\partial\varepsilon}(\varepsilon,u) \\ &= D\mathcal{G}_\varepsilon(g_\varepsilon(u))Dg_\varepsilon(u)Y(\varepsilon,u) + Dg_\varepsilon(u)\frac{\partial Y}{\partial\varepsilon}(\varepsilon,u) \\ &= D\mathcal{G}_\varepsilon(g_\varepsilon(u))\mathcal{Z}_\varepsilon(g_\varepsilon(u)) + Dg_\varepsilon(u)\frac{\partial Y}{\partial\varepsilon}(\varepsilon,u),\end{aligned}$$

while for the right hand side we have

$$\begin{aligned}\frac{d}{d\varepsilon}\left(\mathcal{Z}_\varepsilon(g_\varepsilon(u))\right) &= \frac{\partial\mathcal{Z}}{\partial\varepsilon}(\varepsilon, g_\varepsilon(u)) + D\mathcal{Z}_\varepsilon(g_\varepsilon(u))\frac{\partial g}{\partial\varepsilon}(\varepsilon,u) \\ &= \frac{\partial\mathcal{Z}}{\partial\varepsilon}(\varepsilon, g_\varepsilon(u)) + D\mathcal{Z}_\varepsilon(g_\varepsilon(u))\mathcal{G}_\varepsilon(g_\varepsilon(u)).\end{aligned}$$

Therefore we obtain (we use also (2))

$$\begin{aligned}Dg_\varepsilon(u)\frac{\partial Y}{\partial\varepsilon}(\varepsilon,u) &= -D\mathcal{G}_\varepsilon(g_\varepsilon(u))\mathcal{Z}_\varepsilon(g_\varepsilon(u)) + D\mathcal{Z}_\varepsilon(g_\varepsilon(u))\mathcal{G}_\varepsilon(g_\varepsilon(u) \\ &\quad + \frac{d\mathcal{Z}_\varepsilon}{d\varepsilon}(g_\varepsilon(u)) \\ &= \left([\mathcal{G}_\varepsilon, \mathcal{Z}_\varepsilon] + \frac{d\mathcal{Z}_\varepsilon}{d\varepsilon}\right)(g_\varepsilon(u)).\end{aligned}$$

This establishes (30). □

Therefore we obtain that the desired conjugation (26) is equivalent to the *cohomological equation*

$$[\mathcal{G}_\varepsilon, \mathcal{Z}_\varepsilon] = -\mathcal{R} \quad (\text{or } [\mathcal{Z}_\varepsilon, \mathcal{G}_\varepsilon] = \mathcal{R}),$$

which is now a *linear* equation for $\mathcal{G}_\varepsilon$.

4.1. Cohomological equation for Hamiltonian systems

Assume that we have a symplectic form Ω and let $J(z)$ be the associated matrix defining the vector Hamiltonian vector field (see (4), (5)). We assume that we have an ε-dependent family of Hamiltonians $Z_\varepsilon(z) = Z(\varepsilon, z)$ inducing a ε-dependent family of Hamiltonian vector fields $\mathcal{Z}_\varepsilon$.

We would like to find a family of symplectic transformations (with respect to the form Ω) $z = h_\varepsilon(u)$ such that

$$h_\varepsilon^* Z_\varepsilon = Z_\varepsilon \circ h_\varepsilon = Z_0.$$

We will seek h_ε as the time shift ε along the trajectory of some ε-dependent Hamiltonian H_ε.

Lemma 15. *Assume that $Z(\varepsilon, z)$ is C^1 and $\mathcal{Z}_\varepsilon(z) = J(z)(DZ_\varepsilon(z))^\top$. Assume that $h(\varepsilon, z) \in C^2$ and $H(\varepsilon, z) \in C^2$ are such that $h_0 = \mathrm{Id}$ and*

$$\frac{d}{d\varepsilon} h_\varepsilon = J \cdot (DH_\varepsilon)^\top \circ h_\varepsilon.$$

Then

$$\frac{d}{d\varepsilon}\left(h_\varepsilon^* Z_\varepsilon\right) = h_\varepsilon^* \left(\frac{\partial Z_\varepsilon}{\partial \varepsilon} - DH_\varepsilon \cdot \mathcal{Z}_\varepsilon\right).$$

Proof. We have

$$\begin{aligned}
\frac{d}{d\varepsilon} Z_\varepsilon(h_\varepsilon(u)) &= \frac{\partial Z}{\partial \varepsilon}(\varepsilon, h_\varepsilon(u)) + DZ_\varepsilon(h_\varepsilon(u)) \cdot \frac{d}{d\varepsilon} h_\varepsilon(u) \\
&= \frac{\partial Z_\varepsilon}{\partial \varepsilon}(h_\varepsilon(u)) + DZ_\varepsilon(h_\varepsilon(u)) \cdot \left(J(h_\varepsilon(u))(DH_\varepsilon(h_\varepsilon(u)))^\top\right) \\
&= \frac{\partial Z_\varepsilon}{\partial \varepsilon}(h_\varepsilon(u)) - \left(J(h_\varepsilon(u))(DZ_\varepsilon(h_\varepsilon(u)))^\top\right)^\top \cdot (DH_\varepsilon(h_\varepsilon(u)))^\top \\
&= \frac{\partial Z_\varepsilon}{\partial \varepsilon}(h_\varepsilon(u)) - (\mathcal{Z}_\varepsilon(h_\varepsilon(u)))^\top \cdot (DH_\varepsilon(h_\varepsilon(u)))^\top \\
&= \frac{\partial Z_\varepsilon}{\partial \varepsilon}(h_\varepsilon(u)) - DH_\varepsilon(h_\varepsilon(u)) \cdot \mathcal{Z}_\varepsilon(h_\varepsilon(u)). \quad \square
\end{aligned}$$

Therefore the cohomological equation to be solved in the Hamiltonian context is

$$DH_\varepsilon \cdot \mathcal{Z}_\varepsilon - \frac{\partial Z_\varepsilon}{\partial \varepsilon} = 0. \tag{32}$$

Observe that equation (32) does not depend on the particular formula for Ω. The only way Ω factors in (32) is the relation between Z and $\mathcal{Z}$.

Remark 16. In fact in the proof of the above lemma we obtained

$$\frac{d}{d\varepsilon} Z_\varepsilon(h_\varepsilon(u)) = \left(\frac{dZ_\varepsilon}{d\varepsilon} + \{Z_\varepsilon, H_\varepsilon\}\right)(h_\varepsilon(u)),$$

where the *Poison bracket* is given by

$$\{Z_\varepsilon, H_\varepsilon\} = \Omega(I(dZ_\varepsilon), I(dH_\varepsilon)),$$

where $\mathcal{Z}_\varepsilon = I(dZ_\varepsilon) = J \cdot (dZ_\varepsilon)^\top$ and $\mathcal{H}_\varepsilon = I(dH_\varepsilon) = J \cdot (dH_\varepsilon)^\top$ are defined by (4).

5. Strategy for solving the cohomological equation

In this section we represent the points of our phase space by $z = (x, y, c)$, where c are the center variables and (x, y) the saddle variables, x unstable and y stable.

We set (compare to (19)–(20))

$$\mathcal{Z}_\varepsilon(z) = \big(A(x, y) + \mathcal{N}_{x,y}(z) + \varepsilon\mathcal{R}(z), B(c) + \mathcal{N}_c(z)\big). \tag{33}$$

From equation (30) in Lemma 14 (and (2)) it follows that to bring (33) to the normal form (i.e. to remove $\mathcal{R}$) we need first to solve for $\mathcal{G}_\varepsilon$ the following *cohomological equation*

$$D\mathcal{G}_\varepsilon \cdot \mathcal{Z}_\varepsilon = D\mathcal{Z}_\varepsilon \cdot \mathcal{G}_\varepsilon - \mathcal{R}, \tag{34}$$

where with some abuse of notation we redefined $\mathcal{R}$ by setting $\mathcal{R}(z) = (\mathcal{R}(z), 0)$.

In the symplectic setting we assume that the Hamiltonian function Z_ε is of the following form

$$Z_\varepsilon(z) = A(z) + N(z) + \varepsilon R(z), \tag{35}$$

where (compare (23)) $A(z) = H_{2,c}(c) + H_2(x, y)$, $N(z) = N_c(c) + N_z(x, y, c) + R_1(x, y, c)$ and $R(z) = R_2(x, y, c)$.

The cohomological equation for a Hamiltonian G_ε will be (see (32))

$$DG_\varepsilon \cdot \mathcal{Z}_\varepsilon = R. \tag{36}$$

In both equations (34) and (36) the variable ε plays a role of a parameter. In fact we have a family of equations parameterized by ε, but we want to construct the solution which will depend smoothly on ε. Observe that the solution of (34) (or (36)) cannot be unique, since we can always add $f(\varepsilon)\mathcal{Z}_\varepsilon$ (or $f(\varepsilon)Z_\varepsilon$) for any function $f(\varepsilon)$ to get a new solution.

In the case of general vector field the strategy to construct the solution of the cohomological equation is as follows (for Hamiltonian system the procedure is analogous)

- Preparation of "compact data"; after this step we will have a modified vector field $\tilde{\mathcal{Z}}_\varepsilon$ and $\tilde{\mathcal{R}}$, coinciding with the original one in some neighborhood of the origin, but with a linear vector field $\tilde{\mathcal{Z}}_\varepsilon$ far from the origin and a remainder $\tilde{\mathcal{R}}$ with compact support, see section 5.1. After this step, we will drop the tildes in the modified equations.
- Straightening of the local center-stable and center-unstable manifolds via a transformation T_ε (see Section 5.2). We obtain a new vector field $\mathcal{Z}_\varepsilon = T_{\varepsilon*}\mathcal{Z}_\varepsilon$ and a new remainder $\mathcal{R}_\varepsilon = T_{\varepsilon*}\mathcal{R}$.
- Splitting of the remainder term, $T_*\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$, where $\mathcal{R}_1 = O_{\ell_1+1}(|x|^{\ell_1+1})$ and $\mathcal{R}_2 = O_{\ell_2+1}(|y|^{\ell_2+1})$ (see Sec. 5.3).
- Resolution of (34) (or (36)) with remainders $\mathcal{R}_1$ and $\mathcal{R}_2$ (see Section 6.2), obtaining $\mathcal{G}_1(z)$ and $\mathcal{G}_2(z)$, respectively.

- Finally, the solution is given by $\mathcal{G} = \left(T^{-1}\right)_* (\mathcal{G}_1 + \mathcal{G}_2)$.

The whole process of construction of solution is smooth in ε and other parameters if they are present in $\mathcal{R}$ and $\mathcal{Z}_\varepsilon$. This is important, because the solution of the cohomological equation defines a vector field (depending on ε), for which the time shift from $\varepsilon = 0$ to $\epsilon = 1$ along a trajectory defines the desired transformation (with ε being the time variable).

5.1. Preparation of compact data

Consider system (19)–(20) obtained at the end of the first part of the proof of Theorem 3 in Section 3. We intend to modify the vector field, away from the origin so that all nonlinearities will have a compact support.

Recall (see (11)) that our phase-space is $\mathcal{P} = \left\{(z \in \mathbb{R}^{k_A} \times \mathbb{C}^{l_A}, c \in \mathbb{C}^m)\right\}$. We have the splitting $z = (x, y)$, where $x \in \mathbb{R}^{n_u}$ and $y \in \mathbb{R}^{n_s}$ with $k_A + 2l_A = n_u + n_s$.

Let $\eta : \mathcal{P} \to [0, 1]$ be a C^∞ function such that for some $0 < r_0 < r_1$ it holds that

$$
\begin{aligned}
\eta(z) = 1, &\quad \|z\| \le r_0,\\
\eta(z) = 0, &\quad \|z\| \ge r_1.
\end{aligned}
$$

Let the constants K_η, K_N be such that

$$
\begin{aligned}
\|D\eta(z)\| \le K_\eta, \quad \|D^2\eta(z)\| \le K_\eta \quad \forall z \in \mathcal{P}\\
\|N(z)\| \le K_N \|z\|^2, \quad \text{if} \quad \|z\| \le r_1,\\
\|DN(z)\| \le K_N \|z\|, \quad \text{if} \quad \|z\| \le r_1,\\
\|D^2N(z)\| \le K_N, \quad \text{if} \quad \|z\| \le r_1.
\end{aligned}
$$

Let $\sigma > 0$ be some small number. Let us set

$$
\widetilde{N}(x, y, c) = N(x, y, c)\eta((x, y, c)/\sigma), \quad \widetilde{\mathcal{R}}(x, y, c) = \mathcal{R}(x, y, c)\eta((x, y, c)/\sigma).
$$

The modified system is

$$
\begin{aligned}
\dot{x} &= A_u x + \widetilde{N}_x(x, y, c) + \varepsilon\widetilde{\mathcal{R}}_x(x, y, c) = A_u x + \widetilde{M}_x(\varepsilon, x, y, c), &(37)\\
\dot{y} &= A_s y + \widetilde{N}_y(x, y, c) + \varepsilon\widetilde{\mathcal{R}}_y(x, y, c) = A_s y + \widetilde{M}_y(\varepsilon, x, y, c),\\
\dot{c} &= Bc + \widetilde{N}_c(x, y, c). &(38)
\end{aligned}
$$

Let $\widetilde{M}(\varepsilon, x, y, c) = (\widetilde{M}_x(\varepsilon, x, y, c), \widetilde{M}_y(\varepsilon, x, y, c), 0)$. Obviously, we have (see (21)) (where all the $O(\cdot)$ terms are for their arguments converging to 0)

$$
\begin{aligned}
&\widetilde{M}(\varepsilon, 0, 0, c) = 0, \quad \widetilde{N}_{x,y}(0, 0, c) = 0, &(39)\\
&\widetilde{M}(\varepsilon, x, y, c) = O_2(|(x, y, c)|^2), \quad \widetilde{N}(x, y, c) = O_2(|(x, y, c)|^2),\\
&\widetilde{\mathcal{R}}(x, y, c) = O_{Q+1}(|(x, y)|^{Q+1}).
\end{aligned}
$$

Note that we have defined $\widetilde{M}$ as a notation for the nonlinearities appearing in the saddle directions. Indeed, $\widetilde{M}$ also depends on ε, but since we have bounds that are uniform for $\varepsilon \in [0,1]$, we will usually omit this dependence on ε in the sequel.

We have for $\sigma \leq 1$ and any z (for suitable constant K'_N)

$$\|\widetilde{N}(z)\| \leq \sup_{\|z\|\leq \sigma r_1} \|N(z)\| \leq K_N \sigma^2 r_1^2 \leq K'_N \sigma^2, \tag{40}$$

$$\begin{aligned}\|D\widetilde{N}(z)\| &\leq \sup_{\|z\|\leq \sigma r_1} \left(\frac{1}{\sigma} \|D\eta(z/\sigma)\| \cdot \|N(z)\| + |\eta(z/\sigma)| \cdot \|DN(z)\| \right) \\ &\leq \sigma^{-1} K_\eta K_N \sigma^2 r_1^2 + K_N \sigma r_1 = \sigma (K_\eta K_N r_1^2 + K_N r_1) \leq K'_N \sigma, \end{aligned} \tag{41}$$

$$\begin{aligned}\|D^2\widetilde{N}(z)\| \leq \sup_{\|z\|\leq \sigma r_1} &\left(\frac{1}{\sigma^2} \|D^2\eta(z/\sigma)\| \cdot \|N(z)\| + \frac{2}{\sigma} \|D\eta(z/\sigma)\| \cdot \|DN(z)\| \right. \\ &\left. + \eta(z/\sigma) \|D^2 N(z)\| \right) \leq K_\eta K_N r_1^2 + 2K_\eta K_N r_1 + K_N = K'_N.\end{aligned}$$

Analogously, for some constant $K_{\mathcal{R}}$, we obtain that for any $w = (x, y, c)$

$$\begin{aligned}\|\widetilde{\mathcal{R}}(w)\| &\leq K_{\mathcal{R}} \sigma^{Q+1}, \\ \|D\widetilde{\mathcal{R}}(w)\| &\leq K_{\mathcal{R}} \sigma^{Q}, \\ \|D^2\widetilde{\mathcal{R}}(w)\| &\leq K_{\mathcal{R}} \sigma^{Q-1}. \end{aligned} \tag{42}$$

For the Hamiltonian system we do the following modification of the Hamilton function Z (compare (35)):

$$\widetilde{Z}(z) = A(z) + N(z)\eta(z/\sigma) + \varepsilon R(z)\eta(z/\sigma) = A(z) + \tilde{N}(z) + \varepsilon \tilde{R}(z). \tag{43}$$

It is easy to see that the induced differential equation can be written as (37)–(38) satisfying conditions (39)–(42) for the nonlinear terms.

Let us denote

$$W^c = \{z = 0\}.$$

With some abuse of notation we will also treat W^c as the $\mathbb{C}^r$ component of $\mathcal{P}$, for example in a lemma below we write $[0,1] \times \mathbb{R}^{n_u} \times W^c$ as a domain of some function or in $D(\delta) = \overline{B}_{n_u}(\delta) \times \overline{B}_{n_s}(\delta) \times W^c$ to represent some neighborhood of W^c.

Lemma 17. *Let us consider the system (37)–(38) and denote by $\varphi(t, x, y, c)$ the induced flow. Let $L < 1$.*

Then there exists $\sigma_0 > 0$ so that, for any $0 < \sigma \leq \sigma_0$, it holds that

- *There exist functions $y^u : [0,1] \times \mathbb{R}^{n_u} \times W^c \to \mathbb{R}^{n_s}$ and $x^s : [0,1] \times \mathbb{R}^{n_s} \times W^c \to \mathbb{R}^{n_u}$ in C^{q-1}, such that for any $\varepsilon \in [0,1]$*

$$W^{cu} = \{(x, y^u(\varepsilon, x, c), c), \quad (x, c) \in \mathbb{R}^{n_u} \times W^c\},$$
$$W^{cs} = \{(x^s(\varepsilon, y, c), y, c), \quad (y, c) \in \mathbb{R}^{n_s} \times W^c\},$$

and for any $\varepsilon_1, \varepsilon_2$, c_1, c_2, x_1, x_2 and y_1, y_2, it holds that

$$\|y^u(\varepsilon_1, x_1, c_1) - y^u(\varepsilon_2, x_2, c_2)\| \leq L\left(|\varepsilon_1 - \varepsilon_2| + \|c_1 - c_2\| + \|x_1 - x_2\|\right), \tag{44}$$
$$\|x^s(\varepsilon_1, y_1, c_1) - x^s(\varepsilon_2, y_2, c_2)\| \leq L\left(|\varepsilon_1 - \varepsilon_2| + \|c_1 - c_2\| + \|y_1 - y_2\|\right). \tag{45}$$

- *There exists a constant $K_W = K_W(L, \sigma)$, such that for all $j = 1, \ldots, q-1$,*

$$\|D^j x^s\| \leq K_W, \quad \|D^j y^u\| \leq K_W. \tag{46}$$

  This function $K(L, \sigma)$ is non-decreasing with respect to L and σ.
- *For any $\delta > 0$, introduce the set $D(\delta) = \overline{B}_{n_u}(\delta) \times \overline{B}_{n_s}(\delta) \times W^c$. Then if $(x, y, c) \notin W^{cs}$, then there exists $t_0 \geq 0$ such that $\varphi(t, x, y, c) \notin D(\delta)$ for $t \geq t_0$ and $\varepsilon \in [0, 1]$; and if $(x, y, c) \notin W^{cu}$, then there exists $t_0 \leq 0$ such that $\varphi(t, x, y, c) \notin D(\delta)$ for $t \leq t_0$ and $\varepsilon \in [0, 1]$.*

Proof. We will begin by proving that for any $\delta > 0$ the set $D(\delta) = \overline{B}_{n_u}(\delta) \times \overline{B}_{n_s}(\delta) \times W^c$ is an *isolating block* (see Definition 16 in Appendix B.2) for this system for any $\varepsilon \in [0, 1]$. To prove this, we have to check exit and entry conditions, i.e. (157) and (158), from the Appendix.

