Convergence to Normal Forms of Integrable PDEs

Dario Bambusi & Laurent Stolovitch

Communications in Mathematical Physics

ISSN 0010-3616
Volume 376
Number 2

Commun. Math. Phys. (2020)
376:1441-1470
DOI 10.1007/s00220-019-03661-8

Springer

Commun. Math. Phys. 376, 1441–1470 (2020)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-019-03661-8

Communications in Mathematical Physics

Convergence to Normal Forms of Integrable PDEs

Dario Bambusi[1], **Laurent Stolovitch**[2]

[1] Dipartimento di Matematica "Federigo Enriques", Università degli Studi di Milano, Via Saldini 50, 20133 Milan, Italy. E-mail: dario.bambusi@unimi.it
[2] CNRS, LJAD, Université Côte d'Azur, Parc Valrose, 06108 Nice Cedex 02, France. E-mail: stolo@unice.fr

Received: 27 March 2019 / Accepted: 2 November 2019
Published online: 3 January 2020 –

Abstract: In an infinite dimensional Hilbert space we consider a family of commuting analytic vector fields vanishing at the origin and which are nonlinear perturbations of some fundamental linear vector fields. We prove that one can construct by the method of Poincaré normal form a local analytic coordinate transformation near the origin transforming the family into a normal form. The result applies to the KdV and NLS equations and to the Toda lattice with periodic boundary conditions. One gets existence of Birkhoff coordinates in a neighborhood of the origin. The proof is obtained by directly estimating, in an iterative way, the terms of the Poincaré normal form and of the transformation to it, through a rapid convergence algorithm.

1. Introduction

In a Hilbert space H, consider a family $\{X^i\}$ of (germs of) analytic vector fields defined in a neighborhood of a common singular point, say the origin. We assume that they are pairwise commuting with respect to the Lie bracket. Consider the Taylor expansion

$$X^i = E^i + F^i \tag{1.1}$$

of the fields at the origin, with E^i the linear part. It is known since Poincaré, that, in finite dimension, each one of these vector fields can be transformed, by a **formal** change of variables $\hat{\mathcal{T}}_i$ into a Poincaré normal form $\hat{X}^i = (\hat{\mathcal{T}}_i)_* X^i := D\hat{\mathcal{T}}_i(\hat{\mathcal{T}}_i^{-1}) X^i(\hat{\mathcal{T}}_i^{-1})$. By definition, it means that the Lie bracket $[E^i, \hat{X}^i] = 0$ vanishes. We then say that $\hat{\mathcal{T}}_i$ *normalizes* X_i. Since the family is Abelian, i.e. $[X^i, X^j] = 0$ for all i, j, then one can show that there is a single $\hat{\mathcal{T}}$ that normalizes *simultaneously* the X^i's in the sense that $[E^i, \hat{\mathcal{T}}_* X^j] = 0$, for all i, j.

Research of L. Stolovitch was supported by ANR-FWF Grant "ANR-14-CE34-0002-01" for the Project "Dynamics and CR geometry" and by ANR Grant "ANR-15-CE40-0001-03" for the Project "BEKAM".

In the same spirit, if H is a symplectic space, one can study a family $\{\mathcal{H}^i = \mathcal{H}_2^i + h.o.t\}$ of (germs of) analytic Hamiltonian functions which are higher order perturbations of quadratic Hamiltonians $\mathcal{H}_2^i$ and which are pairwise commuting with respect to the Poisson bracket associated to a symplectic form ω. The normal forms of the Hamiltonians $\hat{\mathcal{T}}^*\mathcal{H}^i := \mathcal{H}^i \circ \hat{\mathcal{T}}$ are then called *Birkhoff normal form*. We have $\{\hat{\mathcal{T}}^*\mathcal{H}^i, \mathcal{H}_2^j\} = 0$ for all i, j and $\hat{\mathcal{T}}$ is a formal *symplectomorphism*, i.e. $\hat{\mathcal{T}}^*\omega = \omega$.

A classical and fundamental problem in dynamics is to know under which assumption the normalizing transformation is not only formal, but also analytic. The motivation is to understand *on the normal forms* themselves many dynamical and geometrical properties which are not tractable directly on the original system. For instance, an analytic differential equation of the form $\dot{x} = 2x + f(x, y)$; $\dot{y} = y + g(x, y)$ with f, g vanishing at order ≥ 2 at the origin is analytically conjugate to $\dot{x} = 2x + \alpha y^2$; $\dot{y} = y$ for some number α. It is very easy to describe completely the dynamics of the latter system, while it is almost impossible to do it directly for the former.

In finite dimension, Vey proved two distinct results on this problem. On the one hand, he considered in [Vey78] a family of n commuting Hamiltonian vector fields in $\mathbb{C}^{2n}$, whose linear parts are linearly independent. On the other hand, he considered in [Vey79] a family of $n-1$ commuting volume preserving vector fields in $\mathbb{C}^n$ whose linear parts are linearly independent. In both cases, he proved the existence of an analytic transformation to a normal form of the family near the origin. A byproduct of Vey Theorem for Hamiltonian systems is the (local) existence of Birkhoff coordinates, namely a sort of Cartesian action angle variables which are regular until the equilibrium point. In the Hamiltonian case, Ito [Ito89, Ito92] improved the results by essentially removing the condition of independence of the linear parts. Zung [Zun05, Zun02] generalized Vey's Hamiltonian approach by considering m "linearly independent" vector fields having $n-m$ "functionally independent" analytic first integrals in $\mathbb{C}^n$. He proved there the convergence of the transformation to normal forms. All these results have been unified in [Sto00, Sto05] (see also [Sto08]): it is proved that if the formal normal form of the family has a very peculiar structure (called "completely integrable"), and if the family of linear parts does not have "bad small divisors", then one can normalize analytically the family through a Newton scheme that simultaneously normalizes all the vector fields. One of the key points connecting the previous results with the later is that, preserving a structure such as a symplectic or a volume form, automatically implies that the formal normal form of the family is "completely integrable".

The approach of [Sto00, Sto05] is similar in spirit to the one by Rüssmann [Rüs67] and Brjuno [Bru72] who, in the case of a single Hamiltonian, proved convergence of the normal form in the case where the formal Birkhoff normal form has the very special form $\hat{F}(\mathcal{H}_2)$, with a formal power series $\hat{F}(E) = E + h.o.t$ of the single variable E (a different proof of the same result, avoiding superconvergence has been given in [LM98], by developing the methods of [GL97]).

In the infinite dimensional case the situation is much more complicated, the reason being that even homogeneous polynomials are defined through an infinite series, therefore even defining the formal normal form can be nontrivial. For this reason, even finite order normalization is far from trivial and only partial results are known [Nik86, Bam05].

Even in the case of integrable Hamiltonian systems, there are not general results ensuring the existence of action angle coordinates, but only a quite general technique to introduce variables of this kind. This technique is due to Kappeler and coworkers [BBGK95, KP03, HK08, GK14, KLTZ09]: the idea is to consider the square of the spectral gaps associated to the Lax pair and to use them as the complete sequence of integrals of motion needed to apply Arnold Liouville procedure [Arn76]. Then, one still has to regularize the singularity at the origin of action angle variables, and thus one gets the Birkhoff coordinates.

A generalization of a part of Vey theorem to infinite dimension Hamiltonian systems has been obtained by Kuksin and Perelman [KP10]. Such a result ensures that if one is able to construct a noncanonical set of coordinates Ψ_j, such that $\left|\Psi_j\right|^2$ pairwise Poisson commute, then one can introduce the Birkhoff variables. However, the construction of the coordinates Ψ_j is far from trivial and has to be performed case by case.

The aim of the present article is to devise a normalizing scheme for an infinite sequence of pairwise commuting analytic vector fields on some (infinite dimensional) Hilbert space. We consider a sequence of commuting vector fields of the form (1.1) which are analytic in a neighborhood of the origin and enjoy a suitable summability property; we assume that the linear parts E^i of the vector fields are diagonal, enjoy a summability property and a further property that we call to be "small divisor free". We first prove that the system can be put in normal form at any order, then we prove that, if the normal form has a special form, namely it is completely integrable, then the transformation of the family to a normal form is convergent in a neighborhood of the origin.

Our goal is to finally apply our result to some concrete PDE. Indeed, we prove that it can be applied to some integrable Hamiltonian PDEs, including KdV and NLS. As a byproduct we get existence of Birkhoff variables near the origin for these equations. The application to the Hamiltonian case is obtained by considering the sequence of Hamiltonians $\{\mathcal{H}^i\}$ given by the square of the spectral gaps of the Lax operator, then the corresponding Hamiltonian vector fields $\{X^i := X_{\mathcal{H}^i}\}$ pairwise commute and fulfill our assumptions. Furthermore, as in finite dimension, the fact that all the X^i's are Hamiltonian implies that their formal normal form is "completely integrable" (see Definition 2.15 below). Actually our Hamiltonian application is done by proving that the hypotheses of Kuksin Perelman's theorem (as generalized in [BM16,Mas18]) imply our assumptions.

We emphasize that at present we are not able to use as starting points for our construction the "Hierarchies of integrable PDEs" such as defined in [Mag78,Tre01,Dic03] since they do not have the good analyticity properties that we need.

We also emphasize that, while Kuksin-Perelman's and Kappeler's approaches are intrinsically based on Hamiltonian techniques, in principle our result is applicable also to non Hamiltonian systems. Unfortunately, at present our only concrete applications are to integrable Hamiltonian PDEs,

From the technical point of view, we point out that, in order to work in the present infinite dimensional context, we have to face several difficulties: the first one is to find a norm suitable to measure the size of a family of analytic vector fields, and the second one is the Lemma 5.3 which allows us to estimate the solution of the "perturbed cohomological equation" without any small divisor problem. The last difficulty are located in Lemmas 3.1 and 3.2, which allow us to estimate the remainder and flows.

We expect that our technique can be generalized to the case of systems with other symmetries such as preserving other structures, e.g. a volume form or being reversible. Here we did not develop this because we are not aware of meaningful examples to which such a theory would apply.

We recall that it is known how to put a system in normal form up to some reminder in a neighborhood of a nonresonant fixed point (see e.g. [Bam03,BG06,BDGS07,Bam08]), under some assumptions on the small divisors and on the structure of the nonlinearity, however the technique we use here is completely different from the one of these papers, and we do not think that the ideas of those papers, applied to integrable PDEs could lead to the convergence result that we prove here.

Finally we remark that normal form results are often a fundamental starting points for studying the stability of perturbed integrable PDEs (see e.g. [KP03,MP18,BKM18]).

2. Main Results

2.1. Families of normally analytic vector fields. Having fixed two sequences of weights $w_j^{(2)} \geq w_j^{(1)} > 0$, $j \geq 1$, we define the Hilbert spaces $H = \ell^2_{w^{(1)}}$ and $H^+ := \ell^2_{w^{(2)}}$ where $\ell^2_{w^{(n)}}$ is the Hilbert spaces of the complex sequences $z := \left\{z_j\right\}_{j\in\mathbb{Z}^*}$, $\mathbb{Z}^* := \mathbb{Z} - \{0\}$ s.t.

$$\|z\|^2_{w^{(n)}} := \sum_{j\in\mathbb{Z}^*} w^{(n)}_{|j|}|z_j|^2 < \infty. \tag{2.2}$$

In the following we will denote the norms simply by $\|z\| := \|z\|_{w^{(1)}}$, and $\|z\|_+ := \|z\|_{w^{(2)}}$, and by $B_r \subset H$ the open ball in H centered at the origin and having radius r. Furthermore, we will denote by $\mathbf{e} := \{\vec{e}_j\}_{j\in\mathbb{Z}^*}$ the vectors with components $(\vec{e}_j)_k \equiv \delta_{j,k}$, the Kronecker symbol.

Let $Q \equiv (..., q_{-k}, ..., q_{-1}, q_1, ..., q_k, ...) \in \mathbf{N}^{\mathbb{Z}^*}$ be an integer vector with finite support, then we write

$$z^Q := ...z_{-k}^{q_{-k}}...z_{-1}^{q_{-1}} z_1^{q_1}....z_k^{q_k}..., \quad |Q| := \sum_{k\in\mathbb{Z}^*} |q_k|.$$

Let $\mathcal{U} \subset H$ be a neighborhood of the origin in H, and let $X : \mathcal{U} \to H^+$ be an analytic vector field, then it can be expanded in series

$$X(z) = \sum_{k\geq 0} \sum_{i,|Q|=k} X_{Q,i} z^Q \vec{e}_i, \quad X_{Q,i} \in \mathbb{C} \tag{2.3}$$

or simply

$$X(z) = \sum_{i,Q} X_{Q,i} z^Q \vec{e}_i, \tag{2.4}$$

(and the series converge in any ball $B_r \subset \mathcal{U}$). The vector field $X_k := \sum_{i,|Q|=k} X_{Q,i} z^Q \vec{e}_i$ is a homogeneous polynomial vector field of degree k, it is often called the homogeneous component or part of (the Taylor expansion at the origin of) X.

