Finite dimensional invariant KAM tori for tame vector fields

L. Corsi*, **R. Feola****, **M. Procesi**†

* Georgia Institute of Technology, Altanta, lcorsi6@math.gatech.edu;

** SISSA, Trieste, rfeola@sissa.it;

† Università di Roma Tre, procesi@mat.uniroma3.it

July 21, 2017

Abstract

We discuss a Nash-Moser/ KAM algorithm for the construction of invariant tori for *tame* vector fields. Similar algorithms have been studied widely both in finite and infinite dimensional contexts: we are particularly interested in the second case where tameness properties of the vector fields become very important. We focus on the formal aspects of the algorithm and particularly on the minimal hypotheses needed for convergence. We discuss various applications where we show how our algorithm allows to reduce to solving only linear forced equations. We remark that our algorithm works at the same time in analytic and Sobolev class.

Contents

1 Introduction

The aim of this paper is to provide a general and flexible approach to study the existence, for finite or infinite dimensional dynamical systems of of finite dimensional invariant tori carrying a quasi-periodic flow. To this purpose we discuss an iterative scheme for finding invariant tori for the dynamics of a vector field $F = N_0 + G$, where N_0 is a linear vector field which admits an invariant torus and G is a *perturbation.*
By an *invariant torus* of $\dot{u} = F(u)$, with $u \in E$ a Banach space, we mean an embedding $\mathbb{T}^d \to E$, which is invariant under the dynamics of F, i.e. the vector field F is tangent to the embedded torus. Similarly an *analytic* invariant torus is a map $\mathbb{T}^d_s \to E$ (with $s > 0$) where $\mathbb{T}^d_s$ is a *thickened* torus[1].

It is very reasonable to work in the setting of the classical Moser scheme of [1], namely F acts on a product space (θ, y, w) where $\theta \in \mathbb{T}^d_s, y \in \mathbb{C}^{d_1}$ while $w \in \ell_{a,p}$ some separable scale of Hilbert spaces. The variables θ appear naturally as a parametrization for the invariant torus of N_0. The y variables are constants of motion for N_0, in applications they naturally appear as "conjugated" to θ, for instance in the Hamiltonian setting they come from the symplectic structure. The variables w describe the dynamics in the directions orthogonal to the torus. The main example that we have in mind is[2]

$$N_0 = \omega^{(0)} \cdot \partial_\theta + \Lambda^{(0)} w \partial_w \tag{1.1}$$

where $\omega^{(0)} \in \mathbb{R}^d$ is a constant vector while $\Lambda^{(0)}$ is a block-diagonal skew self–adjoint operator, independent of θ. Note that N_0 has the invariant torus $y = 0, w = 0$, where the vector field reduces to $\dot{\varphi} = \omega^{(0)}$.

Regarding the normal variables w, we do not need to specify $\ell_{a,p}$ but only give some properties, see Hypothesis 2.1, which essentially amount to requiring that $\ell_{a,p}$ is a weighted sequence space[3] where $a \geq 0$ is an exponential weight while $p > 0$ is polynomial, for example

$$w = \{w_j\}_{j \in \mathcal{I} \subseteq \mathbb{N}}\,, \quad w_j \in \mathbb{C}\,, \quad \|w\|^2_{a,p} := \sum_{j \in \mathcal{I} \subseteq \mathbb{N}} \lambda_j^{2p} e^{2a\lambda_j} |w_j|^2\,, \quad 0 < \lambda_j \leq \lambda_{j+1} \ldots \quad \lambda_i \to \infty.$$

Note that if $\mathcal{I}$ is a finite set our space is finite dimensional.

The existence of an invariant torus in the variables (θ, y, w) means the existence of a map $\mathbb{T}^d_s \xrightarrow{h} \mathbb{C}^{d_1} \times \ell_{a,p}$ of the form $\theta \mapsto h(\theta) = (h^{(y)}(\theta), h^{(w)}(\theta))$ such that

$$\mathcal{F}(F, h) \equiv \mathcal{F}(h) := F^{(\mathtt{v})}(\theta, h^{(y)}(\theta), h^{(w)}(\theta)) - \partial_\theta h^{(\mathtt{v})} \cdot F^{(\theta)}(\theta, h^{(y)}(\theta), h^{(w)}(\theta)) = 0, \quad \mathtt{v} = y, w, \tag{1.2}$$

hence it coincides with the search for zeros of the functional $\mathcal{F}$ with unknown h. Here h lives in some Banach space, say $H^q(\mathbb{T}^d_s; \mathbb{C}^{d_1} \times \ell_{a,p})$ on which $\mathcal{F}$ is at least differentiable. Note moreover that $\mathcal{F}(N_0, 0) = 0$ trivially. Since we are in a perturbative setting, once one has the torus embedding, one can study the dynamics of the variables θ restricted to the torus and look for a change of variables $\varphi \mapsto \theta(\varphi)$ which conjugates the dynamics

[1] we use the standard notation $\mathbb{T}^d_s := \left\{\theta \in \mathbb{C}^d \,:\, \mathrm{Re}(\theta) \in \mathbb{T}^d,\ \max_{h=1,\ldots,d} |\mathrm{Im}\,\theta_h| < s\right\}$
[2] we use the standard notation for vector fields $F = \sum_{\mathtt{v}=\theta,y,w} F^{(\mathtt{v})} \partial_{\mathtt{v}}$
[3] a good example is to consider spaces of Sobolev or analytic functions on compact manifolds

of θ to the linear dynamics $\dot\varphi = \omega$ with $\omega \sim \omega^{(0)}$ a rationally independent vector. Obviously this could be done directly by looking for a *quasi-periodic solution* i.e. a map

$$h : \varphi \to (h^{(\theta)}(\varphi), h^{(y)}(\varphi), h^{(w)}(\varphi)), \quad h \in H^q(\mathbb{T}^d_s; \mathbb{C}^d \times \mathbb{C}^{d_1} \times \ell_{a,p}).$$

which solves the functional equation

$$\mathcal{F}(h) := F(h^{(\theta)}(\varphi), h^{(y)}(\varphi), h^{(w)}(\varphi)) - \omega \cdot \partial_\varphi h = 0.$$

If we take $\ell_{a,p} = \emptyset$, $d_1 = d$, this is the classical KAM framework of Kolmogorov [2], Arnold [3] and Moser [4]; see also [5, 6, 7].

Even in the simplest setting, equation (1.2) cannot be solved by classical Implicit Function Theorem. In fact typically the operator $\mathcal{F}$ in (1.2) linearized at $F = N_0, h = 0$ is not invertible on $H^q(\mathbb{T}^d_s; \mathbb{C}^{d_1} \times \ell_{a,p})$ since it has a spectrum which accumulates to zero (the so-called small divisors). To overcome this problem one can use a *Nash-Moser* iterative scheme in order to find a sequence of approximate solutions rapidly converging to the true solution. The fast convergence is used to control the loss of regularity due to the small divisors. Such schemes are adaptations to Banach spaces of the Newton method to find zeros of functions, see [8].

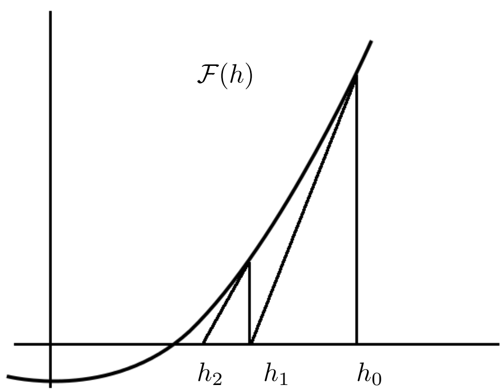

Figure 1.1: Three steps of the Newton algorithm $h_{n+1} := h_n - (d_h\mathcal{F}(h_n))^{-1}[\mathcal{F}(h_n)]$

In order to run an algorithm of this type one must be able to control the linearized operator in a neighborhood of the expected solution; see Figure 1.1. Due to the presence of small divisors it is not possible to invert such operator as a functional from a Sobolev space to itself (not even the operator linearized about zero). However, since the Newton scheme is quadratic, one may accept that $d_h\mathcal{F}^{-1}$ is well defined as an unbounded "tame" operator provided that one has a good control on the loss of regularity.

Of course, if on the one hand in order to achieve such control it is in general not sufficient to give lower bounds on the eigenvalues, on the other hand one surely needs to avoid zero or "too small" eigenvalues. To this end one typically uses parameter modulation. Precisely one assumes that $\omega^{(0)}, \Lambda^{(0)}$ depend (non trivially and with some regularity) on some parameters $\xi \in \mathbb{R}^k$ for some k. Unfortunately, equations from physics may come with no such external parameters, so that one needs to extract them from the initial data: however we shall not address this issue here.

An equivalent approach to the Nash-Moser scheme is to find a change of coordinates

$$(y, w) \rightsquigarrow (\tilde y + h^{(y)}(\theta), \tilde w + h^{(w)}(\theta))$$

such that the push forward of the vector field has an invariant torus at $\tilde y = 0$, $\tilde w = 0$. Clearly the equation for h is always (1.2), hence the solvability conditions on the inverse of the linearized operator appear also in this context.

Following this general strategy one can look for different types of changes of variables such that the push forward of the vector field has a particularly simple form, and the existence of an invariant torus follows trivially. For instance in classical KAM Theorems the idea is to look for an *affine* change of variables such that not only $y = 0, w = 0$ is an invariant torus but the vector filed F linearized in the directions normal to the torus is diagonal. This means that in the quadratic scheme one not only performs a traslation but also a linear change of variables which approximately diagonalizes the linearized operator at each step. Naturally this makes the inversion of $d_h\mathcal{F}$ much simpler. It is well known that it is possible to diagonalize a finite dimensional matrix if it has distinct eigenvalues. Then, in order to diagonalize the linearized operator at an approximate solution, one asks for lower bounds on the differences of the eigenvalues (the so called "Second Mel'nikov" conditions). Under these assumptions the bounds on the inverse follow by just imposing lower bounds on the eigenvalues (the so called the First Mel'nikov conditions). This requirement together with some structural hypotheses on the system (Hamiltonianity, reversibility, ...) provides existence and linear stability of the possible solution. More in general if one wants to cancel non linear terms in the vector fields one needs to add some more restrictive conditions on the linear combinations of the eigenvalues (higher order Mel'nikov conditions). Naturally such conditions are not necessary for the invertibility, actually in many applications they cannot be imposed (already in the case of the NLS on the circle the eigenvalues are double) and typically in a Nash-Moser scheme they are not required.

Quadratic algorithms for the construction of finite dimensional invariant tori have been used in the literature both in finite, see for instance [9] and references therein, or infinite dimensional setting. Starting from [10, 11], this problem has been widely studied in the context of *Hamiltonian* PDEs; a certainly non exaustive list of classical results could be for instance [12, 13, 14, 15, 16, 17, 18, 19, 20, 21], in which either the KAM scheme or the Lyapunov-Schmidt decomposition and the Newton iteration method are used.

The aforementioned literature is mostly restricted to semilinear PDEs on the circle. More recently also other extensions have been considered, such as cases where the spatial variable is higher dimensional, or when the perturbation is unbounded. Regarding the higher dimensional cases, besides the classical papers by Bourgain, we mention also [22, 23, 24, 25, 26, 27, 28, 29, 30], where also manifolds other than the torus are considered.

The first results with unbounded perturbations can be found in [18, 31, 32, 33, 34, 35]; in these papers the authors follows the classical KAM approach, which is based on the so-called "second Mel'nikov" conditions. Unfortunately their approach fails in the case of *quasi-linear* PDEs (i.e. when the nonlinearity contains derivatives of the same order as the linear part). This problem has been overcome in [36, 37, 38], for the periodic case, and in [39, 40], first for forced and then also for autonomous cases, see also [41, 42, 43, 44]. The key idea of such papers is to incorporate into the reducibility scheme techniques coming from pseudo-differential calculus. The extensions for the autonomous cases are based on the ideas developed in [45], where the Hamiltonian structure is exploited in order to extend results on forced equations to autonomous cases.

In all the results mentiond so far, the authors deal with two problems: the convergence of the iterative scheme, and the invertibility of the linearized operator in a neighborhood of the expected solution. Typically these problems are faced at the same time, by giving some non-degeneracy conditions on the spectrum of the linearized operator in order to get estimates on its inverse. However, while the bounds on the linearized operator clearly depend on the specific equation one is dealing with, the convergence of the scheme is commonly believed to be adjustable case by case.

Our purpose is to separate the problems which rely only on abstract properties of the vector fields from those depending on the particular equation under study.

Our point of view is to look for a change of coordinates, say Ψ, such that the push-forward of the vector field has an invariant torus at the origin. In fact all the results described above can be interpreted in this way. Typically one chooses a priori a *group* $\mathcal{G}$ of changes of coordinates in which one looks for Ψ. Then, for many choices of $\mathcal{G}$, one may impose smallness conditions on the perturbation (depending on the choice of $\mathcal{G}$) and perform an iterative scheme which produces Ψ, provided that the parameters ξ belong to some *Cantor like* set (again depending on the choice of $\mathcal{G}$). In this paper we impose some mild conditions on the group $\mathcal{G}$ such that an iterative algorithm can be performed. Then we explicitly state the *smallness conditions* and the conditions on the parameters. Then, in Section 4 we show some particularly relevant choices of $\mathcal{G}$ (in fact we always describe the *algebra* which generates $\mathcal{G}$) and the resulting *Cantor like* sets. In this way, in order to apply our theorem to a particular vector field, one has to: first choose a group, then verify the smallness conditions and finally check that the *Cantor like* set is not empty. As it might be intuitive, the simpler the structure of $\mathcal{G}$ the more complicated is to prove that the resulting *Cantor like* set in not empty.