We have for $(x, y, c) \in D(\delta)$ (we use (154), (155) and (41), (42)) that

$$(A_u x, x) \geq m_l(A_u)\|x\|^2,$$
$$(A_s y, y) \leq \mu_{log}(A_s)\|y\|^2,$$
$$\|\widetilde{N}(x, y, c)\| = \|\widetilde{N}(x, y, c) - \widetilde{N}(c, 0, 0)\| \leq \|D\widetilde{N}\|\sqrt{2}\delta \leq \sqrt{2}K'_N\sigma\delta,$$
$$\|\widetilde{\mathcal{R}}(x, y, c)\| = \|\widetilde{\mathcal{R}}(x, y, c) - \widetilde{\mathcal{R}}(c, 0, 0)\| \leq \|D\widetilde{\mathcal{R}}\|\sqrt{2}\delta \leq \sqrt{2}K_{\mathcal{R}}\sigma^Q\delta.$$

Hence, if $(x, y, c) \in D(\delta)^-$ (i.e. $\|x\| = \delta$)

$$(\dot{x}, x) = (A_u x, x) + (\widetilde{N}(x, y, c), x) + \varepsilon(\widetilde{\mathcal{R}}(x, y, c), x),$$
$$\geq \delta^2\left(m_l(A_u) - \sqrt{2}K'_N\sigma - \sqrt{2}K_{\mathcal{R}}\sigma^Q\right),$$

and for $(x, y, c) \in D(\delta)^+$ (i.e. $\|y\| = \delta$)

$$(\dot{y}, y) = (A_s y, y) + (\widetilde{N}(x, y, c), y) + \varepsilon(\widetilde{\mathcal{R}}(x, y, c), y)$$
$$\leq -\delta^2\left(-\mu_{log}(A_s) - \sqrt{2}K'_N\sigma - \sqrt{2}K_{\mathcal{R}}\sigma^Q\right).$$

Hence if σ is such that

$$m_l(A_u) > \sqrt{2}K'_N\sigma + \sqrt{2}K_{\mathcal{R}}\sigma^Q,$$
$$-\mu_{log}(A_s) > \sqrt{2}K'_N\sigma + \sqrt{2}K_{\mathcal{R}}\sigma^Q,$$

then conditions (157) and (158) are satisfied and therefore $D(\delta)$ is an isolating block. In view of (22), this holds if σ is small enough.

To deal with W^{cu} and W^{cs} and their dependence on ε we add to system (37)–(38) the equation

$$\dot{\varepsilon} = 0. \tag{47}$$

This allows us to obtain the dependence with respect to the parameter ε from the general arguments of [6,7] on the existence of a normally hyperbolic invariant manifold (NHIM) and its center-unstable and center-stable manifolds, which are recalled in the Appendix B. Note that the loss of one unit in the degree of differentiability of W^{cu} and W^{cs} comes in general situation from the loss of differentiability of the center manifold W^c.

Now we check the rate conditions (see Definition 18 and Theorem 52 in Appendix B.2) for system (37)–(38), (47) up to order $q-1$. To be in agreement with the notation used in Theorem 52 we set $\Lambda = [0, \varepsilon] \times W^c$. Hence the center direction denoted there by λ is now (ε, c).

On $\tilde{D}(\delta) = [0, \varepsilon] \times D(\delta)$ we have

$$\begin{aligned}
\overrightarrow{\mu_{s,1}} &= \mu_{log}(A_s) + O(\sigma) + \frac{1}{L}O(\sigma), & \overrightarrow{\mu_{s,2}} &= \mu_{log}(A_s) + O(\sigma) + LO(\sigma),\\
\overrightarrow{\xi_{u,1}} &= m_l(A_u) - O(\sigma) - \frac{1}{L}O(\sigma), & \overrightarrow{\xi_{u,1,P}} &= m_l(A_u) - O(\sigma) - \frac{1}{L}O(\sigma)\\
\overrightarrow{\mu_{cs,1}} &= \mu_{log}(B) + O(\sigma) + LO(\sigma), & \overrightarrow{\mu_{cs,2}} &= \mu_{log}(B) + O(\sigma) + \frac{1}{L}O(\sigma)\\
\overrightarrow{\xi_{cu,1}} &= -O(\sigma) - LO(\sigma), & \overrightarrow{\xi_{cu,1,P}} &= -O(\sigma) - LO(\sigma),\\
\overrightarrow{\xi_{u,2}} &= m_l(A_u) - O(\sigma) - LO(\sigma), & \overrightarrow{\xi_{cu,2}} &= -O(\sigma) - \frac{1}{L}O(\sigma)
\end{aligned}$$

It is easy to see that for any k and L, there exists $\sigma_0 = \sigma_0(k, L)$, such that for $\sigma \le \sigma_0$ the rate conditions of order k are satisfied (we need $k = q - 1$).

Therefore y^u and x^s, which in the notation of Theorem 52 are w^{cu} and w^{cs}, respectively, satisfy the Lipschitz condition with constant L. This establishes inequalities (44), (45).

Since all partial derivatives up to the order of the regularity class of our modified vector field are globally bounded, we obtain condition (46) of Theorem 52.

The above arguments apply to local center-stable and center-unstable manifolds inside $D(\delta)$. But since δ is arbitrary and the bounds on derivatives of our vector field are global, then all the above estimates also apply to the global manifolds W^{cs} and W^{cu}. □

Lemma 18. *There exists $\sigma_0 > 0$ so that, for any $0 < \sigma \le \sigma_0$, the following holds:*

There exist constants E and C, such that for any $R > \delta > 0$ and for any $\varepsilon \in [0, 1]$ $z = (x, y, c) \in D(R)$, holds

- *if $\|x\| \ge \|y\|$, $(x, y) \ne 0$ and $\varphi(t, x, y, c) \in D(R)$ for $t \in [0, T]$, then $\|\pi_x \varphi(t, x, y, c)\| > \|\pi_y \varphi(t, x, y, c)\|$ for $t \in (0, T]$ and*

$$\frac{d}{dt}\|x(t)\| \ge E\|x(t)\|, \quad t \in (0, T]. \tag{48}$$

In particular, $T \leq \frac{1}{E} \ln\left(\frac{R}{\|x(0)\|}\right)$.

- *if* $\|x(t)\| \leq \|y(t)\|$, *and* $\varphi(t,x,y,c) \in D(R)$ *for* $t \in [0,T]$, *then* $\|y(t)\| \leq \|y(0)\| e^{-tC}$.
  In particular, if $t > \frac{1}{E} \ln\left(\frac{\|y(0)\|}{\delta}\right)$, *then* $\varphi(t,x,y,c) \in D(\delta)$.

Proof. To prove the first assertion we need to establish two facts. First, the forward invariance of the cone $\|x\| \geq \|y\|$ and second, the expansion in this cone. We have for $\|x\| \geq \|y\|$ (we use (154), (155)) (compare to the proof of Lemma 17).

We have

$$\begin{aligned}\frac{1}{2}\frac{d\|x\|^2}{dt} &= (\dot{x}, x) = (A_u x, x) + \left(\widetilde{N}(x,y,c)\varepsilon\widetilde{\mathcal{R}}(x,y,c), x\right) \\ &\geq m_l(A_u)\|x\|^2 - \left(\|D\widetilde{N} + \|D\widetilde{\mathcal{R}}\|\right) \cdot \|(x,y)\| \cdot \|x\|\end{aligned}$$

and

$$\begin{aligned}\frac{1}{2}\frac{d\|y\|^2}{dt} &= (\dot{y}, y) = (A_s y, y) + \left(\widetilde{N}(x,y,c) + \varepsilon\widetilde{\mathcal{R}}(x,y,c), y\right) \\ &\leq \mu_{log}(A_s)\|y\|^2 + \left(\|D\widetilde{N} + \|D\widetilde{\mathcal{R}}\|\right) \cdot \|(x,y)\| \cdot \|x\|.\end{aligned}$$

Therefore for (x,y,c) such that $\|x\| \geq \|y\|$, $x \neq 0$ (we use (41), (42)) holds

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left(\|x\|^2 - \|y\|^2\right) &\geq m_l(A_u)\|x\|^2 - \mu_{log}(A_s)\|y\|^2 - 2\left(\|D\widetilde{N} + \|D\widetilde{\mathcal{R}}\|\right) \cdot \|(x,y)\| \cdot \|x\| \\ &\geq \left(m_l(A_u) - \mu_{log}(A_s) + 2\sqrt{2}\left(K'_N\sigma + K_{\mathcal{R}}\sigma^Q\right)\right)\|x\|^2 > 0\end{aligned}$$

if $\sigma \leq \sigma_0$. Therefore the cone $\|x\| \geq \|y\|$ is forward invariant.

Now we deal with the expansion in the x-direction. From previous computations we obtain in the cone $\|x\| \geq \|y\|$

$$\frac{1}{2}\frac{d\|x\|^2}{dt} \geq m_l(A_u)\|x\|^2 - \sqrt{2}(K'_N\sigma + K_{\mathcal{R}}\sigma^Q)\|x\|^2 \geq E\|x\|^2$$

where $E = m_l(A_u) - \sqrt{2}(K'_N\sigma + K_{\mathcal{R}}\sigma_0^Q) > 0$. We chose σ_0, small enough for E to be positive. Observe that this implies (48).

Now we turn to second assertion about the decay in the cone $\|x\| \leq \|y\|$. In this cone from previous computations we obtain

$$\frac{1}{2}\frac{d\|y\|^2}{dt} \leq \left(\mu_{log}(A_s) + \sqrt{2}K'_N\sigma + \sqrt{2}K_{\mathcal{R}}\sigma^Q\right)\|y\|^2.$$

Since $\mu_{log}(A_s) < 0$ we can find σ_0, such that $C = -\mu_{log}(A_s) - \sqrt{2}\left(K'_N\sigma + \sqrt{2}K_{\mathcal{R}}\right) > 0$. This implies the second assertion. □

We fix $0 < \sigma \leq \sigma_0$ so that the assertions of Lemmas 17 and 18 hold true and from now on we work with the modified vector field, but dropping tildes.

We will need some extra bounds for the functions y^{u} and x^{s} as well as for their derivatives up to order two. It is easy to see that if B does not contain any nontrivial Jordan blocks, then $y^{\mathrm{u}}(x,c)=0$ and $x^{\mathrm{s}}(y,c)=0$ if $|c|>O(\sigma r_1)$. This is an immediate consequence of the following fact: as $|\pi_c\varphi(t,x,y,c)=e^{Bt}c|>O(\sigma r_1)$ for $t\in\mathbb{R}$, the trajectory remains in the domain where the dynamics is just linear; $x'=A_u x$, $y'=A_s y$, $c'=Bc$.

However, we want to include the situation with nontrivial Jordan blocks in the center direction and we also need information about the derivatives of y^{u} and x^{s} with respect to x and y, respectively, when $|x|\to\infty$ and $|y|\to\infty$.

Lemma 19. *In the context of Lemma 17,*

$$|y^u(x,c)|\le L\sigma r_1\min\left(1,\left(\frac{|x|}{\sigma r_1}\right)^{\frac{\mu_{log}(A_s)}{\mu_{log}(A_u)}}\right),\quad \forall(x,c), \tag{49}$$

$$|x^s(y,c)|\le L\sigma r_1\min\left(1,\left(\frac{|y|}{\sigma r_1}\right)^{\frac{\mu_{log}(-A_u)}{\mu_{log}(-A_s)}}\right),\quad \forall(y,c).$$

Moreover, for some $w_c>1$ and $w_z>0$ and $w=\min(w_c,w_z)$,

$$\left|\frac{\partial y^u}{\partial x}(x,c)\right|\le\min(L,LO(\sigma^w)(|c|^{w_c}+|x|^{w_z})^{-1}), \tag{50}$$

$$\left|\frac{\partial x^s}{\partial y}(y,c)\right|\le\min(L,LO(\sigma^w)(|c|^{w_c}+|y|^{w_z})^{-1}),$$

and

$$\left|\frac{\partial^2 y^u}{\partial^2 x}(x,c)\right|\le\min(M,O(\sigma^w)(|c|^{w_c}+|x|^{w_z})^{-1}))$$

$$\left|\frac{\partial^2 x^s}{\partial^2 y}(y,c)\right|\le\min(M,O(\sigma^w)(|c|^{w_c}+|y|^{w_z})^{-1})).$$

Proof. It is enough to consider just y^{u}, since the argument for x^{s} is analogous, just reversing the time direction and changing x to y.

Let us denote by S the support of the nonlinearities in system (37)–(38). We know that

$$S\subset\overline{B}_{n_u}(0,\sigma r_1)\times\overline{B}_{n_s}(0,\sigma r_1)\times\overline{B}_{2m}(0,\sigma r_1).$$

Let $(x,y^{\mathrm{u}}(x,c),c)\in W^{cu}$ be such that in the x-direction we have $|x|=\sigma r_1$. Then $|y^{\mathrm{u}}(x,c)|\le L\sigma r_1$ and its image forward in time does not enter S, hence is equal to $(e^{tA_u}x,e^{tA_s}y^{\mathrm{u}}(x,c),e^{tB}c)\in W^{cu}$, which means that

$$e^{tA_s}y^{\mathrm{u}}(x,c)=y^{\mathrm{u}}(e^{tA_u}x,e^{tB}c).$$

Since the graph transform (used in proving the existence of W^{cu}) of W^{cu} is equal to W^{cu}, therefore for $(\bar{x},\bar{c})$ with $|\bar{x}|>\sigma r_1$, there exists (x,c) such that

$$(e^{tA_u}x, e^{tA_s}y^{\mathrm{u},e^{tB}c,}(x,c)) = (\bar{x}, y^{\mathrm{u}}(\bar{x},\bar{c})),$$

and therefore we get (because $\|e^{A_s t}\| \le e^{\mu_{log}(A_s)t} < 1$ for $t > 0$)

$$|y^{\mathrm{u}}(x,c)| \le L\sigma r_1, \quad \forall (x,c).$$

On the other hand, for $|x| > \sigma r_1$ the time to reach $|x|$ from σr_1 is estimated from below by $\frac{1}{\mu_{log}(A_u)} \ln\left(\frac{|x|}{\sigma r_1}\right)$, so we get

$$|y^{\mathrm{u}}(x,c)| \le L\sigma r_1 \left(\frac{|x|}{\sigma r_1}\right)^{\frac{\mu_{log}(A_s)}{\mu_{log}(A_u)}}.$$

Now we proceed to find the estimates for the derivatives of y^{u}. It follows immediately from (44) in Lemma 17 that

$$\left|\frac{\partial y^{\mathrm{u}}}{\partial c}(x,c)\right|, \left|\frac{\partial y^{\mathrm{u}}}{\partial x}(x,c)\right| \le L, \quad \forall (x,c).$$

Consider a point $p = (x, y^{\mathrm{u}}(x,c), c) \in W^{cu}$, $p \notin S$. Let $t(p) > 0$ be such that for $t \in [0, t(p)]$ $\varphi(-t,p) \notin S$. Then $\varphi(-t,p) = (e^{-tA_u}x, e^{-tA_s}y^{\mathrm{u}}(x,c), e^{-tB}c) \in W^{cu}$, therefore

$$e^{-t(p)A_s}y^{\mathrm{u}}(c,x) = y^{\mathrm{u}}(e^{-t(p)A_u}x, e^{-t(p)B}c),$$

and finally

$$y^{\mathrm{u}}(x,c) = e^{t(p)A_s}y^{\mathrm{u}}(e^{-t(p)A_u}x, e^{-t(p)B}c). \tag{51}$$

Observe that $t(p)$ is also good for all points in the neighborhood of p, so when differentiating the above equation we can treat $t(p)$ as a constant, if needed.

Depending on whether $|c|$ or $|x|$ are larger than σr_1 we can take the following expressions $t_x(p)$ or $t_c(p)$ for $t(p)$

$$t_x(p) = \frac{1}{\mu_{log}(A_u)} \ln\left(\frac{|x|}{\sigma r_1}\right), \quad t_c(p) = \frac{1}{\mu_{log}(B)} \ln\left(\frac{|c|}{\sigma r_1}\right).$$

By differentiation of (51)

$$\frac{\partial y^{\mathrm{u}}}{\partial x}(x,c) = e^{t(p)A_s}\left(\frac{\partial y^{\mathrm{u}}}{\partial c}\left(e^{-t(p)A_u}x, e^{-t(p)B}c\right)\right) \cdot e^{-t(p)A_u},$$

hence

$$\left|\frac{\partial y^{\mathrm{u}}}{\partial x}(x,c)\right| \le Le^{t(p)(\mu_{log}(A_s)+\mu_{log}(-A_u))}.$$

For the second derivatives we obtain the following bounds (we use (46))

$$\left|\frac{\partial^2 y^{\mathrm{u}}}{\partial^2 x}(x,c)\right| \leq K_W e^{t(p)(\mu_{log}(A_s)+2\mu_{log}(-A_u))}.$$

Observe that all exponents w appearing in $e^{t(p)\cdot w}$ are negative (if $\mu_{log}(-B)$ is small enough).

Let us evaluate exponentials of the type $e^{t(p)\cdot w}$ for $t(p) = t_x(p), t_c(p)$. We have

$$e^{t_x(p)w} = (\sigma r_1)^{-\frac{w}{\mu_{log}(A_u)}} |x|^{\frac{w}{\mu_{log}(A_u)}},$$
$$e^{t_c(p)w} = (\sigma r_1)^{-\frac{w}{\mu_{log}(B)}} |c|^{\frac{w}{\mu_{log}(B)}},$$

since for each exponent $w \leq \mu_{log}(A_s) + 2\mu_{log}(-B) < 0$, hence $w_c = -\frac{w}{\mu_{log}(B)} > 1$ provided $\mu_{log}(B)$ is small enough.

Regarding the decay with increasing $|c|$ we see that the first order derivatives will be less than or equal to $LO(\sigma^{w_c})|c|^{-w_c}$, and for the second order we have a bound $O(\sigma^{w_c})|c|^{-w_c}$.

Let us see now what will be the decay in x direction.

- $\frac{\partial y^{\mathrm{u}}}{\partial x}(x,c)$. We have $w = \mu_{log}(A_s) + \mu_{log}(-A_u)$, so $w_z = \frac{-\mu_{log}(A_s)}{\mu_{log}(A_u)} - \frac{\mu_{log}(-A_u)}{\mu_{log}(A_u)}$,
- $\frac{\partial^2 y^{\mathrm{u}}}{\partial x_i x_j}(x,c)$. We have $w = \mu_{log}(A_s) + 2\mu_{log}(-A_u)$, so $w_z = \frac{-\mu_{log}(A_s)}{\mu_{log}(A_u)} - 2\frac{\mu_{log}(-A_u)}{\mu_{log}(A_u)}$.

In both cases it can happen that $w_z < 1$. For example if $\mu_{\min} < \cdots < \mu_{\max}$ are eigenvalues of A_u, then $-\mu_{log} - A_u = \mu_{\min}$, and $w_z = \lambda_{\min}/\mu_{\max} + \mu_{\min}/\mu_{\max}$. □

5.2. *Straightening the invariant manifolds*

We continue with the modified vector field obtained in Section 5.1. The center-unstable and center-stable invariant manifolds can be straightened in suitable coordinates. The following transformation does this

$$T_\varepsilon(x,y,c) = (x - x^{\mathrm{s}}(\varepsilon,y,c), y - y^{\mathrm{u}}(\varepsilon,x,c), c). \tag{52}$$

Lemma 20. *Under the same assumptions as in Lemma 17, $T_\varepsilon(x,y,c)$ is C^{q-1} and for every $\varepsilon \in [0,1]$ the transformation $T_\varepsilon : \mathcal{P} \to \mathcal{P}$ is a diffeomorphism. Moreover,*

$$T_\varepsilon^{-1}(x,y,c) = (x + O(L\sigma), y + O(L\sigma), c). \tag{53}$$

Proof. The regularity follows from Lemma 17. We need to show that T_ε is a bijection. To show that it is onto we fix c and observe that for $\|x\| = R$ and $\|y\| \leq R$ it holds that (we use (45))

$$\|\pi_x T_\varepsilon(x,y,c)\| \geq \|x\| - \|x^{\mathrm{s}}(\varepsilon,y,c)\| \geq R - LR = (1-L)R,$$

and analogously for $\|x\| \leq R$ and $\|y\| = R$, from (44) we get

$$\|\pi_y T_\varepsilon(x,y,c)\| \geq (1-L)R.$$

From this inequality, using the local Brouwer degree argument, we obtain that

$$\overline{B}_{n_u}((1-L)R) \times \overline{B}_{n_s}((1-L)R) \subset \pi_{x,y} T_\varepsilon(\overline{B}_{n_u}(R), \overline{B}_{n_s}(R), c).$$

Therefore T_ε is onto.