In the following we will denote by $\mathcal{X}(\mathcal{U}, H^+)$ the space of the vector fields analytic in $\mathcal{U}$ with values in H^+.

Definition 2.1. We will say that a vector field $\mathcal{X}(\mathcal{U}, H^+)$ is $O(M)$ if $\forall k < M$ its Taylor components X_k vanish identically.

Remark 2.2. If $\lim_{j\to\infty} \frac{w_j^{(1)}}{w_j^{(2)}} = 0$, then a vector field in $\mathcal{X}(\mathcal{U}, H^+)$ is smoothing. The use of such vector fields, which, as shown for the first time in [KP10] actually occur in the theory of integrable equations will allow to obtain a result according to which the coordinate transformation putting the system in normal form is actually a smoothing perturbation of the identity. This is sometimes useful in the applications.

Let $X \in \mathcal{X}(\mathcal{U}, H^+)$ be a analytic vector field, and consider the vector field

$$\underline{X}(z) := \sum_{Q,i} |X_{Q,i}| z^Q \vec{e}_i, \tag{2.5}$$

which in general is defined only on a dense subset of $\mathcal{U}$.

Definition 2.3. Let $X \in \mathcal{X}(\mathcal{U}, H^+)$ be an analytic vector field vanishing at the origin. We shall say that X is normally analytic in a ball of radius r if $\underline{X} \in \mathcal{X}(B_r; H^+)$, namely it is analytic in a ball of radius r. In this case we will write $X \in \mathcal{N}_r$.

Remark 2.4. Definition 2.3 immediately extends to the case of applications from H to a general Banach space. In particular we will use it in Sect. 2.3 for the case where the target space is the space $B(H, H^+)$ of bounded linear operators from H to H^+.

Remark 2.5. If X is a (non homogeneous) polynomial, then, it is analytic in an open set $\mathcal{U}$, if and only if it is analytic in the whole H. Thus, in particular, if a polynomial vector field $X \in \mathcal{N}_r$, for some $r > 0$ then $X \in \mathcal{N}_r\ \forall r > 0$.

In what follows, all *analytic* vector fields will be considered as defined in a neighborhood (precised or not) of the origin of H with values in H^+.

A norm on $\mathcal{N}_r$ is given by

$$\left\|X\right\|_r := \sup_{\|z\| \le r} \left\|\underline{X}(z)\right\|_+ . \tag{2.6}$$

Let X, Y be normally analytic vector fields. We shall say that Y dominates X and we shall write $X \prec Y$, if $|X_{Q,i}| \le |Y_{Q,i}|$ for all indices.

Remark 2.6. In particular, if $X \prec Y$, then $\left\|\underline{X}\right\|_r \le \left\|\underline{Y}\right\|_r$ for any positive r.

Definition 2.7. A family $\mathbf{F} = \{F^i\}_{i \ge 1}$ of normally analytic vector fields will be said to be **summably analytic** if the vector field

$$\underline{\mathbf{F}} := \sum_i \underline{F^i} \tag{2.7}$$

is normally analytic in a ball of radius r, i.e. $\underline{\mathbf{F}} \in \mathcal{N}_r$. In this case we will say that $\mathbf{F} \in \mathcal{NF}_r$.

Remark 2.8. Writing

$$F^i(z) = \sum_{Q,j} F^i_{Q,j} z^Q \vec{e}_j,$$

one has

$$\underline{\mathbf{F}}(z) = \sum_{Q,j} \left(\sum_i \left| F^i_{Q,j} \right| \right) z^Q \vec{e}_j, \tag{2.8}$$

so that, for any $r > 0$, the norm $\left\|\underline{\mathbf{F}}\right\|_r$ bounds the norm of each one of the vector fields of the family, that is $\left\|\underline{F^i}\right\|_r \le \left\|\underline{\mathbf{F}}\right\|_r$.

2.2. Normal forms. Consider the family $\mathbf{E} = \{E^i\}_{i\geq 1}$ of linear diagonal vector fields

$$E^i = \sum_{j\in\mathbb{Z}^*} \mu^i_j z_j \vec{e}_j \tag{2.9}$$

Definition 2.9. We shall say that $\mathbf{E}$ is **summable** if

$$\sup_j \sum_k \left|\mu^k_j\right| < \infty. \tag{2.10}$$

Remark 2.10. In such a case, the vector field $\underline{\mathbf{E}} = \sum_{j\in\mathbb{Z}^*}\sum_k \left|\mu^k_j\right| z_j\vec{e}_j$, is well defined and analytic as a map from H to H. Since, if $\frac{w^{(2)}_j}{w^{(1)}_j} \to \infty$ such a map is not analytic as a map taking values in H^+, then $\mathbf{E}$ is not summably analytic.

Definition 2.11. We shall say that $\mathbf{E}$ is **locally finitely supported** if the support of each E^i intersects only a finite number of supports of the E^j's, that is, for any fixed j there exist at most finitely many values of i such that $\mu^i_j \neq 0$.

In order to define the normal form transformation, one has to solve the so called cohomological equation. Precisely, given a family of homogeneous analytic vector fields $F = \{F^i\}_{i\geq 1}$, one has to determine U such that

$$[E^i, U] = F^i, \tag{2.11}$$

furthermore we are interested in the case where the vector fields $X^i := E^i + F^i$ commute, namely $[X^i; X^j] = 0\ \forall i, j$, which implies

$$[E^i, F^j] = [E^j, F^i], \quad i, j \geq 1. \tag{2.12}$$

This justifies the following definition:

Definition 2.12. A family $\mathbf{F} = \{F^i\}_{i\geq 1}$ of normally analytic polynomial vector field is called a **cocycle** with respect to the family $\mathbf{E} = \{E^i\}_{i\geq 1}$, if it satisfies (2.12)

Let F be a cocycle and write

$$F^j = \sum_{Q,i} F^j_{Q,i} z^Q \vec{e}_i.$$

Then equation (2.12) reads,

$$\left((Q, \mu^i) - \mu^i_l\right) F^j_{Q,l} = \left((Q, \mu^j) - \mu^j_l\right) F^i_{Q,l}, \quad \forall Q, i, j \text{ s.t. } F^j_{Ql} \neq 0. \tag{2.13}$$

In turn, this implies, provided the two denominators are both different from zero

$$\frac{F^j_{Q,l}}{(Q, \mu^j) - \mu^j_l} = \frac{F^i_{Q,l}}{(Q, \mu^i) - \mu^i_l}.$$

This formula essentially ensures that one function U solves all the equation (2.11) simultaneously. The solution of the equation (2.12) requires the computation of the above quantities, the idea is to select, for any Q, l the index i corresponding to which the divisor does not vanish.

This justifies the following definition:

Definition 2.13. The family $\mathbf{E} = \{E^i\}$ of linear vector field is said be **small divisors free** if there exists a positive constant c, such that for each $Q \in \mathbb{N}^{\mathbb{Z}^*}$ with $|Q| \geq 2$ and $j \geq 1$, there is $i(Q, j) \geq 1$ such that $|(Q, \mu^{i(Q,j)}) - \mu_j^{i(Q,j)}| > c^{-1}$, unless $(Q, \mu^i) = \mu_j^i$ for all i.

In order to state our main theorem we still need a couple of definitions.

Definition 2.14. An analytic vector field X (resp. a family $\mathbf{X} = \{X^j\}_{j\geq 1}$) is said be a **normal form** with respect to $\mathbf{E}$ if $[E^i, X] = 0$ for all $i \geq 1$ (resp. $[E^i, X^j] = 0$ for all $i, j \geq 1$).

Definition 2.15. A family of normally analytic vector fields $\mathbf{X} := \{X^j\}_{j\geq 1}$ in normal form is said to be **completely integrable w.r.t. E** $:= \{E^i\}_{i\geq 1}$ if it can be written as $X^i = \sum_{j\geq 1} a_{ij} E^j$ where a_{ij} are normally analytic functions, invariant w.r.t $\mathbf{E}$, i.e. $E^i(a_j) = 0$, for all $i, j \geq 1$.

Remark 2.16. In finite dimension, say $\mathbb{R}^{2n}$, let us consider an single analytic Hamiltonian $\mathcal{H} = \mathcal{H}_2 + \mathcal{H}_3 + \cdots$ which is an higher order perturbation of the quadratic Hamiltonian $\mathcal{H}_2 = \sum_{i=1}^n \lambda_i(x_i^2 + y_i^2)$. It is a (Birkhoff) normal form if $\mathcal{H} = F(x_1^2 + y_1^2, \ldots, x_n^2 + y_n^2)$. This normal form is completely integrable w.r.t. $\mathcal{H}_2$ if $\mathcal{H} = F(\mathcal{H}_2) = F(\sum_{i=1}^n \lambda_i(x_i^2 + y_i^2))$. On the other hand, a n-commutative family $\{\mathcal{H}^1, \ldots \mathcal{H}^n\}$ of such Hamiltonians, with a resonance free, linearly independent family quadratic parts $\mathcal{H}_\mathbf{2} := \{\mathcal{H}_2^i\}$ is also completely integrable w.r.t $\mathcal{H}_\mathbf{2}$ (see for instance [Sto00,Sto05,Zun05]).

The first result of this paper is the following theorem, which is just the extension to the present context of Poincaré Dulac normal form theorem which is well known in finite dimensions. We recall that in infinite dimensions, some formulations are also known [Bam05,Nik86], but a general result is not available, and it is not expected to hold.

Theorem 2.1. *Let $\mathbf{E} = \{E^i\}_{i\geq 1}$ be a summable family of linear diagonal vector fields (see Definition 2.11). Consider a family of analytic vector fields of the form*

$$X^i = E^i + F^i, \quad i \geq 1. \tag{2.14}$$

Assume that

0. *the family of linear vector fields* $\mathbf{E}$ *is small divisor free.*
1. *the family* $\mathbf{F} \equiv \{F^i\}_{i\geq 1}$ *is summably analytic,*
2. *there exists c_0 s.t. for r_0 small enough, one has* $\left\|\underline{\mathbf{F}}\right\|_{r_0} \leq c_0 r_0^2$
3. $[X^i; X^j] \equiv 0$, $\forall i, j$.

Then, $\forall k$, there exist constant $r_k > 0$, c_k, a neighborhood $\mathcal{U}_k \supset B_{r_k}$ of the origin and an analytic coordinate transformation $\mathcal{T}_k : \mathcal{U} \to H$ s.t.

$$\mathcal{T}_{k*}X^i = E^i + N_k^i + O(k+1), \quad \forall i \geq 1, \tag{2.15}$$

where $\mathcal{NF}_{r_k} \ni \mathbf{N}_k \equiv \left\{N_k^i\right\}_{i\geq 1}$ is a normal form and is a polynomial vector field of degree $\leq k$.

Furthermore, $\forall r < r_k$ the following estimates hold:

i. $\|\mathbf{N}_k\|_{r_k} \leq c_k r^2$, *ii.* $\sup_{\|z\|\leq r_k} \|z - \mathcal{T}_k(z)\|_+ \leq c_k r_k^2$.

Our main result is the following theorem

Theorem 2.2. *Under the assumptions of Theorem 2.1, assume furthermore that family* **E** *is locally finitely supported (see Definition 2.11) and that* $\forall k$ *the partial normal form* N_k^i *is completely integrable w.r.t.* **E**.

Then there exist constants $r_* > 0$, c_2, c_3, *a neighborhood* $\mathcal{U} \supset B_{r_*}$ *of the origin and an analytic coordinate transformation* $\mathcal{T} : \mathcal{U} \to H$ *s.t.*

$$\mathcal{T}_* X^i = E^i + N^i, \quad \forall i \geq 1, \tag{2.16}$$

where $\mathcal{NF}_{r_*} \ni \mathbf{N} \equiv \left\{N^i\right\}_{i\geq 1}$ *is a completely integrable normal form w.r.t.* **E**.

Furthermore, $\forall r < r_*$ *the following estimates hold:*

i. $\|\mathbf{N}\|_r \leq c_2 r^2$,
ii. $\sup_{\|z\|\leq r} \|z - \mathcal{T}(z)\|_+ \leq c_3 r^2$.

Remark 2.17. From the proof it is clear that if one endows the Hilbert space by the symplectic structure $\mathrm{i}dz_{-k} \wedge dz_k$ the vector fields E^l have the form

$$E^l := -iz_l\vec{e}_l + iz_{-l}\vec{e}_{-l}$$

and F^l are Hamiltonian for any l, then the transformation $\mathcal{T}$ is canonical. In the Hamiltonian case the normal form is automatically completely integrable, so that the assumptions of Theorem 2.1 imply the conclusions of Theorem 2.2. If the vector fields are not Hamiltonian, then the normal form could contain some monomyals which are not completely integrable, so this is not true.

Remark 2.18. The Hilbert spaces considered can chosen to be more general. For instance, it could be be spaces of sequences indexed over $\mathbb{Z}^d\setminus\{0\}$.

The following result considers the extreme case of a sole vector field. In some sense, it is the infinite-dimensional counter-part of Rüssmann-Brjuno theorem when there are no small divisors.