The present paper is mainly inspired by the approach in [45], but we follow a strategy more similar to the one of [1]. In particular this allows us to cover also non-Hamiltonian cases, which require different techniques w.r.t. [45]; compare Subsections 4.2 and 4.3. We essentially produce an algorithm which interpolates a Nash-Moser scheme and KAM scheme. On the one hand we exploit the functional approach of the Nash-Moser scheme, which allows to use and preserve the "PDE structure" of the problems; on the other hand we leave the freedom of choosing a convenient set of coordinates during the iteration (which is typical of a KAM scheme). This allows us to deal with more general classes of vector fields and with analytic nonlinearities. This last point is particularly interesting in applications to quasi-linear PDEs, where the only results in the literature are for Sobolev regularity; see [46]. In fact, we develop a formalism which allows us to cover cases with both analytic or Sobolev regularity, by exploiting the properties of *tame* vector fields, introduced in Definition 2.13 and discussed in Appendix B.

Another feature of our algorithm is related to the "smallness conditions"; see Constraints 2.21 and D.1 and the assumption (2.54). Clearly in every application the smallness is given by the problem, and one needs adjusts the algorithm accordingly. Again, while this is commonly believed to be easy to achieve, our point of view is the opposite, i.e. finding the mildest possible condition that allow the algorithm to converge, and use them only when it is necessary. Of course if on the one hand this makes the conditions intricated, on the other hand it allows more flexibility to the algorithm.

Description of the paper. Let us discuss more precisely the aim of the present paper. We consider vector fields of the form

$$\begin{cases} \dot\theta = F^{(\theta)} := \omega^{(0)}(\xi) + G^{(\theta)}(\theta, y, w) \\ \dot y = F^{(y)} := G^{(y)}(\theta, y, w) \\ \dot w = F^{(w)} := \Lambda^{(0)}(\xi) w + G^{(w)}(\theta, y, w) \end{cases} \tag{1.3}$$

where $N_0 = \omega^{(0)} \cdot \partial_\theta + \Lambda^{(0)} w \partial_w$ and G is a *perturbation*, i.e. (1.3) admits an approximately invariant torus. Note that this does not necessarily mean that G is small, but only that it is approximately tangent to the torus. Recall that $\xi \in \mathcal{O}_0 \in \mathbb{R}^k$ is a vector of parameters.

Then the idea, which goes back to Moser [1], is to find a change of coordinates such that in the new coordinates the system (1.3) takes the form

$$\begin{cases} \dot\theta = \omega(\xi) + \tilde G^{(\theta)}(\theta, y, w) \\ \dot y = \tilde G^{(y)}(\theta, y, w) \\ \dot w = \Lambda^{(0)}(\xi) w + \tilde G^{(w)}(\theta, y, w) \end{cases} \tag{1.4}$$

with $\omega \sim \omega^{(0)}$, the average of $\tilde G(\theta, 0, 0)$ is zero and $\tilde G^{(\mathtt{v})}(\theta, 0, 0) \equiv 0$ for $\mathtt{v} = y, w$. More precisely in our main Theorem 2.25 we prove the convergence of an iterative algorithm which provides a change of variables transforming (1.3) into (1.4), for all choices of ξ in some explicitly defined set $\mathcal{O}_\infty$ (which however might be empty).

The changes of variables are not defined uniquely, and one can specify the problem by – for instance – identifying further terms in the Taylor expansion of $\tilde G$ w.r.t. the variables y and w which one wants to set to zero. Of course different choices of changes of variables modify the set $\mathcal{O}_\infty$ so that in the applications it is not obvious to understand which is the best choice. In fact finding the setting in which one is able to prove that $\mathcal{O}_\infty$ is non empty and possibly of positive measure is the most difficult part in the applications. We do not address this problem at all. Our aim is instead to study very general classes of changes of variables and find general Hypotheses on the functional setting, the vector field under study and the terms of the Taylor series that one wants to cancel, under which such an algorithm can be run, producing an explicit set $\mathcal{O}_\infty$.

In particular in our phase space $\mathbb{T}^d_s \times \mathbb{C}^{d_1} \times \ell_{a,p}$ we *do not distinguish* the cases where either s or a are equal to zero (Sobolev cases) from the analytic cases. In the same spirit we do not require that the vector field is analytic but only that it is C^q for some large finite q. The key ingredients of the paper are the following.

Tame Vector fields. We require that F is C^k*-tame up to order* q, see Definition 2.13, namely it is tame together with its Taylor expansion up to finite order k w.r.t. y, w and it is regular up to order q in θ, see Subsection 2.2. We make this definition quantitative by denoting a *tameness constant* for G by $C_{\vec{v},p}(G)$, here $\vec{v}$ contains all the information relative to the domain of definition of G, while p gives us the Sobolev regularity. In Appendix B we describe some properties of tame vector fields which we believe are interesting for themselves. Finally our vector fields are not necessarily bounded, instead they may lose some regularity, namely we allow

$$F : \mathbb{T}^d_s \times \left\{ y \in \mathbb{C}^{d_1} \, : \, |y|_1 < r^{\mathtt{s}} \right\} \times \left\{ w \in \ell_{a,p+\nu} \, : \, \|w\|_{a,\mathtt{p}_1} < r \right\} \to \mathbb{C}^d \times \mathbb{C}^{d_1} \times \ell_{a,p} \tag{1.5}$$

for some fixed $\nu \geq 0$. The properties we require are very general and are satisfied by large class of PDEs, for instance it is well known that these properties are satisfied by a large class of composition operators on Sobolev spaces; see [4, 47].

The $(\mathcal{N}, \mathcal{X}, \mathcal{R})$ decomposition. We choose a subspace of polynomials of maximal degree $\mathtt{n}$, which we call $\mathcal{X}$, containing all the terms we wants to "cancel out" from G. This space contains the algebra of the changes of variables we shall apply. Clearly the subspace $\mathcal{X}$ must be chosen so that a vector field with no terms in $\mathcal{X}$ possesses an invariant torus. In order to identify the part of F belonging to $\mathcal{X}$ we Taylor-expand it about $y = 0$, $w = 0$: since F is assumed to be a C^q vector field, this obviously requires that q is larger than $\mathtt{n}$ i.e. the maximal degree of the monomials in $\mathcal{X}$. With some abuse of notation (see Definition 2.11 and comments before it) we denote this operation as a projection $\Pi_{\mathcal{X}} F$.

We also define a space of polynomial vector fields $\mathcal{N}$ (which does not intersect $\mathcal{X}$) such that $N_0 \in \mathcal{N}$. We allow a lot of freedom on the choice of $\mathcal{N}$, provided that it satisfies some rather general hypotheses, in particular all vector fields in $\mathcal{N}$ should have an invariant torus at zero, and $\mathcal{N}$ should contain the unperturbed vector field N_0; in fact we shall require a stronger condition on N_0, i.e. that it is diagonal; see the example in (1.1) and Definition 2.20 for a precise formulation.

Our space of C^k-tame vector fields is then decomposed uniquely as $\mathcal{X} \oplus \mathcal{N} \oplus \mathcal{R}$, and we may write

$$(\mathbb{1} - \Pi_{\mathcal{X}})F = N + R$$

where $N = \Pi_{\mathcal{N}} F$ while $R = \Pi_{\mathcal{R}} F$ is a remainder.

The Invariant subspace $\mathcal{E}$. We choose a space of vector fields $\mathcal{E}$ (see Definition 2.19) where we want our algorithm to run. Such space appears naturally in the applications where usually one deals with special structures (such as Hamiltonian or reversible structure) that one wishes to preserve throughout the algorithm. As one might expect, the choice of the space $\mathcal{E}$ influences the set $\mathcal{O}_\infty$. In the applications to PDEs the choice of the space $\mathcal{E}$ is often quite subtle: we give some examples in Section 4.

Regular vector fields. We choose a subspace of polynomial vector fields, which we denote by *regular vector fields* and endow with a Hilbert norm $|\cdot|_{\vec{v},p}$ with the only restriction that they should satisfy a set of properties detailed in Definition 2.18, for instance we require that all regular vector fields are tame with tameness constant equal to the norm $|\cdot|_{a,p}$ and moreover that for $p = \mathtt{p}_1$ this tameness constant is *sharp*. Throughout our algorithm we shall apply close to identity changes of variables which preserve $\mathcal{E}$ and are generated by such vector fields (this is the group $\mathcal{G}$ of changes of variables). This latter condition can probably be weakened but, we believe, not in a substantial way: on the other hand it is very convenient throughout the algorithm.

Our Goal. *We fix any decomposition $(\mathcal{N}, \mathcal{X}, \mathcal{R})$, any space $\mathcal{E}$ and any space of regular vector fields $\mathcal{A}$ provided that they satisfy Definitions 2.17, 2.19 and 2.18.*

We assume that $F = N_0 + G$ belongs to $\mathcal{E}$, is $C^{\mathtt{n}+2}$ tame (the value of $\mathtt{n}$ being fixed by $\mathcal{E}$) and that $\Pi_{\mathcal{X}} F = \Pi_{\mathcal{X}} G$ is appropriately small while $(\mathbb{1} - \Pi_{\mathcal{X}})F$ is "controlled". We look for a change of coordinates $\mathcal{H}_\infty$ such that for *all $\xi \in \mathcal{O}_\infty$ one has*[4]

$$\Pi_{\mathcal{X}} (\mathcal{H}_\infty)_* F \equiv 0\,.$$

At the purely formal level, a change of coordinates Φ generated by a bounded vector field g, transforms F into

$$\Phi_*(F) \sim F + [g, F] + O(g^2) \sim e^{[g,\cdot]} F. \tag{1.6}$$

We find the change of variable we look for via an iterative algorithm; at each step we need Φ to be such that $\Pi_{\mathcal{X}}(\Phi_* F) = 0$ up to negligible terms, so we need to find g such that

$$\Pi_{\mathcal{X}}(F + [g, F]) = 0\,;$$

in other words we need to invert the operator $\Pi_{\mathcal{X}}[F, \cdot]$. Since one expects $g \sim X := \Pi_{\mathcal{X}} F$ (which is assumed to be small) then, at least formally, the term $[g, \Pi_{\mathcal{X}} F]$ is negligible and one needs to solve

$$\Pi_{\mathcal{X}}([(1 - \Pi_{\mathcal{X}})F, g]) - X := u \tag{1.7}$$

with $g \sim X$ and $u \sim X^2$. Equation (1.7) is called *homological equation* and in order to solve it one needs the "approximate invertibility" for the operator

$$\mathfrak{A}(\cdot) := \Pi_{\mathcal{X}}[(1 - \Pi_{\mathcal{X}})F, \cdot]. \tag{1.8}$$

Then the iteration is achieved by setting $\Psi_0 := \mathbb{1}$ and

$$\Psi_n := \Phi^1_{g_n} \circ \Psi_{n-1}\,, \quad F_n := (\Psi_n)_* F \tag{1.9}$$

where $\Phi^1_{g_n}$ is the time-1 flow map generated by the vector field g_n which in turn solves the homological equation

$$\Pi_{\mathcal{X}}([(1 - \Pi_{\mathcal{X}})F_{n-1}, g_n]) - X_{n-1} := u_n. \tag{1.10}$$

Since we need to preserve the structure, namely we need that at each step $F_n \in \mathcal{E}$, then at each step the change of variables $\Phi^1_{g_n}$ should preserve $\mathcal{E}$. In fact we require that g_n is a *regular* vector field,

In order to pass from the formal level to the convergence of the scheme we need to prove that g_n and u_n satisfy appropriate bounds.

Homological equation. We say that a set of parameters $\mathcal{O}$ satisfies the homological equation (for $(F, K, \vec{v}^0, \rho)$) if there exist g, u which satisfy (1.7) with appropriate bounds (depending on the parameters $K, \vec{v}^0, \rho$; $\vec{v}^0$ controls the domain of definition, ρ controls the size of the change of variables and K is an *ultravioled cut-off*), see Definition 2.23. Since F is a merely differentiable vector field the bounds are delicate since expressions like (1.6) may be meaningless in the sense that –apparently– the new vector field F_{n+1} is less regular than F_n, and hence it is not obvious that one can iterate the procedure. Indeed the commutator loses derivatives and thus one cannot use Lie series expansions in order to describe the change of variables. However one can use Lie series expansion formula on polynomials, such as $\Pi_{\mathcal{X}} \Phi_* F$. We show that, provided that X is small while $R, N - N_0$ are appropriately bounded, we obtain a converging KAM algorithm.

Note that the smallness of X implies the existence of a sufficiently good approximate solution; on the other hand, we only need very little control on $(\mathbb{1} - \Pi_{\mathcal{X}})F_n$, which results on very weak (but quite cumbersome) assumptions on R and especially on $N - N_0$; see (1.12) and Remark 2.22.

Compatible changes of variables. It is interesting to notice that at each step of the iterative scheme (1.9) we may apply any change of variables $\mathcal{L}_n$ with the only condition that it does not modify the bounds, i.e. $\Pi_{\mathcal{X}}(\mathcal{L}_n)_\star F_{n-1} \sim \Pi_{\mathcal{X}} F_{n-1}$ (and the same for the other projections). We formalize this idea in Definition 2.24 where we introduce the changes of variables compatible with $(F, K, \vec{v}^0, \rho)$. Then we may set

$$\mathcal{H}_n = \Phi^1_{g_n} \circ \mathcal{L}_n \circ \mathcal{H}_{n-1}\,, \quad F_n := (\mathcal{H}_n)_* F \tag{1.11}$$

[4]given a diffeomorphism Φ one defines the push-forward of a vector field F as $\Phi_* F = d\Phi(\Phi^{-1})[F(\Phi^{-1})]$.

and the algorithm is still convergent.
This observation is essentially tautological but it might well be possible that in a new set of coordinates it is simpler to invert the operator $\mathfrak{A}_n := \Pi_{\mathcal{X}}\mathrm{ad}((1-\Pi_{\mathcal{X}})F_{n-1})$ (for instance $\mathfrak{A}_n$ may be diagonal up to a negligible remainder, see Subsection 4.4). Note that since the approximate invertibility of $\mathfrak{A}_n$ is in principle independent from the coordinates (provided the change does not lose regularity), if one knows that a change of variables simplifying $\mathfrak{A}_n$ exists, there might be no need to apply it in order to deduce bounds on the approximate inverse. This is in fact the strategy of the papers [39, 40], where the authors study fully nonlinear equations and prove existence of quasi-periodic solutions with Sobolev regularity. On the other hand one might modify the definition of the subspace $\mathcal{X}$ in such a way that $\mathcal{L}_n$ is the time one flow of a regular bounded vector field in $\mathcal{X}$. The best strategy clearly depends on the application, so we leave the $\mathcal{L}_n$ as an extra degree of freedom.
We can summarize our result as follows, for the notations we refer to the informal Definitions written above.