Now we prove the injectivity. Take any $(x_1, y_1) \neq (x_2, y_2)$. We have (we use (45), (44))

$$\|\pi_x (T_\varepsilon(x_1, y_1, c) - T_\varepsilon(x_2, y_2, c))\| \geq \|x_1 - x_2\| - L\|y_1 - y_2\|,$$

$$\|\pi_y (T_\varepsilon(x_1, y_1, c) - T_\varepsilon(x_2, y_2, c))\| \geq \|y_1 - y_2\| - L\|x_1 - x_2\|.$$

Since $L < 1$, if $\|x_1 - x_2\| \geq \|y_1 - y_2\|$, then $\|x_1 - x_2\| - L\|y_1 - y_2\| > 0$ and analogously in the other case. Hence T_ε is an injection.

Condition (53) follows immediately from (49) in Lemma 19. □

5.2.1. *Cohomological equation after straightening*

Observe that the transformation T_ε given by (52) gives new variables in terms of the old ones, therefore to transform vector fields and Hamiltonians we use the pushforward $T_{\varepsilon*}$. We will often drop ε in the sequel.

Cohomological equation (34) becomes

$$[T_{\varepsilon*}\mathcal{Z}, T_{\varepsilon*}\mathcal{G}] = -T_{\varepsilon*}\mathcal{R}$$

or in the Hamiltonian case, (36) becomes

$$(D(T_{\varepsilon*}G)) \cdot T_{\varepsilon*}\mathcal{Z} = T_{\varepsilon*}R.$$

Hence the cohomological equation retains its form, but the vector field $\mathcal{Z}$ and the remainder $\mathcal{R}$ (or R) change.

In the Hamiltonian case T might not be not symplectic in the presence of center variables. Therefore the transformed symplectic form $T_*\Omega$ no longer has the standard form. This influences the relationship between the Hamiltonian function and the induced vector field.

5.2.2. *The vector field after straightening*

Lemma 21. *For a general vector field, $T_{\varepsilon*}\mathcal{Z}_\varepsilon$ and $T_{\varepsilon*}\mathcal{R}$ are C^{q-2} as functions of (ε, x, y, c), whereas in the Hamiltonian case $T_{\varepsilon*}Z_\varepsilon$ and $T_{\varepsilon*}R$ are C^{q-1}.*

The vector field $\mathcal{Z}_\varepsilon$ after straightening, i.e. $T_{\varepsilon*}\mathcal{Z}_\varepsilon$, becomes

$$\dot{x} = g_x(\varepsilon, x, y, c)x, \tag{54}$$

$$\dot{y} = g_y(\varepsilon, x, y, c)y, \tag{55}$$

$$\dot{c} = Bc + M_c(\varepsilon, x, y, c), \tag{56}$$

where $g_x, g_y \in C^{q-3}$, M_c has compact support and

$$M_c(\varepsilon, x, y, c) = O_2(|(x, y, c)|^2), \tag{57}$$

$$(g_x, g_y)(\varepsilon, x, y, c) = (A_u, A_s) + O(L) + O(\sigma) + (O(L)y, O(L)x). \tag{58}$$

All $O()$ terms are uniform with respect to $\varepsilon \in [0,1]$ in (57) and with respect to $(\varepsilon, x, y, c) \in [0,1] \times \mathcal{P}$.

Moreover,

$$T_{\varepsilon *}\mathcal{R}(x, y, c) = O_{Q+1}(|(x, y)|^{Q+1}), \tag{59}$$

or $T_{\varepsilon *}R(x, y, c) = O_{Q+1}(|(x, y)|^{Q+1})$ in the Hamiltonian case.

Proof. Recall that in Theorem 3 we assumed that

$$q \geq Q + 3. \tag{60}$$

The regularity of $T_{\varepsilon *}\mathcal{Z}_\varepsilon$ and $T_{\varepsilon *}\mathcal{R}$ is an immediate consequence of Lemma 20.

The factorization of x in (54) and y in (55) is due to the fact that in the new coordinates $W^{cs} = \{x = 0\}$ and $W^{cu} = \{y = 0\}$, respectively. For this we need that the transformed vector field is at least of class C^1, so we need $q - 2 \geq 1$, which is granted by (60).

The straightening implies that in the new variables we obtain the equations

$$\begin{aligned}
\dot{x} &= f_x(\varepsilon, x, y, c) = \left(\int_0^1 \frac{\partial f_x}{\partial x}(\varepsilon, c, tx, y) dt \right) \cdot x, \\
\dot{y} &= f_y(\varepsilon, x, y, c) = \left(\int_0^1 \frac{\partial f_y}{\partial y}(\varepsilon, c, x, ty) dt \right) \cdot y, \\
\dot{c} &= f_c(\varepsilon, x, y, c),
\end{aligned}$$

where

$$(f_x, f_y, f_c)(\varepsilon, x, y, c) = DT_\varepsilon(T_\varepsilon^{-1}(x, y, c))\mathcal{Z}_\varepsilon(T_\varepsilon^{-1}(x, y, c)).$$

In the sequel we drop ε because we have bounds which are valid for all ε.

Our first goal is the computation of $\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial c}(T^{-1}(x, y, c))\right)$, $\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial x}(T^{-1}(x, y, c))\right)$ as well as of $\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x, y, c))\right)$.

From (44), (45) in Lemma 17 it follows that

$$\left\| \frac{\partial x^{\mathrm{s}}}{\partial y} \right\|, \left\| \frac{\partial x^{\mathrm{s}}}{\partial c} \right\|, \left\| \frac{\partial y^{\mathrm{u}}}{\partial x} \right\|, \left\| \frac{\partial y^{\mathrm{u}}}{\partial c} \right\| \leq L. \tag{61}$$

Recall that L is a parameter, which can be chosen arbitrarily small.

From (61) and (52) it follows that for any (x, y, c) it holds that

$$DT = \begin{bmatrix} I_x, & O(L), & O(L) \\ O(L) & I_y, & O(L) \\ 0, & 0, & I_c \end{bmatrix}, \qquad DT^{-1} = \begin{bmatrix} I_x + O(L), & O(L), & O(L) \\ O(L) & I_y + O(L), & O(L) \\ 0, & 0, & I_c \end{bmatrix}, \tag{62}$$

where by I_c, I_x, I_y we denote the identity matrix acting on c, x, y variables, respectively.

From Lemma 19 we can obtain better bounds for $\frac{\partial T_y^{-1}}{\partial x}(x,y,c)$ (and $\frac{\partial T_x^{-1}}{\partial y}(x,y,c)$). Let us denote, for $L, \sigma, c_1, z_1 \in \mathbb{R}$,

$$E(L,\sigma,c_1,z_1) = \min(L, L\sigma^{w_z}(c_1^{w_c} + z_1^{w_z})^{-1}),$$

where w_z and w_c are as Lemma 19. Then from (50) in Lemma 19 we have, using the formula $(I-N)^{-1} = I + N + N^2 + \cdots$ with $\frac{\partial T_{x,y}}{\partial(x,y)} = I_{x,y} + N$ and $N = \begin{bmatrix} 0 & \frac{\partial x^s}{\partial y} \\ \frac{\partial y^u}{\partial x} & 0 \end{bmatrix}$,

$$\left|\frac{\partial T_y^{-1}}{\partial x}(x,y,c)\right| = O(E(L,\sigma,|c|,|x|)). \tag{63}$$

We compute $\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial c}(T^{-1}(x,y,c))\right)$. From (52) we see that $\frac{\partial T_c^{-1}}{\partial x} = 0$ and $\frac{\partial^2 T_x}{\partial c \partial x} = 0$, hence from (46), (62) and (63) we obtain

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial c}(T^{-1}(x,y,c))\right) &= \frac{\partial^2 T_x}{\partial c \partial c}(T^{-1}(x,y,c)) \cdot \frac{\partial T_c^{-1}}{\partial x}(x,y,c) \\
&\quad + \frac{\partial^2 T_x}{\partial c \partial x}(T^{-1}(x,y,c)) \cdot \frac{\partial T_x^{-1}}{\partial x}(x,y,c) \\
&\quad + \frac{\partial^2 T_x}{\partial c \partial y}(T^{-1}(x,y,c)) \cdot \frac{\partial T_y^{-1}}{\partial x}(x,y,c) \\
&= -\frac{\partial^2 x^s}{\partial c \partial y}(T_{y,c}^{-1}(x,y,c)) \cdot \frac{\partial T_y^{-1}}{\partial x}(x,y,c) \\
&= K_W O(E(L,\sigma,|c|,|x|)) = O(E(L,\sigma,|c|,|x|)).
\end{aligned}$$

Using (46) and (63) we obtain

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\right) &= \frac{\partial^2 T_x}{\partial y \partial c}(T^{-1}(x,y,c)) \cdot \frac{\partial T_c^{-1}}{\partial x}(x,y,c)(=0) \\
+ \frac{\partial^2 T_x}{\partial y \partial x}(T^{-1}(x,y,c))(=0) \cdot \frac{\partial T_x^{-1}}{\partial x}(x,y,c) &+ \frac{\partial^2 T_x}{\partial y \partial y}(T^{-1}(x,y,c)) \cdot \frac{\partial T_y^{-1}}{\partial x}(x,y,c) \\
= -\frac{\partial^2 x^s}{\partial y \partial y}(T_{y,c}^{-1}(x,y,c)) \cdot \frac{\partial T_y^{-1}}{\partial x}(x,y,c) &= K_W O(E(L,\sigma,|c|,|x|)) = O(E(L,\sigma,|c|,|x|)).
\end{aligned}$$

Summarizing, we get

$$\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial c}(T^{-1}(x,y,c))\right) = O(E(L,\sigma,|c|,|x|)), \tag{64}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial x}(T^{-1}(x,y,c))\right)=0, \tag{65}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\right)=O(E(L,\sigma,|c|,|x|)). \tag{66}$$

In order to compute $T_*\mathcal{Z}(x,y,c))$ we will treat separately the linear part of $\mathcal{Z}$ and the nonlinear one given by $\widetilde{M}$. The linear part is given by $\mathcal{L}(x,y,c)=(A_u x, A_s y, Bc)$, and it is immediate that

$$(T_*\mathcal{L})_c(x,y,c)=Bc.$$

For the x-component (for y-component the computations are analogous) we get

$$\begin{aligned}(T_*\mathcal{L})_x(x,y,c)=&\frac{\partial T_x}{\partial c}(T^{-1}(x,y,c))Bc+\frac{\partial T_x}{\partial x}(T^{-1}(x,y,c))(A_u T_x^{-1}(x,y,c))\\ &+\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))(A_s T_y^{-1}(x,y,c)).\end{aligned}$$

We are interested in the derivative of $(T_*\mathcal{L})_x(x,y,c)$ with respect to x. Let us investigate each term of the above sum separately. For the first term we have following estimate (we use (64))

$$\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial c}(T^{-1}(x,y,c))Bc\right)=O(|c|E(L,\sigma,|c|,|x|))=O(L).$$

For the second term we have (we use (62), (65))

$$\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial x}(T^{-1}(x,y,c))(A_u(T_x^{-1}(x,y,c)))\right)=A_u\frac{\partial T_x^{-1}}{\partial x}(x,y,c)=A_u+O(L).$$

For the third term we get

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))(A_s T_y^{-1}(x,y,c))\right)=&\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\right)\cdot A_s T_y^{-1}(x,y,c)\\ &+\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\cdot A_s\frac{\partial T_y^{-1}}{\partial x}(x,y,c),\end{aligned}$$

and from (66) and (53) we have

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\right)\cdot A_s T_y^{-1}(x,y,c)&=O(E(L,\sigma,|c|,|x|))A_s(y+O(\sigma L))\\ &=O(L)\cdot y+O(\sigma^2L^2),\\ \frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\cdot A_s\frac{\partial T_y^{-1}}{\partial x}(x,y,c)&=-\frac{\partial x^s}{\partial y}(T_{c,y}^{-1}(x,y,c))A_s O(L)=O(L^2)\end{aligned}$$

Therefore we obtain

$$\frac{\partial}{\partial x}((T_*\mathcal{L})_x(x,y,c)) = O(L) + (A_u + O(L)) + (O(L)y + O(L^2)) = A_u + O(L)y + O(L).$$

From bounds (40), (41), (42) it easily follows that

$$T_*\widetilde{M}(x,y,c) = O(\sigma^2),$$
$$DT_*\widetilde{M}(x,y,c) = O(\sigma),$$

and $T_*\widetilde{M}(x,y,c))$ has compact support.

This finishes the proof of (54), (55), (56), (57), (58).

It remains to prove (59). We will start with the function R, because the result is valid also for components of the vector field $\mathcal{R}$.

From (52) and inequalities (46) in Lemma 17 one can easily infer that there exists a constant K_1, such that

$$\|D^jT\| \le K_1, \quad \|D^jT^{-1}\| \le K_1, \quad j = 1,\dots,q-1. \tag{67}$$

Since $T^{-1}(c,0,0) = (c,0,0)$ we obtain from (62) that

$$T^{-1}(x,y,c) = (x + O(L)\cdot(x,y), y + O(L)\cdot(x,y), c),$$

hence there exists a constant K_2, such that

$$\|\pi_{x,y}T^{-1}(x,y,c)\| \le K_2\|(x,y)\|. \tag{68}$$

From this it is immediate that

$$|T_*R(x,y,c)| \le K\|\pi_{x,y}T^{-1}(x,y,c)\|^{Q+1} \le KK_2^{Q+1}\|(x,y)\|^{Q+1}.$$

Observe that the derivatives of T_*R are of the following form, for $j = 1,\dots,Q+1$,

$$D^jT_*R = D^jR()(DT^{-1})^j + D^{j-1}R()(DT^{-1})^j(D^2T^{-1}) + \cdots + DR()(D^jT^{-1}).$$

Using (67) and (68) we see that (59) holds for the function R.

From this we also have that in the case of the vector field $\mathcal{R}$

$$\mathcal{R}(T^{-1}(x,y,c)) = O_{Q+1}(|(x,y)|^{Q+1}).$$

The x-component of $T_*\mathcal{R}$ is given by (recall that $\mathcal{R}_c = 0$)

$$(T_*\mathcal{R})_x(x,y,c) = \mathcal{R}(T^{-1}(x,y,c))_x + \frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))\mathcal{R}(T^{-1}(x,y,c))_y.$$

Observe that due to (67) the function $\frac{\partial T_x}{\partial y}(T^{-1}(x,y,c))$ has bounded derivatives up to order $q-2$ (condition (60) gives us the needed regularity). Hence we can infer in the same way as in the case of functions that $(T_*\mathcal{R})_x(x,y,c) = O_{Q+1}(|(x,y)|^{Q+1})$. The same holds for the y-component. This finishes the proof of (59) for the vector field $\mathcal{R}$. □

Definition 3. Let $\tilde{r}_1$ be such that the support of $T_*\mathcal{R}$ (or T_*R in Hamiltonian case) is contained in $E(\tilde{r}_1) = \overline{B}_{n_u}(0,\tilde{r}_1) \times \overline{B}_{n_s}(0,\tilde{r}_1) \times \overline{B}_{2m}(0,\tilde{r}_1)$ for all $\varepsilon \in [0,1]$. Then we denote by χ the *indicator function* for $E(\tilde{r}_1)$: $\chi(z) = 1$ if $z \in D(\tilde{r}_1)$ and $\chi(z) = 0$ otherwise.

5.2.3. Consequences of straightening of invariant manifolds

Let $\mathcal{Z}$ be the vector field obtained after straightening and let $\varphi(t,z)$ be the local flow induced by it. Let us define

$$A_u(\varepsilon,c) = g_x(\varepsilon,0,0,c), \quad A_s(\varepsilon,c) = g_s(\varepsilon,0,0,c),$$

where the functions g_x and g_y defined in Lemma 21.

From Lemma 21 it follows that there exists functions $h_1(L) = O(L)$ and $h_2(\sigma) = O(\sigma)$, such that for all $\varepsilon \in [0,1]$ it holds that

$$\begin{aligned}
m_l(A_u) - h_1(L) - h_2(\sigma) &\le \inf_{c\in W^c} m_l(A_u(\varepsilon,c)) \le \sup_{c\in W^c} \mu_{log}(A_u(\varepsilon,c)) \\
&\le \mu_{log}(A_u) + h_1(L) + h_2(\sigma) \\
m_l(A_s) - h_1(L) - h_2(\sigma) &\le \inf_{c\in W^c} m_l(-A_s(\varepsilon,c)) \le \sup_{c\in W^c} \mu_{log}(-A_s(\varepsilon,c)) \\
&\le \mu_{log}(-A_s) + h_1(L) + h_2(\sigma)
\end{aligned}$$

where $m_l(A)$ and $\mu_{log}(A)$ are the logarithmic norms defined in Appendix A in Definition 6.

We assume that the numbers $\mu_{\min}, \mu_{\max}, \lambda_{\min}, \lambda_{\max}$ are such that (we decrease L and σ if needed)

$$0 < \mu_{\min} \le m_l(A_u) - h_1(L) - h_2(\sigma), \qquad \mu_{log}(A_u) + h_1(L) + h_2(\sigma) \le \mu_{\max}, \tag{69}$$

$$0 < \lambda_{\min} \le m_l(A_s) - h_1(L) - h_2(\sigma), \qquad \mu_{log}(-A_s) + h_1(L) + h_2(\sigma) \le \lambda_{\max}. \tag{70}$$

Observe that the numbers $\mu_{\min}, \mu_{\max}, \lambda_{\min}, \lambda_{\max}$ can be made arbitrarily close to the numbers $\mu_{\min}(A)$, $\mu_{\max}(A)$, $\lambda_{\min}(A)$, $\lambda_{\max}(A)$ given in Definition 2 if we take L and σ small enough.

Definition 4. Let $N \subset \mathcal{P}$. For $z \in N$ we define $T_N^{\pm}(z)$ by

$$\begin{aligned}
T_N^+(z) &= \sup\{T \in \mathbb{R}_+ \cup \{\infty\}: \quad \varphi(t,z) \in N, \quad t \in [0,T)\}, \\
T_N^-(z) &= \inf\{T \in \mathbb{R}_- \cup \{-\infty\}: \quad \varphi(t,z) \in N, \quad t \in (T,0]\}.
\end{aligned}$$

Lemma 22. *For any $\delta > 0$ let $N = D(\delta) = \overline{B}_{n_u}(\delta) \times \overline{B}_{n_s}(\delta) \times W^c$ and let μ_1' be any constant such that $0 < \mu_1' < \mu_{\min}$ (see (69)).*

Then there exists a constant $C_x(\mu_1', N)$ such that for all $z \in N$ it holds that

$$|\pi_x \varphi(t,z)| \le C_x(\mu_1', N)|x|e^{\mu_1' t}, \quad t \in (T_N^-(z), 0]. \tag{71}$$

Proof. From Lemma 21 we have

$$\dot{x} = g_x(\varepsilon, x, y, c)x = (A_u + O(L) + O(\sigma) + O(L)y)x. \tag{72}$$

Let us fix δ. If we take $0 < \delta' \leq \delta$, such that $m_l(A_u + O(L) + O(\sigma) + O(L)y) > \mu_1'$, then we obtain for all $z \in D(\delta')$

$$|\pi_x \varphi(t, z)| \leq |x| e^{\mu_1' t}, \quad t \in (T^-_{D(\delta')}(z), 0]. \tag{73}$$

Observe that the set $D(\delta')$ isolates W^c (i.e., it is a maximal invariant set in N), hence all backward orbits starting in N either converge to W^c or leave it. From Lemma 18 it follows that each backward orbit starting in $z \in N$ can be split into at most three parts: 1) for $t \in [0, T_1]$ $\varphi(-t, z) \notin D(\delta')$ and $\varphi(-T_1, z) \in D(\delta')$, 2) for $t \in [T_1, T_2)$ with $T_2 = \infty$ for $z \in W^{cu}$ $\varphi(-t, z) \in D(\delta')$, 3) if $T_2 < \infty$, then there exist T_3 such that for $t \in (T_2, T_3]$ $\varphi(-t, z) \in N \setminus D(\delta')$ and $\varphi(-t - \eta, z) \notin N$ for $\eta > 0$ arbitrary small.