Corollary 1. *Let E be a single summable linear diagonal vector field. Consider a vector field $X = E + F$ such that*

0. *the linear vector field* **E** *is small divisor free.*
1. *the vector fields F is normally analytic,*
2. *there exists c_0 s.t. for r_0 small enough, one has $\|F\|_{r_0} \leq c_0 r_0^2$*

Then the same conclusion of Theorem 2.1 holds. Assume also that, $\forall k \geq 2$, the normal form N_k have the form $N_k = a_{k-1}E$ for some polynomial function a_{k-1} of degree $\leq k-1$ fulfilling $E(a_{k-1}) = 0$, then the same conclusion of Theorem 2.2 holds. In particular the normalized vector field has the form

$$\mathcal{T}_* X = (1 + a)E,$$

with a analytic in B_{r_*} fulfilling $E(a) = 0$.

We are now going to give a more precise statement for the Hamiltonian case, showing in particular that the transformation $\mathcal{T}$ introduces Birkhoff coordinates for the integrable Hierarchy associated to the fields $\{X^i\}$.

Thus, in the space H, we introduce the symplectic form $\mathrm{i}dz_{-k}\wedge dz_k$. Given an analytic function $\mathcal{H} \in C^\omega(H, \mathbb{R})$, we define the corresponding Hamiltonian vector field $X_\mathcal{H}$ as the vector field with k-th component

$$(X_\mathcal{H})_k(z) := -i(\mathrm{sgn}k)\frac{\partial \mathcal{H}}{\partial z_{-k}}.$$

Given also a second function $\mathcal{K} \in C^\omega(H, \mathbb{R})$ we define their Poisson Bracket by

$$\{\mathcal{H}, \mathcal{K}\}(z) := d\mathcal{H}(z)X_\mathcal{K}(z).$$

It is well known that such a quantity can fail to be well defined, nevertheless in all the cases we will consider it will be well defined.

Consider now a sequence of analytic Hamiltonians $\mathcal{H}^i$ of the form

$$\mathcal{H}^i = \mathcal{H}_2^i + \mathcal{K}^i, \quad \mathcal{H}_2^i := z_i z_{-i},$$

and $\mathcal{K}^i$ having a zero of order at least 3 at the origin.

Corollary 2. *Assume that the vector fields $X^i := X_{\mathcal{H}^i}$ fulfill the assumptions of Theorem 2.2, then the coordinate transformation $\mathcal{T}$ is canonical. Furthermore, given any analytic Hamiltonian $\mathcal{H}$ with a zero of order 2 at the origin, such that*

$$\{\mathcal{H}, \mathcal{H}^i\} \equiv 0, \quad \forall i \geq 1, \tag{2.17}$$

one has that $\mathcal{H} \circ \mathcal{T}^{-1}$ is a function of $\{(z_j z_{-j})\}_{j\geq 1}$ only.

Remark 2.19. The smoothness assumption are given on the Hamiltonian vector fields instead of the Hamiltonian. This is standard, since, in case of weak symplectic forms (the only one appearing when dealing with partial differential equations) in general the smoothness of the Hamiltonian does not imply any smoothness property on the field. As far as we know, the first to remark this fact was Kuksin who was the first to use this kind of assumptions on vector fields [Kuk87].

Proof of Corollary 2. The proof follows the proof of Corollary 2.13 of [BM16]. First, it is clear that E^j is the Hamiltonian vector field of $\mathcal{H}_2^j$. Denote $\tilde{\mathcal{H}} := \mathcal{H} \circ \mathcal{T}^{-1}$, thus, from the property that $\mathcal{T}_* X^j$ is in normal form one has that

$$[\mathcal{T}_* X_\mathcal{H}, E^j] = X_{\{\tilde{\mathcal{H}}, \mathcal{H}_2^j\}} = 0, \tag{2.18}$$

from which $\{\tilde{\mathcal{H}}, \mathcal{H}_2^j\} = c^j$. However, since both $\tilde{\mathcal{H}}$ and $\mathcal{H}_2$ have a zero of order 2 at the origin, the constants must vanish. Expand now $\tilde{\mathcal{H}}$ in Taylor series, one has

$$\tilde{\mathcal{H}}(z) = \sum_{\substack{r\geq 2,\\ |\alpha|+|\beta|=r}} H^r_{\alpha,\beta} z_+^\alpha z_-^\beta,$$

where we denoted $z_+ := \{z_j\}_{j\geq 1}$ and $z_- := \{z_{-j}\}_{j\geq 1}$. Then $\left\{\tilde{\mathcal{H}}; \mathcal{H}_2^i\right\} = 0$ imply that, in each term of the summation, $\alpha = \beta$, therefore $\tilde{\mathcal{H}}$ is a function of $z_j z_{-j}$ only. $\square$

2.3. Kuksin–Perelman's Theorem. In this section we recall the Vey type theorem obtained by Kuksin and Perelman in [KP10] (see also [BM16,Mas18]) and prove that it can be obtained as a corollary of Theorem 2.2.

We come to the assumptions of the Kuksin-Perelman's Theorem.

Consider an analytic map Ψ of the form

$$\Psi = \mathrm{id} + G, \tag{2.19}$$

with $G \in \mathcal{N}_r$ (with some $r > 0$) having a zero of second order at the origin. For $j > 0$ consider also the functions $I_j(z) := \Psi_j(z)\Psi_{-j}(z)$ and the Hamiltonian vector fields $X^j := X_{I_j}$.

Remark 2.20. We recall that the verification of these assumptions was obtained by Kuksin and Perelman in their paper [KP10] for the KdV equation, and by [BM16,Mas18] for the Toda lattice and for the NLS equation. It is based on the application of Kato perturbation theory of eigenspaces to the Lax operator of the system. Some very delicate estimates are needed.

Assume that the following Hypotheses hold:

(KP1) The functions $I_j(z)$ pairwise commute, namely $\left\{I_j; I_k\right\} \equiv 0$ for all $j, k \geq 1$.
(KP2) the maps $\underline{dG}$ and $\underline{dG^*}$ are analytic as maps from B_R to $B(H, H^+)$. Here dG^* is the adjoint of dG with respect to the ℓ^2 metric.

Theorem 2.3. *Assume that (KP1) and (KP2) hold, then the same conclusions of Theorem 2.2 and Corollary 2 hold.*

Proof. It is enough to show that the assumptions (KP1-KP2) imply the assumptions of Theorem 2.2 with the fields $X^j := X_{I_j}$. First remark that assumption 3 of Theorem 2.1 follows from (KP1), while assumption of Theorem 2.2 follows from the fact that the fields X^i are Hamiltonian. Assumption 0 of Theorem 2.1 follows from the structure (2.19) of the function Ψ.

In order to verify assumptions 1, compute explicitly the components of the vector fields X_{I_l}. For $k \geq 1$ its $k-th$ component is given by

$$\mathrm{i}\left(X_{I_l}\right)_k = z_k\delta_{lk} + G^k\delta_{lk} + z_l\frac{\partial G^{-l}}{\partial z_{-k}} + G^l\frac{\partial G^{-l}}{\partial z_{-k}} + z_{-l}\frac{\partial G^l}{\partial z_{-k}} + G^{-l}\frac{\partial G^l}{\partial z_{-k}}\ ; \tag{2.20}$$

the first term contribute to E^l, while all the other ones contribute to F^l. From (2.20) we have that the k-th component of $\underline{\mathbf{F}}$, $(k \geq 1)$ is given by

$$\underline{G^k} + \sum_{l\geq 1} z_l\frac{\partial \underline{G^{-l}}}{\partial z_{-k}} + \sum_{l\geq 1}\underline{G^l}\,\frac{\partial \underline{G^{-l}}}{\partial z_{-k}} + \sum_{l\geq 1} z_{-l}\frac{\partial \underline{G^l}}{\partial z_{-k}} + \sum_{l\geq 1}\underline{G^{-l}}\,\frac{\partial \underline{G^l}}{\partial z_{-k}}. \tag{2.21}$$

We have to show that each one of the terms of this expression define the k-th component of an analytic vector field. For the first term this is a trivial consequence of the fact that $G \in \mathcal{N}_r$. Consider the second term. In order to see that it is analytic we write it in terms of $\underline{dG^*}$. To this end define the involution $(Iz)_k := z_{-k}$ and the truncation operator $(Tz)_k = z_k$ if $k \geq 1$ and zero otherwise. Then the second term of the above expression is the $-k$-th component of $\underline{dG^*}(ITz)$, which belongs to $\mathcal{N}_r$ by assumption (KP2). All the other terms can be dealt with in the same way getting that assumption 1 is fulfilled. Assumption 2 is a direct consequence of the fact that G has a zero of order 2 at the origin. $\square$

3. Flow of normally analytic vector fields

In this section we study the flow Φ_U^t of a vector field $U \in \mathcal{N}_r$. This is needed since all the coordinate transformations we will use are constructed as flows of such vector fields.

Lemma 3.1. *Assume that $U \in \mathcal{N}_r$ for some $r > 0$ fulfills $\epsilon := \|U\|_r < \frac{\delta}{4e}$ and let $\mathbf{F} \in \mathcal{NF}_r$ and $\delta < r$. family $(\Phi^{-1})^*\mathbf{F} \equiv \{(\Phi^{-1})^* F^i\}_{i\geq 1}$ is summably analytic and, defining $S^i := (\Phi^{-1})^* F^i - F^i$ and $\tilde{S}^i := (\Phi^{-1})^* F^i - F^i - [U, F^i]$, one has*

$$\left\|\underline{\mathbf{S}}\right\|_{r-\delta} \leq \frac{4}{\delta} \|\mathbf{F}\|_r \, \epsilon, \quad \left\|\underline{\tilde{\mathbf{S}}}\right\|_{r-\delta} \leq \frac{8e}{\delta^2} \|\mathbf{F}\|_r \, \epsilon^2. \tag{3.22}$$

Proof. In this proof we will omit the index U from Φ. To start with, we remark that, since $\sup_{\|z\|<r} \|U(z)\|_+ \leq \left\|\underline{U}\right\|_r$, $\forall\, |t| \leq 1$, one has

$$\left\| \Phi^t(z) - z \right\| = \left\| \int_0^t U(\Phi^s(z)) ds \right\| \leq \left\| \int_0^t U(\Phi^s(z)) ds \right\|_+ \leq \delta$$

and therefore $z \in B_{r-\delta}$ implies $\Phi^t(z) \in B_r$ i.e. $\Phi^t(B_{r-\delta}) \subset B_r$ (B_r denoting the ball in H of radius r centered at zero). Thus the flow is well defined and analytic at least up to $|t| = 1$. By Taylor expanding in t at $t = 0$, one has

$$(\Phi^{-t})^* F^i = \sum_{k\geq 0} \frac{t^k Ad_U^k}{k!} F^i, \tag{3.23}$$

where $Ad_U G := [U, G]$. To estimate this family remark first that

$$\underline{Ad_U F^i} \preceq \underline{DU\, F^i} + \underline{DF^i\, U} =: AAd_{\underline{U}} \underline{F^i}.$$

Summing over i one gets

$$\sum_i \underline{Ad_U F^i} \preceq AAd_{\underline{U}} \underline{\mathbf{F}},$$

and, by induction on k

$$\sum_i \underline{Ad_U^k F^i} \preceq AAd_{\underline{U}}^k \underline{\mathbf{F}}.$$

Thus we have

$$\sum_i \left(\underline{(\Phi^{-1})^* F^i} - \underline{F^i} \right) \preceq \sum_{k\geq 1} \frac{1}{k!} AAd_{\underline{U}}^k \underline{\mathbf{F}}. \tag{3.24}$$

In order to estimate the r.h.s. remark first that, for any family $\mathbf{G} \in \mathcal{NF}_{r-\delta-\delta_1}$ (for some $\delta, \delta_1 \geq 0$), we have, by Cauchy estimate

$$\left\| AAd_{\underline{U}} \underline{\mathbf{G}} \right\|_{r-\delta-\delta_1-\delta_2} \leq \frac{2}{\delta_2} \left\| \underline{U} \right\|_r \left\| \underline{\mathbf{G}} \right\|_{r-\delta-\delta_1}. \tag{3.25}$$

Fix now some $k \geq 0$, define $\delta' := \delta/k$ and look for constants $C_l^{(k)}$, $0 \leq l \leq k$ s.t.

$$\left\| AAd_{\underline{U}}^l \underline{\mathbf{F}} \right\|_{r-l\delta'} \leq C_l^{(k)}.$$

Of course, by (3.25) they can be recursively defined by

$$C_l^{(k)} = \frac{2}{\delta'} C_{l-1}^{(k)} \left\|\underline{U}\right\|_r , \quad C_0^{(k)} := \left\|\underline{\mathbf{F}}\right\|_r ,$$

which gives

$$C_l^{(k)} = \left(\frac{2}{\delta'}\left\|\underline{U}\right\|_r\right)^l \left\|\underline{\mathbf{F}}\right\|_r ;$$

taking $l = k$ this produces an estimate of the general term of the r.h.s. of (3.24):

$$\left\|\frac{\underline{AAd_U^k \mathbf{F}}}{k!}\right\|_{r-\delta} \leq \frac{k^k}{k!}\left\|\underline{\mathbf{F}}\right\|_r \left(\frac{2}{\delta}\left\|\underline{U}\right\|_r\right)^k \leq \frac{\left\|\underline{\mathbf{F}}\right\|_r}{e}\left(\frac{2e}{\delta}\left\|\underline{U}\right\|_r\right)^k , \tag{3.26}$$

where we used $k! \geq k^k e^{-k+1}$. Summing over $k \geq 1$ or $k \geq 2$, one gets the thesis. □

Although the family $\mathbf{E}$ is not summably normally analytic, its composition with the flow has the following remarkable property.