Theorem. *Fix $\nu \geq 0$ as in (1.5), and fix a decomposition $(\mathcal{N}, \mathcal{X}, \mathcal{R})$, a subspace $\mathcal{E}$ and a space of regular vector fields $\mathcal{A}$. Fix parameters $\varepsilon_0, \mathtt{R}_0, \mathtt{G}_0, \mathfrak{p}_2$ satisfying appropriate constraints. Let N_0 be a diagonal vector field and consider a vector field*

$$F_0 := N_0 + G_0 \in \mathcal{E}$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$.

Fix $\gamma_0 > 0$ and assume that

$$\gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(G_0) \leq \mathtt{G}_0\,, \quad \gamma_0^{-1} C_{\vec{v}_0,\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G_0) \leq \mathtt{R}_0\,, \quad \gamma_0^{-1} |\Pi_{\mathcal{X}} G_0|_{\vec{v}_0,\mathfrak{p}_1} \leq \varepsilon_0\,, \quad \gamma_0^{-1} |\Pi_{\mathcal{X}} G_0|_{\vec{v}_0,\mathfrak{p}_2} \leq \mathtt{R}_0\,, \tag{1.12}$$

here $C_{\vec{v},p}(G)$ is a tameness constant, while $|\cdot|_{\vec{v},p}$ is the norm on regular vector fields.

For all $n \geq 0$ we define recursively changes of variables $\mathcal{L}_n, \Phi_n$ and compact sets $\mathcal{O}_n$ as follows.
Set $\mathcal{H}_{-1} = \mathcal{H}_0 = \Phi_0 = \mathcal{L}_0 = \mathbb{1}$, and for $0 \leq j \leq n-1$ set recursively $\mathcal{H}_j = \Phi_j \circ \mathcal{L}_j \circ \mathcal{H}_{j-1}$ and $F_j := (\mathcal{H}_j)_* F_0 := N_0 + G_j$. Let $\mathcal{L}_n$ be any compatible change of variables *for $(F_{n-1}, K_{n-1}, \vec{v}_{n-1}, \rho_{n-1})$ and $\mathcal{O}_n$ be any compact set*

$$\mathcal{O}_n \subseteq \mathcal{O}_{n-1}\,,$$

which satisfies the homological equation *for $((\mathcal{L}_n)_* F_{n-1}, K_{n-1}, \vec{v}^{o}_{n-1}, \rho_{n-1})$, let g_n be the solution of the homological equation and Φ_n the time-1 flow map generated by g_n.*
Then the sequence $\mathcal{H}_n$ converges for all $\xi \in \mathcal{O}_0$ to some change of variables

$$\mathcal{H}_\infty = \mathcal{H}_\infty(\xi) : D_{a_0,p}(s_0/2, r_0/2) \longrightarrow D_{\frac{a_0}{2},p}(s_0, r_0).$$

such that defining $F_\infty := (\mathcal{H}_\infty)_* F_0$ one has

$$\Pi_{\mathcal{X}} F_\infty = 0 \quad \forall \xi \in \mathcal{O}_\infty := \bigcap_{n\geq 0} \mathcal{O}_n$$

and

$$\gamma_0^{-1} C_{\vec{v}_\infty,\mathfrak{p}_1}(\Pi_{\mathcal{N}} F_\infty - N_0) \leq 2\mathtt{G}_0, \quad \gamma_0^{-1} C_{\vec{v}_\infty,\mathfrak{p}_1}(\Pi_{\mathcal{R}} F_\infty) \leq 2\mathtt{R}_0$$

with $\vec{v}_\infty := (\gamma_0/2, \mathcal{O}_\infty, s_0/2, a_0/2)$.

While the scheme is quite general, as a drawback the set of parameters ξ for which the invariant torus exists is defined in a very complicated way, in terms of the approximate invertibility of the operators $\mathfrak{A}_n$.

In order to get a simpler description of the good parameters we may require that $(\mathcal{N}, \mathcal{X}, \mathcal{R})$ is a "triangular decomposition" i.e. there exists a decomposition $\mathcal{X} = \oplus_{j=1}^{\mathtt{b}} \mathcal{X}_j$ such that for all $N \in \mathcal{N}$ the action of the operator $\mathfrak{N} := \Pi_{\mathcal{X}}[N, \cdot]$ is block diagonal while for all $R \in \mathcal{R}$ the action of $\mathfrak{R} := \Pi_{\mathcal{X}}[R, \cdot]$ is strictly upper triangular, see Definition 3.1. In Proposition 3.5 we show that under such hypotheses the problem of solving the homological equation (1.10) is reduced to proving the approximate invertibility of $\mathfrak{N}$, the so called *Melnikov* *conditions*, which is typically much simpler to analyze in the applications. Indeed solving (1.10) amounts to inverting $\mathfrak{A}$ but since $\mathfrak{R}$ is upper triangular and $\mathfrak{N}$ is diagonal, the Neumann series

$$\mathfrak{A}^{-1} = (\mathfrak{N} + \mathfrak{R})^{-1} = \mathfrak{N}^{-1} \sum_j (-1)^j (\mathfrak{N}^{-1}\mathfrak{R})^j$$

is a finite sum.

Note that in order to produce a triangular decomposition one can associate degrees to w, y (say resp. 1 and $\mathtt{d} > 0$) in order to induce a degree decomposition (see Remark 3.2) which gives to the space of polynomial vector fields a graded Lie algebra structure. By Lemma 3.3 triangularity is achieved when $\mathcal{N}$ contains only terms with degree 0 while $\mathcal{R}$ contains only terms with degree > 0 while $\mathcal{X} = \oplus_{j=1}^{\mathtt{b}} \mathcal{X}_j$ is the decomposition of $\mathcal{X}$ in subspaces of increasing degree. Note that this can be done for any choice for $\mathtt{d}$.

Another natural way to understand the nature of the subspaces $\mathcal{X}, \mathcal{N}, \mathcal{R}$ is through "rescalings". This means introducing a special degree decomposition where the degree of y is the same $\mathtt{s}$ as the one used in the definition of the domains in the phase space, see (1.5). This latter degree decomposition separates terms which behave differently under rescaling of the order on magnitude of the domains $r \to \lambda r$. Note that in this way $\mathcal{X}$ contains terms with negative scaling, $\mathcal{N}$ contains the terms with scaling zero and finally $\mathcal{R}$ contains terms with positive scaling.

Typically the invertibility of $\mathfrak{N}$ relies on non degeneracy conditions on the eigenvalues which provides a lower bounds on the small divisors. The size γ_0 of such denominators is essentially given by the problem. A vector field is considered a perturbation if its size is small with respect to the size of the small divisor. Hence in giving the smallness conditions on G one must find a modulation between the size of the remainder R, which can be made small by a rescaling, and the size of $\mathcal{X}$ which grows under the rescaling. The polynomial vector fields in N are more delicate. Indeed some such terms do not change under rescaling. Of course $\Pi_{\mathcal{N}} G$ should be much smaller with respect to N_0 but in fact one does not need that $\Pi_{\mathcal{N}} G$ is small with respect to γ_0 provided that $\Pi_{\mathcal{X}} G$ is sufficiently small. This further justifies why the smallness conditions on the vector field G are imposed separately on each term. We refer the reader to Paragraph 4.6 for more details.

In Section 4 we discuss various applications and examples and we show how our algorithm allows to interpolate between the Nash-Moser algorithm and the classical KAM one. Example 1. is the classic Nash-Moser approach, where one fixes the subspace $\mathcal{X}$ to be as simple as possible.
In Example 2. we study Hamiltonian vector fields, and exploit the Hamiltonian structure in order to simplify the Melnikov conditions; this is a reinterpretation to our setting of the strategy of [45]. We conclude subsection 4.2 by collecting our results into Theorem 4.11.
In Example 3. we only assume that our vector fiels is reversible and we simplify the Melnikov conditions by making an appropriate choice of the sets $\mathcal{X}, \mathcal{N}, \mathcal{R}$, this is a new result which we believe should enable us to prove existence of quasi-periodic solutions in various settings, see [46] for the case of the fully nonlinear NLS on the circle. We conclude subsection 4.3 by collecting our results into Theorem 4.16.
In Examples 2. and 3. our formulation essentially decouples the dynamics on the approximate invariant torus, which is given by the equation for θ and y, and the dynamics in the normal direction w. More precisely in these cases the invertibility of $\mathfrak{N}_n$ follow from the conditions:

- The frequency vector $\omega_n(\xi) := \langle F_n^{(\theta)}(\theta, 0, 0) \rangle$ needs to be diophantine;
- The operator $\mathfrak{L}_n := \omega_n \cdot \partial_\theta + d_w F^{(w)}(\theta, 0, 0)$ acting on $H^p(\mathbb{T}_s^d, \ell_{a,p})$ must be "approximatively invertible".

Note that $\mathfrak{L}_n$ has the same form of the linearized operator of a forced equation, namely the case where $F^\theta = \omega \cdot \partial_\theta$ and the frequency vector ω plays the rôle of an external parameter. Moreover if F is a vector field coming from a PDE (possibly after one step of Birkhoff Normal Form), then $\mathfrak{L}_n$ differs from the linearization of a composition operator by a finite rank term, this is an essential property in the study of quasi-linear PDEs.
In Example 4 we prove a KAM theorem for a class of Hamiltonian vector fields corresponding to the classical paper [15], but requiring only finite differentiability and imposing milder smallness conditions, comparable to

those of [48]. We conclude subsection 4.4 by stating a result on the existence of reducible invariant tori; this is the finitely differentiable version of [15, 48].

Finally in subsection 4.5 we discuss how to apply Examples 2 and 4 to an NLS with Fourier multipliers, respectively in the case of a Lie group and of $[0, \pi]$ with Dirichlet boundary conditions.

Acknowledgements. We thank L. Biasco, F. Giuliani, J. Massetti and R. Montalto for useful comments on the manuscript. This research was supported by the European Research Council under FP7 “Hamiltonian PDEs and small divisor problems: a dynamical systems approach”, by PRIN2012 “Variational and perturbative aspects of non linear differential problems" and McMaster University.

2 Functional setting and main result

In this paragraph we introduce all the relevant notation and tools we need. In particular we define our phase space, the vector fields we will deal with and the type of change of variables we need in order to perform our KAM algorithm, as explained in the introdution.

2.1 The Phase Space

Our starting point is an infinite dimensional space with a product structure $V_{a,p} := \mathbb{C}^d \times \mathbb{C}^{d_1} \times \ell_{a,p}$. Here $\ell_{a,p}$ is a scale of separable Hilbert spaces endowed with norms $\| \cdot \|_{a,p}$, in particular this means that $\|f\|_{a,p} \le \|f\|_{a',p'}$ if $(a,p) \le (a',p')$ lexicographically.

Hypothesis 2.1. *The space $\ell_{0,0}$ is endowed with a bilinear scalar product*

$$f, g \in \ell_{0,0} \mapsto f \cdot g \in \mathbb{C}.$$

The scalar product identifies the dual $\ell_{a,p}^*$ with $\ell_{-a,-p}$ and is such that

$$\|w\|_{0,0}^2 = w \cdot \overline{w}\,, \quad |g \cdot f| \le \|g\|_{a,p} \|f\|_{-a,-p} \quad |g \cdot f| \le \|g\|_{0,0} \|f\|_{0,0} \le \|g\|_{a,p} \|f\|_{0,0}. \tag{2.1}$$

We denote the set of variables $\mathtt{V} := \{\theta_1, \dots, \theta_d, y_1, \dots, y_{d_1}, w\}$. Moreover we make the following assumption on the scale $\ell_{a,p}$. We assume that there is a non-decreasing family $(\ell_K)_{K \ge 0}$ of closed subspaces of $\ell_{a,p}$ such that $\cup_{K \ge 0} \ell_K$ is dense in $\ell_{a,p}$ for any $p \ge 0$, and that there are projectors

$$\Pi_{\ell_K} : \ell_{0,0} \to \ell_K, \quad \Pi_{\ell_K}^{\perp} := \mathbb{1} - \Pi_{\ell_K}, \tag{2.2}$$

such for all $p, \alpha, \beta \ge 0$ there exists a constant $\mathtt{C} = \mathtt{C}(a, p, \alpha, \beta)$ such that one has

$$\|\Pi_{\ell_K} w\|_{a+\alpha, p+\beta} \le \mathtt{C} e^{\alpha K} K^{\beta} \|w\|_{a,p} \ \ \forall w \in \ell_{a,p}, \tag{2.3a}$$

$$\|\Pi_{\ell_K}^{\perp} w\|_{a,p} \le \mathtt{C} e^{-\alpha K} K^{-\beta} \|w\|_{a+\alpha, p+\beta}, \quad \forall w \in \ell_{a+\alpha, p+\beta}. \tag{2.3b}$$

We shall need two parameters, $\mathfrak{p}_0 < \mathfrak{p}_1$. Precisely $\mathfrak{p}_0 > d/2$ is needed in order to have the Sobolev embedding and thus the algebra properties, while $\mathfrak{p}_1$ will be chosen very large and is needed in order to define the phase space.