Moreover, the time spent in the first and third parts is bounded from above by some constant T_{13} and due to (72) we would have the following estimate on first part (if present)

$$|\pi_x \varphi(-t, z)| \leq |x| e^{\mu_{log}(-g_x(N))t} \quad t \in [0, T_1],$$

and analogously for the third part (if needed). This combined with (73) plus bounds on T_1 and $T_3 - T_2$ gives us (71) for a suitable constant $C_x(\mu_1', N)$. □

Lemma 23. *Let us take any $\delta > 0$. Then for any μ_1', λ_1' such that $0 < \mu_1' < \mu_{\min}$ and $0 < \lambda_1' < \lambda_{\min}$, there exist constants $C_x(\mu_1')$ and $C_y(\lambda_1')$, such that for all $z \in D(\delta)$,*

$$\begin{aligned} |\pi_x \varphi(t, z) \chi(\varphi(t, z))| &\leq C_x(\mu_1') |x| e^{\mu_1' t}, \quad t \in (-\infty, 0], \\ |\pi_y \varphi(t, z) \chi(\varphi(t, z))| &\leq C_y(\lambda_1') |y| e^{-\lambda_1' t}, \quad t \in [0, \infty). \end{aligned} \tag{74}$$

Proof. We use Lemma 22 with $N = D(\tilde{r}_1)$. The proof of (74) is analogous. □

5.3. *Preparation of R_1 and R_2*

We will not make any difference between the vector field case, where the remainder $\mathcal{R}(z)$ is a vector, and the Hamiltonian case, where $R(z)$ is a real number. In both cases we will use the same letter R. Below our R is $T_{\varepsilon *} R$.

Lemma 24. *Assume that $R(\varepsilon, z) \in C^q$, and Q, q, ℓ_1, ℓ_2 are such that*

$$q \geq Q + 1, \tag{75}$$

$$R(\varepsilon, z) = O_{Q+1}(|(x, y)|^{Q+1}),$$

$$\ell_1 + \ell_2 \leq Q, \tag{76}$$

$$2\ell_2 + 1 \leq Q. \tag{77}$$

Then there exist functions $R_1(\varepsilon, z)$ and $R_2(\varepsilon, z)$ such that

$$\begin{aligned} &R_1, R_2 \in C^{q-\ell_2}, \\ &R_1(\varepsilon, z) = O_{\ell_1+1}(|x|^{\ell_1+1}), \quad R_2(\varepsilon, z) = O_{\ell_2+1}(|y|^{\ell_2+1}), \\ &R(\varepsilon, z) = R_1(\varepsilon, z) + R_2(\varepsilon, z). \end{aligned} \tag{78}$$

Moreover, the supports of R_1 and R_2 are contained in the support of R.

Proof. Below we will not write ε, c in the arguments of R and its derivatives.

For a given x consider a Taylor formula with an integral remainder for the function $y \mapsto R(x, \cdot)$

$$R(x, y) = \sum_{j=0}^{\ell_2} \frac{1}{j!} D_y^j R(x, 0)(y)^j + \mathrm{Rem}(x, y)(y)^{\ell_2+1}, \tag{79}$$

where $\mathrm{Rem}(x, y)(y)^{\ell_2+1}$ is given by

$$\mathrm{Rem}(z) = \left(\int_0^1 \frac{(1-t)^{\ell_2}}{\ell_2!} D_y^{\ell_2+1} R(x, ty) dt \right) (y)^{\ell_2+1},$$

where $D_y^{\ell_2+1} R(x, p)(y)^{\ell_2+1}$ is the $(\ell_2 + 1)$-linear map $D_y^{\ell_2+1} R(x, p)$ at $y = p$ applied to the argument $(y, \dots, y)$ with y appearing $\ell_2 + 1$ times.

We have for $s = 0, \dots, \ell_2$

$$D_y^s R(x, 0)(y)^s \in C^{q-s}, \quad D_y^s R(x, 0)(y)^s = O_{Q+1-s}\left(|x|^{Q+1-s} \right).$$

We set

$$R_1(x, y) = R(x, 0) + D_y R(x, 0) y + \dots + \frac{1}{\ell_2!} D_y^{\ell_2} R(x, 0)(y)^{\ell_2}.$$

Then

$$R_1 \in C^{q-\ell_2}, \qquad R_1(x, y) = O_{Q+1-\ell_2}(|x|^{Q+1-\ell_2}).$$

Since from (76) we have $Q + 1 - \ell_2 \geq \ell_1 + 1$, then $R_1(x, y) = O_{\ell_1+1}(|x|^{\ell_1+1})$.

Since $R_2 = R - R_1$ we get

$$R_2(x, y) = \mathrm{Rem}(x, y) \in C^{q-\ell_2}.$$

It remains to show that $R_2(x, y) = O_{\ell_2+1}(|y|^{\ell_2+1})$. Observe that

$$R_2(x, y) = I(x, y)(y)^{\ell_2+1}, \qquad I(x, y) \in C^{q-\ell_2-1}.$$

From (75), (77) we have that $q - \ell_2 - 1 \geq \ell_2 + 1$, so that $I(x, y) \in C^{\ell_2+1}$. From this we obtain the second condition in (78). $\square$

6. Solving the cohomological equation

Before we begin let us comment that to solve the cohomological equation we use the method of characteristics. The only difficulty is around W^c, but we ask for $\mathcal{R}$ (or R) to vanish on W^c, which makes it possible to solve it. Once a solution in the neighborhood of W^c is obtained we can in a unique way extend it to some neighborhoods of W^{cs} and W^{cu}.

To deal with the cases of a general vector field and Hamiltonians in a uniform way, from now on we will not use different fonts for the vector fields. The nature of each object will be clear from the context.

We consider the equation (see (34), (36))

$$DG(z)Z(z) = M(z)G(z) + R(z), \tag{80}$$

where $z \in \mathbb{R}^n$, $G : \mathbb{R}^n \to \mathbb{R}^m$, $Z : \mathbb{R}^n \to \mathbb{R}^n$, $M(z) \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^m)$ and $R : \mathbb{R}^n \to \mathbb{R}^m$. Z, M, R are given functions, and we are looking for G.

In our case we have either $n = m$ and $M(z) = DZ(z)$ for a general vector field or $m = 1$ and $M(z) = 0$ for the Hamiltonian case.

We assume that

$$Z \in C^{q(Z)}, \quad R \in C^{q(R)}.$$

We apply the *method of characteristics* with characteristic lines given as solutions of

$$\dot{z} = Z(z). \tag{81}$$

Let φ be a local flow induced by (81). It is well known that $\varphi(t, z)$ is in $C^{q(Z)}$.

Let $S(t, z) \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^m)$ be a solution of

$$\frac{d}{dt} S(t, z) = -S(t, z)M(\varphi(t, z)),$$

with initial condition $S(0, z) = I$. In the Hamiltonian case $S(t, z) = 1$.

Observe that

$$S(t, z) = D\varphi(t, z)^{-1}, \tag{82}$$

where $D\varphi(t, z)$ satisfies the following equation

$$\frac{d}{dt} D\varphi(t, z) = M(\varphi(t, z))D\varphi(t, z).$$

Assume that G satisfies (80), then

$$\frac{d}{dt}\left(S(t, z)G(\varphi(t, z))\right) = -S(t, z)M(\varphi(t, z))G(\varphi(t, z)) + S(t, z)DG(\varphi(t, z))\frac{d}{dt}\varphi(t, z)$$

$$= S(t,z)\,(DG(\varphi(t,z)Z(\varphi(t,z)) - M(\varphi(t,z))G(\varphi(t,z)))$$
$$= S(t,z)\,(DG\cdot Z - M\cdot G)\,(\varphi(t,z)) = S(t,z)R(\varphi(t,z)),$$

hence we obtain

$$\frac{d}{dt}\,(S(t,z)G(\varphi(t,z))) = S(t,z)R(\varphi(t,z)) \tag{83}$$

and therefore

$$S(t_2,z)G(\varphi(t_2,z)) - S(t_1,z)G(\varphi(t_1,z)) = \int_{t_1}^{t_2} S(w,z)R(\varphi(w,z))dw. \tag{84}$$

Observe that if G is smooth and G satisfies (83) or (84), then G solves (80).

6.1. Estimates on $\frac{\partial^r S}{\partial z^r}$ and $\frac{\partial^r \varphi}{\partial z^r}$

The goal of this section is to derive bounds on $\frac{\partial^r S}{\partial z^r}$ and $\frac{\partial^r \varphi}{\partial z^r}$.
Note that from (82) it follows that (for the case of a general vector field)

$$S(t,z) = \frac{\partial\varphi}{\partial z}(-t,z),$$

therefore estimates for $S(t,z)$ and its derivatives can be obtained from those on $\frac{\partial\varphi}{\partial z}$.

Lemma 25. *Let $N \subset \mathcal{P}$ be any convex set. Assume that*

$$\mu_{log}(DZ(z)) \le \alpha \neq 0, \quad z \in N. \tag{85}$$

Then for any $z \in N$ and for any $k \in \mathbb{Z}_+$, such that $1 \le k \le q(Z)$, it holds that

- *if $\alpha > 0$*

$$\left|\frac{\partial^k\varphi}{\partial z^k}(t,z)\right| \le C_k e^{k\alpha t}, \quad 0 \le t \le T_N^+(z),$$

- *if $\alpha < 0$*

$$\left|\frac{\partial^k\varphi}{\partial z^k}(t,z)\right| \le C_k e^{\alpha t}, \quad 0 \le t \le T_N^+(z),$$

for some constants C_k, with $C_1 = 1$.

Proof. Observe that $D^j Z$ for $1 \leq j \leq q(Z)$ are bounded on N.

The variational equations of increasing orders are

$$\frac{d}{dt}\frac{\partial \varphi}{\partial z}(t,z) = DZ(\varphi(t,z))\frac{\partial \varphi}{\partial z}(t,z)$$
$$\frac{d}{dt}\frac{\partial^2 \varphi}{\partial z^2}(t,z) = DZ(\varphi(t,z))\frac{\partial^2 \varphi}{\partial z^2}(t,z) + D^2 Z(\varphi(t,z))\left(\frac{\partial \varphi}{\partial z}(t,z), \frac{\partial \varphi}{\partial z}(t,z)\right).$$

For higher orders we have

$$\frac{d}{dt}\frac{\partial^k \varphi}{\partial z^k}(t,z) = DZ(\varphi(t,z))\frac{\partial^k \varphi}{\partial z^k}(t,z) + R_k(t,z),$$

where R_k is the sum of terms of the following form

$$R_{k,j}(t,z) = C D^m Z(\varphi(t,z))\left(\frac{\partial^{k_1} \varphi}{\partial z^{k_1}}(t,z), \frac{\partial^{k_2} \varphi}{\partial z^{k_2}}(t,z), \dots, \frac{\partial^{k_m} \varphi}{\partial z^{k_m}}(t,z)\right),$$

where C is a constant, $k_i > 0$ and $\sum_{i=1}^m k_i = k$.

We will use formula (153) from Theorem 43 in the Appendix A applied to the variational equations of k-th order to estimate $\left\| \frac{\partial^k \varphi}{\partial z^k}(t,z) \right\|$.

For $\frac{\partial \varphi}{\partial z}(t,y)$ we obtain from (85) and Theorem 41

$$\left\| \frac{\partial \varphi}{\partial z}(t,z) \right\| \leq e^{\alpha t}$$

and then inductively (observe that $D^j Z$ for $1 \leq j \leq q(Z)$ are bounded on N and the initial condition for $\frac{\partial^k \varphi}{\partial z^k}$ vanishes for $t = 0$)

$$\left\| \frac{\partial^k \varphi}{\partial z^k}(t,z) \right\| \leq \sum_j \int_0^t e^{\alpha(t-s)} \left\| R_{k,j}(s,z) \right\| ds.$$

For $\alpha < 0$, from our induction assumption it follows that each term R_{kj} will give a contribution bounded by

$$\tilde{C} \int_0^t e^{\alpha(t-s)} \left(e^{\alpha s}\right)^k ds \leq \tilde{C} e^{\alpha t} \int_0^t e^{\alpha(k-1)s} ds \leq \frac{\tilde{C} e^{\alpha t}}{-\alpha(k-1)}.$$

For $\alpha > 0$ this contribution will be bounded by

$$\tilde{C} \int_0^t e^{\alpha(t-s)} \left(e^{k\alpha s}\right) \leq \tilde{C} e^{\alpha t} \int_0^t e^{\alpha(k-1)s} ds < \frac{\tilde{C} e^{k\alpha t}}{\alpha(k-1)}. \quad \square$$

Remark 26. If the vector field $\mathcal{Z}$ depends smoothly on parameters, the above bounds also apply to partial derivatives involving parameters at the price of taking arbitrary $\alpha' > \alpha$, $\alpha' > 0$ and increasing constants C_k.

Proof. We add equation $\dot{\varepsilon} = 0$ to the system. In a suitable norm on the extended phase space $\mathcal{P} \times \Lambda$, where Λ is some compact set in parameter space we will have $\alpha = \sup_\varepsilon \sup_{z\in N} \mu_{log}(DZ_\varepsilon(z))$ to be approximately equal to $\alpha' = \sup_{\varepsilon} \sup_{z\in N} \mu_{log}\left(\frac{\partial Z_\varepsilon(z)}{\partial(z,\varepsilon)}(z,\varepsilon)\right)$. □

Remark 27. If $N = \overline{B}_{n_u}(\delta) \times \overline{B}_{n_s}(\delta) \times W^c$ is a neighborhood of W^c, then in the above lemma instead of $\alpha = \sup_{z\in N} \mu_{log}(DZ(z))$ we can take any $\alpha' > \sup_{z\in W^c} \mu_{log}(DZ(z))$ at the price of increasing the constants C_k.

Proof. The argument is the same as in the proof of Lemma 23. □

6.2. Solving for particular remainder—backward in time

We solve (80) with $R_1(z) = O_{\ell+1}(|x|^{\ell+1})$. We set (in the case of a general vector field)

$$M(z) = DZ(z), \qquad M(z) \in C^{q(Z)-1},$$

and in the Hamiltonian case

$$M(z) = 0.$$

We are in the setting discussed at the beginning of this section and we use (84) with $t_2 = 0$ and $t_1 \to -\infty$. Let us assume that

$$S(t_1, z)G(\varphi(t_1, z)) \to 0, \quad t_1 \to -\infty, \tag{86}$$

which has to be verified later, and then we obtain

$$G(z) = \int_{-\infty}^{0} S(t, z)R(\varphi(t, z))dt. \tag{87}$$

Lemma 28. *Assume that $Z = Z_\varepsilon$ and $R = R_\varepsilon = R_1$ are as obtained after straightening of invariant manifolds and the decomposition of T_*R described in Lemma 24. Let $\mu_{\min}, \mu_{\max}, \lambda_{\min}, \lambda_{\max}$ satisfy (69)–(70).*

Assume that for some positive integer ℓ it holds that

$$q(R) \geq \ell + 1, \tag{88}$$

$$R(z) = O_{\ell+1}(|x|^{\ell+1}), \tag{89}$$

$$Z(z) \in C^{q(Z)}.$$

Assume that for the general vector field

$$\ell + 1 \leq q(Z) - 1, \tag{90}$$

and in the Hamiltonian case

$$\ell + 1 \leq q(Z). \tag{91}$$

Assume that $k \geq 1$ satisfies

- *for the general vector field*

$$k < \min\left(\frac{\mu_{\min}(\ell+1) - \mu_{\max}}{\mu_{\min} + \lambda_{\max}}, \frac{\mu_{\min}(\ell+1) - \mu_{\max}}{\mu_{\max}}\right), \tag{92}$$

- *for the Hamiltonian vector field*

$$k < \frac{\mu_{\min}(\ell+1)}{\mu_{\min} + \lambda_{\max}}. \tag{93}$$

Let us define G_ε by (87) for all $(\varepsilon, z) \in [0,1] \times \mathcal{P}$. Then $G(\varepsilon, z) \in C^k$ and G_ε is a solution of (80).

Moreover,

$$G(\varepsilon, z) = O_k(|x|^{\ell+1}), \quad \text{for } |x| \to 0, \tag{94}$$

$$\|G(\varepsilon, z)\| \leq K|x|^{\ell+1}, \quad \forall z, \tag{95}$$

$$\|DG(\varepsilon, z)\| \leq K|x|^{\ell}, \quad \forall z. \tag{96}$$

Proof. In this proof we consider ε as a part of the phase space, hence it enters as one of the components in the variable z and with the vector field Z, G, R extended so that it vanishes in the ε direction. On this extended space the cohomological equation is still a cohomological equation with respect to variables (ε, z).

Let $\xi = 1$ for a general vector field and $\xi = 0$ in the Hamiltonian case.

Let $\alpha \approx \mu_{\min}$, $\gamma \approx \lambda_{\max}$ be such that

$$0 < \alpha < \mu_{\min}, \quad \gamma > \lambda_{\max}, \quad \beta > \mu_{\max}.$$

Let γ be a bound on the logarithmic norm for the backward in time evolution close to W^c, which should be close to $\mu_{log}(-A_s(c)) = \lambda_{\max} \approx \mu_{log}(-A_s) = \lambda_{\max}(A)$, $\gamma > \lambda_{\max}$ and let β be a logarithmic norm for forward in time evolution close to W^c, which should be $> \mu_{\max}$ but close.

For any z, from Lemmas 23 and 25 (see also Remarks 26 and 27) we have the following bounds for $t \in (-\infty, 0]$:

$$\|\pi_x\varphi(t,z)\|\chi(\varphi(t,z)) \le C_\alpha |x| e^{\alpha t}, \tag{97}$$

$$\left\| \frac{\partial^j \varphi}{\partial z^j}(t,z) \right\| \le C_{\gamma,j} e^{-j\gamma t}, \quad j=1,\ldots,q(Z), \tag{98}$$

$$\left\| \frac{\partial^j S}{\partial z^j}(t,z) \right\| \le C'_{\beta,j} e^{-(j+1)\beta t}, \quad j=0,\ldots,q(Z)-1, \tag{99}$$

where χ is the indicator function for a ball containing support of R (see Definition 3 and Lemma 23).

Observe that from (92), (93) and assumption $k \ge 1$ it follows that

$$\xi\mu_{\max} + \lambda_{\max} < \ell\mu_{\min}. \tag{100}$$

From (100) for $\beta \to \mu_{\max}$, $\gamma \to \lambda_{\max}$ and $\alpha \to \mu_{\min}$ we have

$$\xi\beta + \gamma < \ell\alpha. \tag{101}$$

From (92), (93) we should also have

$$k < \frac{\alpha(\ell+1) - \xi\beta}{\alpha+\gamma}, \tag{102}$$

and for the case of a general vector field additionally the following holds

$$k < \frac{\alpha(\ell+1) - \beta}{\beta}. \tag{103}$$

Since R has support in $\{|y| \le \tilde{r}_1\}$, then the integral in (87) is as smooth as R and S for $z=(x,y,c)$ with $y \neq 0$, as each backward trajectory spends only a finite interval of time in the set $\{|y| \le \tilde{r}_1\}$.

Since

$$G(z) = \lim_{T\to-\infty} \int_T^0 S(t,z)R(\varphi(t,z))dt,$$

we will show that $\int_T^0 S(t,z)R(\varphi(t,z))dt$ converges uniformly in C^k-norm as $T \to -\infty$.

From (89) and since R has support contained in $D(\tilde{r}_1)$ it follows that

$$\|D^s R(\varepsilon,z)\| \le K_s \|x\|^{\ell+1-s}, \quad s=0,\ldots,\ell+1, \quad \forall(\varepsilon,z). \tag{104}$$

From (97) and (104) it follows that for constants $K_{\alpha,s} = C_\alpha^{\ell+1-s} K_s$ and for any z it holds that

$$\left|(D^s R)(\varphi(t,z))\right| \le K_{\alpha,s} |x|^{\ell+1-s} \mathrm{e}^{(\ell+1-s)\alpha t} \text{ for } t \le 0, \quad s=0,\ldots,\ell+1. \tag{105}$$

From the above expression with $s=0$ and from (99) with $j=0$ (recall that for the Hamiltonian case we have $S=I$) we have for $t \le 0$ and some constant C

$$|S(t,z)R(\varphi(t,z))| \leq C e^{((\ell+1)\alpha - \xi\beta)t} |x|^{\ell+1}.$$

Since from (101) it follows that

$$\alpha(\ell+1) - \xi\beta > 0, \tag{106}$$

we see that the improper integral defining $G(z)$ in eq. (87) is convergent. Moreover, we have the following bound for some constant K, valid for all z

$$|G(z)| \leq K|x|^{\ell+1}. \tag{107}$$

To complete the proof that G is a solution of cohomological equation (80), we need to show that $G \in C^1$ and that condition (86) holds.

First we deal with (86). We have from (99), (97), (107), for $t < 0$

$$\begin{aligned}|S(t,z)G(\varphi(t,z))| &\leq K|S(t,z)| \cdot |\pi_x \varphi(t,z)|^{\ell+1} \\ &\leq K e^{-\xi\beta t}|x|^{\ell+1} e^{(\ell+1)\alpha t} \leq K|x|^{\ell+1} e^{-(\xi\beta - (\ell+1)\alpha)t},\end{aligned}$$

hence it converges to 0 for $t \to -\infty$ due to (106).