Lemma 3.2. *Assume that* $U \in \mathcal{N}_r$ *for some* $r > 0$ *fulfills* $\epsilon := \|U\|_r < \frac{\delta}{8e}$ *with* $0 < \delta < r$*; then the family* $\mathbf{T} \equiv \{(\Phi^{-1})^* E^i - E^i - [U, E^i]\}_{i \geq 1}$ *is summably normally analytic and one has*

$$\left\|\underline{\mathbf{T}}\right\|_{r-\delta} \leq \frac{8Cr}{e\delta}\left(\frac{4e\epsilon}{\delta}\right)\epsilon, \tag{3.27}$$

where $C > 0$ *depends only on* $\mathbf{E}$*.*

Proof. We proceed as in the proof of the previous Lemma except that we compute explicitly the first term of the expansion (3.23).

Since the family $\mathbf{E}$ is summable, there exist a positive constant C such that one has $D\underline{U}\,\underline{\mathbf{E}} \prec C(D\underline{U}(z))z$ and $\sum_i (D\underline{E^i})\underline{U} \prec C\underline{U}$, so we get (for any $\delta' < r$),

$$\left\|\underline{[U, \mathbf{E}]}\right\|_{r-\delta'} \leq C\left(\frac{r}{\delta'}+1\right)\left\|\underline{U}\right\|_r \leq \frac{2Cr}{\delta'}\left\|\underline{U}\right\|_r . \tag{3.28}$$

So, by (3.26),

$$\frac{1}{(k-1)!}\left\|\underline{AAd_U^k \mathbf{E}}\right\|_{r-2\delta'} = \frac{1}{(k-1)!}\left\|\underline{AAd_U^{k-1}[U, \mathbf{E}]}\right\|_{r-2\delta'} \leq \frac{2Cr}{e\delta'}\left(\frac{2e}{\delta'}\epsilon\right)^{k-1}\epsilon,$$

thus

$$\begin{aligned}\sum_{k\geq 2}\frac{1}{k!}\left\|\underline{AAd_U^k \mathbf{E}}\right\|_{r-2\delta'} &\leq \sum_{k\geq 2}\frac{1}{k}\frac{2Cr}{e\delta'}\left(\frac{2e}{\delta'}\epsilon\right)^{k-1}\epsilon \\ &\leq \frac{2Cr}{e\delta'}\epsilon\sum_{k\geq 1}\left(\frac{2e}{\delta'}\epsilon\right)^k \leq \frac{2Cr}{e\delta'}\epsilon 2\left(\frac{2e}{\delta'}\epsilon\right).\end{aligned}$$

Taking $\delta' = \delta/2$ one gets the thesis. □

Finally we need a Lemma on the composition of flows.

Lemma 3.3. *Let $U_1, U_2 \in \mathcal{N}_r$ be two polynomial normally analytic vector fields. Then, for any k, $\exists R_{k+1} \in \mathcal{N}_r$ and a polynomial $U_{(k)} \in \mathcal{N}_r$ which are normally analytic, s.t.*

$$\Phi^1_{U_1} \circ \Phi^1_{U_1} = \Phi^1_{U_{(k)}} + R_{k+1}, \tag{3.29}$$

furthermore $R_{k+1} = O(k+1)$.

Its elementary proof is left to the reader.

4. Proof Theorem 2.1

We start by some terminology.

Let $\mathcal{N}^{res}$ be the centralizer of the family $\mathbf{E}$, that is

$$\mathcal{N}^{res} := \{F \in \mathcal{N} \mid [E^i, F] = 0, \ \forall i\}.$$

By the definition of $\mathbf{E}$, we have

$$[E^i, z^Q \vec{e}_j] = \left(\sum_{l \geq 1} q_l \mu^i_l - \mu^i_j \right) z^Q \vec{e}_j =: \left((Q, \mu^i) - \mu^i_j \right) z^Q \vec{e}_j.$$

Hence, any function $F \in \mathcal{N}^{res}$ is obtained as the (possible infinite) linear combination of the monomyals $z^Q \vec{e}_j$ for which $\left((Q, \mu^i) - \mu^i_j\right) = 0$ for all $i \geq 1$.

Let $\mathcal{N}^{nres}$ be the subspace of $\mathcal{N}$ generated by monomyals $z^Q \vec{e}_j$ for which $\left((Q, \mu^i) - \mu^i_j\right) \neq 0$ for some $i \geq 1$.

So, any vector field $F \in \mathcal{N}$ can be uniquely decomposed as

$$F = F^{res} + F^{nres}, \quad F^{res} \in \mathcal{N}^{res}, \ F^{nres} \in \mathcal{N}^{nres}.$$

A vector field $F \in \mathcal{N}^{res}$ will be called *resonant*, while a vector field $F \in \mathcal{N}^{nres}$ will be called *non resonant*. When speaking of the vector field U which generates a coordinate transformation we shall say that it is *normalized* if $U^{res} = 0$.

The same notation and terminology will be used also for families of vector fields, and in such a case we will write $\mathcal{NF}^{nres}$ for a nonresonant family, namely a family composed by nonresonant vector fields and similarly for $\mathcal{NF}^{res}$.

The following lemma allows to solve and estimate the solution of the cohomological equation.

Lemma 4.1. *Let E^i be small divisor free. Let $\mathbf{B} = \{B^i\} \in \mathcal{NF}_r$ be a polynomial summably analytic family of nonresonant cocycles. Then there exists a unique normalized $U \in \mathcal{N}_r$ which solves the cohomological equation*

$$[E^i; U] = B^i. \tag{4.30}$$

Furthermore, one has

$$\|U\|_r \leq c \left\| \underline{\mathbf{B}} \right\|_r, \tag{4.31}$$

with c the constant in the definition of "small divisor free".

Proof. Write

$$B^i(z) = \sum_{Q,j} B^i_{Qj} z^Q \vec{e}_j, \quad U(z) = \sum_{Q,j} U_{Qj} z^Q \vec{e}_j,$$

then (cf. (2.13)) (4.30) is equivalent to

$$B^i_{Qj} = [(Q, \mu^i) - \mu^i_j] = U_{Qj}.$$

Since E^i is small divisor free, then, $\forall Q, j$ there exists $i = i(Q, j)$, s.t.

$$\left|(Q, \mu^i) - \mu^i_j\right| \geq c^{-1},$$

thus define

$$U(z) := \sum_{Qj} \frac{B^{i(Q,j)}_{Qj}}{(Q, \mu^i) - \mu^i_j} z^Q \vec{e}_j$$

and one has

$$\begin{aligned}\|U(z)\| &\leq \sup_{x:\|x\|=\|z\|} \left\| \sum_{Qj} \frac{|B^{i(Q,j)}_{Qj}|}{c^{-1}} x^Q \vec{e}_j \right\| \\ &\leq \sup_{x:\|x\|=\|z\|} c \left\| \sum_{iQj} |B^i_{Qj}| x^Q \vec{e}_j \right\| \leq \sup_{x:\|x\|=\|z\|} c \left\| \underline{\mathbf{B}}(x) \right\| .\end{aligned}$$

Theorem 2.1 is an immediate consequence of the following iterative lemma. □

Lemma 4.2. *Consider a family of vector fields of the form*

$$X^i_{(k)} = E^i + N^i_k + R^i_{k+1}, \tag{4.32}$$

with $\mathcal{NF}_{r_k} \ni \mathbf{N}_k = \left\{N^i_k\right\} = O(2)$ *a polynomial normal form of degree* k*, and* $\mathcal{NF}_{r_k} \ni \mathbf{R}_{k+1} = \left\{R^i_{k+1}\right\} = O(k+1)$*. Assume* $[X^i_{(k)}; X^j_{(k)}] = 0$, $\forall j, i$*, and assume that* E^i *is small divisor free and summable. Then there exists* $U_{(k)} \in \mathcal{N}_{r_k}$*, which is a homogeneous polynomial of degree* $k+1$*, s.t.*

$$(\Phi^{-1}_{U_{(k)}})^* X^i_{(k)} = X^i_{(k+1)},$$

with $X^i_{(k+1)}$ *fulfilling the assumptions of the Lemma with* $k+1$ *in place of* k*.*

Proof. Let B^i_{k+1} be the homogeneous Taylor polynomial of degree $k+1$ of R^i_{k+1} and decompose $B^i_{k+1} = B^i_{nonres} + B^i_{res}$; define $U_{(k)}$ as the normalized solution of (4.30) (with B^i_{nonres} in place of B^i) and then define

$$\begin{aligned} N^i_{k+1} &:= N^i_k + B^i_{res}, \\ R^i_{k+2} &:= (\Phi^{-1}_{U_{(k)}})^*(R^i_{k+1} - B^i_{k+1}) + \left((\Phi^{-1}_{U_{(k)}})^* B^i_{k+1} - B^i_{k+1}\right) \\ &\quad + \left((\Phi^{-1}_{U_{(k)}})^* N^i_k - N^i_k + \left((\Phi^{-1}_{U_{(k)}})^* E^i - E^i - [E^i; U_{(k)}]\right)\right), \end{aligned}$$

then of course $(\Phi^{-1}_{U_{(k)}})^* X^i_{(k)} = X^i_{(k+1)}$. Furthermore, by Lemmas 3.1, 3.2, the transformed vector field fulfills the assumptions of the Lemma with $k+1$ in place of k. □

5. Proof of the Main Theorem

In order to prove Theorem 2.2 we will use an algorithm which is different from that of Lemma 4.2, for this reason we need to know that the normal form we will get through the new algorithm is completely integrable if the normal form obtained by the previous algorithm (or any other one) is completely integrable.

Lemma 5.1. *Under the assumptions of Theorem 2.1; assume that there exist $U \in \mathcal{N}_r$, $\tilde{U} \in \mathcal{N}_{\tilde{r}}$, s.t.*

$$X^i_{(k)} := (\Phi_U^{-1})^* X^i = E^i + N^i + R^i_{k+1}, \tag{5.33}$$

$$\tilde{X}^i_{(k)} := (\Phi_{\tilde{U}}^{-1})^* X^i = E^i + \tilde{N}^i + \tilde{R}^i_{k+1}, \tag{5.34}$$

with $R^i_{k+1} = O(k+1)$. Assume that N^i is completely integrable, then also $\tilde{N}^i$ is completely integrable.

Proof. Assume for definiteness that $r < \tilde{r}$ then, by Lemma 3.3, there exists $U_{(k)} \in \mathcal{N}_r$ s.t.

$$X^i_{(k)} = (\Phi_{U_{(k)}}^{-1})^* \tilde{X}^i_{(k)} + O(k+1).$$

First we prove that, if $\{N^j\}$ and $\{(\Phi_{U_{(k)}}^{-1})^* N^j\}$ are normal forms then $[U_{(k)}, E^i] = 0$ for all i. *From now on we omit the index* (k) *from* U. We have

$$(\Phi_U^{-1})^* N^j = N^j + [U, N^j] + \frac{1}{2}[U, [U, N^j]] + \cdots$$

Taking the bracket with E^i and using Jacobi identity, we obtain

$$[E^i, (\Phi_U^{-1})^* N^j] = -[N^j, [E^i, U]] + \frac{1}{2}[E^i, [U, [U, N^j]]] + \cdots$$

Let U_{d_0} be lowest order term in the Taylor expansion of U at origin. Then, one has $[E^j, [E^i, U_{d_0}]] = 0$ for all i, j. Hence, $[E^i, U_{d_0}]$ belongs to both the range and the kernel of the semi-simple map $[E_i, .]$. Hence, $[E^i, U_{d_0}] = 0$ for all i. Hence, the lowest order term of $[N^j, [E^i, U]]$ is $[E^j, [E^i, U_{d_0+1}]]$. On the other hand, since the bracket of resonant vector fields is still resonant, we have, for $k \geq 2$, $[E^i, ad_U^k(N^j)] = [E^i, ad_U^{k-1}([U - U_{d_0}, N^j])$ which is of order $\geq d_0 + 1 + (k-1)(d_0 - 1) \geq 2d_0 > d_0 + 1$. Hence, the lowest order term of $[E^i, (\Phi_U^{-1})^* N^j]$ is $[E^j, [E^i, U_{d_0+1}]] = 0$ and we proceed by induction on the order.