Definition 2.1 (Phase space)**.** *Given $\mathfrak{p}_1$ large enough, we consider the toroidal domain*

$$\mathbb{T}_s^d \times D_{a,p}(r) := \mathbb{T}_s^d \times B_{r^2} \times \boldsymbol{B}_{r,a,p,\mathfrak{p}_1}\,, \subset V_{a,p} \tag{2.4}$$

where

$$\mathbb{T}_s^d := \left\{ \theta \in \mathbb{C}^d \,:\, \mathrm{Re}(\theta) \in \mathbb{T}^d, \ \max_{h=1,\dots,d} |\mathrm{Im}\, \theta_h| < s \right\},$$

$$B_{r^{\mathtt{s}}} := \left\{ y \in \mathbb{C}^{d_1} \,:\, |y|_1 < r^{\mathtt{s}} \right\}, \qquad \boldsymbol{B}_{r,a,p,\mathfrak{p}_1} := \left\{ w \in \ell_{a,p} \,:\, \|w\|_{a,\mathfrak{p}_1} < r \right\},$$

and we denote by $\mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d$ the d-dimensional torus.

Remark 2.2. *Note that $\boldsymbol{B}_{r,a,p,\mathfrak{p}_1}$ is the intersection of $\ell_{a,p}$ with the ball of radius r in $\ell_{a,\mathfrak{p}_1}$, thus our phase space clearly depends on the parameter $\mathfrak{p}_1$. We drop it in the notations since it will be fixed once and for all, while a, p, s, r vary throughout the algorithm and we carefully need to keep track of them.*

Fix some numbers $s_0, a_0 \geq 0$ and $r_0 > 0$. Given $s \leq s_0$, $a, a' \leq a_0$, $r \leq r_0$, $p, p' \geq \mathfrak{p}_0$. We endow the space $V_{a,p}$ with the following norm. For $u = (u^{(\theta)}, u^{(y)}, u^{(w)}) \in \mathbb{T}^d_s \times D_{a,p}(r)$

$$\|u\|_{V_{a,p}} = \frac{1}{\max\{1, s_0\}}\|u^{(\theta)}\|_{\mathbb{C}^d} + \frac{1}{r_0^{(\mathtt{s})}}\|u^{(y)}\|_{\mathbb{C}^{d_1}} + \frac{1}{r_0}\|u^{(w)}\|_{\ell_{a,p}}, \tag{2.5}$$

Now consider maps

$$\begin{aligned} f : \mathbb{T}^d_s \times D_{a',p'}(r) &\to V_{a,p} \\ (\theta, y, w) &\to (f^{(\theta)}(\theta, y, w), f^{(y)}(\theta, y, w), f^{(w)}(\theta, y, w)), \end{aligned} \tag{2.6}$$

with

$$f^{(\mathtt{v})}(\theta, y, w) = \sum_{l \in \mathbb{Z}^d} f_l^{(\mathtt{v})}(y, w) e^{\mathrm{i} l \cdot \theta}, \quad \mathtt{v} \in \mathtt{V},$$

where $f^{(\mathtt{v})}(\theta, y, w) \in \mathbb{C}$ for $\mathtt{v} = \theta_i, y_i$ while $f^{(w)}(\theta, y, w) \in \ell_{a,p}$. We shall use also the notation $f^{(\theta)}(\theta, y, w) \in \mathbb{C}^d$, $f^{(y)}(\theta, y, w) \in \mathbb{C}^{d_1}$.

Remark 2.3. *We think of these maps as families of torus embeddings from $\mathbb{T}^d_s$ into $V_{a,p}$ depending parametrically on $y, w \in D_{a',p'}(r)$, and this is the reason behind the choice of the norm; see below.*

We define a norm (pointwise on y, w) by setting

$$\|f\|^2_{s,a,p} := \|f^{(\theta)}\|^2_{s,p} + \|f^{(y)}\|^2_{s,p} + \|f^{(w)}\|^2_{s,a,p} \tag{2.7}$$

where

$$\|f^{(\theta)}\|_{s,p} := \begin{cases} \dfrac{1}{s_0} \sup\limits_{i=1,\dots,d} \|f^{(\theta_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d_s)} & s \leq s_0 \neq 0 \\ \sup\limits_{i=1,\dots,d} \|f^{(\theta_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d_s)}, & s = s_0 = 0 \end{cases} \tag{2.8}$$

$$\|f^{(y)}\|_{s,p} := \frac{1}{r_0^{\mathtt{s}}} \sum_{i=1}^{d_1} \|f^{(y_i)}(\cdot, y, w)\|_{H^p(\mathbb{T}^d_s)} \tag{2.9}$$

$$\|f^{(w)}\|_{s,a,p} := \frac{1}{r_0}\left(\|f^{(w)}\|_{H^p(\mathbb{T}^d_s;\ell_{a,\mathfrak{p}_0})} + \|f^{(w)}\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\ell_{a,p})}\right), \tag{2.10}$$

where $H^p(\mathbb{T}^d_s) = H^p(\mathbb{T}^d_s; \mathbb{C})$ is the standard space of analytic functions in the strip of size s which are Sobolev on the boundary with norm

$$\|u(\cdot)\|^2_{H^p(\mathbb{T}^d_s)} := \sum_{l \in \mathbb{Z}^d} |u_l|^2 e^{2s|l|} \langle l \rangle^{2p}, \qquad \langle l \rangle := \max\{1, |l|\}. \tag{2.11}$$

If $s = 0$ clearly one has that $H^p(\mathbb{T}^d_s) = H^p(\mathbb{T}^d)$ is the standard Sobolev space. More in general given a Banach space E we denote by $H^p(\mathbb{T}^d_s; E)$ the space of analytic functions in the strip of size s which are Sobolev in θ on the boundary with values in E endowed with the natural norm. Note that trivially $\|\partial_\theta^{p'} u\|_{H^p(\mathbb{T}^d_s)} = \|u\|_{H^{p+p'}(\mathbb{T}^d_s)}$.

Remark 2.4. *We can interpret (2.10) as follows. We associate $f^{(w)}$ with a function of θ defined as*

$$\mathtt{f}_p(\theta) := \sum_{l \in \mathbb{Z}^d} \|f_l^{(w)}(y, w)\|_{a,p} e^{\mathrm{i}\theta \cdot l}, \tag{2.12}$$

and then we have

$$\|f^{(w)}\|_{s,a,p} = \|\mathtt{f}_{\mathfrak{p}_0}\|_{H^p(\mathbb{T}^d_s)} + \|\mathtt{f}_p\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s)}.$$

Remark 2.5. *If $\ell_{a,p} = H^p(\mathbb{T}^r_a)$ then fixing $\mathfrak{p}_0 \geq (d+r)/2$ we have that $\|\cdot\|_{s,a,p}$ in* (2.10) *is equivalent to* $\|\cdot\|_{H^p(\mathbb{T}^d_s \times \mathbb{T}^r_a)}$

By recalling the definition in (2.5) we have that

$$\|f\|_{s,a,p} \sim \|f\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d;V_{a,p}) \cap H^p(\mathbb{T}^d;V_{a,\mathfrak{p}_0})} := \|f\|_{H^p(\mathbb{T}^d;V_{a,\mathfrak{p}_0})} + \|f\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d;V_{a,p})} \tag{2.13}$$

Remark 2.6. *Formula* (2.7) *depends on the point (y,w), hence it is* not *a norm for vector fields and this is very natural in the context of* Sobolev regularity. *Indeed in the scale of domains $\mathbb{T}^d_s \times D_{a,p}(r)$ one controls only the $\mathfrak{p}_1$ norm of w (see Definition 2.1), and hence there is no reason for which one may have*

$$\sup_{(y,w)\in D_{a,p}(r)} \|f\|_{s,a,p} < \infty\,.$$

Naturally if one fixes $p = \mathfrak{p}_1$ one may define as norm of F the quantity $\sup_{(y,w)\in D_{a,\mathfrak{p}_1}(r)} \|F\|_{s,a,\mathfrak{p}_1}$.

The motivation for choosing the norm (2.7) *is the following. Along the algorithm we need to control commutators of vector fields. In the analytic case, i.e. if $s_0 \neq 0$, one may keep p fixed and control the derivatives via* Cauchy estimates *by reducing the analyticity, so the phase space can be defined in terms of the fixed p. However, since we do not want to add the hypothesis $s_0 \neq 0$, we have to leave p as a parameter and use* tameness *properties of the vector field (see Definition 2.13) as in the Sobolev Nash-Moser schemes.*

It is clear that any f as in (2.6) can be identified with "unbounded" vector fields by writing

$$f \leftrightarrow \sum_{\mathtt{v}\in\mathtt{V}} f^{(\mathtt{v})}(\theta,y,w)\partial_{\mathtt{v}}, \tag{2.14}$$

where the symbol $f^{(\mathtt{v})}(\theta,y,w)\partial_{\mathtt{v}}$ has the obvious meaning for $\mathtt{v} = \theta_i, y_i$ while for $\mathtt{v} = w$ is defined through its action on differentiable functions $G : \ell_{a,p} \to \mathbb{C}$ as

$$f^{(w)}(\theta,y,w)\partial_w G := d_w G[f^{(w)}(\theta,y,w)].$$

Similarly, provided that $|f^{(\theta)}(\theta,y,w)|$ is small for all $(\theta,y,w) \in \mathbb{T}^d_s \times D_{a,p}(r)$ we may lift f to a map

$$\Phi := (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)}) : \mathbb{T}^d_s \times D_{a',p'} \to \mathbb{T}^d_{s_1} \times \mathbb{C}^{d_1} \times \ell_{a,p}\,, \quad \text{for some } s_1 \geq s\,, \tag{2.15}$$

and if we set $\|\theta\|_{s,a,p} := 1$ we can write

$$\|\Phi^{(\mathtt{v})}\|_{s,a,p} = \|\mathtt{v}\|_{s,a,p} + \|f^{(\mathtt{v})}\|_{s,a,p}\,,\ \mathtt{v} = \theta, y, w\,.$$

Note that

$$\|y\|_{s,a,p} = r_0^{-\mathtt{s}}|y|_1\,, \quad \|w\|_{s,a,p} = r_0^{-1}\|w\|_{a,p}.$$

Remark 2.7. *There exists $\mathtt{c} = \mathtt{c}(d)$ such that if $\|f\|_{s,a,\mathfrak{p}_1} \leq \mathtt{c}\rho$ one has*

$$\Phi : \mathbb{T}^d_s \times D_{a+\rho a_0,p}(r) \to \mathbb{T}^d_{s+\rho s_0} \times D_{a,p}(r+\rho r_0).$$

We are interested in vector fields defined on a scale of Hilbert spaces; precisely we shall fix $\rho,\nu,q \geq 0$ and consider vector fields

$$F : \mathbb{T}^d_s \times D_{a+\rho a_0,p+\nu}(r) \times \mathcal{O} \to V_{a,p}\,, \tag{2.16}$$

for some $s < s_0$, $a + \rho a_0 \leq a_0$, $r \leq r_0$ and all $p + \nu \leq q$. Moreover we require that $\mathfrak{p}_1$ in Definition 2.1 satisfies $\mathfrak{p}_1 \geq \mathfrak{p}_0 + \nu + 1$.

Definition 2.8. *Fix $0 \le \rho, \varrho \le 1/2$, and consider two differentiable maps $\Phi = \mathbb{1} + f$, $\Psi = \mathbb{1} + g$ as in* (2.15) *such that for all $p \ge \mathfrak{p}_0$, $2\rho s_0 \le s \le s_0$, $2\rho r_0 \le r \le r_0$ and $0 \le a \le a_0(1-2\varrho)$ one has*

$$\Phi, \Psi : \mathbb{T}^d_{s-\rho s_0} \times D_{a+\varrho a_0,p}(r - \rho r_0) \to \mathbb{T}^d_s \times D_{a,p}(r). \tag{2.17}$$

If

$$\begin{aligned} \mathbb{1} = \Psi \circ \Phi : \mathbb{T}^d_{s-2\rho s_0} \times D_{a+2\varrho a_0,p}(r - 2\rho r_0) \quad &\longrightarrow \mathbb{T}^d_s \times D_{a,p}(r) \\ (\theta, y, w) \quad &\longmapsto (\theta, y, w) \end{aligned} \tag{2.18}$$

we say that Ψ is a left inverse of Φ and write $\Phi^{-1} := \Psi$.

Moreover fix $\nu \ge 0$, $0 \le \varrho' \le 1/2$. Then for any $F : \mathbb{T}^d_s \times D_{a+\varrho' a_0, p+\nu}(r) \to V_{a,p}$, with $0 \le a \le a_0(1-2\varrho-\varrho')$, we define the "pushforward" of F as

$$\Phi_* F := d\Phi(\Phi^{-1})[F(\Phi^{-1})] : \mathbb{T}^d_{s-2\rho s_0} \times D_{a+(2\varrho+\varrho')a_0, p+\nu}(r - 2\rho r_0) \to V_{a,p}\,. \tag{2.19}$$

We need to introduce parameters $\xi \in \mathcal{O}_0$ a compact set in $\mathbb{R}^{\mathfrak{d}}$. Given any compact $\mathcal{O} \subseteq \mathcal{O}_0$ we consider Lipschitz families of vector fields

$$F : \mathbb{T}^d_s \times D_{a',p'}(r) \times \mathcal{O} \to V_{a,p}\,, \tag{2.20}$$

and say that they are *bounded* vector fields when $p = p'$ and $a = a'$. Given a positive number γ we introduce the weighted Lipschitz norm

$$\|F\|_{\vec{v},p} = \|F\|_{\gamma,\mathcal{O},s,a,p} := \sup_{\xi \in \mathcal{O}} \|F(\xi)\|_{s,a,p} + \gamma \sup_{\xi \ne \eta \in \mathcal{O}} \frac{\|F(\xi) - F(\eta)\|_{s,a,p-1}}{|\xi - \eta|}\,. \tag{2.21}$$

and we shall drop the labels $\vec{v} = (\gamma, \mathcal{O}, s, a)$ when this does not cause confusion. More in general given E a Banach space we define the Lipschitz norm as

$$\|F\|_{\gamma,\mathcal{O},E} := \sup_{\xi \in \mathcal{O}} \|F(\xi)\|_E + \gamma \sup_{\xi \ne \eta \in \mathcal{O}} \frac{\|F(\xi) - F(\eta)\|_E}{|\xi - \eta|}\,. \tag{2.22}$$

Remark 2.9. *Note that in some applications one might need to assume a higher regularity in ξ. In this case it is convenient to define the weighted q_1-norm*

$$\|F\|_{\vec{v},p} = \|F\|_{\gamma,\mathcal{O},s,a,p} := \sum_{\substack{h \in \mathbb{N}^{\mathfrak{d}} \\ |h| \le q_1}} \gamma^{|h|} \sup_{\xi \in \mathcal{O}} \|\partial^h_\xi F(\xi)\|_{s,a,p-|h|}.$$

Where the derivatives are in the sense of Whitney.