The derivatives of G (of order $k \leq \ell+1$) will be given by expressions of the following form

$$D^k G(z) = \int_{-\infty}^{0} \sum_{j=0}^{k} C_j D_z^j S(t,z) \cdot D_z^{k-j} \left(R(\varphi(t,z))\right) dt, \tag{108}$$

where for the Hamiltonian case only the term with $j=0$ is present and $S = I$.

Bounds for $\left\| D_z^j S(t,z) \right\|$ are given by (99). We need to find estimates for $D_z^{k-j}\left(R(\varphi(t,z))\right)$.

It is easy to see that $D_z^j\left(R(\varphi(t,z))\right)$ is the sum of terms of the following form, with $s = 1, \ldots, j$

$$(D^s R)(\varphi(t,z)) \left(\frac{\partial^{k_1}\varphi}{\partial z^{k_1}}(t,z), \frac{\partial^{k_2}\varphi}{\partial z^{k_2}}(t,z), \ldots, \frac{\partial^{k_m}\varphi}{\partial z^{k_m}}(t,z) \right),$$

where $k_i > 0$ and $\sum_{i=1}^{s} k_i = j$.

From the above using (105), (98) it is easy to see that for $t \leq 0$ we have, for some constant C,

$$\left\| D_z^j \left(R(\varphi(t,z))\right) \right\| \leq C|x|^{\ell+1-j} e^{\alpha(\ell+1-j)t} e^{-j\gamma t}, \quad j = 0, 1, \ldots, \ell+1. \tag{109}$$

By combining (108), (99), (109) we obtain the following bound for the case of a general vector field for $k \leq \ell+1$

$$\|D^k G(z)\| \leq C \int_{-\infty}^{0} \sum_{j=0}^{k} e^{-\beta(j+1)t} |x|^{\ell+1-(k-j)} e^{\alpha(\ell+1-(k-j))t} e^{-(k-j)\gamma t} dt. \tag{110}$$

In the Hamiltonian case only the term with $j = 0$ is present in (110) with $\beta = 0$, hence for the convergence of the above improper integral we need that

$$\alpha(\ell + 1) - k(\alpha + \gamma) > 0,$$

which is implied by (102).

In the case of a general vector field we need that for $j = 0, \dots, k$ it holds that

$$\alpha(\ell + 1 - (k - j)) - \beta(j + 1) - \gamma(k - j) = \alpha(\ell + 1) - k(\alpha + \gamma) - \beta - j(\beta - (\alpha + \gamma)) > 0.$$

It is easy to see that this is enough to satisfy the above inequality for $j = 0$ and $j = k$. We obtain

$$\alpha(\ell + 1) - k(\alpha + \gamma) - \beta > 0, \tag{111}$$

$$\alpha(\ell + 1) - (k + 1)\beta > 0. \tag{112}$$

Observe that (111) is implied by (102) and condition (112) by (103). □

6.3. Solving for a particular remainder—forward in time

In this section we discuss solving equation (80) but this time we assume that the remainder term R decays with y.

In (84) we pass to the limit $t_2 \to \infty$ and set $t_1 = 0$ to get

$$G(z) = -\int_0^\infty S(w, z) R(\varphi(w, z)) dw, \tag{113}$$

provided the above integral is convergent and

$$S(t_2, z) G(\varphi(t_2, z)) \to 0, \quad t_2 \to \infty.$$

Lemma 29. *Assume that Z and $R = R_2$ are as obtained after straightening of invariant manifolds.*

Let $\mu_{\min}, \mu_{\max}, \lambda_{\min}, \lambda_{\max}$ satisfy (69)–(70). Assume that

$$q(R) \geq \ell + 1, \tag{114}$$

$$R(z) = O_{\ell+1}(|y|^{\ell+1}),$$

$$Z(z) \in C^{q(Z)}.$$

Assume that

$$\ell + 1 \leq q(Z) - 1, \tag{115}$$

and in the Hamiltonian case

$$\ell + 1 \leq q(Z). \tag{116}$$

Assume that $k \geq 1$ satisfies

- *for a general vector field*

$$k < \min\left(\frac{\lambda_{\min}(\ell+1) - \lambda_{\max}}{\lambda_{\min} + \mu_{\max}}, \frac{\lambda_{\min}(\ell+1) - \lambda_{\max}}{\lambda_{\max}}\right),$$

- *for a Hamiltonian vector field*

$$k < \frac{\lambda_{\min}(\ell+1)}{\lambda_{\min} + \mu_{\max}}.$$

Let us define $G_\varepsilon(z)$ by (113) for all $(\varepsilon, z) \in [0,1] \times \mathcal{P}$. Then $G(\varepsilon, z) \in C^k$ and G is a solution of (80).

Moreover,

$$G(\varepsilon, z) = O_k(|y|^{\ell+1}), \quad \textit{for } |x| \to 0, \tag{117}$$

$$\|G(\varepsilon, z)\| \leq K|y|^{\ell+1}, \quad \forall z, \tag{118}$$

$$\|DG(\varepsilon, z)\| \leq K|y|^{\ell}, \quad \forall z.. \tag{119}$$

The proof is the same as for Lemma 28 simply reversing the direction of time, and exchanging $x \leftrightarrow y$, $\lambda \leftrightarrow \mu$.

6.4. Conclusion of the proof of Theorem 3

6.4.1. Inequalities related to the regularity

We need to show that the inequalities relating k (the regularity class of G, which is the solution of the cohomological equation) to q (the regularity of the vector field, $\mathcal{Z}, \mathcal{R} \in C^q$) and to Q, ℓ_1, ℓ_2, can be satisfied. These inequalities appear in Lemmas 21, 24, 28 and 29.

Let us notice that in the Hamiltonian case the solution of cohomological equation gives us a Hamiltonian defining the vector field for the deformation method, hence $G \in C^{k+1}$. For the same reason we have $R, Z \in C^{q+1}$ and $R(z,c) = O_{Q+2}(|z|^{Q+2})$, (see (24)), hence when applying Lemmas 28 and 29 in the Hamiltonian setting we should replace there k by $k+1$ and Q by $Q+1$.

From Lemma 21 we obtain that $T_*\mathcal{Z}, T_*\mathcal{R} \in C^{q-2}$, and in the Hamiltonian case, $T_*R \in C^{q-1}$. In the pure saddle case, because the stable and unstable manifolds are as smooth as the vector field, we have $T_*\mathcal{Z}, T_*\mathcal{R} \in C^{q-1}$, and $T_*R \in C^q$ in the Hamiltonian case.

For Lemma 24 applied to the vector field after straightening of invariant manifolds to obtain R_1, R_2 in $C^{q-2-\ell_2}$ and in $C^{q-1-\ell_2}$ in Hamiltonian case, and $C^{q-1-\ell_2}$ in pure saddle case ($C^{q-\ell_2}$ in Hamiltonian case), we need the following inequalities for the general vector field case (in the pure saddle case the first inequality will be $q-1 \geq Q+1$)

$$q - 2 \geq Q + 1, \tag{120}$$

$$\ell_1 + \ell_2 \leq Q, \tag{121}$$

$$2\ell_2 + 1 \leq Q, \tag{122}$$

while in the Hamiltonian case we have the following inequalities (in the pure saddle case the first inequality will be $q \geq Q + 2$)

$$q - 1 \geq Q + 2, \tag{123}$$

$$\ell_1 + \ell_2 \leq Q + 1, \tag{124}$$

$$2\ell_2 + 1 \leq Q + 1. \tag{125}$$

Observe that from (120), (121) it follows that conditions (88), (90), (91) and (114), (115), (116) from Lemmas 28 and 29, respectively, are satisfied.

From Lemmas 28 and 29 we obtain the following inequalities for the general vector field

$$k < \min\left(\frac{\mu_{\min}(\ell_1+1)-\mu_{\max}}{\mu_{\min}+\lambda_{\max}}, \frac{\mu_{\min}(\ell_1+1)-\mu_{\max}}{\mu_{\max}}\right), \tag{126}$$

$$k < \min\left(\frac{\lambda_{\min}(\ell_2+1)-\lambda_{\max}}{\lambda_{\min}+\mu_{\max}}, \frac{\lambda_{\min}(\ell_2+1)-\lambda_{\max}}{\lambda_{\max}}\right), \tag{127}$$

and for the Hamiltonian case

$$k + 1 < \frac{\mu_{\min}(\ell_1+1)}{\mu_{\min}+\lambda_{\max}}, \tag{128}$$

$$k + 1 < \frac{\lambda_{\min}(\ell_2+1)}{\lambda_{\min}+\mu_{\max}}. \tag{129}$$

The order of setting parameters is as follows. We pick any $k \geq 1$. Then we find the minimal ℓ_1 satisfying (126),

$$\ell_1(k) = \max\left(k + \left[k\frac{\lambda_{\max}}{\mu_{\min}} + \frac{\mu_{\max}}{\mu_{\min}}\right], \left[(k+1)\frac{\mu_{\max}}{\mu_{\min}}\right]\right),$$

and in the Hamiltonian case we use (128)

$$\ell_1^H(k) = k + 1 + \left[(k+1)\frac{\lambda_{\max}}{\mu_{\min}}\right].$$

Analogously we find the minimal ℓ_2 satisfying (127).

$$\ell_2(k) = \max\left(k + \left[k\frac{\mu_{\max}}{\lambda_{\min}} + \frac{\lambda_{\max}}{\lambda_{\min}}\right], \left[(k+1)\frac{\lambda_{\max}}{\lambda_{\min}}\right]\right)$$

and for the Hamiltonian case from (129)

$$\ell_2^H(k) = k + 1 + \left[(k+1)\frac{\mu_{\max}}{\lambda_{\min}}\right].$$

From (121) for general vector field we obtain

$$\begin{aligned} Q \geq Q_0 &= \ell_1(k) + \ell_2(k) \\ &= \max\left(k + \left[k\frac{\lambda_{\max}}{\mu_{\min}} + \frac{\mu_{\max}}{\mu_{\min}}\right], \left[(k+1)\frac{\mu_{\max}}{\mu_{\min}}\right]\right) + \\ &\quad \max\left(k + \left[k\frac{\mu_{\max}}{\lambda_{\min}} + \frac{\lambda_{\max}}{\lambda_{\min}}\right], \left[(k+1)\frac{\lambda_{\max}}{\lambda_{\min}}\right]\right) \end{aligned}$$

For the Hamiltonian case from (124) we obtain

$$\begin{aligned} Q \geq Q_0 = \ell_1^H(k) + \ell_2^H(k) - 1 &= 2k + 1 + \left[(k+1)\frac{\lambda_{\max}}{\mu_{\min}}\right] + \left[(k+1)\frac{\mu_{\max}}{\lambda_{\min}}\right] \\ &= 2k + 1 + 2\left[(k+1)\frac{\mu_{\max}}{\mu_{\min}}\right]. \end{aligned}$$

Observe that we ignored condition (122) (and (125)), but we can always rearrange the decomposition $R = R_1 + R_2$, so that $\ell_2 \leq \ell_1$ and then from (120) (and (123)) it follows that $q_0 = Q_0 + 3$ and in the pure saddle case $q_0 = Q_0 + 2$.

This finishes the proof of Remark 10.

6.4.2. Transformation removing the remainder

Consider the general vector field case. From the solution of cohomological equation for $\varepsilon \in [0, 1]$ we obtain the vector field $\mathcal{G}(\varepsilon, z) = T_{\varepsilon*}^{-1}(G_1(\varepsilon, z) + G_2(\varepsilon, z))$, which is at least C^k in (ε, z).

$\mathcal{G}(\varepsilon, z)$ is defined for $\varepsilon \in [0, 1]$ and arbitrary z and satisfies for all $\varepsilon \in [0, 1]$ the following estimates, see (94), (95), (96) in Lemma 28 and (117), (118), (119) in Lemma 29

$$\mathcal{G}(\varepsilon, z) = O_k(|(x, y)|^{\ell+1}), \quad (x, y) \to 0, \tag{130}$$

$$\|\mathcal{G}(\varepsilon, z)\| \leq K\|(x, y)\|^{\ell+1}, \quad \forall z, \tag{131}$$

$$\|D\mathcal{G}(\varepsilon, z)\| \leq K\|(x, y)\|^{\ell}, \quad \forall z, \tag{132}$$

where $k \leq \ell + 1$, and $\ell = \min(\ell_1(k), \ell_2(k))$ with $\ell_i(k)$ defined in the previous section.

Note that $T_{\varepsilon*}^{-1}$ preserves conditions (130), (131), (132), which were proved for $G_1 + G_2$.

The differential equation defining $g(\varepsilon, z)$ is

$$\frac{d}{d\varepsilon} g(\varepsilon, z) = \mathcal{G}(\varepsilon, g(\varepsilon, z)). \tag{133}$$

From (130) it follows that the points from the center manifold are fixed points of g_1 and $Dg_1 = Id$ for such points.

In view of (131) the solution may be defined only for a finite time (i.e., ε) interval. To estimate the length of this interval we proceed as follows. As

$$\frac{d\|\pi_{xy} g(\varepsilon, z)\|}{d\varepsilon} \leq K\|\pi_{xy} g(\varepsilon, z)\|^{\ell+1},$$

we compare $\|\pi_{xy} g(\varepsilon, z)\|$ to the solution of the equation

$$v' = Kv^{\ell+1}, \quad v(0) = \|\pi_{xy} z\|,$$

which is given by (we assume that $v(0) > 0$)

$$v(t) = \frac{v(0)}{(1 - K\ell t v(0)^\ell)^{1/\ell}},$$

and we obtain

$$\|\pi_{x,y} g(\varepsilon, z)\| \leq \frac{\|z\|}{(1 - K\ell t \|\pi_{xy} z\|^\ell)^{1/\ell}}.$$

Therefore if

$$\|\pi_{xy} z\| \leq \frac{1}{(K\ell)^{1/\ell}},$$

then $g(1, z)$ is defined.

The same reasoning applies to the inverse map of g_1, which is a solution of (133) moving backward in time.

This finishes the proof of Theorem 3.

7. Sign symmetry

Assume that (compare (11)) $\mathcal{P} = \{(z, c), z = (z_+, z_-) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, c \in \mathbb{C}^{2m}\}$, i.e., we identify $\mathbb{R}^{k_A} \times \mathbb{C}^{l_A}$ with $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

For $l = 1, \dots, m$ we define the hyperplane W_l as

$$W_l = \{(z_+, z_-, c) \in \mathcal{P} : c_l = 0\},$$

and $W_\pm$ as

$$W_+ = \{(z_+, z_-, c) \in \mathcal{P} : z_+ = 0\},$$
$$W_- = \{(z_+, z_-, c) \in \mathcal{P} : z_- = 0\}.$$

We also define the following subspaces

$$V_j = \{(z_+, z_-, c) \in \mathcal{P} :\ z_\pm = 0, c_l = 0,\ l \neq j\} = \bigcap_{l \neq j} W_l \cap W_- \cap W_+,$$

$$V_+ = \{(z_+, z_-, c) \in \mathcal{P} :\ z_- = 0, c_l = 0,\ 1 \leq l \leq m\} = \bigcap_{l=1}^{m} W_l \cap W_-,$$

$$V_- = \{(z_+, z_-, c) \in \mathcal{P} :\ z_+ = 0, c_l = 0,\ 1 \leq l \leq m\} = \bigcap_{l=1}^{m} W_l \cap W_+.$$

Observe that our phase space has the following decomposition

$$\mathcal{P} = V_+ \oplus V_- \oplus V_1 \oplus \cdots \oplus V_m.$$

Let X and Y be the unstable and stable manifolds of the linearization at 0. We assume that we have decompositions $X = X_+ \oplus X_-$ and $Y = Y_+ \oplus Y_-$. We also assume that

$$V_+ = X_+ \oplus Y_+, \quad V_- = X_- \oplus Y_-.$$

We introduce the variables $x_\pm$, $y_\pm$ by $z_+ = (x_+, y_+) \in X_+ \oplus Y_+$ and $z_- = (x_-, y_-) \in X_- \oplus Y_-$. We will use coordinates $x_{i\pm}$ and $y_{i\pm}$ in $X_\pm$ and $Y_\pm$, so that z_- is given by x_{i-} and y_{i-}, etc. We will use multiindices $\alpha \in \mathbb{N}^{k_1}$ and $\beta \in \mathbb{N}^{k_2}$, to define $z_+^\alpha z_-^\beta$ and $\dfrac{\partial^{|\alpha|R}}{\partial z_+^\alpha}$, etc.

Definition 5. For any choice of signs $s_+, s_-, s_j \in \{-1, 1\}$ consider the map $\mathcal{S}_s$ be given by

$$\mathcal{S}_s(z_+, z_-, c_1, \ldots, c_m) = (s_+ z_+, s_- z_-, s_1 c_1, \ldots, s_m c_m).$$

We say that a differential equation or map or symplectic form has a *sign-symmetry* if it is $\mathcal{S}_s$-symmetric for any choice of signs $\{s = (s_+, s_-, s_1, \ldots, s_m)\}$.

In the Hamiltonian case we assume that

$$\Omega = \sum_j dx_{j-} \wedge dy_{j-} + \sum_j dx_{j+} \wedge dy_{j+} + \frac{i}{2} \sum_k dc_k \wedge d\bar{c}_k. \tag{134}$$

Observe that if Ω has sign-symmetry, $\Omega(\mathcal{S}_s(z, c)) = \Omega(z, c)$ for any s and $(z, c) \in \mathcal{P}$.

Observe that the sign-symmetry implies the invariance of the spaces W_l for a vector field or map and also the subspaces V_l and $V_\pm$ are invariant. Moreover, the origin 0 is a fixed point of an o.d.e. or map under consideration.

It is easy to see that

- the inverse map of G with sign-symmetry has also this property
- the composition of two maps with sign-symmetry has also this property,
- a vector field with sign-symmetry induces a flow for which any time shift has also these properties.

Our goal is to establish the following extension of Theorem 3.

Theorem 30. *Consider system* (10) *with the same assumptions as in Theorem 3. Then the change of coordinates bringing it to polynomial normal form* (12) *has sign-symmetry if system* (10) *has sign-symmetry and is symplectic if the transformed system was Hamiltonian.*

For the proof of Theorem 30 we show that each step in the construction of coordinate change in Theorem 3 is performed so that the sign-symmetry are preserved. Obviously the regularity of the coordinate change is the same as obtained in Theorem 3. Corollary 11 is also valid in the case of sign-symmetry.

7.1. The toy model system for transfer of energy in NLSE

The motivation of the present work was the toy model system derived in [8] in the context of the transfer of energy in cubic defocusing non-linear Schrödinger equation (NLSE), see also [11] and references given there.

The phase space is described by $z = (x_-, y_-, x_+, y_+, c) \in \mathbb{R}^4 \times \mathbb{C}^N$. The toy model system is

$$\begin{aligned}
\dot{x}_- &= \lambda x_- + O_2(|z|^2),\\
\dot{y}_- &= -\lambda y_- + O_2(|z|^2),\\
\dot{x}_+ &= \lambda x_+ + O_2(|z|^2),\\
\dot{y}_+ &= -\lambda y_+ + O_2(|z|^2),\\
\dot{c}_l &= i\nu_k c_l + O(|z)|^2), \quad l = 1, \dots, N
\end{aligned} \tag{135}$$

where $\lambda > 0$, $\nu_l > 0$. We assume that $\{(x_-, y_-, x_+, y_+) = 0\}$ is invariant and the system has the sign-symmetry.

The following result follows from Corollary 11 adapted to include the sign-symmetry.

Lemma 31. *For any $k \geq 1$ if system (135) is C^q with q sufficiently large, then there exists a C^k change of variables in a neighborhood of the origin, transforming system (135) to the system*

$$\begin{aligned}
\dot{x}_- &= \lambda x_- + N_{x_-}(z),\\
\dot{y}_- &= -\lambda y_- + N_{y_-}(z),\\
\dot{x}_+ &= \lambda x_+ + N_{x_+}(z),\\
\dot{y}_+ &= -\lambda y_+ + N_{y_+}(z),\\
\dot{c}_\ell &= i\nu_l c_\ell + O_2(z), \quad \ell = 1, \dots, N
\end{aligned}$$

where for any saddle variable $v \in \{x_-, y_-, x_+, y_+\}$ we have

$$N_v(z) = \sum_{m \in M_{1,v}} g_{v,m}(c_*) z^m + \sum_{m \in M_{2,v}} g_{v,m}(z) z^m,$$

where $g_{v,m}$ are continuous functions, $M_{1,v}, M_{2,v}$ are finite *sets of indices, and any z^m is a resonant monomial for the saddle variables, satisfying on the one hand $m_s := m_{x_-} + m_{y_-} + m_{x_+} + m_{y_+} \geq 3$ and $m_c = 0$ if $m = \left(m_{x_-}, m_{y_-}, m_{x_+}, m_{y_+}, m_c\right) \in M_{1,v}$, and on the other hand $m_s = 1$ and $m_c \geq 3$ if $m \in M_{2,v}$.*

This change of coordinates preserves the sign-symmetry and if the original system is Hamiltonian, then this change is symplectic.