Assume that the the family $\{N^j\}$ is completely integrable. Transform it to another normal form $\tilde{N}^j$ by a transformation Φ_U. According to the first point, U commutes with each E^i. Hence, it commutes with each N^i since $[U, \sum_j a_{i,j} E^j] = \sum_j a_{i,j}[U, E^j] + U(a_{i,j})E^j$. On the other hand, $E^k(U(a_{i,j})) = [E^k, U](a_{i,j}) = 0$ for all k. So that

$$(\Phi_U^{-1})^* N^i = \sum_j (\sum_k \frac{U^k(a_{i,j})}{k!}) E^j.$$

$\square$

5.1. Perturbed cohomological equation. Assume the Abelian family $\mathbf{X} = \{X^i\}$ is normalized up to order $m = 2^k$:

$$X^i = E^i + N^i_{\leq m} + R^i_{\geq m+1}.$$

where $\mathbf{N}_{\leq m} \in \mathcal{NF}_{R_m}$ is a completely integrable normal form of degree m; we shall write $NF^i_{\leq m} = E^i + N^i_{\leq m} = \sum_{j\geq 1}(\delta_{i,j} + a_{i,j})E^j$ where $a_{i,j}$ are polynomials of degree $\leq m-1$ that are common first integrals of $\mathbf{E}$. Let us Taylor expand $R^i_{\geq m+1} = B^i_{\leq 2m} + \tilde{R}^i_{\geq 2m+1}$ up to degree $2m$. We shall (mostly) omit the dependence on m in this section. Since X^i and X^j are pairwise commuting, then

$$\begin{aligned}0 = [X^i, X^j] = [NF^i, B^j] - [NF^j, B_i] + [NF^i, \tilde{R}^j_{\geq 2m+1}]\\ -[NF^j, \tilde{R}^i_{\geq 2m+1}] + [\tilde{R}^i_{\geq 2m+1}, \tilde{R}^j_{\geq 2m+1}].\end{aligned}$$

Therefore, the truncation at degree $\leq 2m$ gives

$$0 = J^{2m}([NF^i, B^j] - [NF^j, B^i]), \tag{5.35}$$

where $J^{2m}(V)$ denotes the $2m$-jet of V.

Lemma 5.2. *Let* $\mathbf{B}$ *be a nonresonant family and* $\mathbf{N}$ *a completely integrable normal form. Assume that they fulfill* (4.1). *Then there exists a unique* U *normalized (i.e. no resonant term in expansion) such that for all* j *one has* $J^{2m}([NF^j, U]) = B^j$.

Proof. We give here a direct proof although a more conceptual proof involving spectral sequences can be found in [Sto00] [proposition 7.1.1] for the finite dimensional case. For any integer $m+1 \leq k \leq 2m$, the homogeneous polynomial of degree k of Eq. (5.35) is

$$\sum_{p=1}^{k-m} [NF^i_p, B^j_{k-p+1}] = \sum_{p=1}^{k-m} [NF^j_p, B^i_{k-p+1}] \tag{5.36}$$

Let us prove, by induction on the integer $m+1 \leq k \leq 2m$, that there exists a unique normalized polynomial V_k homogeneous of degree k, such that

$$\forall 1 \leq i,\ [E^i, V_k] = B^i_k + \sum_{p=2}^{k-m} [V_{k-p+1}, NF^i_p], \tag{5.37}$$

that is $J^k\left([NF^i, \sum_{p=m+1}^k V_p]\right) = J^k(B^i)$.

For $k = m+1$, the Eq. (5.36) leads to $[E^i, Z^j_{m+1}] = [E^j, Z^i_{m+1}]$. According to the Lemma 4.1, there exists a unique normalized V_{m+1} homogeneous of degree $m+1$ such that, for all $1 \leq i$, $[E^i, V_{m+1}] = Z^i_{m+1}$.

Let us assume that the result holds for all integers $q < k$. Let $2 \leq p \leq k-m$ be an integer, then $m+1 \leq k-p+1 < k$. Let us first recall that, by assumptions, $[NF^i_k, NF^j_{k'}] = 0$ for all integers $1 \leq i, j$ and $1 \leq k, k'$.

Thus, by Jacobi Identity, we have

$$[NF^i_p, [E^j, V_{k-p+1}]] = -[E^j, [V_{k-p+1}, NF^i_p]]$$

$$[NF^i_p, [V_{k-p-q+2}, NF^j_q]] = -[NF^j_q, [NF^i_p, V_{k-p-q+2}]] \quad \forall\, 2 \leq q \leq k-p+1-m$$

With these remarks as well as (5.37), it follows, by induction, that

$$[NF^i_p, Z^j_{k-p+1}] = \left[NF^i_p, [E^j, V_{k-p+1}] - \sum_{q=2}^{k-p+1-m} [V_{k-p-q+2}, NF^j_q]\right]$$
$$= -[E^j, [V_{k-p+1}, NF^i_p]] + \sum_{q=2}^{k-p+1-m} [NF^j_q, [NF^i_p, V_{k-p-q+2}]]$$

Since $[NF^j_q, [NF^i_p, V]] = [NF^i_p, [NF^j_q, V]]$, then exchanging j and i leads to

$$[NF^i_p, Z^j_{k-p+1}] + [E^j, [V_{k-p+1}, NF^i_p]] = [NF^j_p, Z^i_{k-p+1}] + [E^i, [V_{k-p+1}, NF^j_p]]$$

Summing over $2 \le p \le k$ and using the compatibility condition (5.36) leads to

$$\left[E^j, Z^i_k + \sum_{p=2}^{k-m} [V_{k-p+1}, NF^i_p]\right] = \left[E^i, Z^j_k + \sum_{p=2}^{k-m} [V_{k-p+1}, NF^j_p]\right]$$

But, the same argument as in the proof of the first point of this proposition will show that, $\{\sum_{p=2}^{k-m} [V_{k-p+1}, NF^i_p]\}$ is a non-resonant family of homogeneous vector fields of degree k. Therefore, according to Lemma 4.1, there exists a unique normalized V_k such that, for all $i \ge 1$,

$$[E^i, V_k] = Z^i_k + \sum_{p=2}^{k-m} [V_{k-p+1}, NF^i_p],$$

which ends the proof of the induction and the proposition. $\square$

Let us construct and estimate the unique nonresonant solution U (i.e. with $U^{res} \equiv 0$), of order $\ge m+1$ and degree $\le 2m$ of the *perturbed cohomological equation*, namely

$$J^{2m}([NF^i, U]) = B^i_{nres}, \tag{5.38}$$

where B^i_{nres} denotes the nonresonant projection of B^i.

Lemma 5.3. *Assume that, for all* $\|z\| \le r$, *for all* $v \in H$

$$\|D\underline{\mathbf{N}}(z).v\|_+ \le \frac{1}{2}\|v\|, \quad \|\underline{\mathbf{R}}_{\ge m+1}\|_r \le \epsilon.$$

Then (5.38) *has a unique nonresonant solution* U *which satisfies*

$$\|\underline{U}\|_r \le 4\epsilon.$$

Proof. Let us write (5.38) as

$$[NF^i, U] = B^i_{nres} + Z^i_{\ge 2m+1} =: F^i \tag{5.39}$$

where $Z^i_{\ge 2m+1} := J^{2m}[NF^i, U] - [NF^i, U]$.

Let $\lambda_i^{(d)}$ be an eigenvalue of the operator $[E^i, \cdot]$ in the space of formal homogeneous polynomial vector fields of degree d. Let $h_{i,\lambda_i^{(d)}}$ be the associated eigenspace.

Remark 5.4. In the Hamiltonian case, the family $\mathbf{E}$ is defined by $E^i = z_i\vec{e}_i - z_{-i}\vec{e}_{-i}$ so that we have $\lambda_i^{(d)} = q_i - q_{-i} - s_i$ where q_i denotes the ith component of a multi index $Q = (\cdots q_{-i-1}, q_{-i}, \cdots, q_{-1}, q_1, \cdots q_i, q_{i+1}, \cdots)$ with modulus d and s_i is 1,-1 or 0. Indeed, we have $[E^i, z^Q\vec{e}_k] = (q_i - q_{-i} - s_i)z^Q\vec{e}_k$ with $s_i = 1$ if $k = i$, $s_i = -1$ if $k = -i$ and $s_i = 0$ otherwise.

Let $\lambda^{(d)} = (\lambda_1^{(d)}, \lambda_2^{(d)}, \ldots)$ be a collection of such eigenvalues. We shall say that $\lambda^{(d)}$ is a generalized eigenvalue of degree d. If $\lambda^{(d)} \neq 0$, then only a finite number of its components are non zero; this is a consequence of the locally finiteness support assumption on $\mathbf{E}$. Let us denote $\mathrm{Supp}(\lambda^{(d)})$, the support of $\lambda^{(d)}$, that is the set of indexes j such that $\lambda_j^{(d)} \neq 0$. From now on, we shall write λ for $\lambda^{(d)}$, if there is no confusion.

We remark that, given $U \in \cap_{i\geq 1} h_{i,\lambda_i^{(d)}}$, and any function a which is a common first integral of the family E^i, namely s.t. $E^k(a) = 0$, $\forall k$, one has

$$[E^i, aU] = \lambda_i^{(d)} aU,$$

thus it is convenient to denote

$$H_{\lambda^{(d)}} := \left\{U \in \mathcal{N} \ : \ [E^i, U] = \lambda_i^{(d)} U\right\}. \tag{5.40}$$

We now show that $[NF^i, .]$ leaves invariant H_λ (where we omitted the index d from λ). We have

$$[NF^i, U] = [E^i, U] + \sum_{j\geq 1} a_{i,j}[E^j, U] + U(a_{i,j})E^j. \tag{5.41}$$

Here, $U(a_{i,j})$ denotes the Lie derivative of $a_{i,j}$ along U. Since the E^i's are pairwise commuting and since the $a_{i,j}$'s are first integrals of $\mathbf{E}$, we have

$$\begin{aligned}\left[E^l, [E^i, U] + \sum_{j\geq 1} a_{i,j}[E^j, U]\right] &= [E^i, [E^l, U] + \sum_{j\geq 1} a_{i,j}[E^j, [E^l, U]] \\ &= \lambda_l\left([E^i, U] + \sum_{j\geq 1} a_{i,j}[E^j, U]\right).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\left[E^l, \sum_{j\geq 1} U(a_{i,j})E^j\right] &= \sum_{j\geq 1} E^l(U(a_{i,j}))E^j = \sum_{j\geq 1}[E^l, U](a_{i,j})E^j \\ &= \lambda_l \sum_{j\geq 1} U(a_{i,j})E^j.\end{aligned}$$

From which the invariance of H_λ follows.

Let U_λ (resp. F^i_λ) be the projection onto H_λ of U (resp. F^i). Therefore, the projection onto H_λ of Eq. (5.39) reads

$$[NF^i, U_\lambda] = F^i_\lambda. \tag{5.42}$$

Using (5.41), this equation reads

$$\left(\lambda_i + \sum_{j\geq 1} a_{i,j}\lambda_j\right) U_\lambda + \sum_{j\geq 1} U_\lambda(a_{i,j})E^j = F^i_\lambda.$$

Let ϵ_i be the sign of λ_i, if $i \in \mathrm{Supp}(\lambda)$. Let us multiply the ith-equation by ϵ_i and then let us sum up over $i \in \mathrm{Supp}(\lambda)$. We obtain

$$\left(|\lambda| + \sum_{j\geq 1}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i a_{i,j}\lambda_j\right) U_\lambda + \sum_{j\geq 1}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i U_\lambda(a_{i,j})E^j = \sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i F^i_\lambda =: \tilde{F}_\lambda. \tag{5.43}$$

Let us define

$$b_\lambda := |\lambda| + \sum_{j\in\mathrm{Supp}(\lambda)}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i a_{i,j}\lambda_j =: |\lambda| + c_\lambda. \tag{5.44}$$

Remark that it is an analytic function whose value at 0 is $|\lambda|$; furthermore one has $E^j(b_\lambda) = 0, \forall j$. Let us consider the operator

$$P_\lambda \ : U_\lambda \mapsto \sum_{j\geq 1}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i U_\lambda(a_{i,j})E^j.$$

We have $P_\lambda^2 = 0$. Indeed, since the $a_{i,j}$ are first integrals of $\mathbf{E}$, we have

$$\begin{aligned}
P_\lambda(P_\lambda(U_\lambda)) &= \sum_{j\geq 1}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i P_\lambda(U_\lambda)(a_{i,j})E^j \\
&= \sum_{j\geq 1}\sum_{k\geq 1}\sum_{i\in\mathrm{Supp}(\lambda)} \epsilon_i \sum_{l\in\mathrm{Supp}(\lambda)} \epsilon_l U_\lambda(a_{l,k})E^k(a_{i,j})E^j \\
&= 0.
\end{aligned}$$

Similarly one has $P_\lambda(P_\lambda(./b_\lambda)) = 0$. As a consequence, the nonresonant solution of Eq. (5.43) is

$$U_\lambda = (I - \frac{1}{b_\lambda}P_\lambda)\left(\frac{\tilde{F}_\lambda}{b_\lambda}\right). \tag{5.45}$$