Throughout the paper we shall always use the Lipschitz norm (2.21)*, although all the properties hold verbatim also for the q_1-norm.*

Definition 2.10. *We shall denote by $\mathcal{V}_{\vec{v},p}$ with $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ the space of vector fields as in* (2.16) *with $\varrho = 0$. By slight abuse of notation we denote the norm $\|\cdot\|_{\gamma,\mathcal{O},s,a,p} = \|\cdot\|_{\vec{v},p}$.*

2.2 Polynomial decomposition

In $\mathcal{V}_{\vec{v},p}$ we identify the closed *monomial* subspaces

$$\begin{aligned} &\mathcal{V}^{(\mathtt{v},0)} := \{f \in \mathcal{V}_{\vec{v},p} \,:\, f = f^{(\mathtt{v},0)}(\theta)\partial_{\mathtt{v}}\}\,, \quad \mathtt{v} \in \mathtt{V} \\ &\mathcal{V}^{(\mathtt{v},\mathtt{v}')} := \{f \in \mathcal{V}_{\vec{v},p} \,:\, f = f^{(\mathtt{v},\mathtt{v}')}(\theta)[\mathtt{v}']\partial_{\mathtt{v}}\}\,, \quad \mathtt{v} \in \mathtt{V}, \quad \mathtt{v}' = y_1, \dots, y_{d_1}, w \\ &\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)} := \{f \in \mathcal{V}_{\vec{v},p} \,:\, f = f^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)}(\theta)[\mathtt{v}_1, \dots, \mathtt{v}_k]\partial_{\mathtt{v}}\}\,, \quad \mathtt{v} \in \mathtt{V}, \quad \mathtt{v}_i = y_1, \dots, y_{d_1}, w\,, \end{aligned} \tag{2.23}$$

where $\mathtt{v}_1,\ldots,\mathtt{v}_k \in \mathtt{U} := \{y_1,\ldots,y_{d_1},w\}$ are ordered and $f^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_k)}$ is a multilinear form symmetric w.r.t. repeated variables.

As said after (2.6) it will be convenient to use also vector notation so that, for instance

$$f^{(y,y)}(\theta)y\cdot\partial_y \in \mathcal{V}^{(y,y)} = \bigoplus_{i\le j=1,\ldots,d_1} \mathcal{V}^{(y_i,y_j)}$$

with $f^{(y,y)}(\theta)$ a $d_1\times d_1$ matrix.

Note that the polynomial vector fields of (maximal) degree k are

$$\mathcal{P}_k := \bigoplus_{\mathtt{v}\in\mathtt{V}}\bigoplus_{j=0}^{k}\bigoplus_{(\mathtt{v}_1,\ldots,\mathtt{v}_j)\in\mathtt{U}} \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_j)}\,,$$

so that, given a polynomial $F\in\mathcal{P}_k$ we may define its "projection" onto a monomial subspace $\Pi_{\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_j)}}$ in the natural way.

Since we are not working on spaces of polynomials, but on vector fields with finite regularity, we need some more notations. Given a C^{k+1} vector field $F\in\mathcal{V}_{\vec v,p}$, we introduce the notation

$$F^{(\mathtt{v},0)}(\theta) := F^{(\mathtt{v})}(\theta,0,0),\quad F^{(\mathtt{v},\mathtt{v}')}(\theta)[\cdot] := d_{\mathtt{v}'}F^{(\mathtt{v})}(\theta,0,0)[\cdot],\quad \mathtt{v}\in\mathtt{V}\,,\quad \mathtt{v}'=y_1,\ldots,y_{d_1},w \tag{2.24}$$

$$F^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_k)}(\theta)[\cdot,\ldots,\cdot] = \frac{1}{\alpha(\mathtt{v}_1,\ldots,\mathtt{v}_k)!}\Big(\prod_{i=1}^{k} d_{\mathtt{v}_i}\Big)F^{(\mathtt{v})}(\theta,0,0)[\cdot,\ldots,\cdot]\,,\quad \mathtt{v}\in\mathtt{V}\,,\quad \mathtt{v}_i=y_1,\ldots,y_{d_1},w\,,$$

where we assume that $\mathtt{v}_1,\ldots\mathtt{v}_k$ are ordered and the $(d+1)$-dimensional vector $\alpha(\mathtt{v}_1,\ldots,\mathtt{v}_k)$ denotes the multiplicity of each component.

By Taylor approximation formula any vector field in $\mathcal{V}_{\vec v,p}$ which is C^{k+1} in y,w may be written in a unique way as sum of its Taylor polynomial in $\mathcal{P}_k$ plus a C^{k+1} (in y,w) vector field with a zero of order at least $k+1$ at $y=0$, $w=0$. We think of this as a direct sum of vector spaces and introduce the notation

$$\Pi_{\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_k)}}F := F^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_k)}(\theta)[\mathtt{v}_1,\ldots,\mathtt{v}_k]\,, \tag{2.25}$$

we refer to such operators as *projections.*

Definition 2.11. *We identify the vector fields in $\mathcal{V}_{\vec v,p}$ which are C^{k+1} in y,w, with the direct sum*

$$\mathcal{W}^{(k)}_{\vec v,p} = \mathcal{P}_k\oplus\mathcal{R}_k\,,$$

where $\mathcal{R}_k$ is the space of C^{k+1} (in y,w) vector fields with a zero of order at least $k+1$ at $y=0$, $w=0$. On $\mathcal{W}^{(k)}_{\vec v,p}$ we induce the natural norm for direct sums, namely for

$$f=\sum_{\mathtt{v}\in\mathtt{V}}\sum_{j=0}^{k}\sum_{(\mathtt{v}_1,\ldots,\mathtt{v}_j)\in\mathtt{U}} f^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_j)}(\theta)[\mathtt{v}_1,\ldots,\mathtt{v}_j]\partial_{\mathtt{v}} + f_{\mathcal{R}_k}\qquad f_{\mathcal{R}_k}\in\mathcal{R}_k\,,$$

we set

$$\|f\|^{(k)}_{\vec v,p} := \sum_{\mathtt{v}\in\mathtt{V}}\sum_{j=0}^{k}\sum_{(\mathtt{v}_1,\ldots,\mathtt{v}_j)\in\mathtt{U}} \|f^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_j)}(\cdot)[\mathtt{v}_1,\ldots,\mathtt{v}_j]\|_{\vec v,p} + \|f_{\mathcal{R}_k}\|_{\vec v,p}\,. \tag{2.26}$$

Remark 2.12. *Note that with this definition if $k=\infty$ we are considering analytic maps with the norm*

$$\sum_{\mathtt{v}\in\mathtt{V}}\sum_{j=0}^{\infty}\sum_{(\mathtt{v}_1,\ldots,\mathtt{v}_j)\in\mathtt{U}} \|f^{(\mathtt{v},\mathtt{v}_1,\ldots,\mathtt{v}_j)}(\cdot)[\mathtt{v}_1,\ldots,\mathtt{v}_j]\|_{\vec v,p}.$$

We can and shall introduce in the natural way the polynomial subspaces and the norm (2.26) also for maps $\Phi = (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)})$ with

$$\Phi : \mathbb{T}^d_s \times D_{a',p'}(r) \times \mathcal{O} \to \mathbb{T}^d_{s_1} \times D_{a,p}(r_1)\,,$$

since the Taylor formula holds also for functions of this kind.

We also denote

$$\begin{aligned} \langle \mathcal{V}(\mathtt{v}, \mathtt{v}_1, \cdots, \mathtt{v}_k)\rangle &:= \{f \in \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)} : f = \langle f^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)}\rangle \cdot \partial_{\mathtt{v}}\}, \\ \mathcal{V}_0^{(\mathtt{v},\mathtt{v}_1,\cdots,\mathtt{v}_k)} &:= \{f \in \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\cdots,\mathtt{v}_k)} : f = (f^{(\mathtt{v},\mathtt{v}_1,\cdots,\mathtt{v}_k)} - \langle f^{(\mathtt{v},\mathtt{v}_1,\cdots,\mathtt{v}_k)}\rangle) \cdot \partial_{\mathtt{v}}\}\,, \end{aligned} \tag{2.27}$$

where $\langle f\rangle := \fint_{\mathbb{T}^d} f(\theta)d\theta$.

Tame vector fields. We now define vector fields behaving "tamely" when composed with maps Φ. Let us fix a degree $\mathtt{n} \in \mathbb{N}$ and thus a norm

$$\|f\|_{\vec{v},p} = \|f\|^{(\mathtt{n})}_{\vec{v},p}\,. \tag{2.28}$$

Definition 2.13. *Fix a large $q \geq \mathfrak{p}_1$, $k \geq 0$ and a set $\mathcal{O}$. Consider a $C^{k+\mathtt{n}+1}$ vector field*

$$F \in \mathcal{W}^{(\mathtt{n})}_{\vec{v},p}, \qquad \vec{v} = (\gamma, \mathcal{O}, s, a, r)\,.$$

We say that F is C^k-tame *(up to order q) if there exists a scale of constants $C_{\vec{v},p}(F)$, with $C_{\vec{v},p}(F) \leq C_{\vec{v},p_1}(F)$ for $p \leq p_1$, such that the following holds.*

For all $\mathfrak{p}_0 \leq p \leq p_1 \leq q$ consider any $C^{\mathtt{n}+1}$ map $\Phi = (\theta + f^{(\theta)}, y + f^{(y)}, w + f^{(w)})$ with $\|f\|_{\vec{v}',\mathfrak{p}_1} < 1/2$ and

$$\Phi : \mathbb{T}^d_{s'} \times D_{a_1,p_1}(r') \times \mathcal{O} \to \mathbb{T}^d_s \times D_{a,p+\nu}(r)\,, \quad \text{for some} \quad r' \leq r\,, s' \leq s;$$

and denote $\vec{v}' = (\gamma, \mathcal{O}, s', a, r')$. Then for any $m = 0, \dots, k$ and any m vector fields

$$h_1, \dots, h_m : \ \mathbb{T}^d_{s'} \times D_{a_1,p_1}(r') \times \mathcal{O} \to V_{a,p+\nu},$$

one has

$$(T_m) \quad \|d^m_{\mathtt{U}} F(\Phi)[h_1, \dots, h_m]\|_{\vec{v}',p} \quad \leq \quad \big(C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)\|\Phi\|_{\vec{v}',p+\nu}\big) \prod_{j=1}^m \|h_j\|_{\vec{v}',\mathfrak{p}_0+\nu} + C_{\vec{v},\mathfrak{p}_0}(F) \sum_{j=1}^m \|h_j\|_{\vec{v}',p+\nu} \prod_{i\neq j} \|h_i\|_{\vec{v}',\mathfrak{p}_0+\nu}$$

for all $(y, w) \in D_{a_1,p_1}(r')$ and $p \leq q$. Here $d_{\mathtt{U}}F$ is the differential of F w.r.t. the variables $\mathtt{U} := \{y_1, \dots, y_{d_1}, w\}$ and the norm is the one defined in (2.28).

We say that a bounded *vector field F is tame if the conditions (T_m) above hold with $\nu = 0$. We call $C_{\vec{v},p}(F)$ the p*-tameness constants *of F.*

Remark 2.14. *Note that in the definition above appear two regularity indices: k being the maximum regularity in y, w and q the one in θ.*

Remark 2.15. *Definition 2.13 is quite natural if one has to deal with functions and vector fields which are merely differentiable. In order to clarify what we have in mind we consider the following example. Let L be a linear operator*

$$L : H^p(\mathbb{T}^d) \to H^p(\mathbb{T}^d)\,.$$

In principle there is no reason for L to satisfy a bound like

$$\|Lu\|_p \lessdot \|L\|_{\mathcal{L},p}\|u\|_{\mathfrak{p}_0} + \|L\|_{\mathcal{L},\mathfrak{p}_0}\|u\|_p \tag{2.29}$$

where $\|\cdot\|_{\mathcal{L},p}$ is the H^p-operator norm. However if $L = M_a$ is a multiplication operator, i.e. $M_a u = au$ for some $a \in H^p(\mathbb{T}^d)$ then it is well known that

$$\|M_a u\|_p \leq \kappa_p(\|a\|_p\|u\|_{\mathfrak{p}_0} + \|a\|_{\mathfrak{p}_0}\|u\|_p)$$

which is (2.29) since $\|a\|_p = \|M_a\|_{\mathcal{L},p}$. In this case we may set for all $p \le q$ $C_p(M_a) = \kappa_q \|a\|_p$, where q is the highest possible regularity. This is of course a trivial (though very common in the applications) example in which the tameness constants and the operator norm coincide; we preferred to introduce Definition 2.13 since it is the most general class in which we are able to prove our result.

Remark 2.16. *It is trivial to note that, given a sequence $C_{\vec{v},p}(F)$ of tameness constants for the field F, then any increasing sequence $\tilde{C}_{\vec{v},p}(F)$ such that $C_{\vec{v},p}(F) \le \tilde{C}_{\vec{v},p}(F)$ for any p is a possible choice of tameness constants for F. This leads to the natural question of finding a sharp sequence which then could be used as norm. Throughout the paper we shall write $C_{\vec{v},p}(F) \le C$ to mean that the tameness constants of F can (and shall) be chosen in order to satisfy the bound.*

2.3 Normal form decomposition

Definition 2.17 (**$(\mathcal{N}, \mathcal{X}, \mathcal{R})$-decomposition).** *Let $\mathcal{N}, \mathcal{X} \subseteq \mathcal{P}^{(\mathtt{n})}$ have the following properties:*

(i) if $\mathcal{N} \cap \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_j)} \neq \emptyset$ then either $\mathcal{N} \cap \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_j)} = \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_j)}$ or $\mathcal{N} \cap \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_j)} = \langle \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_j)} \rangle$ for all $j \le \mathtt{n}$;

(ii) one has $\mathcal{V}^{(\mathtt{v},0)} \subset \mathcal{X}$ for $\mathtt{v} = y, w$.