Proof. We take any $P = Q + 2$, this gives us $P + 1 - Q = 3$ which gives $m_c \geq 3$ in the second sum. We have $m_s \geq 3$ because there are no resonant terms of order 2. □

Let us see what q we need to guarantee $k = 2$. For this we use Remark 10. We have (see Definition 2)

$$\mu_{\min} = \mu_{\max} = \lambda_{\min} = \lambda_{\max} = \lambda.$$

Hence from (13) for a general vector field we obtain

$$Q_0(k) = 4k + 2, \quad q_0(k) = 4k + 5,$$

while for a Hamiltonian system (see (14)) we obtain

$$Q_0(k) = 4k + 3, \quad q_0(k) = 4k + 6.$$

For a general vector field, we need $q \geq q_0$, $q - 1 \geq P \geq Q$, $q \geq Q + 3$ (see the statement of Theorem 3). We take $P = Q_0(k) + 2$ to obtain

$$P = 4k + 4, \quad q \geq 4k + 5.$$

Hence to obtain $k = 2$ we need $q \geq 13$. In the Hamiltonian case we need $q \geq 14$.

7.2. Conditions of sign-symmetry for Hamiltonian systems

For Hamiltonians with the symplectic form (134) the sign-symmetry is

$$R(z) = R(\mathcal{S}_s(z)) \tag{136}$$

for any choice of signs $\{s\}$. It is easy to see that the Hamiltonian vector field induced by R using the symplectic form (134) also has sign-symmetry.

However after straightening center-stable and center-unstable manifolds the symplectic form $T_*\Omega$ is no longer in standard form (134) and the relation between Hamiltonian and its induced vector field given by (4) is more complicated than (6). However, we will show that the straightening transformation T_ε has also sign-symmetry, as well as $T_{\varepsilon *}\Omega$ and the condition on the sign-symmetry of $T_* R$ will be still (136).

8. Preservation of sign symmetry for all steps leading to polynomial normal forms

8.1. The sign-symmetry for the Taylor expansion in saddle directions

The goal of this section is to show that if a map or vector field or Hamiltonian function has the sign-symmetry, then each term in the Taylor expansion has the sign-symmetry. This is a content of Lemma 32. This result is needed to justify that removing the non-resonant terms can be done preserving the sign-symmetry.

Assume $G \in C^{Q+1}$ (vector field or map). The Taylor formula with respect to saddle directions (z_+, z_-) with an integral remainder might be written as follows

$$G_{\pm,l}(z_+, z_-, c) = \sum_{i=0}^{Q} \sum_{|\alpha|+|\beta|=i} a^{\alpha,\beta} \frac{\partial^i G_{\pm,l}}{\partial z_+^\alpha \partial z_-^\beta}(z_\pm = 0, c) z_+^\alpha z_-^\beta + \sum_{|\alpha|+|\beta|=Q+1} d_{\pm,l}^{\alpha,\beta}(z_+, z_-, c) z_+^\alpha z_-^\beta, \tag{137}$$

where $a^{\alpha,\beta}$ are some constants and $d_{\pm,l}^{\alpha,\beta}(z_+, z_-, c)$ are coefficients of the $(Q+1)$-linear map representing the remainder

$$G_{\pm l,\mathrm{Rem}}(z_+, z_-, c) = \left(\int_0^1 \frac{(1-t)^Q}{Q!} D^{(Q+1)}_{(z_+,z_-)} G_{\pm,l}(tz_+, tz_-, c) dt \right) (z_+, z_-)^{Q+1},$$

i.e., up to the constant depending on α, β and Q,

$$d_{\pm,l}^{\alpha,\beta}(z_+, z_-, c) = a^{\alpha,\beta} \int_0^1 (1-t)^Q \frac{\partial^{Q+1} G_{\pm,l}}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c) dt.$$

For each term (137) we can define a map as follows. For example the term $a^{\alpha,\beta} \frac{\partial^i G_+}{\partial z_+^\alpha \partial z_-^\beta}(0, 0, c) z_+^\alpha z_-^\beta$ induces the map $T(z_+, z_-, c) = (T_+(z_+, z_-, c), T_-(z_+, z_-, c), T_1(z_+, z_-, c), \dots, T_m(z_+, z_-, c))$, with

$$T_+(z_+, z_-, c) = a^{\alpha,\beta} \frac{\partial^i G_+}{\partial z_+^\alpha \partial z_-^\beta}(0, 0, c) z_+^\alpha z_-^\beta, \quad T_- \equiv 0, \quad T_l \equiv 0,$$

and analogously for the other terms.

Lemma 32. *Assume that $G \in C^{Q+1}$ has the sign-symmetry. Then each non-vanishing term (i.e. not identically equal to 0) in the Taylor expansion (137) also has the sign-symmetry.*

Proof. The sign-symmetry of all the terms with respect to the changes of sign on c_k is immediate. Hence it is enough to consider the changes of signs of z_+ (for z_- the proof is analogous).

Consider G_l (for G_- the argument is the same). The sign-symmetry with respect z_+ means that $G_l(z_+, z_-, c) = G_l(-z_+, z_-, c)$. This implies that

$$\frac{\partial^{|\alpha|+|\beta|} G_l}{\partial z_+^\alpha \partial z_-^\beta}(z_+ = 0, z_-, c) = 0, \quad \text{if } |\alpha| \text{ is odd}.$$

Therefore all terms of order less than or equal to Q have the symmetry $z_+ \to -z_+$.

Now we look at the remainder. Since for all t it holds that

$$\frac{\partial^{|\alpha|+|\beta|} G_l}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c) = (-1)^{|\alpha|} \frac{\partial^{|\alpha|+|\beta|} G_l}{\partial z_+^\alpha \partial z_-^\beta}(t(-z_+), tz_-, c),$$

we see that

$$\frac{\partial^{|\alpha|+|\beta|}G_l}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c)z_+^\alpha = (-1)^{|\alpha|}\frac{\partial^{|\alpha|+|\beta|}G_l}{\partial z_+^\alpha \partial z_-^\beta}(t(-z_+), tz_-, c)(-z_+)^\alpha,$$

and therefore

$$d_l^{\alpha,\beta}(-z_+, z_-, c) = d_l^{\alpha,\beta}(z_+, z_-, c).$$

This establishes the sign-symmetry with respect to z_+ of all the remainder terms for G_l. It remains to consider G_+. The sign-symmetry with respect z_+ means that $G_+(z_+, z_-, c) = -G_+(z_+, z_-, c)$. This implies that

$$\frac{\partial^{|\alpha|+|\beta|}G_+}{\partial z_+^\alpha \partial z_-^\beta}(z_+ = 0, z_-, c) = 0, \quad \text{if } |\alpha| \text{ is even.}$$

Therefore all terms of order less than or equal than Q have the symmetry $z_+ \to -z_+$.

Now we look at the remainder. Since

$$\frac{\partial^{|\alpha|+|\beta|}G_+}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c) = (-1)^{|\alpha|+1}\frac{\partial^{|\alpha|+|\beta|}G_+}{\partial z_+^\alpha \partial z_-^\beta}(t(-z_+), tz_-, c),$$

we see that

$$\frac{\partial^{|\alpha|+|\beta|}G_+}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c)z_+^\alpha = (-1)^{|\alpha|+1}\frac{\partial^{|\alpha|+|\beta|}G_+}{\partial z_+^\alpha \partial z_-^\beta}(t(-z_+), tz_-, c)(-z_+)^\alpha,$$

and therefore

$$d_+^{\alpha,\beta}(-z_+, z_-, c) = -d_+^{\alpha,\beta}(z_+, z_-, c).$$

This establishes the sign-symmetry with respect to z_+ of all the remainder terms for G_+. □

8.1.1. Sign-symmetry for Taylor expansions for Hamiltonians

The Taylor formula with an integral remainder might be written as follows

$$G(z_+, z_-, c) = \sum_{i=0}^{Q}\sum_{|\alpha|+|\beta|=i} a^{\alpha,\beta}\frac{\partial^i G}{\partial z_+^\alpha \partial z_-^\beta}(0,0,c)z_+^\alpha z_-^\beta + \sum_{|\alpha|+|\beta|=Q+1} d^{\alpha,\beta}(z_+, z_-, c)z_+^\alpha z_-^\beta \tag{138}$$

where $a^{\alpha,\beta}$ are some constants and $d_{\alpha,\beta}(z_+, z_-, c)$ are the coefficients of the $(Q+1)$-linear form of the remainder

$$G_{\text{Rem}}(z_+, z_-, c) = \left(\int_0^1 \frac{(1-t)^Q}{Q!} D^{(Q+1)}_{(z_+,z_-)}G(tz_+, tz_-, c)dt\right)(z_+, z_-)^{Q+1},$$

i.e., up to the constant depending on α and β,

$$d^{\alpha,\beta}(z_+, z_-, c) = a^{\alpha,\beta} \int_0^1 (1-t)^Q \frac{\partial^{Q+1} G}{\partial z_+^\alpha \partial z_-^\beta}(tz_+, tz_-, c) dt.$$

Lemma 33. *For $Q \geq 2$ assume that $G \in C^{Q+2}$ has the sign-symmetry. Then every non-vanishing term (i.e. not identically equal to* 0*) in the Taylor expansion (138) also has the sign-symmetry*

We omit an easy proof similar to the proof of Lemma 32.

8.2. The sign-symmetry for the terms in the Taylor expansion in y-direction

The goal of this section is to show that all the terms defined in the decomposition obtained in Lemma 24 have the sign-symmetry if the original vector field (or map) has it.

In this section we use coordinates (x, y, c), $x = (x_+, x_-)$, $y = (y_+, y_-)$, so that $z_+ = (x_+, y_+)$ and $z_- = (x_-, y_-)$.

Given a vector field (or map) R we write it as (compare with (79))

$$R(x, y, c) = R(x, 0, c) + D_y R(x, 0, c)y + \cdots + \frac{1}{\ell!} D_y^\ell R(x, 0, c)(y)^\ell + \text{Rem}(x, y, c)(y)^{\ell+1} \tag{139}$$

Lemma 34. *Assume that $R \in C^{q(R)}$ has the sign-symmetry and $\ell + 1 \leq q(R)$. Then every term in (139) also has the sign-symmetry.*

We omit an easy proof.

8.2.1. The sign-symmetry for the terms in the Taylor expansion in the y-direction for Hamiltonians

Recall that the condition for the sign-symmetry for a Hamiltonian G is given by (see (136))

$$G(z) = G(\mathcal{S}_s(z))$$

for any choice of signs $\{s\}$.

Given a Hamiltonian R we write it as (compare with (79))

$$R(x, y, c) = R(x, 0, c) + D_y R(x, 0, c)y + \cdots + \frac{1}{\ell!} D_y^\ell R(x, 0, c)(y)^\ell + \text{Rem}(x, y, c)(y)^{\ell+1}. \tag{140}$$

Lemma 35. *Assume that $R \in C^{q(R)}$ and $\ell + 1 \leq q(R)$. If R has the sign-symmetry, then every non-vanishing term (i.e. not identically equal to* 0*) in the Taylor expansion (140) also has the sign-symmetry.*

We omit an easy proof.

8.3. The sign-symmetry is preserved in the preparation of the remainder

Assume that R has the sign-symmetry.

The decomposition $R = R_1 + R_2$ is performed using Lemma 24, then from Lemma 34 for general vector fields and Lemma 35 for Hamiltonians we know that R_1 and R_2 have the sign-symmetry.

8.4. The transformation removing non-resonant terms preserves the sign-symmetry

Assume that vector field $\mathcal{Z}$ (or the Hamiltonian Z) has the sign-symmetry.

For a general vector field, the transformation bringing $\mathcal{Z}$ to the normal form is a composition of maps of the type Id plus some non-resonant terms (see [1]). From Lemma 32 it follows that each such term has the desired geometric properties, and the induced transformation also has.

For Hamiltonians the transformation removing non-resonant terms (see for example [13,12]) is a composition of symplectic maps, which are the time shift 1 of the Hamiltonian flow induced by the polynomial non-resonant term in the Hamiltonian to be removed. From Lemma 33 it follows that each such term has the sign-symmetry. So the induced time-1 also has sign-symmetry.

For reference purposes, we formulate the above as a lemma.

Lemma 36. *The transformation of removing non-resonant terms can be performed in such a way that the sign-symmetry is preserved. Moreover, in the Hamiltonian case it is symplectic.*

8.5. The sign-symmetry and preparation of compact data

When preparing “compact data” in Section 5.1 we modified the vector fields and Hamiltonians.

The question is: if $\mathcal{Z}$ and $\mathcal{R}$ have the sign-symmetry, then $\tilde{\mathcal{Z}}$ and $\tilde{\mathcal{R}}$ introduced in (37)–(38) will also have the sign-symmetry? Analogously for the Hamiltonian $\tilde{H}$ introduced in (43). The answer is yes if the function η used in both cases is constructed as follows.

Let $\alpha : \mathbb{R} \to [0, 1]$ be C^∞ with compact support and such that $\alpha(x) = \alpha(-x)$, so that $\alpha'(0) = 0$. We define η, which will not depend on c, by

$$\eta(x_{1\pm}, x_{2\pm}, \ldots, y_{1\pm}, y_{2\pm}, \ldots) = \Pi_i \alpha(x_{i+}) \cdot \Pi_i \alpha(x_{i-}) \Pi_i \alpha(y_{i+}) \cdot \Pi_i \alpha(y_{i-}). \tag{141}$$

Observe that

$$\frac{\partial \eta}{\partial x_{i\pm}}(p) = 0, \quad \text{if } x_{i\pm}(p) = 0, \qquad \frac{\partial \eta}{\partial y_{i\pm}}(p) = 0, \quad \text{if } y_{i\pm}(p) = 0.$$

The following lemma is obvious

Lemma 37. *Let η be as in (141). Assume that the vector field $\mathcal{Z}$ and $\mathcal{R}$ have the sign-symmetry. Then $\tilde{\mathcal{Z}}$ and $\tilde{\mathcal{R}}$ have it too. The same for the Hamiltonian $\tilde{Z}$ and $\tilde{R}$, as long as Z and R have the sign-symmetry.*

8.6. The transformation straightening the invariant manifolds preserves the sign-symmetry

Let $(x_-, x_+, y^u_-(x_-, x_+, c), y^u_+(x_-, x_+, c), c)$ be the center-unstable manifold and $(x^s_-(y_-, y_+, c), x^s_+(y_-, y_+, c), y_-, y_+, c)$ be the center-stable manifold as obtained in Lemma 17.

We have the following result about some properties of the functions $y^u_\pm$ and $x^s_\pm$.

Lemma 38. *If the vector field (33) has the sign-symmetry, then for any* $s_+, s_- \in \{-1, 1\}$ *it holds that*

$$y^u(x, c_k, c_*) = y^u(x, -c_k, c_*), \tag{142}$$

$$x^s(y, c_k, c_*) = x^s(y, -c_k, c_*), \tag{143}$$

$$y^u_+(s_+x_+, s_-x_-, c) = s_+y^u_+(x_+, x_-, c), \tag{144}$$

$$y^u_-(s_+x_+, s_-x_-, c) = s_-y^u_-(x_+, x_-, c), \tag{145}$$

$$x^s_+(s_+y_+, s_-y_-, c) = s_+x^s_+(y_+, y_-, c)$$

$$x^s_-(s_+y_+, s_-y_-, c) = s_-x^s_-(y_+, y_-, c).$$

Proof. It is enough to consider the equalities involving the function y^u, because analogous symmetric arguments give us the statements for x^s.

The center-unstable manifold is the graph of the function $(x_-, x_+, c) \mapsto (y^u_-(x_-, x_+, c), y^u_+(x_-, x_+, c))$. The sign-symmetry $(c_k \to -c_k)$ implies that $(x, y^u(x, c_k, c_*), c_k, c_*) \in W^{cu}$ iff $(x, y^u(x, c_k, c_*), -c_k, c_*) \in W^{cu}$. Therefore

$$y^u(x, c_k, c_*) = y^u(x, -c_k, c_*).$$

This establishes (142) (and by symmetry (143)).

The sign-symmetry $(z_- \mapsto -z_-)$ implies that if $(x_+, x_-, y^u_+(x_+, x_-, c), y^u_-(x_+, x_-, c), c) \in W^{cu}$, then $(x_+, -x_-, y^u_+(x_+, x_-, c), -y^u_-(x_+, x_-, c), c) \in W^{cu}$. Therefore

$$y^u_+(x_+, -x_-, c) = y^u_+(x_+, x_-, c), \quad y^u_-(x_+, -x_-, c) = -y^u_-(x_+, x_-, c).$$

This argument and an analogous one for the sign-symmetry $z_+ \mapsto -z_+$ establishes (144) and (145). □

From the above Lemma we immediately obtain the following result.

Lemma 39. *If the vector field (33) has the sign-symmetry then, for every* $\varepsilon \in [0, 1]$*, the transformation* T_ε *defined by (52) has the sign-symmetry.*

8.7. The solutions of the cohomological equation have the sign-symmetry

Recalling that

- $\dfrac{\partial \varphi}{\partial z}(t, z)$ and $S(t, z)$ were defined in Section 6,

- G_1 and G_2 were defined by (87) and (113), respectively,

it is obvious that G_1 and G_2 have the sign-symmetry.

8.8. Conclusion of the proof of Theorem 30

We have shown that all stages in the construction of the coordinate change of Theorem 3 have the sign-symmetry. This finishes the proof of Theorem 30.

Appendix A. Logarithmic norms and related topics

In this section we state some facts about logarithmic norms and their applications to ODEs [9,14–16].

We define for a linear map $A : V \to V$ of a normed space

$$m(A) = \inf_{\|x\|=1} \|A(x)\|. \tag{146}$$

For an interval matrix $\mathbf{A} \subset \mathbb{R}^{k\times n}$ we set

$$m(\mathbf{A}) = \inf_{A\in\mathbf{A}} m(A). \tag{147}$$

Definition 6. For a square matrix $A \in \mathbb{R}^{n\times n}$ we define the logarithmic norm of A denoted by $\mu_{\log}(A)$ by [9,14–16]

$$\mu_{\log}(A) = \lim_{h\to 0^+} \frac{\|I + hA\| - \|I\|}{h}$$

and the logarithmic minimum of A [7]

$$m_l(A) = \lim_{h\to 0^+} \frac{m(I + hA) - \|I\|}{h}.$$

It is known that $\mu_{\log}(A)$ is a continuous and convex function. Moreover, see [7, Lem. 3],

$$m_l(A) = -\mu_{\log}(-A).$$

In the following theorem a bound on the distance between solutions of an ODE is established in terms of the logarithmic norm. The proof of this result can be found in [14] (the part involving $\mu_{\log}$) and in [7, Thm. 5] for the lower bound involving m_l.

Theorem 40. *Consider an ODE*

$$x' = f(t, x), \tag{148}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is C^1.

Let $x(t)$ and $y(t)$ for $t \in [t_0, t_0 + T]$ be two solutions of (148). Let $W \subset \mathbb{R}^n$ such that for each $t \in [t_0, t_0 + T]$ the segment connecting $x(t)$ and $y(t)$ is contained in W. Let

$$L = \sup_{x \in W, t \in [t_0, t_0+T]} \mu_{\log}\left(\frac{\partial f}{\partial x}(t, x)\right),$$

$$l = \inf_{x \in W, t \in [t_0, t_0+T]} m_l\left(\frac{\partial f}{\partial x}(t, x)\right).$$

Then for $t \in [0, T]$ it holds that

$$\exp(lt)\left\|x(t_0) - y(t_0)\right\| \le \left\|x(t_0 + t) - y(t_0 + t)\right\| \le \exp(Lt)\left\|x(t_0) - y(t_0)\right\|.$$

From the above result one easy derives the following theorem.

Theorem 41. *Consider an ODE*

$$x' = f(x), \tag{149}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^1.