Summing up over the set of generalized eigenvalues λ of degree $m+1 \leq d \leq 2m$, and applying J^{2m} we obtain

$$U = J^{2m}\left(\sum_\lambda \frac{\tilde{F}_\lambda}{b_\lambda} - \sum_\lambda \frac{1}{b_\lambda}P_\lambda\left(\frac{\tilde{F}_\lambda}{b_\lambda}\right)\right). \tag{5.46}$$

Since U is of degree $\leq 2m$, we can substitute B_λ to F_λ, thus we are led to the final definition of U, namely

$$U = J^{2m}\left(\sum_\lambda \frac{\tilde{B}_\lambda}{b_\lambda} - \sum_\lambda \frac{1}{b_\lambda}P_\lambda\left(\frac{\tilde{B}_\lambda}{b_\lambda}\right)\right), \tag{5.47}$$

where $\tilde{B}_\lambda := \sum_{i\in \mathrm{Supp}(\lambda)} \epsilon_i B^i_\lambda$. We now estimate such a quantity. Remark first that one has

$$\underline{\left(\frac{1}{b_\lambda}\right)} = \underline{\left(\frac{1}{|\lambda|-c_\lambda}\right)} \preceq \frac{1}{|\lambda|}\sum_{k\geq 0}\left(\frac{\underline{c_\lambda}}{|\lambda|}\right)^k \preceq \frac{1}{|\lambda|-\underline{c_\lambda}},$$

so that we have

$$\underline{\sum_\lambda \frac{\tilde{B}_\lambda}{b_\lambda}} \prec \sum_\lambda \frac{\sum_{i\in \mathrm{Supp}(\lambda)} \underline{B^i_\lambda}}{|\lambda|-\underline{c_\lambda}}.$$

On the other hand, given an orthonormal basis $\mathbf{e}$ of H^+, a sequence $\{G_\lambda\}$ of vectors with nonnegative coordinates on $\mathbf{e}$ and a bounded sequence $\{g_i\}$ of nonnegative numbers, we have

$$\left\|\sum_\lambda g_\lambda G_\lambda\right\|_+^2 = \sum_{\lambda,\lambda'} g_\lambda g_{\lambda'} (G_\lambda, G_{\lambda'})_+ \leq (\sup_{\lambda,\lambda'} g_\lambda g_{\lambda'}) \left\|\sum_\lambda G_\lambda\right\|_+^2 \leq (\sup_\lambda g_\lambda)^2 \left\|\sum_\lambda G_\lambda\right\|_+^2.$$

Evaluating at a point near the origin in the domain, we can apply this with $g_\lambda = \frac{1}{|\lambda|-\underline{c_\lambda}}$ and $G_\lambda = \sum_{i\in \mathrm{Supp}(\lambda)} \underline{B^i_\lambda}$ Hence, we obtain

$$\left\|\sum_\lambda \frac{\sum_{i\in \mathrm{Supp}(\lambda)} \underline{B^i_\lambda}}{|\lambda|-\underline{c_\lambda}}\right\|_+ \leq \sup_\lambda \left|\frac{1}{|\lambda|-\underline{c_\lambda}}\right| \left\|\sum_\lambda \sum_{i\in \mathrm{Supp}(\lambda)} \underline{B^i_\lambda}\right\|_+ \leq \sup_\lambda \left|\frac{1}{|\lambda|-\underline{c_\lambda}}\right| \|\underline{\mathbf{B}}\|_+. \tag{5.48}$$

In order to estimate $\underline{c_\lambda}$, remark first that according to (5.44), we have

$$\underline{c_\lambda} \prec \sum_{j\in \mathrm{Supp}(\lambda)} \sum_{i\in \mathrm{Supp}(\lambda)} \underline{a_{i,j}} |\lambda_j| \prec \sum_{j\in \mathrm{Supp}(\lambda)} |\lambda_j| \left(\sum_i \underline{a_{i,j}}\right). \tag{5.49}$$

To estimate $\beta_j := \sum_i \underline{a_{ij}}$, we proceed as follows. According to (2.9), we have $\underline{N^i} = \sum_{j\in\mathbb{Z}^*} \underline{a_{i,j}} z_j \vec{e}_j$ so that $\underline{\mathbf{N}} = \sum_{j\in\mathbb{Z}^*} \left(\sum_i \underline{a_{i,j}}\right) z_j \vec{e}_j = \sum_j \beta_j z_j \vec{e}_j$. Hence, we have

$$\frac{\partial \underline{\mathbf{N}}}{\partial z_k} = \sum_{j\in\mathbb{Z}^*} \frac{\partial \beta_j}{\partial z_k} z_j \vec{e}_j + \beta_k \vec{e}_k.$$

Since the previous equality involves only vectors with nonnegative coefficients, we have

$$\beta_k e_k \prec \frac{\partial \underline{\mathbf{N}}}{\partial z_k} \tag{5.50}$$

$$\sum_{j\in\mathbb{Z}^*} \frac{\partial \beta_j}{\partial z_k} z_j e_j \prec \frac{\partial \underline{\mathbf{N}}}{\partial z_k}. \tag{5.51}$$

So, $\forall v \in H$ and for all $\|z\| \leq r$, we have

$$\left(\frac{1}{2}\|v\|\right)^2 = \frac{1}{4}\sum_k w_k^1(1)\,|v_k|^2 \geq \left\| D\underline{\mathbf{N}}(z)v\right\|_+^2 = \left\|\sum_k \frac{\partial \underline{\mathbf{N}}}{\partial z_k} v_k\right\|_+^2$$
$$\geq \left\|\sum_k v_k \beta_k \vec{e}_k\right\|_+^2 = \sum_k w_k^{(2)} \beta_k^2 v_k^2 = \sum_k w_k^{(1)} \frac{w_k^{(2)}}{w_k^{(1)}} \beta_k^2 v_k^2.$$

Taking $v := v_k \vec{e}_k = 1/\sqrt{w_k^{(1)}}\vec{e}_k$, which has norm 1, one gets

$$\frac{1}{4} \geq \frac{w_l^{(2)}}{w_l^{(1)}}\beta_l^2 \geq \beta_l^2 = \left(\sum_i \underline{a}_{il}\right)^2.$$

Inserting in (5.49) one gets

$$\left|\underline{c}_\lambda\right| \leq |\lambda|\,\frac{1}{2},$$

hence

$$\left|\frac{1}{|\lambda| - \underline{c_\lambda(z)}}\right| \leq \frac{2}{|\lambda|}.$$

Since the family $\mathbf{E}$ is small divisor free, then we always have $1 \leq |\lambda|$ (we have set $c = 1$ for simplicity), then by (5.48)

$$\sup_{\|z\|\leq r_m} \left\|\underline{\sum_\lambda \frac{\tilde{B}_\lambda}{b_\lambda}}\right\| \leq \sup_\lambda \frac{2\epsilon}{|\lambda|} \leq 2\epsilon \tag{5.52}$$

as soon as $\|\underline{\mathbf{B}}\|_+ \leq \epsilon$. On the other hand, we have

$$\sum_\lambda \frac{1}{b_\lambda} P_\lambda\left(\frac{\tilde{B}_\lambda}{b_\lambda}\right) \prec \sum_\lambda \frac{1}{(|\lambda| - \underline{c_\lambda})^2} \sum_{j\geq 1}\sum_{i\in \mathrm{Supp}(\lambda)} \underline{Da_{i,j}}\left(\sum_{l\in\mathrm{Supp}(\lambda)} \underline{B_\lambda^l}\right)\underline{E^j}.$$

According to (5.51), we have

$$\sum_{j\geq 1}\sum_{i\in \mathrm{Supp}(\lambda)} \underline{Da_{i,j}}\left(\sum_{l\in\mathrm{Supp}(\lambda)} \underline{B_\lambda^l}\right)\underline{E^j} \prec D\underline{\mathbf{N}}.\left(\sum_{l\in\mathrm{Supp}(\lambda)} \underline{B_\lambda^l}\right).$$

As in (5.48), we have

$$\left\|\underline{\sum_\lambda \frac{1}{b_\lambda} P_\lambda\left(\frac{\tilde{B}_\lambda}{b_\lambda}\right)}\right\| \leq \sup_\lambda \frac{1}{(|\lambda| - \underline{c_\lambda})^2}\left\|D\underline{\mathbf{N}}.\left(\sum_\lambda \sum_{l\in\mathrm{Supp}(\lambda)} \underline{B_\lambda^l}\right)\right\|.$$

Hence, for $\|z\| \leq r$,

$$\left\| \underline{\sum_{\lambda} \frac{1}{b_\lambda} P_\lambda \left(\frac{\tilde{B}_\lambda}{b_\lambda} \right)} \right\| \leq 2\epsilon. \tag{5.53}$$

Collecting estimates (5.52) and (5.53), we obtain

$$\sup_{\|z\| \leq r} \left\| \underline{\sum_{\lambda} \frac{\tilde{B}_\lambda}{b_\lambda} - \frac{1}{b_\lambda} P_\lambda \left(\frac{\tilde{B}_\lambda}{b_\lambda} \right)} \right\| \leq 4\epsilon,$$

and remarking that, for functions of class $\mathcal{N}$ the projector J^{2m} does not increase the norm, one gets

$$\sup_{\|z\| \leq r} \|\underline{U}(z)\| \leq 4\epsilon. \tag{5.54}$$

□

5.2. Iteration. Without loos of generality we can assume that $C \geq 1$. We use U to generate a change of variables which is the time 1 flow, Φ of the system $\dot{z} = U(z)$. We have

$$\begin{aligned} \Phi_*^{-1} X_i &= X_i + [-U, X_i] + O(2m+1) \\ &= NF^i_{\leq m} + B^i_{nres} + B^i_0 + [NF^i_{\leq m}, -U] + O(2m+1) \\ &= \underbrace{NF^i_{\leq m} + B^i_0}_{=:NF^i_{\leq 2m}} + O(2m+1). \end{aligned} \tag{5.55}$$

By assumption, $B^i_0 = \sum_{j \geq 1} \tilde{a}_{i,j} E^j$ where $\tilde{a}_{i,j}$ are polynomials of degree $\leq 2m - 1$ that are common first integrals of $\mathbf{E}$.

By assumption, we have $\mathbf{X} = \mathbf{E} + \mathbf{F}$ and there exists c_0 such that

$$\|\underline{\mathbf{F}}\|_{r_0} \leq c_0 r_0^2 \tag{5.56}$$

for some small parameter r_0. We also fix two large constants c_1 and $b \geq 1$ (we will track the dependence of everything on such constants). Their precise value will be decided along the procedure.

We denote $m := 2^k$, $k \geq 0$ then the sequences we are interested in are defined by

$$q_m := m^{-\frac{b}{m}}, \quad m = 2^k, \quad k \geq 0 \tag{5.57}$$

$$\epsilon_k := \frac{\epsilon_0}{4^k}, \quad k \geq 0 \tag{5.58}$$

$$\epsilon_0 = \frac{c_0}{C} r_0^2 \tag{5.59}$$

$$\delta_0 := \frac{r_0}{2}, \tag{5.60}$$

$$\delta := \frac{1}{Cc_1} r_0, \quad \delta_k := \frac{\delta}{4^k}, \quad k \geq 1 \tag{5.61}$$

$$r_1 := \frac{1}{4C}(r_0 - \delta_0) = \frac{1}{8C} r_0, \quad r_{k+1} := q_{2^k} (r_k - \delta_k), \quad k \geq 1. \tag{5.62}$$

In the appendix we will prove that the following properties hold

$$d_k := \prod_{l=0}^{k-1} q_{2^l} = \frac{1}{4^{b\left(1-\frac{k+1}{2^k}\right)}} \geq 4^{-b}, \tag{5.63}$$

$$r_k \geq \frac{1}{4^b} r_1 - \frac{\delta}{3} = \frac{1}{4^b C}\left(\frac{1}{8} - \frac{4^b}{3}\frac{1}{c_1}\right) r_0 =: r_\infty \geq \frac{r_0}{C 4^{b+2}}, \tag{5.64}$$

provided $c_1 \geq 4^{b+2}/3$. Actually we take

$$c_1 = \frac{4^{b+2}}{3}, \implies r_\infty = \frac{r_0}{4^{b+2} C}. \tag{5.65}$$

We will also prove that

$$\sum_{l=0}^{k-1} \epsilon_l \leq \frac{4}{3}\epsilon_0, \quad \sum_{l=0}^{k-1} \frac{\epsilon_l}{r_l - r_{l+1}} \leq \frac{8C}{8C-1}\frac{\epsilon_0}{r_0} + \frac{\epsilon_0}{r_\infty} 2^{b/2}. \tag{5.66}$$

Consider the following inequalities (with $m = 2^k$)

$$\left\| \underline{\mathbf{R}_{\geq m+1}} \right\|_{r_k} \leq \epsilon_k, \tag{5.67}$$

$$\left\| \underline{\mathbf{N}_{\leq m}} \right\|_{r_k} \leq \begin{cases} 0 & \text{if } k = 0 \\ \sum_{l=0}^{k-1} \epsilon_l & \text{if } k \geq 1 \end{cases} \tag{5.68}$$

$$\sup_{\|z\| \leq r_k} \left| D\underline{\mathbf{N}_{\leq m}}(z) \right|_{\mathcal{B}(H,H^+)} \leq \begin{cases} 0 & \text{if } k = 0 \\ \sum_{l=0}^{k-1} \frac{\epsilon_l}{r_l - r_{l+1}} & \text{if } k \geq 1 \end{cases}. \tag{5.69}$$

Lemma 5.5. *Assume*

$$b \geq 8 + \frac{2\ln(3 \cdot 16C)}{\ln 2},$$

take

$$c_1 = \frac{4^{b+2}}{3}, \tag{5.70}$$

$$r_0 < \min\left\{ \frac{3C}{136 c_0}; \frac{1}{32 e c_0 c_1}; \frac{1}{7 \cdot 2^9 c_0 c_1^2}; \frac{1}{2c_0(\frac{8}{7} + 4^{b+2} 2^{b/2})}; \right.$$
$$\left. \frac{c_1 C - \frac{1}{4}}{c_0\left((2^4+6)c_1 + 2^9 C c_1^2 + 4\frac{9}{7} + 2^{b/2} 4^{b+3}\right)} \right\}. \tag{5.71}$$

Assume that the inequalities (5.67)–(5.69) hold with some $k \geq 0$. Let Φ_m be the flow generated by U_m defined by (5.46). It conjugates the family $X^i_m = E^i + N^i_{\leq m} + R^i_{\geq m+1}$ to the family $X^i_{2m} := E^i + N^i_{\leq 2m} + R^i_{\geq 2m+1}$ and (5.67), (5.68), (5.69) hold for the new N and R with $k+1$ in place of k.