We then decompose

$$\mathcal{W}_{\vec{v},p} = C^{2\mathtt{n}+3} \cap \mathcal{W}^{(\mathtt{n})}_{\vec{v},p} := \mathcal{N} \oplus \mathcal{X} \oplus \mathcal{R}$$

where $C^{2\mathtt{n}+3}$ is the set of vector fields with $(2\mathtt{n}+3)$-regularity in y, w, $\mathcal{R}$ contains all of $\mathcal{R}_{\mathtt{n}}$ and all the polynomials generated by monomials not in $\mathcal{N} \oplus \mathcal{X}$. We shall denote $\Pi_{\mathcal{R}} := \mathbb{1} - \Pi_{\mathcal{N}} - \Pi_{\mathcal{X}}$ and more generally for $\mathcal{S} = \mathcal{N}, \mathcal{X}, \mathcal{R}$ we shall denote $\Pi^{\perp}_{\mathcal{S}} := \mathbb{1} - \Pi_{\mathcal{S}}$.

The following Definition is rather involved since we are trying to make our result as general as possible. However in the applications we have in mind, it turns out that one can choose $\mathcal{A}_{s,a,p}$ satisfying the properties of the Definition below in an explicit and natural way; see Section 4.

Definition 2.18 (**Regular vector fields).** *Given a subset $\mathcal{A}_{s,a,p} \subset \mathcal{P}^{(\mathtt{n})}$ of polynomial vector fields $f : \mathbb{T}^d_s \times D_{a,p+\nu}(r) \to V_{a,p}$ we say that $\mathcal{A}$ is a space of regular vector fields if the following holds.*

Given a compact set $\mathcal{O} \in \mathcal{O}_0$ we denote by $\mathcal{A}_{\vec{v},p}$ with $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ the set of Lipschitz families $\mathcal{O} \to \mathcal{A}_{s,a,p}$. We require that $\mathcal{A}_{s,a,p}$ is a scale of Hilbert spaces w.r.t a norm $|\cdot|_{s,a,p} = |\cdot|_{s,a,p,\nu}$ and we denote by $|\cdot|_{\vec{v},p}$ the corresponding γ-weighted Lipschitz norm.

1. *$\mathcal{V}^{(\mathtt{v},0)} \in \mathcal{A}_{\vec{v},p}$ for $\mathtt{v} = y, w$ while either $\mathcal{A}^{(\theta)}_{\vec{v},p} = \emptyset$ or $\mathcal{A}^{(\theta)}_{\vec{v},p} = \mathcal{V}^{(\theta,0)}_0$.*

2. *All $f \in \mathcal{A}_{\vec{v},q}$ are C^k tame up to order q for all k, with tameness constants*

$$C_{\vec{v},p}(f) = \mathtt{C}\, |f|_{\vec{v},p}. \tag{2.30}$$

for some C depending on $\mathfrak{p}_0$ on the dimensions d, d_1 and on the maximal regularity q. Moreover $|\cdot|_{\mathfrak{p}_1}$ is a sharp *tameness constant, namely there exists a $\mathtt{c}$ depending on $\mathfrak{p}_0, \mathfrak{p}_1, d, d_1$ such that*

$$|f|_{\vec{v},\mathfrak{p}_1} \le \mathtt{c}\, C_{\vec{v},\mathfrak{p}_1}(f) \tag{2.31}$$

for any f and any tameness constant $C_{\vec{v},\mathfrak{p}_1}(f)$.

3. *For $K > 1$ there exists smoothing projection operators $\Pi_K : \mathcal{A}_{\vec{v},p} \to \mathcal{A}_{\vec{v},p}$ such that $\Pi_K^2 = \Pi_K$, for $p_1 \ge 0$, one has*

$$|\Pi_K F|_{\vec{v},p+p_1} \le \mathtt{C} K^{p_1} |F|_{\vec{v},p} \tag{2.32}$$

$$|F - \Pi_K F|_{\vec{v},p} \le \mathtt{C} K^{-p_1} |F|_{\vec{v},p+p_1} \tag{2.33}$$

finally if $C_{\vec{v},p}(F)$ is any tameness constant for F then we may choose a tameness constant such that

$$C_{\vec{v},p+p_1}(\Pi_K F) \le \mathtt{C} K^{p_1} C_{\vec{v},p}(F) \tag{2.34}$$

We denote by $E^{(K)}$ the subspace where $\Pi_K E^{(K)} = E^{(K)}$.

4. *Let $\mathcal{B}$ be the set of bounded vector fields in $\mathcal{A}_{s,a,p} \ni f : \mathbb{T}^d_s \times D_{a,p}(r) \to V_{a,p}$ with the corresponding norm $|\cdot|_{s,a,p,0}$. For all $f \in \mathcal{B}$ such that*

$$|f|_{\vec{v},\mathtt{p}_1} \le \mathtt{c}\rho, \tag{2.35}$$

with $\rho > 0$ small enough, the following holds:

(i) The map $\Phi := \mathbb{1} + f$ is such that

$$\Phi : \mathbb{T}^d_s \times D_{a,p}(r) \times \mathcal{O} \longrightarrow \mathbb{T}^d_{s+\rho s_0} \times D_{a,p}(r + \rho r_0). \tag{2.36}$$

(ii) There exists a vector field $h \in \mathcal{B}$ such that

- *$|h|_{\vec{v}_1,p} \le 2|f|_{\vec{v},p}$, the map $\Psi := \mathbb{1} + h$ is such that*

$$\Psi : \mathbb{T}^d_{s-\rho s_0} \times D_{a,p}(r - \rho r_0) \times \mathcal{O} \to \mathbb{T}^d_s \times D_{a,p}(r). \tag{2.37}$$

- *for all $(\theta, y, w) \in \mathbb{T}^d_{s-2\rho s_0} \times D_{a,\mathtt{p}_1}(r - 2\rho r_0)$ one has*

$$\Psi \circ \Phi(\theta, y, w) = (\theta, y, w). \tag{2.38}$$

5. *Given any regular bounded vector field $g \in \mathcal{B}$, $p \ge \mathtt{p}_1$ with $|g|_{\vec{v},\mathtt{p}_1} \le \mathtt{c}\rho$ then for $0 \le t \le 1$ there exists $f_t \in \mathcal{B}$ such that the time$-t$ map of the flow of g is of the form $\mathbb{1} + f_t$ moreover we have $|f_t|_{\vec{v},p} \le 2|g|_{\vec{v}_1,p}$ where $\vec{v}_1 = (\lambda, \mathcal{O}, s - \rho s_0, a, r)$.*

Definition 2.19. *Consider $\mathcal{E}$ a subspace[5] of $\mathcal{V}_{\vec{v},p}$. We say that $\mathcal{E}$ is compatible with the $(\mathcal{N}, \mathcal{X}, \mathcal{R})$-decomposition if*

(i) any $F \in \mathcal{E} \cap \mathcal{X}$ is a regular vector field;

(ii) for any $F \in \mathcal{E} \cap \mathcal{P}_{\mathtt{n}}$ one has $\Pi_{\mathcal{U}} F \in \mathcal{E}$ for $\mathcal{U} = \mathcal{N}, \mathcal{X}, E^{(K)}$;

(iii) denoting

$$\mathcal{B}_{\mathcal{E}} := \Big\{ f \in \mathcal{X} \cap \mathcal{B} \,:\, \Phi^t_f \text{ is } \mathcal{E} \text{ preserving for all } t \in [0,1] \Big\} \subset \mathcal{X} \cap \mathcal{B}\,, \tag{2.39}$$

one has

$$\forall g \in \mathcal{B}_{\mathcal{E}}, F \in \mathcal{E} : \quad [g, F] \in \mathcal{E}\,, \quad \forall g, h \in \mathcal{B}_{\mathcal{E}} : \quad \Pi_{\mathcal{X}}[g, h] \in \mathcal{B}_{\mathcal{E}}\,. \tag{2.40}$$

Definition 2.20 (Normal form). *We say that $N_0 \in \mathcal{N} \cap \mathcal{E}$ is a* diagonal *vector field if for all $K > 1$*

$$\mathrm{ad}(N_0)\Pi_{E^{(K)}}\Pi_{\mathcal{X}} = \Pi_{E^{(K)}}\Pi_{\mathcal{X}}\mathrm{ad}(N_0)\,, \quad \text{on } \mathcal{B}_{\mathcal{E}}. \tag{2.41}$$

[5]For instance $\mathcal{E}$ may be the subspace of Hamiltonian vector fields.

2.4 Main result

Let us fix once and for all a $(\mathcal{N},\mathcal{X},\mathcal{R})$-decomposition. Before stating the result we need to introduce parameters (which shall depend on the application) fulfilling the following constraints.

Constraint 2.21 (The exponents). *We fix parameters $\varepsilon_0, \mathtt{R}_0, \mathtt{G}_0, \mu, \nu, \eta, \chi, \alpha, \kappa_1, \kappa_2, \kappa_3, \mathfrak{p}_2$ such that the following holds.*

- $0 < \varepsilon_0 \le \mathtt{R}_0 \le \mathtt{G}_0$ *with* $\varepsilon_0 \mathtt{G}_0^3, \varepsilon_0 \mathtt{G}_0^2 \mathtt{R}_0^{-1} < 1$.
- $\mu, \nu, \kappa_3 \ge 0$, $\mathfrak{p}_2 > \mathfrak{p}_1$, $0 \le \alpha < 1$, $1 < \chi < 2$ *such that* $\alpha\chi < 1$.
- *Setting* $\kappa_0 := \mu + \nu + 4$ *and* $\Delta\mathfrak{p} := \mathfrak{p}_2 - \mathfrak{p}_1$ *one has*

$$\kappa_1 > \max(\frac{\kappa_0+\kappa_3}{\chi}, \frac{\kappa_0}{\chi-1}), \tag{2.42a}$$

$$\kappa_2 > \max\big(\frac{2\kappa_0}{2-\chi}, \frac{1}{1-\alpha\chi}((1+\alpha)\kappa_0 + 2\max(\kappa_1,\kappa_3) - \chi\kappa_1)\big), \tag{2.42b}$$

$$\eta > \mu + (\chi-1)\kappa_2 + 1, \tag{2.42c}$$

$$\Delta\mathfrak{p} > \max\big(\kappa_0 + \chi\kappa_2 + \max(\kappa_1,\kappa_3), \frac{1}{1-\alpha}(\kappa_0 + (\chi-1)\kappa_2 + \max(\kappa_1,\kappa_3))\big), \tag{2.42d}$$

$$\alpha\Delta\mathfrak{p} \le \kappa_2 + \chi\kappa_1 - \kappa_0 - \max(\kappa_1,\kappa_3). \tag{2.42e}$$

- *there exists* $K_0 > 1$ *such that*

$$\log K_0 \ge \frac{1}{\log\chi} C, \tag{2.43}$$

with C a given function of $\mu, \nu, \eta, \alpha, \kappa_1, \kappa_2, \kappa_3, \mathfrak{p}_2$ and moreover

$$\mathtt{G}_0^2 \mathtt{R}_0^{-1} \varepsilon_0 K_0^{\kappa_0} \max(1, \mathtt{R}_0 \mathtt{G}_0 K_0^{\kappa_0 + (\chi-1)\kappa_2}) < 1, \tag{2.44a}$$

$$\max(K_0^{\kappa_1}, \varepsilon_0 K_0^{\kappa_3}) K_0^{\kappa_0 - \Delta\mathfrak{p} + (\chi-1)\kappa_2} \mathtt{G}_0 \varepsilon_0^{-1} \max\Big(1, \mathtt{R}_0, \varepsilon_0 \mathtt{G}_0 K_0^{\alpha\Delta\mathfrak{p}}\Big) \le 1, \tag{2.44b}$$

$$\max(K_0^{\kappa_1}, \varepsilon_0 K_0^{\kappa_3}) K_0^{\kappa_0 - \chi\kappa_1} \mathtt{G}_0 \mathtt{R}_0^{-1} \max\Big(\mathtt{R}_0, \varepsilon_0 \mathtt{G}_0 K_0^{\alpha\Delta\mathfrak{p}}\Big) \le 1. \tag{2.44c}$$

Remark 2.22. *In the applications the constants $\mathtt{G}_0, \mathtt{R}_0, \varepsilon_0$ in Constraint 2.21 are given by the problem under study and typically they depend parametrically on* $\mathrm{diam}(\mathcal{O}_0) \sim \gamma$*; then one wishes to show that for γ small enough it is possible to choose all other parameters in order to fulfill Constraint 2.21. Often this implies requiring that $K_0 \to \infty$ as $\gamma \to 0$. In order to highlight this dependence one often uses ε_0 as parameter and introduces $\mathtt{g}, \mathtt{r}$ such that*

$$\mathtt{G}_0 \sim \varepsilon_0^{\mathtt{g}}, \quad \mathtt{R}_0 \sim \varepsilon_0^{\mathtt{r}}, \quad \text{with} \quad \mathtt{g} \le \mathtt{r} \le 1, \quad \min\{1+3\mathtt{g}, 1+2\mathtt{g}-\mathtt{r}\} > 0. \tag{2.45}$$

Then given κ_0, κ_3 one looks for $\alpha, \chi, \kappa_1, \kappa_2, \mathfrak{p}_2$ satisfying (2.42) *and, setting $K_0 = \varepsilon_0^{-\mathtt{a}}$, the constraints* (2.44) *become constraints on $\mathtt{a}$. Another typical procedure is to write $\mathtt{G}_0, \mathtt{R}_0, \varepsilon_0$ as powers of K_0, see paragraph 4.6. Note that in* (2.45) *we need $1+3\mathtt{g} > 0$ but in principle we allow $\mathtt{g} < 0$; same for $\mathtt{r}$. This means in particular that $\mathtt{G}_0$ and $\mathtt{R}_0$ might be very large.*

Definition 2.23 (Homological equation). *Let $\gamma > 0$, $K \ge K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider a vector field $F \in \mathcal{W}_{\vec{v}^o, p}$ i.e.*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p},$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We say $\mathcal{O}$ satisfies the homological equation, for $(F, K, \vec{v}^o, \rho)$ if the following holds.