Let $x(t)$ be a solution of (149) and let $W \subset \mathbb{R}^n$ such that $x(t) \in W$ for each $t \in [0, T]$. Let

$$L = \sup_{x \in W} \mu_{\log}\left(\frac{\partial f}{\partial x}(t, x)\right),$$

$$l = \inf_{x \in W} m_l\left(\frac{\partial f}{\partial x}(t, x)\right)$$

Then for $t \in [0, T]$ it holds that

$$e^{lt} \le m\left(\frac{\partial \varphi}{\partial x}(t, x)\right), \qquad \left\|\frac{\partial \varphi}{\partial x}(t, x)\right\| \le \exp(Lt), \quad t \in [0, T]$$

The following theorem follows from Lemma 4.1 in [15]

Theorem 42. *Assume that either*

- *$z : [0, T] \to \mathbb{R}^k$ is a solution of the equation*

$$z'(t) = A(t) \cdot z(t) + \delta(t), \tag{150}$$

where $\delta : [0, T] \to \mathbb{R}^k$ and $A : [0, T] \to \mathbb{R}^{k \times k}$ are continuous.

or

- *$z : [0, T] \to \mathbb{R}^{k \times k}$ is a solution of the equation*

$$z'(t) = A(t) \cdot z(t) + \delta(t), \tag{151}$$

or the equation

$$z'(t) = z(t) \cdot A(t) + \delta(t), \tag{152}$$

where $\delta : [0, T] \to \mathbb{R}^{k\times k}$ *and* $A : [0, T] \to \mathbb{R}^{k\times k}$ *are continuous.*

Assume that the continuous functions $J : [0, T] \to \mathbb{R}$ *and* $C : [0, T] \to \mathbb{R}_+$ *satisfy the following inequalities for all* $t \in [0, T]$

$$\mu_{\log}(A(t) \leq J(t), \quad |\delta(t)| \leq C(t).$$

Then

$$|z(t)| \leq y(t)$$

where $y : [0, T] \to \mathbb{R}^n$ *is a solution of the problem*

$$y'(t) = J(t)y(t) + C(t), \qquad y(0) = |z(0)|.$$

While the result for (150) is indeed a direct consequence of Lemma 4.1 in [15], the statements for (151) and (152) are obtained by the same reasoning which led to Lemma 4.1 in [15].

From Theorem 42 we immediately obtain the following result.

Theorem 43. *Under the same assumptions about* $z(t)$, $A(t)$, $\delta(t)$, $C(t)$ *as in Theorem 42, assume that there exists* $\alpha \in \mathbb{R}$ *such that*

$$\mu_{\log}(A(t)) \leq \alpha, \quad t \in [0, T].$$

Then

$$|z(t)| \leq e^{\alpha t}|z(0)| + \int_0^t e^{\alpha(t-s)}C(s)ds. \tag{153}$$

In the previous results, the choice of norms has been arbitrary. We now apply these results to the case where the norm is Euclidean. In this case we have the following formula for the logarithmic norm of a matrix.

Lemma 44. *If we choose the Euclidean norm for the computation of* $\mu_{\log}(A)$*, then*

$$\mu_{\log}(A) = \max\{\lambda \in \textit{spectrum of } (A + A^\top)/2\} = \max_{\|x\|=1} (Ax, x) \tag{154}$$

$$m_l(A) = \min\{\lambda \in \textit{spectrum of } (A + A^\top)/2\} = \min_{\|x\|=1} (Ax, x). \tag{155}$$

Appendix B. Some results about normally hyperbolic invariant manifolds

The content of this section is taken by adapting the results of [6] for the maps case and of [7] for the ODE case. In those papers, the normally hyperbolic invariant manifold was $\Lambda = \mathbb{T}^c$, while here we consider $\Lambda = \mathbb{R}^c$.

In [6,7] the covering map for a c-dimensional torus

$$\varphi : \mathbb{R}^c \to \Lambda = \mathbb{R}/((2R_\Lambda) \cdot \mathbb{Z})^c,$$

provided a global coordinate system for the whole torus. Given a map f or a vector field on $\Lambda \times \mathbb{R}^u \times \mathbb{R}^s$, it can be lifted to $\mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s$ and all the assumptions in [6,7] are formulated in terms of the 'lifted' f. The lifted f are periodic in the first variable, but nevertheless, all the techniques and conclusions also apply to the situation when we have uniform bounds in the center direction (i.e., $\mathbb{R}^c$), and are therefore applicable to global maps or vector fields defined on $\mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s$. In addition, some topological conditions that appear in the case of the torus can be dropped in the global case.

Throughout this section we shall use the notation $z = (\lambda, x, y)$ to denote points in $\mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s$. This means that the notation λ will stand for points on Λ, x for points in $\mathbb{R}^u$, and y for points in $\mathbb{R}^s$. We will write f as (f_λ, f_x, f_y), where f_λ, f_x, f_y stand for projections onto Λ, $\mathbb{R}^u$ and $\mathbb{R}^s$, respectively. On $\mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s$ we will use the Euclidian norm.

B.1. NHIMs for maps

Given $U \subset \mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s$ we consider a C^{k+1} map, for $k \geq 1$,

$$f : U \to \mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s.$$

In [6] where we considered tori, a number R_Λ being the radius (in fact period of lifted f in λ-direction) was introduced. Now when considering $\Lambda = \mathbb{R}^n$ we can take $R_\Lambda = \infty$.

For any $R > 0$ (in [6] it was required that $R < \frac{1}{2} R_\Lambda$) denote by $D = D(R)$ the set

$$D(R) = \Lambda \times \overline{B}_u(R) \times \overline{B}_s(R),$$

where $\overline{B}_n(R)$ stands for a closed ball of radius R, centered at zero, in $\mathbb{R}^n$.

We will now define several constants ξ and μ related to the expansion and contraction rates of f on D along several directions. These constants are defined on the set D and depend on L (this will measure slopes) and the notion of $P(z)$ for $z \in D$.

In [6, page 6220] it was asked that $L \in \left(\frac{2R}{R_\Lambda}, 1\right)$, but in our global case we can have arbitrary L and we can set

$$P(q) = D.$$

The restriction on L and the notion of $P(q)$ were introduced in order to have a projection of cones with slopes L or $1/L$ to be on the same chart (defined by φ) on Λ and $P(q)$ was a preimage of all points that can be the same chart as point q. For the same reason we also drop the backward cone condition (Def. 14 in [6]).

Before presenting definition of various constants that describe expansion and contraction rates we need to introduce one more notation.

Definition 7. Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a C^1 function. We define the interval enclosure of the derivative Df on $U \subset \mathbb{R}^n$ as the set $[Df(U)] \subset \mathbb{R}^{k\times n}$, defined as

$$[Df(U)] = \left\{ A = \left(a_{ij}\right)_{\substack{i=1,\dots,k\\ j=1,\dots,n}} : a_{ij} \in \left[\inf_{x\in U} \frac{\partial f_j}{\partial x_j}(x), \sup_{x\in U} \frac{\partial f_j}{\partial x_j}(x) \right] \right\}.$$

Recall that in (146) and (147) we already defined $m(A)$ and $m([DF(U)])$. These notations are used below.

For $L > 0$ and the set D we define (see [6, page 6220])

$$\mu_{s,1} = \sup_{z\in D} \left\{ \left\| \frac{\partial f_y}{\partial y}(z) \right\| + \frac{1}{L} \left\| \frac{\partial f_y}{\partial(\lambda, x)}(z) \right\| \right\},$$

$$\mu_{s,2} = \sup_{z\in D} \left\{ \left\| \frac{\partial f_y}{\partial y}(z) \right\| + L \left\| \frac{\partial f_{(\lambda,x)}}{\partial y}(z) \right\| \right\},$$

$$\xi_{u,1} = \inf_{z\in D} \left\{ m\left(\frac{\partial f_x}{\partial x}(z) \right) - \frac{1}{L} \left\| \frac{\partial f_x}{\partial(\lambda, y)}(z) \right\| \right\},$$

$$\xi_{u,1,P} = \inf_{z\in D} m \left[\frac{\partial f_x}{\partial x}(P(z)) \right] - \frac{1}{L} \sup_{z\in D} \left\| \frac{\partial f_x}{\partial(\lambda, y)}(z) \right\|,$$

$$\xi_{u,2} = \inf_{z\in D} \left\{ m\left(\frac{\partial f_x}{\partial x}(z) \right) - L \left\| \frac{\partial f_{(\lambda,y)}}{\partial x}(z) \right\| \right\},$$

$$\mu_{cs,1} = \sup_{z\in D} \left\{ \left\| \frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z) \right\| + L \left\| \frac{\partial f_{(\lambda,y)}}{\partial x}(z) \right\| \right\},$$

$$\mu_{cs,2} = \sup_{z\in D} \left\{ \left\| \frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z) \right\| + \frac{1}{L} \left\| \frac{\partial f_x}{\partial(\lambda, y)}(z) \right\| \right\},$$

$$\xi_{cu,1} = \inf_{z\in D} \left\{ m\left(\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(z) \right) - L \left\| \frac{\partial f_{(\lambda,x)}}{\partial y}(z) \right\| \right\},$$

$$\xi_{cu,1,P} = \inf_{z\in D} m \left[\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(P(z)) \right] - L \sup_{z\in D} \left\| \frac{\partial f_{(\lambda,x)}}{\partial y}(z) \right\|,$$

$$\xi_{cu,2} = \inf_{z\in D} \left\{ m\left(\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(z) \right) - \frac{1}{L} \left\| \frac{\partial f_y}{\partial(\lambda, x)}(z) \right\| \right\}.$$

Definition 8. [6, Def. 5] We say that f satisfies rate conditions of order $k \geq 1$ if $\xi_{u,1}$, $\xi_{u,1,P}$, $\xi_{u,2}$, $\xi_{cu,1}$, $\xi_{cu,1,P}$, $\xi_{cu,2}$ are strictly positive, and for all $k \geq j \geq 1$ it holds that

$$\mu_{s,1} < 1 < \xi_{u,1,P},$$

$$\frac{\mu_{cs,1}}{\xi_{u,1,P}} < 1, \qquad \frac{\mu_{s,1}}{\xi_{cu,1,P}} < 1,$$
$$\frac{\left(\mu_{cs,1}\right)^{j+1}}{\xi_{u,2}} < 1, \qquad \frac{\mu_{s,2}}{(\xi_{cu,1})^{j+1}} < 1,$$
$$\frac{\mu_{cs,2}}{\xi_{u,1}} < 1, \qquad \frac{\mu_{s,1}}{\xi_{cu,2}} < 1.$$

We introduce the following notation:

$$J_s(z, M) = \left\{(\lambda, x, y) : \left\|(\lambda, x) - \pi_{\lambda,x} z\right\| \leq M \left\|y - \pi_y z\right\|\right\},$$
$$J_u(z, M) = \left\{(\lambda, x, y) : \left\|(\lambda, y) - \pi_{\lambda,y} z\right\| \leq M \left\|x - \pi_x z\right\|\right\}.$$

We shall refer to $J_s(z, M)$ as a stable cone of slope M at z, and to $J_u(z, M)$ as an unstable cone of slope M at z.

Definition 9. We say that a sequence $\{z_i\}_{i=-\infty}^{0}$ is a (full) backward trajectory of a point z if $z_0 = z$, and $f(z_{i-1}) = z_i$ for all $i \leq 0$.

Definition 10. We define the local center-stable set in D as

$$W^{cs}_{\text{loc}} = \left\{z : f^n(z) \in D \text{ for all } n \in \mathbb{N}\right\}.$$

Definition 11. We define the local center-unstable set in D as

$$W^{cu}_{\text{loc}} = \{z : \text{there is a full backward trajectory of } z \text{ in } D\}.$$

Definition 12. We define the maximal invariant set in D as

$$\Lambda^{*}_{\text{loc}} = \{z : \text{there is a full trajectory of } z \text{ in } D\}.$$

Definition 13. Assume that $z \in W^{cs}_{\text{loc}}$. We define the local stable fiber of z as

$$W^{s}_{z,\text{loc}} = \left\{p \in D : f^n(p) \in J_s\left(f^n(z), 1/L\right) \cap D \text{ for all } n \in \mathbb{N}\right\}.$$

Definition 14. Assume that $z \in W^{cu}_{\text{loc}}$. We define the local unstable fiber of z as

$$W^{u}_{z,\text{loc}} = \Big\{p \in D : \exists \text{ backward trajectory } \{p_i\}_{i=-\infty}^{0} \text{ of } p \text{ in } D, \text{ and for any such}$$
$$\text{backward trajectory } \{z_i\}_{i=-\infty}^{0} \text{ of } z \text{ in } D \text{ it holds that } p_i \in J_u(z_i, 1/L) \cap D\Big\}.$$

The definitions of $W^{s}_{z,\text{loc}}$ and $W^{u}_{z,\text{loc}}$ are related to cones, which is a nonstandard approach, as the standard one is through convergence rates. In [6] it is shown that it implies the convergence rate as in the standard theory.

Under above assumptions it will turn out that f is injective on W^{cu}_{loc}. Therefore the backward orbit in the definition of $W^{u}_{z,\text{loc}}$ is unique.

In [6, page 6223] for given $\lambda \in \Lambda$ the set $D_\lambda^{\pm}$ was defined. Now when $\Lambda = \mathbb{R}^c$ we do not need to 'localize' D in a good chart and instead we define the following sets:

$$D^+ = \mathbb{R}^c \times \overline{B}_u(R) \times \partial B_s(R),$$
$$D^- = \mathbb{R}^c \times \partial \overline{B}_u(R) \times B_s(R).$$

This is a 'global' modification of Def. 15 in [6]

Definition 15. We say that f satisfies covering conditions (on D) if there exist a homotopy h

$$h : [0,1] \times D \to \mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s,$$

a $\lambda^* \in \mathbb{R}^c$ and a linear map $A : \mathbb{R}^u \to \mathbb{R}^u$ which satisfy:

1. $h_0 = f|_D$,
2. for any $\alpha \in [0,1]$,

$$h_\alpha\left(D^-\right) \cap D = \emptyset,$$
$$h_\alpha(D) \cap D^+ = \emptyset,$$

3. $h_1(\lambda, x, y) = (\lambda^*, Ax, 0)$,
4. $A(\partial B_u(R)) \subset \mathbb{R}^u \setminus \overline{B}_u(R)$.

The following theorem is Theorem 16 in [6]. The only change is that now $\Lambda = \mathbb{R}^c$, we assume the boundedness of $\|D^j f\|$ for $j = 1, \ldots, k+1$ on D and we drop the backward cone condition.

Theorem 45. *Let $\Lambda = \mathbb{R}^c$, $k \geq 1$ and $f : D \to \Lambda \times \mathbb{R}^u \times \mathbb{R}^s$ be a C^{k+1} map. If f satisfies covering conditions and rate conditions of order k, and $\|D^j f\|$ are bounded on D for $j = 1, \ldots, k$, then W^{cs}_{loc}, W^{cu}_{loc} and Λ^* are C^k manifolds, which are graphs of C^k functions*

$$w^{cs}_{\text{loc}} : \Lambda \times \overline{B}_s(R) \to \overline{B}_u(R),$$
$$w^{cu}_{\text{loc}} : \Lambda \times \overline{B}_u(R) \to \overline{B}_s(R),$$
$$\chi : \Lambda \to \overline{B}_u(R) \times \overline{B}_s(R),$$

meaning that

$$W^{cs}_{\text{loc}} = \left\{\left(\lambda, w^{cs}_{\text{loc}}(\lambda, y), y\right) : \lambda \in \Lambda, y \in \overline{B}_s(R)\right\},$$
$$W^{cu}_{\text{loc}} = \left\{\left(\lambda, x, w^{cu}_{\text{loc}}(\lambda, y)\right) : \lambda \in \Lambda, x \in \overline{B}_u(R)\right\},$$
$$\Lambda^* = \{(\lambda, \chi(\lambda)) : \lambda \in \Lambda\}.$$

Moreover, if f restricted to W^{cu}_{loc} is an injection, w^{cs}_{loc} and w^{cu}_{loc} are Lipschitz with constants L, and χ is Lipschitz with the constant $\sqrt{2}L/\sqrt{1-L^2}$. The manifolds W^{cs}_{loc} and W^{cu}_{loc} intersect transversally on Λ^*, and $W^{cs}_{\text{loc}} \cap W^{cu}_{\text{loc}} = \Lambda^*$.

The manifolds W^{cs}_{loc} and W^{cu}_{loc} are foliated by invariant fibers $W^{s}_{z,\mathrm{loc}}$ and $W^{u}_{z,\mathrm{loc}}$, which are graphs of C^k functions

$$w^{s}_{z,\mathrm{loc}} : \overline{B}_s(R) \to \Lambda \times \overline{B}_u(R),$$
$$w^{u}_{z,\mathrm{loc}} : \overline{B}_u(R) \to \Lambda \times \overline{B}_s(R),$$

meaning that

$$W^{s}_{z,\mathrm{loc}} = \left\{ \left(w^{s}_{z,\mathrm{loc}}(y), y \right) : y \in \overline{B}_s(R) \right\},$$
$$W^{u}_{z,\mathrm{loc}} = \left\{ \left(\pi_\lambda w^{u}_{z,\mathrm{loc}}(x), x, \pi_y w^{u}_{z,\mathrm{loc}}(x) \right) : x \in \overline{B}_u(R) \right\}.$$

The functions $w^{s}_{z,\mathrm{loc}}$ and $w^{u}_{z,\mathrm{loc}}$ are Lipschitz with constants $1/L$. Moreover,

$$W^{s}_{z,\mathrm{loc}} = \left\{ p \in D : f^n(p) \in D \text{ for all } n \geq 0, \text{ and } \exists n_0, \exists C > 0 \text{ (which can depend on p)} \right.$$
$$\left. \text{s.t. for } n \geq n_0, \left\| f^n(p) - f^n(z) \right\| \leq C \mu^n_{s,1} \right\},$$

and if $\{z_i\}_{i=-\infty}^{0}$ is the unique backward trajectory of z in D, then

$$W^{u}_{z,\mathrm{loc}} = \left\{ p \in W^{cu}_{\mathrm{loc}} : \text{such that the unique backward trajectory } \{p_i\}_{i=-\infty}^{0} \text{ of } p \text{ in } D \text{ satisfies that} \right.$$
$$\left. \exists n_0 \geq 0, \exists C > 0 \text{ (which can depend on p) s.t. for } n \geq n_0, \| p_{-n} - z_{-n} \| \leq C \xi^{-n}_{u,1,P} \right\}.$$

Remark 46. The above theorem says nothing about the dependence of $W^{s,u}_{z,\mathrm{loc}}$ on z.

From the proof of Theorem 45 we immediately obtain the following result.

Remark 47. In the context of Theorem 45 the following holds. For $j = 1, \ldots, k$ the derivatives of the functions w^{cu}_{loc}, w^{cs}_{loc}, $w^{u,s}_{z,\mathrm{loc}}$ of order j are bounded by some constants K_j depending on $\|D^s f\|$ for $s = 1, \ldots, j$ and constants L, ξ and μ.

While in [6] there is no explicit general formula for K_j, the explicit bound for second derivatives has been worked out in Theorem 23 in [7].

Observe that we obtain L or $1/L$ as bounds for the Lipschitz constants for the functions w^{cu}_{loc}, w^{cs}_{loc}, w^{u}_{loc}, w^{s}_{loc}. Hence when L is small we get quite big bounds for some slopes. This is clearly an overestimate for the case when $\mathbb{T} \times \{0\} \times \{0\}$ is our NHIM. This a consequence of the choices that have been made when formulating Theorem 45, as the authors in [6] did not introduce different constants for each type of cones, plus several inequalities between them. However, below following [6] we give conditions which allow us to obtain better Lipschitz constants.

Theorem 48. *[6, Thm. 17]*

$$\mu = \sup_{z \in D} \left\{ \left\| \frac{\partial f_y}{\partial y}(z) \right\| + M \left\| \frac{\partial f_y}{\partial(\lambda, x)}(z) \right\| \right\},$$
$$\xi = \inf_{z \in D} m \left(\left[\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(P(z)) \right] \right) - \frac{1}{M} \sup_{z \in D} \left\| \frac{\partial f_{(\lambda,x)}}{\partial y}(z) \right\|.$$

If assumptions of Theorem 45 hold true and also $\xi/\mu > 1$, then the function $w^s_{z,\mathrm{loc}}$ from Theorem 45 is Lipschitz with constant M.

Theorem 49. *[6, Thm. 18]*

$$\xi = \inf_{z\in D} m\left(\left[\frac{\partial f_x}{\partial x}(P(z))\right]\right) - M \sup_{z\in D}\left\|\frac{\partial f_x}{\partial(\lambda, y)}(z)\right\|,$$
$$\mu = \sup_{z\in D}\left\{\left\|\frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z)\right\| + \frac{1}{M}\left\|\frac{\partial f_{(\lambda,y)}}{\partial x}(z)\right\|\right\}.$$

If assumptions of Theorem 45 hold true and also $\xi/\mu > 1$, then the function $w^u_{z,\mathrm{loc}}$ from Theorem 45 is Lipschitz with constant M.