Proof. First we define

$$N^i_{\leq 2m} := N^i_{\leq m} + \left(J^{2m} R^i_{\geq m+1}\right)_{res}, \tag{5.72}$$

so that the estimate (5.68) immediately follows and the estimate (5.69) follows from Cauchy inequality.

Then, an explicit computation gives

$$R^i_{\geq 2m+1} = (\Phi_m^{-1})^* E^i - E^i - [U, E^i] \tag{5.73}$$

$$+ (\Phi_m^{-1})^* N^i_{\leq m} - N^i_{\leq m} - [U, N^i_{\leq m}] \tag{5.74}$$

$$+ (\Phi_m^{-1})^* R^i_{\geq m+1} - R^i_{\geq m+1} \tag{5.75}$$

$$+ \left(\mathrm{I} - J^{2m}\right)\left([U, E^i] + R^i_{\geq m+1} + [U, N^i_{\leq m}]\right). \tag{5.76}$$

We remark that, as it can be seen by a qualitative analysis and we will also see quantitatively, the largest contribution to the estimate of the reminder term comes from the term $[U, E^i]$ in (5.76), followed (in size, but not in terms of order of magnitude) by the term coming from $R_{\geq m+1}$ still in (5.76). All the other terms admit estimates which of higher order.

Let us prove by induction on $k \geq 0$ estimates (5.67), (5.68) and (5.69)

For $k = 0$, one has $N_{\leq 1} \equiv 0$ and $\left\|\underline{\mathbf{R}_{\geq 2}}\right\|_{r_0} \leq c_0 r_0^2 = \epsilon_0$. Hence, inequalities hold true for $k = 0$. Assume that they hold for all $0 \leq l \leq k$ and let us prove the inequality for $m = 2^{k+1}$.

Since r_1 and δ_0 do not follow the induction definition of r_k and δ_k, we have to prove separately the case $k = 1$. Since $N_{\leq 1} \equiv 0$ then (5.74) is not present, as well as the last term in (5.76). Furthermore the perturbed cohomological equation reduces to the linear one, so U can be estimated using Lemma 4.1 with $c = 1$ which gives

$$\|U\|_{r_0} \leq \epsilon_0.$$

We have that, by (3.27), (3.22), the families corresponding to (5.73) and (5.75) are estimated by (with a little abuse of notation)

$$\left\|\underline{(5.73)}\right\|_{r_0-\delta_0} \leq \frac{8 r_0 C}{e\delta_0}\left(\frac{4e\epsilon_0}{\delta_0}\right)\epsilon_0 = 32\left(\frac{r_0}{\delta_0}\right)^2 c_0 r_0 \epsilon_0 = 32 \cdot 4 c_0 r_0 \epsilon_0$$

$$\left\|\underline{(5.75)}\right\|_{r_0-\delta_0} \leq \frac{4}{\delta_0}\epsilon_0\epsilon_0 = \frac{8c_0}{C} r_0 \epsilon_0 \leq 8 c_0 r_0 \epsilon_0.$$

Concerning (5.76), by (3.28) we have

$$\left\|\underline{[U, \mathbf{E}]}\right\|_{r_0-\delta_0} \leq \epsilon_0 \frac{2Cr_0}{\delta_0} = 4C\epsilon_0,$$

and thus

$$\left\|\underline{(5.76)}\right\|_{r_0-\delta_0} \leq \epsilon_0 \frac{2Cr_0}{\delta_0} + \epsilon_0 = (4C+1)\epsilon_0 \leq 5C\epsilon_0.$$

It follows that

$$\left\|\underline{\mathbf{R}_{\geq 3}}\right\|_{r_0-\delta_0} \leq ((32 \cdot 4 + 8) c_0 r_0 + 5C)\epsilon_0 \leq 8C\epsilon_0,$$

provided

$$r_0 \le \frac{3C}{136c_0}. \tag{5.77}$$

From Lemma A.4 it follows that

$$\left\|\underline{\mathbf{R}_{\ge 3}}\right\|_{\frac{1}{4C}(r_0-\delta_0)} \le \frac{1}{4^3C^3}\left\|\underline{\mathbf{R}_{\ge 3}}\right\|_{r_0-\delta_0} \le \frac{1}{4^3C^3}4\frac{r_0}{\delta_0}\epsilon_0 = \frac{1}{8C^2}\epsilon_0 < \frac{\epsilon_0}{4} = \epsilon_1.$$

We also remark that, by Cauchy estimate, we have

$$\left\|\underline{D\mathbf{N}_{\le 2}}\right\|_{r_1} \le \frac{1}{r_0-r_1}\left\|\underline{\mathbf{N}_{\le 2}}\right\|_{r_0} \le \frac{1}{r_0-r_1}\left\|\underline{\mathbf{R}_{\ge 2}}\right\|_{r_0} \le \frac{\epsilon_0}{r_0-r_1}.$$

This concludes the proof of for case $k=1$.

Assume now $k\ge 1$. According to Lemma 5.3, we have $\|\underline{U}(z)\|_{r_k} \le 4\epsilon_k$ as soon as

$$\frac{8C}{8C-1}\frac{\epsilon_0}{r_0} + \frac{\epsilon_0}{r_\infty}2^{b/2} \le \frac{1}{2},$$

that is, since $\frac{\epsilon_0}{r_0} = \frac{c_0 r_0}{C}$ and $\frac{r_0}{r_\infty} = 4^{b+2}C$

$$c_0 r_0\left(\frac{8}{7} + 4^{b+2}2^{b/2}\right) \le \frac{1}{2}. \tag{5.78}$$

Let us assume that

$$\frac{16e\epsilon_k}{\delta_k} = \frac{16e\epsilon_0}{\delta} = 16ec_0c_1r_0 < \frac{1}{2} \iff r_0 < \frac{1}{32ec_0c_1}. \tag{5.79}$$

Then, by (3.27), (3.22), the above families are estimated by

$$\|\underline{(5.73)}\|_{r_k-\delta_k} \le \frac{8r_kC}{e\delta_k}\left(\frac{4e4\epsilon_k}{\delta_k}\right)4\epsilon_k = 2^9C\left(\frac{\epsilon_0}{\delta}\right)^2 r_k \le 2^9Cc_1^2r_0c_0\epsilon_0$$

$$\|\underline{(5.74)}\|_{r_k-\delta_k} \le \frac{8e}{\delta_k^2}\frac{4}{3}\epsilon_0(4\epsilon_k)^2 = 2^9c_0^2e\left(\frac{r_0}{\delta}\right)^2 r_0^2\epsilon_0 = \frac{2^9e}{3}c_0^2c_1^2r_0^2\epsilon_0 < 2^9c_0^2c_1^2r_0^2\epsilon_0,$$

$$\|\underline{(5.75)}\|_{r_k-\delta_k} \le \frac{4}{\delta_k}\epsilon_k4\epsilon_k = \frac{2^4c_0}{4^k}\left(\frac{r_0}{\delta}\right)r_0\epsilon_0 \le 2^4c_0r_0\left(\frac{r_0}{\delta}\right)\epsilon_0 = 2^4c_0r_0c_1\epsilon_0.$$

Concerning (5.76), by (3.28) we have

$$\|\underline{[U,\mathbf{E}]}\|_{r_k-\delta_k} \le 4\epsilon_k\frac{2r_kC}{\delta_k} \le 8C\left(\frac{r_0}{\delta}\right)\epsilon_0 = 8Cc_1\epsilon_0$$

and, by (5.69), (5.68), (5.67) and (5.66), we have

$$\begin{aligned}
\left\|\underline{[U,\mathbf{N}_{\le m}]}\right\|_{r_k-\delta_k} &\le \left\|\underline{DU\mathbf{N}_{\le m}}\right\|_{r_k-\delta_k} + \left\|\underline{D\mathbf{N}_{\le m}U}\right\|_{r_k-\delta_k} \\
&\le \frac{4\epsilon_k}{\delta_k}\frac{4}{3}\epsilon_0 + 4\epsilon_k\sum_{l=0}^{k-1}\frac{\epsilon_l}{r_l-r_{l+1}} \le \frac{\epsilon_0}{\delta}\frac{16}{3}\epsilon_0 + 4\epsilon_k\left(\frac{8C}{7} + 2^{b/2}\frac{r_0}{r_\infty}\right)\frac{\epsilon_0}{r_0} \\
&\le \left[\frac{16}{3}c_1 + 4\frac{8}{7} + 4\cdot 2^{b/2}4^{b+2}\right]c_0r_0\epsilon_0.
\end{aligned}$$

Here, we have used the inequalities $\frac{\epsilon_0}{\delta} = c_1 c_0 r_0$, $\frac{r_0}{r_\infty} = C4^{b+2}$ and $\frac{\epsilon_0}{r_0} = \frac{c_0 r_0}{C}$. Summing up we have

$$\left\| \underline{[U, \mathbf{E}] + [U, \mathbf{N}_{\leq m}] + \mathbf{R}_{\geq m+1}} \right\|_{r_k - \delta_k} \\ \leq \epsilon_0 \left[8Cc_1 + \frac{1}{4^k} + c_0 r_0 \left(\frac{16}{3} c_1 + 4\frac{8}{7} + 4^{b+3} \cdot 2^{b/2} \right) \right],$$

and therefore the same estimate holds for $\left\| \underline{(5.76)} \right\|_{r_k - \delta_k}$. Summing up the different contributions, we have

$$\left\| \underline{\mathbf{R}_{\geq 2m+1}} \right\|_{r_k - \delta_k} \\ \leq \epsilon_0 \left[8Cc_1 + \frac{1}{4^k} + c_0 r_0 \left((2^4 + 6)c_1 + 2^9 C c_1^2 + 4\frac{8}{7} + 4^{b+3} \cdot 2^{b/2} + 2^9 c_0 c_1^2 r_0 \right) \right], \tag{5.80}$$

which, provided

$$r_0 < \frac{4}{7 \cdot 2^9 c_0 c_1^2}, \quad r_0 < (Cc_1 - \frac{1}{4}) \left[c_0 \left((2^4 + 6)c_1 + 2^9 C c_1^2 + 4\frac{9}{7} + 4^{b+3} \cdot 2^{b/2} \right) \right]^{-1}, \tag{5.81}$$

gives

$$\left\| \underline{\mathbf{R}_{\geq 2m+1}} \right\|_{r_k - \delta_k} \leq 9 c_1 C \epsilon_0. \tag{5.82}$$

Now, from Lemma A.4, since $b \geq 2$, one has

$$\left\| \underline{\mathbf{R}_{\geq 2m+1}} \right\|_{r_{k+1}} \leq q_m^{2m+1} 9 c_1 C \epsilon_0 = 3 \cdot 4^{b+2} 2^{-bk\left(2 + \frac{1}{2^k}\right)} C \epsilon_0. \tag{5.83}$$

For $k = 1$ (which corresponds to $m = 2$), we have

$$\left\| \underline{\mathbf{R}_{\geq 5}} \right\|_{r_2} \leq 3C \frac{4^{b+2}}{2^{\frac{5}{2}b}} \epsilon_0 = \frac{3 \cdot 4^2 C}{2^{b/2}} \epsilon_0 \leq \frac{\epsilon_0}{4^2},$$

provided

$$\frac{3 \cdot 4^2 C}{2^{b/2}} < \frac{1}{2^4} \iff \ln(3 \cdot 4^2 C) < (\frac{b}{2} - 4) \ln 2,$$

which is equivalent to

$$b > 8 + \frac{2 \ln(3 \cdot 4^2 C)}{\ln 2}. \tag{5.84}$$

For $k \geq 2$ we have

$$\left\| \underline{\mathbf{R}_{\geq 2m+1}} \right\|_{r_{k+1}} \leq 3 \cdot 4^{b+2} 2^{-2bk} C \epsilon_0 \leq \frac{\epsilon_0}{4^{k+1}},$$

provided

$$3 \cdot 4^2 C < 4^{b(k-1) - (k+1)} \iff \frac{\ln(3 \cdot 4^2 C)}{\ln 4} < b(k-1) - (k+1) \\ \iff b > \frac{k+1}{k-1} + \frac{1}{k-1} \frac{\ln(3 \cdot 4^2 C)}{\ln 4},$$

which, since the r.h.s. is a decreasing function of k, is implied by

$$b > 3 + \frac{\ln(3 \cdot 4^2 C)}{2 \ln 4}, \tag{5.85}$$

which in turn is implied by (5.84). □

From Lemma 5.5, by a completely standard argument, the following Corollary follows

Corollary 3. *The sequence of transformations* $\{\Psi_k\}_{k\geq 1}$ *defined by* $\Psi_k := \Phi^{-1}_{2^{k-1}} \circ \cdots \circ \Phi^{-1}_1$ *converges to an analytic transformation* Ψ *in a neighborhood of the origin and it conjugates the family* $\{X^i\}_{i\geq 1}$ *to a a family of normal forms* $\{NF^i\}_{i\geq 1}$.