1. For all $\xi \in \mathcal{O}$ one has $F(\xi) \in \mathcal{E}$ and $|\Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_2-1} \le \mathtt{C}\, C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G)$.
2. there exist a bounded regular vector field $g \in \mathcal{W}_{\vec{v}^o,p} \cap E^{(K)}$ such that

   (a) $g \in \mathcal{B}_{\mathcal{E}}$ for all $\xi \in \mathcal{O}$,

   (b) one has $|g|_{\vec{v}^o,\mathfrak{p}_1} \le \mathtt{C}|g|_{\vec{v},\mathfrak{p}_1} \le \mathtt{c}\rho$ and for $\mathfrak{p}_1 \le p \le \mathfrak{p}_2$

$$|g|_{\vec{v},p} \le \gamma^{-1} K^{\mu} (|\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v},p} + K^{\alpha(p-\mathfrak{p}_1)} |\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_1} \gamma^{-1} C_{\vec{v},p}(G)) , \tag{2.46}$$

$$|\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}^{\perp} G, g]|_{\vec{v},p-1} \le C_{\vec{v},p+1}(G)|g|_{\vec{v},\mathfrak{p}_1} + C_{\vec{v},\mathfrak{p}_1}(G)|g|_{\vec{v},p+\nu+1}$$

   (c) setting $u := \Pi_K \Pi_{\mathcal{X}}(\mathrm{ad}(\Pi_{\mathcal{X}}^{\perp} F)[g] - F)$, one has

$$\begin{aligned} |u|_{\vec{v},\mathfrak{p}_1} &\le \varepsilon_0 \gamma^{-1} K^{-\eta+\mu} C_{\vec{v},\mathfrak{p}_1}(G) |\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_1} , \\ |u|_{\vec{v},\mathfrak{p}_2} &\le \gamma^{-1} K^{\mu} \Big(|\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_2} C_{\vec{v},\mathfrak{p}_1}(G) + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)} |\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_1} C_{\vec{v},\mathfrak{p}_2}(G) \Big) ; \end{aligned} \tag{2.47}$$

   (d) setting $\vec{v}' = (\gamma, \mathcal{O}, s - \rho s_0, a, r - \rho r_0)$, and let Φ the change of variables generated by g, one has that

$$|\Pi_{\mathcal{X}} \Phi_* F|_{\vec{v}',\mathfrak{p}_2-1} \le C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G) + \mathtt{C}\,(|g|_{\vec{v},\mathfrak{p}_2} C_{\vec{v},\mathfrak{p}_1}(G) + |g|_{\vec{v},\mathfrak{p}_1} C_{\vec{v},\mathfrak{p}_2}(G)) \tag{2.48}$$

Definition 2.24 (Compatible changes of variables). *Let the parameters in Constraint 2.21 be fixed. Fix also* $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$, $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$ *with* $\mathcal{O} \subseteq \mathcal{O}_0$ *a compact set, parameters* $K \ge K_0, \rho < 1$. *Consider a vector field* $F = N_0 + G \in \mathcal{W}_{\vec{v}^o,p}$ *which is* $C^{\mathtt{n}+2}$*-tame up to order* $q = \mathfrak{p}_2 + 2$ *and such that,*

$$F \in \mathcal{E} \quad \forall \xi \in \mathcal{O} , \qquad |\Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_2-1} \le \mathtt{C} C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G).$$

We say that a left invertible $\mathcal{E}$*-preserving change of variables*

$$\mathcal{L}, \mathcal{L}^{-1} : \mathbb{T}^d_s \times D_{a,\mathfrak{p}_1}(r) \times \mathcal{O}_0 \to \mathbb{T}^d_{s+\rho s_0} \times D_{a-\rho a_0,\mathfrak{p}_1}(r + \rho r_0)$$

is compatible *with* $(F, K, \vec{v}, \rho)$ *if the following holds:*

(i) $\mathcal{L}$ is "close to identity", i.e. denoting $\vec{v}^o_1 := (\gamma, \mathcal{O}_0, s - \rho s_0, a - \rho a_0, r - \rho r_0)$ one has

$$\|(\mathcal{L} - \mathbb{1})h\|_{\vec{v}^o_1,\mathfrak{p}_1} \le \mathtt{C}\varepsilon_0 K^{-1} \|h\|_{\vec{v}^o,\mathfrak{p}_1} . \tag{2.49}$$

(ii) $\mathcal{L}_*$ conjugates the $C^{\mathtt{n}+2}$-tame vector field F to the vector field $\hat{F} := \mathcal{L}_* F = N_0 + \hat{G}$ which is $C^{\mathtt{n}+2}$-tame; moreover denoting $\vec{v}_2 := (\gamma, \mathcal{O}, s - 2\rho s_0, a - 2\rho a_0, r - 2\rho r_0)$ one may choose the tameness constants of $\hat{G}$ so that

$$\begin{aligned} &C_{\vec{v}_2,\mathfrak{p}_1}(\hat{G}) \le C_{\vec{v},\mathfrak{p}_1}(G)(1 + \varepsilon_0 K^{-1}) , \\ &C_{\vec{v}_2,\mathfrak{p}_2}(\hat{G}) \le \mathtt{C}\big(C_{\vec{v},\mathfrak{p}_2}(G) + \varepsilon_0 K^{\kappa_3} C_{\vec{v},\mathfrak{p}_1}(G)\big) \\ &|\Pi_{\mathcal{X}} \hat{G}|_{\vec{v}_2,\mathfrak{p}_2-1} \le \mathtt{C}\big(C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G) + \varepsilon_0 K^{\kappa_3} C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp} G)\big) . \end{aligned} \tag{2.50}$$

(iii) $\mathcal{L}_*$ "preserves the $(\mathcal{N}, \mathcal{X}, \mathcal{R})$-decomposition", namely one has

$$\Pi_{\mathcal{N}}^{\perp}(\mathcal{L}_* \Pi_{\mathcal{N}} F) = 0 , \qquad \Pi_{\mathcal{X}}(\mathcal{L}_* \Pi_{\mathcal{X}}^{\perp} F) = 0 . \tag{2.51}$$

Given $\gamma_0 > 0$ we set for $n \geq 0$

$$
\begin{aligned}
&\mathtt{G}_n = \mathtt{G}_0(1 + \sum_{j=1}^{n} 2^{-j}), \quad \mathtt{R}_n = \mathtt{R}_0(1 + \sum_{j=1}^{n} 2^{-j}), \quad K_n = (K_0)^{\chi^n}, \, \gamma_n = \gamma_{n-1}(1 - \frac{1}{2^{n+2}}), \\
&a_n = a_0(1 - \frac{1}{2}\sum_{j=1}^{n} 2^{-j}), \quad r_n = r_0(1 - \frac{1}{2}\sum_{j=1}^{n} 2^{-j}), \quad s_n = s_0(1 - \frac{1}{2}\sum_{j=1}^{n} 2^{-j}), \\
&\Pi_n := \Pi^{(K_n)}, \quad \Pi_n^{\perp} := \mathbb{1} - \Pi_n \, , E_n = E^{(K_n)}, \quad \rho_n := \frac{1}{2^{n+5}}
\end{aligned}
\tag{2.52}
$$

Finally, for all $n \geq 0$ we denote $\vec{v}_n = (\gamma_n, \mathcal{O}_n, s_n, a_n, r_n)$, $\vec{v}_n^{\text{o}} = (\gamma_n, \mathcal{O}_0, s_n, a_n, r_n)$.

We reformulate our main result, stated in the Introduction, in a more precise way. This is useful for applications, where one needs to have information of the sequence of vector fields F_n and on the changes of variables $\mathcal{H}_n$ in order to prove that the set $\mathcal{O}_\infty$ is not empty.

Theorem 2.25 (Abstract KAM). *Fix a decomposition and a subspace $\mathcal{E}$ as in Definitions 2.17 and 2.19. Fix parameters $\varepsilon_0, \mathtt{R}_0, \mathtt{G}_0, \mu, \nu, \eta, \chi, \alpha, \kappa_1, \kappa_2, \kappa_3, \mathfrak{p}_2$ satisfying Constraint 2.21. Let N_0 be a diagonal vector field as in Definition 2.20 and consider a vector field*

$$
F_0 := N_0 + G_0 \in \mathcal{E} \cap \mathcal{W}_{\vec{v}_0, p} \tag{2.53}
$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$.

Fix $\gamma_0 > 0$ and assume that

$$
\gamma_0^{-1} C_{\vec{v}_0, \mathfrak{p}_2}(G_0) \leq \mathtt{G}_0 \, , \quad \gamma_0^{-1} C_{\vec{v}_0, \mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G_0) \leq \mathtt{R}_0 \, , \quad \gamma_0^{-1} |\Pi_{\mathcal{X}} G_0|_{\vec{v}_0, \mathfrak{p}_1} \leq \varepsilon_0 \, , \quad \gamma_0^{-1} |\Pi_{\mathcal{X}} G_0|_{\vec{v}_0, \mathfrak{p}_2} \leq \mathtt{R}_0 \, . \tag{2.54}
$$

For all $n \geq 0$ we define recursively changes of variables $\mathcal{L}_n, \Phi_n$ and compact sets $\mathcal{O}_n$ as follows.

Set $\mathcal{H}_{-1} = \mathcal{H}_0 = \Phi_0 = \mathcal{L}_0 = \mathbb{1}$, and for $0 \leq j \leq n-1$ set recursively $\mathcal{H}_j = \Phi_j \circ \mathcal{L}_j \circ \mathcal{H}_{j-1}$ and $F_j := (\mathcal{H}_j)_* F_0 := N_0 + G_j$. Let $\mathcal{L}_n$ be any change of variables compatible with $(F_{n-1}, K_{n-1}, \vec{v}_{n-1}, \rho_{n-1})$, and $\mathcal{O}_n$ be any compact set

$$
\mathcal{O}_n \subseteq \mathcal{O}_{n-1} \, , \tag{2.55}
$$

which satisfies the homological equation for $((\mathcal{L}_n)_* F_{n-1}, K_{n-1}, \vec{v}_{n-1}^{o}, \rho_{n-1})$. For $n > 0$ let g_n be the regular vector field defined in item (2) of Definition 2.23 and set Φ_n the time-1 flow map generated by g_n.

Then Φ_n is left invertible and $F_n := (\Phi_n \circ \mathcal{L}_n)_* F_{n-1} \in \mathcal{W}_{\vec{v}_n^{o}, p}$ is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. Moreover the following holds.

(i) *Setting $G_n = F_n - N_0$ then*

$$
\begin{aligned}
\Gamma_{n,\mathfrak{p}_1} &:= \gamma_n^{-1} C_{\vec{v}_n, \mathfrak{p}_1}(G_n) \leq \mathtt{G}_n, \quad \Gamma_{n,\mathfrak{p}_2} := \gamma_n^{-1} C_{\vec{v}_n, \mathfrak{p}_2}(G_n) \leq \mathtt{G}_0 K_n^{\kappa_1}, \\
\Theta_{n,\mathfrak{p}_1} &:= \gamma_n^{-1} C_{\vec{v}_n, \mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp} G_n) \leq \mathtt{R}_n, \quad \Theta_{n,\mathfrak{p}_2} := \gamma_n^{-1} C_{\vec{v}_n, \mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp} G_n) \leq \mathtt{R}_0 K_n^{\kappa_1} \\
\delta_n &:= \gamma_n^{-1} |\Pi_{\mathcal{X}} G_n|_{\vec{v}_n, \mathfrak{p}_1} \leq K_0^{\kappa_2} \varepsilon_0 K_n^{-\kappa_2}, \quad \gamma_n^{-1} |\Pi_{\mathcal{X}} G_n|_{\vec{v}_n, \mathfrak{p}_2} \leq \mathtt{R}_0 K_n^{\kappa_1} \\
|g_n|_{\vec{u}_n, \mathfrak{p}_1} &\leq K_0^{\kappa_2} \varepsilon_0 \mathtt{G}_0 K_{n-1}^{-\kappa_2 + \mu + 1}, \quad |g_n|_{\vec{u}_n, \mathfrak{p}_2} \leq \mathtt{R}_0 \mathtt{G}_0^{-1} K_{n-1}^{-\nu - 1 + \chi \kappa_1}
\end{aligned}
\tag{2.56}
$$

where $\vec{u}_n = (\gamma_n, \mathcal{O}_n, s_n + 12\rho_n s_0, a_n + 12\rho_n a_0, r_n + 12\rho_n r_0)$.

(ii) *The sequence $\mathcal{H}_n$ converges for all $\xi \in \mathcal{O}_0$ to some change of variables*

$$
\mathcal{H}_\infty = \mathcal{H}_\infty(\xi) : D_{a_0, p}(s_0/2, r_0/2) \longrightarrow D_{\frac{a_0}{2}, p}(s_0, r_0). \tag{2.57}
$$

(iii) *Defining* $F_\infty := (\mathcal{H}_\infty)_* F_0$ *one has*

$$\Pi_{\mathcal{X}} F_\infty = 0 \quad \forall \xi \in \mathcal{O}_\infty := \bigcap_{n \ge 0} \mathcal{O}_n \tag{2.58}$$

and

$$\gamma_0^{-1} C_{\vec{v}_\infty, \mathfrak{p}_1}(\Pi_{\mathcal{N}} F_\infty - N_0) \le 2\mathtt{G}_0, \quad \gamma_0^{-1} C_{\vec{v}_\infty, \mathfrak{p}_1}(\Pi_{\mathcal{R}} F_\infty) \le 2\mathtt{R}_0$$

with $\vec{v}_\infty := (\gamma_0/2, \mathcal{O}_\infty, s_0/2, a_0/2)$.