Theorem 50. *[6, Thm. 19]*

$$\xi = \inf_{z\in D} m\left[\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(P(z))\right] - M \sup_{z\in D}\left\|\frac{\partial f_{(\lambda,x)}}{\partial y}(z)\right\|,$$
$$\mu = \sup_{z\in D}\left\{\left\|\frac{\partial f_y}{\partial y}(z)\right\| + \frac{1}{M}\left\|\frac{\partial f_y}{\partial(\lambda, x)}(z)\right\|\right\}.$$

If assumptions of Theorem 45 hold true and also $\xi/\mu > 1$, then the function w^{cu}_{loc} from Theorem 45 is Lipschitz with constant M.

Theorem 51. *[6, Thm. 20]*

$$\xi = \inf_{z\in D} m\left[\frac{\partial f_x}{\partial x}(P(z))\right] - \frac{1}{M}\sup_{z\in D}\left\|\frac{\partial f_x}{\partial(\lambda, y)}(z)\right\|,$$
$$\mu = \sup_{z\in D}\left\{\left\|\frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z)\right\| + M\left\|\frac{\partial f_{(\lambda,y)}}{\partial x}(z)\right\|\right\}$$

If assumptions of Theorem 45 hold true and also $\xi/\mu > 1$, then the function w^{cs}_{loc} from Theorem 45 is Lipschitz with constant M.

B.1.1. Comments on the inequalities

Let $J^c_s(z, M)$ and $J^c_u(z, M)$ stand for the complements of $J_s(z, M)$ and $J_u(z, M)$, respectively. We now discuss the meaning of several inequalities on Definition 8 of rate conditions as well as where they are needed (see section 3.3 in [6] for more details):

- $\mu_{cs,1} < \xi_{u,1,P}$: the forward invariance of $J_u(z, 1/L)$. $\xi_{u,1,P} > 1$: the expansion in $J_u(z, 1/L)$ for x - coordinate. This is needed for the proof of the existence of W^{cs}.
- $\xi_{cu,1,P} > \mu_{s,1}$: the forward invariance of $J^c_s(z, 1/L)$. $\mu_{s,1} < 1$: the contraction in y-direction in $J_s(z, 1/L)$. This is needed for the proof of the existence of W^{cu}.
- $\frac{\mu_{s,2}}{(\xi_{cu,1})^{j+1}} < 1$, $j = 1, \dots, k$: the C^k-smoothness of W^{cu}.
- $\frac{(\mu_{cs,1})^{j+1}}{\xi_{u,2}} < 1$, $j = 1, \dots, k$: the C^k-smoothness of W^{cs}.

- $\frac{\mu_{cs,1}}{\xi_{u,1,P}} < 1$: the existence of fibers W_q^u. $\frac{\mu_{cs,2}}{\xi_{u,1}} < 1$: the C^k smoothness of W_q^u.
- $\frac{\mu_{s,1}}{\xi_{cu,1,P}} < 1$: the existence of fibers W_q^s. $\frac{\mu_{s,1}}{\xi_{cu,2}} < 1$: the C^k smoothness of W_q^s.

B.2. NHIMs for ODEs

The content of this section is partially based on Section 5 in [7], where the results from [6] about maps were adapted to the context of ODEs. The goal of this section is to describe how these results extend to the case of $\Lambda = \mathbb{R}^c$.

We consider an ODE

$$q' = f(q), \tag{156}$$

where

$$f : \Lambda \times \mathbb{R}^u \times \mathbb{R}^s \to \mathbb{R}^c \times \mathbb{R}^u \times \mathbb{R}^s.$$

We denote by $\Phi(t, q)$ the flow induced by (156).

We define the notion of an isolating block, which is an ODE version of Definition 15.

Definition 16. [7, Def. 19] We say that $D(r) = \Lambda \times \overline{B}_u(r) \times \overline{B}_s(r)$ is an isolating block for f if

1. Exit. For any $q \in \Lambda \times \partial\overline{B}_u(r) \times \overline{B}_s(r)$,

$$(\pi_x F_\varepsilon(q)|\pi_x q) > 0. \tag{157}$$

2. Entry. For any $q \in \Lambda \times \overline{B}_u(r) \times \partial\overline{B}_s(r)$,

$$\left(\pi_y F_\varepsilon(q)|\pi_y q\right) < 0. \tag{158}$$

Isolating blocks are important constructs in the Conley index theory [17]. Intuitively, in Definition 16 the set $\Lambda \times \partial\overline{B}_u(r) \times \overline{B}_s(r)$ plays the role of the exit set, and $\Lambda \times \overline{B}_u(r) \times \partial\overline{B}_s(r)$ of the entry set.

Definition 17. We define the center-unstable set of ode (156) in D as

$$W^{cu}_{\mathrm{loc},D} = \{q : \Phi(t, q) \in D \text{ for all } t < 0\},$$

and the center-stable set of ode (156) in D as

$$W^{cs}_{\mathrm{loc},D} = \{q : \Phi(t, q) \in D \text{ for all } t > 0\}.$$

We shall consider the set

$$D = \Lambda \times \overline{B}_u(R) \times \overline{B}_s(R).$$

Since the set D will remain fixed throughout the discussion, from now on we will simplify notation by writing W^{cu} instead of $W^{cu}_{\mathrm{loc},D}$.

We will now define the constants corresponding to ξ and μ for the map case.

The rule is that the corresponding constant for ξ will be $\overrightarrow{\xi}$, and analogously for μ.

The definitions are such that if we consider a map $\Phi(h, \cdot)$ i.e., the flow induced by (156) by $h > 0$, then for all constants ξ and μ it holds that (see Thm. 31 in [7])

$$\mu = 1 + h\overrightarrow{\mu} + O(h^2),$$
$$\xi = 1 + h\overrightarrow{\xi} + O(h^2).$$

Therefore following [7] we introduce the following constants ($\mu_{\log}()$ and $m_l()$ are defined in Appendix A)

$$\overrightarrow{\mu_{s,1}} = \sup_{z\in D}\left\{\mu_{\log}\left(\frac{\partial f_y}{\partial y}(z)\right) + \frac{1}{L}\left\|\frac{\partial f_y}{\partial(\lambda, x)}(z)\right\|\right\},$$

$$\overrightarrow{\mu_{s,2}} = \sup_{z\in D}\left\{\mu_{\log}\left(\frac{\partial f_y}{\partial y}(z)\right) + L\left\|\frac{\partial f_{(\lambda,x)}}{\partial y}(z)\right\|\right\},$$

$$\overrightarrow{\xi_{u,1}} = \inf_{z\in D}\left\{m_l\left(\frac{\partial f_x}{\partial x}(z)\right) - \frac{1}{L}\left\|\frac{\partial f_x}{\partial(\lambda, y)}(z)\right\|\right\},$$

$$\overrightarrow{\xi_{u,1,P}} = \inf_{z\in D} m_l\left(\frac{\partial f_x}{\partial x}(P(z))\right) - \frac{1}{L}\sup_{z\in D}\left\|\frac{\partial f_x}{\partial(\lambda, y)}(z)\right\|,$$

$$\overrightarrow{\mu_{cs,1}} = \sup_{z\in D}\left\{\mu_{\log}\left(\frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z)\right) + L\left\|\frac{\partial f_{(\lambda,y)}}{\partial x}(z)\right\|\right\},$$

$$\overrightarrow{\mu_{cs,2}} = \sup_{z\in D}\left\{\mu_{\log}\left(\frac{\partial f_{(\lambda,y)}}{\partial(\lambda, y)}(z)\right) + \frac{1}{L}\left\|\frac{\partial f_x}{\partial(\lambda, y)}(z)\right\|\right\},$$

$$\overrightarrow{\xi_{cu,1}} = \inf_{z\in D}\left\{m_l\left(\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(z)\right) - L\left\|\frac{\partial f_{(\lambda,x)}}{\partial y}(z)\right\|\right\},$$

$$\overrightarrow{\xi_{cu,1,P}} = \inf_{z\in D} m_l\left(\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(P(z))\right) - L\sup_{z\in D}\left\|\frac{\partial f_{(\lambda,x)}}{\partial y}(z)\right\|,$$

$$\overrightarrow{\xi_{u,2}} = \inf_{z\in D}\left\{m_l\left(\frac{\partial f_x}{\partial x}(z)\right) - L\left\|\frac{\partial f_{(\lambda,y)}}{\partial x}(z)\right\|\right\},$$

$$\overrightarrow{\xi_{cu,2}} = \inf_{z\in D} m_l\left(\frac{\partial f_{(\lambda,x)}}{\partial(\lambda, x)}(z)\right) - \frac{1}{L}\sup_{z\in D}\left\|\frac{\partial f_y}{\partial(\lambda, x)}(z)\right\|.$$

We define the rate conditions for ODEs as follows.

Definition 18. [7, Def. 28] We say that the vector field f satisfies rate conditions of order $k \geq 1$ if for all $k \geq j \geq 1$ it holds that

$$\overrightarrow{\mu_{s,1}} < 0 < \overrightarrow{\xi_{u,1,P}},$$

$$\begin{aligned} \overrightarrow{\mu_{cs,1}} &< \overrightarrow{\xi_{u,1,P}}, & \overrightarrow{\mu_{s,1}} &< \overrightarrow{\xi_{cu,1,P}}, \\ \overrightarrow{\mu_{s,2}} &< (j+1)\overrightarrow{\xi_{cu,1}}, & \overrightarrow{\mu_{cs,2}} &< \overrightarrow{\xi_{u,1}}, \\ (j+1)\overrightarrow{\mu_{cs,1}} &< \overrightarrow{\xi_{u,2}} & \overrightarrow{\mu_{s,1}} &< \overrightarrow{\xi_{cu,2}}. \end{aligned} \tag{159}$$

In fact in [7, Def. 28] conditions (159) are missing. Because there the focus was just on W^{cu}, and dealing with W^{cs} requires (159).

There is no ODE-analogue of Theorem 45 in [7], only a part of it related to W^{cu} is stated there as Theorem 30. However, the full result stated below can in the same way be derived from Theorem 45.

Theorem 52. *Let $k \geq 1$ and assume that f is C^{k+1}. Assume that $D = \Lambda \times \overline{B}_u(R) \times \overline{B}_s(R)$ is an isolating block for f and f has bounded derivatives up to order $k+1$ on D and rate conditions of order k are satisfied on D for some constant L.*

Then W^{cs}, W^{cu} and Λ^* are C^k manifolds, which are graphs of C^k functions

$$\begin{aligned} w^{cs} &: \Lambda \times \overline{B}_s(R) \to \overline{B}_u(R), \\ w^{cu} &: \Lambda \times \overline{B}_u(R) \to \overline{B}_s(R), \\ \chi &: \Lambda \to \overline{B}_u(R) \times \overline{B}_s(R), \end{aligned}$$

meaning that

$$\begin{aligned} W^{cs} &= \left\{ \left(\lambda, w^{cs}(\lambda, y), y\right) : \lambda \in \Lambda, y \in \overline{B}_s(R) \right\}, \\ W^{cu} &= \left\{ \left(\lambda, x, w^{cu}(\lambda, y)\right) : \lambda \in \Lambda, x \in \overline{B}_u(R) \right\}, \\ \Lambda^* &= \{(\lambda, \chi(\lambda)) : \lambda \in \Lambda\}. \end{aligned}$$

Moreover, w^{cs} and w^{cu} are Lipschitz with constants L, and χ is Lipschitz with the constant $\sqrt{2}L/\sqrt{1-L^2}$. The manifolds W^{cs} and W^{cu} intersect transversally on Λ^*, and $W^{cs} \cap W^{cu} = \Lambda^*$.

The manifolds W^{cs} and W^{cu} are foliated by invariant fibers W^s_z and W^u_z, which are graphs of C^k functions

$$\begin{aligned} w^s_z &: \overline{B}_s(R) \to \Lambda \times \overline{B}_u(R), \\ w^u_z &: \overline{B}_u(R) \to \Lambda \times \overline{B}_s(R), \end{aligned}$$

meaning that

$$\begin{aligned} W^s_z &= \left\{ \left(w^s_z(y), y\right) : y \in \overline{B}_s(R) \right\}, \\ W^u_z &= \left\{ \left(\pi_\lambda w^u_z(x), x, \pi_y w^u_z(x)\right) : x \in \overline{B}_u(R) \right\}. \end{aligned}$$

The functions w^s_z and w^u_z are Lipschitz with constants $1/L$. Moreover,

$$W_z^s = \{p \in D : \Phi(t,p) \in D \text{ for all } t \geq 0, \text{ and } \exists T_0 \geq 0, \exists C > 0 \text{ (which can depend on p)} \\ \text{s.t. for } t \geq T_0, \|\Phi(t,p) - \Phi(t,z)\| \leq C \exp(t\overrightarrow{\mu_{s,1}})\Big\},$$

and

$$W_z^u = \Big\{p \in W^{cu} : \exists T_0 \leq 0, \exists C > 0 \text{ (which can depend on p)} \\ \text{s.t. for } t \leq T_0, \|\Phi(t,p) - \Phi(t,z)\| \leq C \exp(t\overrightarrow{\xi_{u,1,P}})\Big\}.$$

For $j = 1, \ldots, k$ the derivatives of the functions w^{cu}, w^{cs}, $w_z^{u,s}$ of order j are bounded by some constants $C_j = O(\sup_{s=1,\ldots,j} \sup_{z \in D} \|D^s f\|)$ and constants L, $\overrightarrow{\xi}$ and $\overrightarrow{\mu}$.

Appendix C. Comparison with other works

C.1. Comparison with [3,4]

In [3] we have the following formulas for Q_0 and q_0 (denoted there by K) are given (see [3, page 63, Theorem 3.20 in Chapter II Sec. 2.3] and ([3, II.3.2 page 48])

$$Q_0(k) = \left[\frac{\lambda_{\max}}{\lambda_{min}} + k\left(\frac{\mu_{max}}{\lambda_{min}} + 1\right)\right] + \left[\frac{\mu_{\max}}{\mu_{min}} + k\left(\frac{\lambda_{\max}}{\mu_{min}} + 1\right)\right] + 2, \quad (160)$$
$$q_0(k) = Q_0(k) + k.$$

It is clear that the value of Q_0 obtained by us (i.e., (13)) is smaller than that of (160). The difference seems to be equal to two. Our value of q_0 is also better.

The paper [4], inspired by [5], is an improvement of results from [3], as the authors consider the preservation of volume, symplectic form or contact structure. In terms of Q_0 and q_0 the results appear to be better than the ones from [3], however the case of a general vector field is not considered. For the Hamiltonian vector field they obtain (see [4, Theorem 8.1]) the following inequalities (we use our notation)

$$Q_0(k) + 1 = q_0(k) > \frac{2\lambda_{max}(k+1)}{\lambda_{min}} + 2. \quad (161)$$

Since for Hamiltonian systems $\lambda_{max} = \mu_{max}$ and $\lambda_{min} = \mu_{min}$ we see that (160) gives

$$Q_0(k) \geq \frac{2\lambda_{max}(k+1)}{\lambda_{min}} + 2k,$$

hence (161) is an improvement for Hamiltonian systems. As a result their numbers for Hamiltonian systems are better than ours.

Let us also stress that in [3,4] the authors show also how some resonant terms can be removed by finitely smooth coordinate change, which we are not discussing in our work. In our paper [11] (and also in [5]) we deal with the question of removing the remainder, only.

C.2. Comparison with [5]

Theorem 1.2 in [5] gives (we use our notation) the following. Let

$$A = \frac{\lambda_{min}}{\mu_{max}} \frac{\mu_{min}}{\lambda_{max} + \mu_{min}}$$

and let B for the general vector field be given by

$$B_{no-struct} = \frac{\mu_{max}^2 + \mu_{min}(\lambda_{max} - \lambda_{min})}{\mu_{max}(\lambda_{max} + \mu_{min})}$$

and for Hamiltonian vector fields

$$B_{ham} = 1 - 2A.$$

Then the relation between k and Q claimed by the authors is

$$1 \leq k \leq QA - B.$$

Therefore

$$Q_0(k) \geq \frac{k + B}{A}.$$

Moreover,

$$q_0 \geq Q_0 + 2.$$

It appears that the above bounds are considerably better than the ones obtained by us or in [3,4]. For example when $\mu_{min} = \mu_{max} = \lambda_{min} = \lambda_{max}$ using the above formulas in the general vector field case one obtains $A = B = \frac{1}{2}$, hence

$$Q_0(k) = 2k + 1, \qquad q_0 = 2k + 3.$$

While our formulas give (see Section 7.1)

$$Q_0(k) = 4k + 2.$$

Therefore the technique from [5] appears to produce better constants for the pure saddle case. We have found it difficult at adopt it to the case when center directions are present.

Data availability

No data was used for the research described in the article.

References

[1] V.I. Arnol'd, Geometrical Methods in the Theory of Ordinary Differential Equations, Springer, 1977. Chapt. V.

[2] V.I. Arnol'd, Mathematical Methods of Classical Mechanics, Graduate Texts in Mathematics, vol. 60, Springer, 1978. Appendix: Normal forms of quadratic Hamiltonians.

[3] I.U. Bronstein, A.Ya. Kopanskii, Smooth Invariant Manifolds and Normal Forms, World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises, vol. 7, World Scientific Publishing Co., Inc., River Edge, NJ, 1994, xii+384 pp.

[4] I.U. Bronstein, A.Y. Kopanskii, Normal forms of vector fields satisfying certain geometric conditions, in: H.W. Broer, S.A. van Gils, I. Hoveijn, F. Takens (Eds.), Nonlinear Dynamical Systems and Chaos, in: Progress in Nonlinear Differential Equations and Their Applications, vol. 19, Birkhäuser, Basel, 1996.

[5] A. Banyaga, R. de la Llave, C.E. Wayne, Cohomology equations near hyperbolic points and geometric versions of Sternberg linearization theorem, J. Geom. Anal. 6 (1996) 613–649.

[6] M. Capiński, P. Zgliczyński, Geometric proof for normally hyperbolic invariant manifolds, J. Differ. Equ. 259 (2015) 6215–6286.

[7] M. Capiński, P. Zgliczyński, Beyond the Melnikov method: a computer assisted approach, J. Differ. Equ. 262 (2017) 365–417.

[8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation, Invent. Math. 181 (2010) 39–113.

[9] G. Dahlquist, Stability and Error Bounds in the Numerical Intgration of Ordinary Differential Equations, Almqvist & Wiksells, Uppsala, 1958; Transactions of the Royal Institute of Technology, Stockholm, 1959.

[10] A. Delshams, A. Simon, P. Zgliczyński, Shadowing of non-transversal heteroclinic chains, J. Differ. Equ. 264 (2018) 3619–3663.

[11] A. Delshams, P. Zgliczyński, Shadowing of non-transversal heteroclinic chains in lattices, arXiv preprint https://arxiv.org/abs/2406.00398, 2024.

[12] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, C. Simó, Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem, J. Dyn. Differ. Equ. 77 (1) (1989) 167–198.

[13] F.G. Gustavson, On constructing formal integrals of a Hamiltonian system near an equilibrium point, Astron. J. 71 (1966) 670–686.

[14] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I, Nonstiff Problems, Springer-Verlag, Berlin Heidelberg, 1987.

[15] T. Kapela, P. Zgliczyński, A Lohner-type algorithm for control systems and ordinary differential inclusions, Discrete Contin. Dyn. Syst., Ser. B 11 (2009) 365–385.

[16] S.M. Lozinskii, Error estimates for the numerical integration of ordinary differential equations, part I, Izv. Vysš. Učebn. Zaved., Mat. 6 (1958) 52–90 (Russian).

[17] K. Mischaikow, M. Mrozek, Conley Index. Handbook of Dynamical Systems, vol. 2, North-Holland, Amsterdam, 2002, pp. 393–460.

[18] K.R. Meyer, D. Offin, Introduction to Hamiltonian Dynamical Systems and the N-Body Problem, third edition, Appl. Math. Sci., vol. 90, Springer, 2017.

[19] J. Puig, Reducibility of linear differential equations with quasi-periodic coefficients: a survey, Research Gate preprint, https://www.researchgate.net/publication/2588346, 2002.

[20] J. Williamson, On an algebraic problem concerning the normal forms of linear dynamical systems, Am. J. Math. 58 (1936) 141–163.