Acknowledgements. We thank Michela Procesi for many discussions. We acknowledge the support of Università degli Studi di Milano and of GNFM.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A. A Technical Lemma

Lemma A.1. *Equation* (5.63) *holds.*

Proof. Denote by d_k the l.h.s. of (5.63), one has

$$d_k = \exp\left(\sum_{l=0}^{k-1} \ln m^{-\frac{b}{m}}\right) = \exp\left(-\sum_{l=0}^{k-1} \frac{b}{2^l} \ln 2^l\right) = \exp\left(-\frac{b \ln 2}{2} \sum_{l=0}^{k-1} \frac{l}{2^{l-1}}\right)$$
$$= \exp\left(-\frac{b \ln 2}{2} 4 \left(1 - \frac{k+1}{2^k}\right)\right),$$

where we used the formula

$$\sum_{l=0}^{k-1} \frac{l}{2^{l-1}} = 4\left(1 - \frac{k+1}{2^k}\right).$$

Now, the result immediately follows. □

Lemma A.2. *Equation* (5.64) *holds.*

Proof. We use the discrete analogue of the formula of the Duhamel formula, namely we make the substitution $r_k = d_k s_k$, where d_k was defined in the proof of Lemma A.1. One gets

$$r_{k+1} = d_{k+1} s_{k+1} = q_{2^k} d_k s_{k+1} = q_{2^k}(d_k s_k - \delta_k)$$

and thus

$$s_{k+1} = s_k - \frac{\delta_k}{d_k}, \quad s_1 = \frac{r_1}{d_1} = r_1,$$

from which

$$s_k = s_1 - \sum_{l=1}^{k-1} \frac{\delta_l}{d_l}.$$

Now, one has

$$\sum_{l=1}^{k-1} \frac{\delta_l}{d_l} = \sum_{l=1}^{k-1} \frac{\delta}{4^l} \frac{4^b}{4^{b\frac{l+1}{2^l}}} \leq \sum_{l=1}^{k-1} \frac{\delta}{4^l} 4^b = \frac{4^b}{3}\delta.$$

Thus,

$$r_k \geq d_k \left(\frac{r_1}{d_1} - \frac{4^b}{3}\delta \right).$$

□

Lemma A.3. *Equation* (5.66) *holds.*

Proof. The first inequality is trivial. We discuss the second one. Using the definition of r_{k+1}, we have

$$\frac{\epsilon_k}{r_k - r_{k+1}} = \frac{\epsilon_k}{r_k(1-q_{2^k}) + q_{2^k}\delta_k} \leq \frac{\epsilon_k}{r_k(1-q_{2^k})}\,; \tag{A.86}$$

now, one has

$$1 - q_m = 1 - \exp\left(-\frac{b}{m} \ln m \right),$$

which is of the form $1 - e^{-x}$ with x varying from 0 to $\frac{b}{2}\ln 2$. Remarking that in an interval $[0, x_0]$ one has

$$1 - e^{-x} \geq e^{-x_0} x,$$

we get

$$1 - q_m \geq 2^{-b/2} \left(\frac{b}{m} \ln m \right) = \frac{b}{2^{k+b/2}} \ln 2^k = \frac{k}{2^k} \frac{b}{2^{b/2}} \ln 2,$$

and thus, for $k \geq 1$,

$$\frac{\epsilon_k}{r_k(1-q_{2^k})} \leq \frac{\epsilon_0}{r_\infty} \frac{2^{b/2}}{b \ln 2} \frac{2^k}{k} \frac{1}{4^k} = \frac{\epsilon_0}{r_\infty} \frac{2^{b/2}}{b \ln 2} \frac{1}{k 2^k}.$$

Now one has

$$\sum_{k \geq 1} \frac{x^k}{k} = \sum_{k \geq 1} \int_0^x y^{k-1} dy = \sum_{k \geq 0} \int_0^x y^k dy = \int_0^x \frac{1}{1-y} dy$$
$$= [-\ln|1-y|]_0^x = -\ln|1-x|,$$

which, for $x = 1/2$, gives

$$\sum_{k \geq 1} \frac{1}{k 2^k} = \ln 2,$$

and thus

$$\sum_{l\geq 2}\frac{\epsilon_{l-1}}{r_{l-1}-r_l}\leq\frac{\epsilon_0}{r_\infty}2^{b/2}.$$

adding the first term, namely $\frac{8C}{8C-1}\frac{\epsilon_0}{r_0}$, one gets the thesis immediately follows. □

Lemma A.4. *Let* $\mathbf{F}$ *be a summable normally analytic vector fields with* F^i *having a zero of order m at the origin for all* i. *Let* $0<\alpha\leq 1$, *then*

$$\left\|\underline{\mathbf{F}}\right\|_{\alpha r}\leq\alpha^m\left\|\underline{\mathbf{F}}\right\|_r. \tag{A.87}$$

Proof. Consider the function $\underline{\mathbf{F}}(z)=\sum_{Q,i}F_{Q,i}z^Q\vec{e}_i$; since all the coefficients are positive one has, for any i,

$$\sum_Q F_{Q,i}(\alpha z)^Q=\alpha^m\sum_{Q,i}\alpha^{|Q|-m}F_{Q,i}z^Q\leq\alpha^m\sum_Q F_{Q,i}z^Q$$

Thus one gets

$$\sup_{\|z\|\leq\alpha r}\left\|\underline{\mathbf{F}}(z)\right\|_+=\sup_{\|z\|\leq r}\left\|\underline{\mathbf{F}}(\alpha z)\right\|_+\leq\alpha^m\sup_{\|z\|\leq r}\left\|\underline{\mathbf{F}}(z)\right\|_+.$$

□

References

[Arn76] Arnold, V.I.: Méthodes mathématiques de la mécanique classique. Mir, Moscow (1976)

[Bam03] Bambusi, D.: Birkhoff normal form for some nonlinear PDEs. Commun. Math. Phys. **234**(2), 253–285 (2003)

[Bam05] Bambusi, D.: Galerkin averaging method and Poincaré normal form for some quasilinear PDEs. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **4**(4), 669–702 (2005)

[Bam08] Bambusi, D.: A Birkhoff normal form theorem for some semilinear PDEs. In: Craig, W. (ed.) Hamiltonian Dynamical Systems and Applications. NATO Sci. Peace Secur. Ser. B Phys. Biophys., pp. 213–247. Springer, Dordrecht (2008)

[BBGK95] Bättig, D., Bloch, A.M., Guillot, J.-C., Kappeler, T.: On the symplectic structure of the phase space for periodic KdV, Toda, and defocusing NLS. Duke Math. J. **79**(3), 549–604 (1995)

[BDGS07] Bambusi, D., Delort, J.-M., Grébert, B., Szeftel, J.: Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds. Commun. Pure Appl. Math. **60**(11), 1665–1690 (2007)

[BG06] Bambusi, D., Grébert, B.: Birkhoff normal form for partial differential equations with tame modulus. Duke Math. J. **135**(3), 507–567 (2006)

[BKM18] Berti, M., Kappeler, T., Montalto, R.: Large KAM tori for perturbations of the defocusing NLS equation. Astérisque, **403**, viii+148 (2018)

[BM16] Bambusi, D., Maspero, A.: Birkhoff coordinates for the Toda lattice in the limit of infinitely many particles with an application to FPU. J. Funct. Anal. **270**(5), 1818–1887 (2016)

[Bru72] Bruno, A.D.: Analytical form of differential equations. Trans. Mosc. Math. Soc **25**, 131–288 (1971). 26, 199–239 (1972), 1971–1972

[Dic03] Dickey, L.A.: Soliton Equations and Hamiltonian Systems. Advanced Series in Mathematical Physics, vol. 26, 2nd edn. World Scientific Publishing Co., Inc., River Edge (2003)

[GK14] Grébert, B., Kappeler, T.: The defocusing NLS equation and its normal form. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich (2014)

[GL97] Giorgilli, A., Locatelli, U.: Kolmogorov theorem and classical perturbation theory. Z. Angew. Math. Phys. **48**(2), 220–261 (1997)

[HK08] Henrici, A., Kappeler, T.: Global Birkhoff coordinates for the periodic Toda lattice. Nonlinearity **21**(12), 2731–2758 (2008)

[Ito89] Ito, H.: Convergence of Birkhoff normal forms for integrable systems. Comment. Math. Helv. **64**, 412–461 (1989)
[Ito92] Ito, H.: Integrability of Hamiltonian systems and Birkhoff normal forms in the simple resonance case. Math. Ann. **292**, 411–444 (1992)
[KLTZ09] Kappeler, T., Lohrmann, P., Topalov, P., Zung, N.T.: Birkhoff coordinates for the focusing NLS equation. Commun. Math. Phys. **285**(3), 1087–1107 (2009)
[KP03] Kappeler, T., Pöschel, J.: KdV & KAM, Volume 45 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics (Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics). Springer, Berlin, (2003)
[KP10] Kuksin, S., Perelman, G.: Vey theorem in infinite dimensions and its application to KdV. Discrete Contin. Dyn. Syst. **27**(1), 1–24 (2010)
[Kuk87] Kuksin, S.B.: Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. Funct. Anal. Appl. **21**, 192–205 (1987)
[LM98] Locatelli, U., Meletlidou, E.: Convergence of Birkhoff normal form for essentially isochronous systems. Meccanica **33**(2), 195–211 (1998)
[Mag78] Magri, F.: A simple model of the integrable Hamiltonian equation. J. Math. Phys. **19**(5), 1156–1162 (1978)
[Mas18] Maspero, A.: Tame majorant analyticity for the Birkhoff map of the defocusing nonlinear Schrödinger equation on the circle. Nonlinearity **31**(5), 1981–2030 (2018)
[MP18] Maspero, A., Procesi, M.: Long time stability of small finite gap solutions of the cubic nonlinear Schrödinger equation on $\mathbb{T}^2$. J. Differ. Equ. **265**(7), 3212–3309 (2018)
[Nik86] Nikolenko, N.V.: The method of Poincaré normal form in problems of integrability of equations of evolution type. Russ. Math. Surv. **41**, 63–114 (1986)
[Rüs67] Rüssmann, H.: Über die Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung. Math. Ann. **169**, 55–72 (1967)
[Sto00] Stolovitch, L.: Singular complete integrabilty. Publ. Math. I.H.E.S. **91**, 133–210 (2000)
[Sto05] Stolovitch, L.: Normalisation holomorphe d'algèbres de type Cartan de champs de vecteurs holomorphes singuliers. Ann. Math. **161**, 589–612 (2005)
[Sto08] Stolovitch, L.: Normal forms of holomorphic dynamical systems. In: Craig, W. (ed.) Hamiltonian Dynamical Systems and Applications, pp. 249–284. Springer, Berlin (2008)
[Tre01] Treves, F.: An algebraic characterization of the Korteweg-de Vries hierarchy. Duke Math. J. **108**(2), 251–294 (2001)
[Vey78] Vey, J.: Sur certains systèmes dynamiques séparables. Am. J. Math. **100**, 591–614 (1978)
[Vey79] Vey, J.: Algèbres commutatives de champs de vecteurs isochores. Bull. Soc. Math. France **107**, 423–432 (1979)
[Zun02] Zung, N.T.: Convergence versus integrability in Poincaré–Dulac normal form. Math. Res. Lett. **9**(2–3), 217–228 (2002)
[Zun05] Zung, N.T.: Convergence versus integrability in Birkhoff normal form. Ann. Math. (2) **161**(1), 141–156 (2005)

Communicated by C. Liverani