Proof. The proof of this result is deferred to Section 5. □

Remark 2.26. *Note that if one makes the further assumption that* $s_0 > 0$, *the smallness conditions as well as the definition of the set of parameters in Definition 2.23 simplify drastically: in particular one may choose* $\mathfrak{p}_2 = \mathfrak{p}_1$. *We are not makeing this assumption because our aim was to have a unified proof; however we discuss the time analytic case in Appendix D for completeness.*

3 Triangular decomposition and Mel'nikov conditions

In most applications one may redefine the sets on which one can solve the homological equation in a more direct way, by introducing the so-called *Mel'nikov conditions*.

We start by introducing some notation.

Definition 3.1 (Triangular decomposition). *We say that a decomposition* $(\mathcal{N}, \mathcal{X}, \mathcal{R})$ *is triangular if* $\mathcal{X}$ *admits a block decomposition*

$$\mathcal{X} = \bigoplus_{j=1}^{\mathtt{b}} \mathcal{X}_j \tag{3.1}$$

such that for all $N \in \mathcal{N}, R \in \mathcal{R}$ *setting*

$$\mathfrak{N} := \Pi_{\mathcal{X}} \mathrm{ad}(N), \quad \mathfrak{R} := \Pi_{\mathcal{X}} \mathrm{ad}(R) \tag{3.2}$$

then $\mathfrak{N}$ *is block diagonal and* $\mathfrak{R}$ *is strictly upper triangular, i.e.*

$$\mathfrak{N} : \mathcal{X}_i \to \mathcal{X}_i, \quad \mathfrak{R} : \mathcal{X}_i \to \bigoplus_{j > i} \mathcal{X}_j .$$

Remark 3.2. *In order to construct a triangular decomposition one generally associates some degree to the variables*[6]*:*

$$\deg(\theta) = 0, \quad \deg(w) = 1, \quad \deg(y) = \mathtt{d},$$

this automatically fixes the degree of a monomial vector field as

$$\deg(y^j e^{\mathrm{i}\theta \cdot \ell} w^\alpha \partial_{\mathtt{v}}) = j\mathtt{d} + |\alpha| - \deg(\mathtt{v}),$$

moreover one verifies that if g *has degree* d_1 *and* f *degree* d_2 *then* $[f, g]$ *has degree* $d_1 + d_2$. *Finally we remark that* $\mathcal{V}^{(\mathtt{v},0)}$ *has negative degree for* $\mathtt{v} = y, w$ *and degree equal to zero for* $\mathtt{v} = \theta$, *in the same way* $\mathcal{V}^{(\mathtt{v},\mathtt{v})}$ *has always degree zero for* $\mathtt{v} = y, w$ *while* $\mathcal{V}^{\theta,\mathtt{v}}$ *has positive degree. Then to a polynomial we may associate its minimal and maximal degree. In the same way if a* C^{k+1} *function has zero projection on all spaces* $\mathcal{V}^{\mathtt{v},\mathtt{v}_1,\cdots,\mathtt{v}_h}$, *with* $h \le k$ *and degree* $\le d$ *we call* d *its minimal degree. In many applications it is convenient to place all monomials of degree* ≤ 0 *in* $\mathcal{N} \cap \mathcal{X}$ *and all those of positive degree in* $\mathcal{R}$.

[6] clearly one must give θ degree zero since we do not Taylor expand on it, then by convention we decide to give degree one to w.

Lemma 3.3. *Any decomposition such that $\mathcal{N}$ contains only polyomials of degree zero and $\mathcal{R}$ contains only terms of minimal degree > 0 is triangular w.r.t. the degree decomposition of $\mathcal{X} = \oplus_j \mathcal{X}_j$ where the $\mathcal{X}_j$ are spaces of homogeneous polynomials with increasing degree $\mathtt{d}_j$.*

Proof. Given any polynomial P of positive minimal degree the operator ad$P$ has positive degree, namely its action on any polynomial increases its minimal degree. Now for any $N \in \mathcal{N}$ $\Pi_{\mathcal{X}}\mathrm{ad}N$ preserves the degree and thus maps each $\mathcal{X}_j$ into itself. By definition the maximal degree in $\mathcal{X}$ is given by $\mathtt{d}_b$. Then if we have a tame function f with minimal degree $> \mathtt{d}_b - \max(1, \mathtt{d})$ the operator $\Pi_{\mathcal{X}}\mathrm{ad}f$ on $\mathcal{X}$ is equal to zero, and this is just a property of polynomial subspaces in $\mathcal{R}$. Since all $R \in \mathcal{R}$ have positive degree, then obviously $\Pi_{\mathcal{X}}\mathrm{ad}R$ is upper triangular. □

Once we have a triangular decomposition we introduce the following notion

Definition 3.4 (Mel'nikov conditions). *Let $\gamma, \mu_1 > 0$, $K \geq K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider a vector field $F \in \mathcal{W}_{\vec{v}^o,p}$ i.e.*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}\,,$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We say $\mathcal{O}$ satisfies the Mel'nikov conditions for $(F, K, \vec{v}^o)$ if the following holds.

1. For all $\xi \in \mathcal{O}$ one has $F(\xi) \in \mathcal{E}$ and $|\Pi_{\mathcal{X}} G|_{\vec{v},\mathfrak{p}_2-1} \leq \mathtt{C}\, C_{\vec{v},\mathfrak{p}_2}(\Pi^{\perp}_{\mathcal{N}} G)$.

2. Setting $\mathfrak{N} := \Pi_K \Pi_{\mathcal{X}} \mathrm{ad}(\Pi_{\mathcal{N}} F)$ for all $\xi \in \mathcal{O}$ there exists a block-diagonal operator $\mathfrak{W} : E^{(K)} \cap \mathcal{X} \cap \mathcal{E} \to E^{(K)} \cap \mathcal{B}_{\mathcal{E}}$ such that for any vector field $X \in E^{(K)} \cap \mathcal{X} \cap \mathcal{E}$

(a) one has

$$|\mathfrak{W}X|_{\vec{v},p} \leq \gamma^{-1} K^{\mu_1} (|X|_{\vec{v},p} + K^{\alpha(p-\mathfrak{p}_1)} |X|_{\vec{v},\mathfrak{p}_1} \gamma^{-1} C_{\vec{v},p}(G)). \tag{3.3}$$

(b) setting $u := (\Pi_K \mathrm{ad}(\Pi_{\mathcal{N}} F)[\mathfrak{W}X] - X)$ one has

$$\begin{aligned} |u|_{\vec{v},\mathfrak{p}_1} &\leq \varepsilon_0 \gamma^{-1} K^{-\eta+\mu_1} C_{\vec{v},\mathfrak{p}_1}(G) |X|_{\vec{v},\mathfrak{p}_1}, \\ |u|_{\vec{v},\mathfrak{p}_2} &\leq \gamma^{-1} K^{\mu_1} \Big(|X|_{\vec{v},\mathfrak{p}_2} C_{\vec{v},\mathfrak{p}_1}(G) + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)} |X|_{\vec{v},\mathfrak{p}_1} C_{\vec{v},\mathfrak{p}_2}(G) \Big)\,. \end{aligned} \tag{3.4}$$

Then we have the following result.

Proposition 3.5 (Homological equation). *Let $\gamma > 0$, $K \geq K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider a vector field $F \in \mathcal{W}_{\vec{v}^o,p}$ i.e.*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}\,,$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. Assume that $\gamma \sim \mathrm{diam}\mathcal{O}_0$ and set

$$\Gamma_p := \gamma^{-1} C_{\vec{v},p}(G), \quad \Theta_p := \gamma^{-1} C_{\vec{v},p}(\Pi^{\perp}_{\mathcal{N}} G). \tag{3.5}$$

Assume finally that for any $f \in \mathcal{B}_{\mathcal{E}}$

$$\begin{aligned} |\Pi_{\mathcal{X}}[\Pi_{\mathcal{R}} G, f]|_{\vec{v},p-1} &\leq C_{\vec{v},p+1}(\Pi^{\perp}_{\mathcal{N}} G) |f|_{\vec{v},\mathfrak{p}_1} + C_{\vec{v},\mathfrak{p}_1}(\Pi^{\perp}_{\mathcal{N}} G) |f|_{\vec{v},p+\nu+1}, \\ |\Pi_{\mathcal{X}}[\Pi_{\mathcal{N}} G, f]|_{\vec{v},p-1} &\leq C_{\vec{v},p+1}(G) |f|_{\vec{v},\mathfrak{p}_1} + C_{\vec{v},\mathfrak{p}_1}(G) |f|_{\vec{v},p+\nu+1}, \end{aligned} \tag{3.6}$$

If $\mathcal{O}$ satisfies the Mel'nikov conditions of Definition 3.4 for $(F, K, \vec{v}^o)$ then $\mathcal{O}$ satisfies items 1. and 2.a-b-c of Definition 2.23 provided that we fix

$$\mu = (b+1)(\mu_1 + \nu + 1) + \mathtt{t} \tag{3.7}$$

where $\mathtt{t} > 0$ is such that

$$(1 + \Theta_{\mathfrak{p}_1}(1 + \Gamma_{\mathfrak{p}_1}))^b (1 + \Gamma_{\mathfrak{p}_1}) \leq K_0^{\mathtt{t}}.$$

Proof. The proof is deferred to Appendix C. □

4 Applications

In order to use the *Mel'nikov conditions* in Definition 3.4 instead of the *homological equation* in Definition 2.23 in Theorem 2.25 we need to prove that also item 2.d of Definition 2.23 holds. This latter point depends strongly on the application so we discuss it in various examples.

Clearly the simplest possible case is $\ell_{a,p} = 0$ or a finite dimensional space. In any case we need to work in some subspace $\mathcal{E}$ endowed with some structure (say reversible or Hamiltonian). To this purpose we restrict $\ell_{a,p}$ as follows.

Definition 4.1. *We assume that $\ell_{a,p}$ has a product structure $\ell_{a,p} = h_{a,p} \times h_{a,p}$ with $w = (z^+, z^-)$ and $h_{a,p}$ is a scale of Hilbert spaces w.r.t. a norm $\|\cdot\|_{a,p}$ satisfying* (2.1)*. Moreover we assume that the subspaces ℓ_K have a product structure as well $\ell_K = h_K \times h_K$ with the h_K satisfying Hypothesis 2.1.*

4.1 Example 1: Reversible Nash-Moser.

Let us first discuss the "minimal choice", i.e. where in all the definitions we make the simplest possible choices.

Clearly the minimal choice for $\mathcal{X}$ is

$$\mathcal{X} := \mathcal{V}^{(y,0)} \oplus \mathcal{V}^{(w,0)}, \tag{4.1}$$

whereas for $\mathcal{N}$ one can make for instance the classical choice

$$\mathcal{N} := \mathcal{V}^{(\theta,0)} \oplus \mathcal{V}^{(w,w)} \oplus \mathcal{V}^{(y,y)} \oplus \mathcal{V}^{(y,w)} \oplus \mathcal{V}^{(w,y)}. \tag{4.2}$$

The decomposition (4.1) and (4.2) is *trivially* triangular, see Definition 3.1, provided that we set $\mathtt{b} = 1$ since for any $R \in \mathcal{R}, X \in \mathcal{X}$ one has

$$\Pi_{\mathcal{X}}[R, X] = 0\,. \tag{4.3}$$

Note that it is a degree decomposition with $\mathtt{d}(y) = 1$, where $\mathcal{N}$ is generated by all the monomials of degree zero and $\mathcal{X}$ is generated by all those of negative degree.

We choose the regular vector fields as $\mathcal{A} = \mathcal{X}$, by setting for $f = \big(0, f^{(y)}(\theta), f^{(w)}(\theta)\big)$,

$$|f|_{\vec{v},p} := \|f\|^{(1)}_{\vec{v},p} = \|f\|_{\vec{v},p},$$

with the projectors Π_K defined as

$$\begin{aligned} (\Pi_K f^{(y,0)})(\theta) &:= \sum_{|\ell| \le K} f^{(y,0)}_\ell e^{\mathrm{i}\ell\cdot\theta}, \\ (\Pi_K f^{(w,0)})(\theta) &:= \sum_{|\ell| \le K} \Pi_{\ell_K} f^{(w,0)}_\ell e^{\mathrm{i}\ell\cdot\theta}. \end{aligned} \tag{4.4}$$

Lemma 4.2. *The regular vector fields defined above satisfy all the properties of Definition 2.18; moreover the norm $|f|_{\vec{v},p}$ is a* sharp *tameness constant for all $p \ge \mathfrak{p}_1$.*

Proof. Item (1) is trivial and the bound (2.30) follows essentially by an explicit computation (see the proof of Lemma B.7 for more details). Now by Definition 2.13 we have that the bounds (T_m) hold for any change of variables Φ and any y, w. Hence for $\Phi \equiv \mathbb{1}$ and $y = 0 = w$, for any $p \ge \mathfrak{p}_0$, one has

$$|f|_{\vec{v},p} = \|f\|_{\vec{v},p} = \|f(\Phi)\|_{\vec{v},p} \le C_{\vec{v},p}(f) + C_{\vec{v},\mathfrak{p}_0}(f)\|\Phi\|_{\vec{v},p} \le \mathtt{c}C_{\vec{v},p}(f), \tag{4.5}$$

where the last inequality holds since $\|\Phi\|_{\vec{v},p} \equiv 1$ independently of p (recall that the map Φ is evaluated at $w = y = 0$). This means that $|f|_{\vec{v},p}$ is a sharp tameness constant: this trivially implies that item (2) in

Definition 2.18 hold. Let us check item (3). Recall the definition of the projectors in (4.4) and of the norm in (2.7); for $\mathtt{v} = \theta, y$ one has

$$\begin{aligned}\|\Pi_K f\|^2_{s+s_1,a,p+p'} &= \sum_{|l|\le K} |f_l^{(\mathtt{v},0)}|^2 \langle l\rangle^{2(p+p')} e^{2|l|(s+s_1)} \\ &\le CK^{2p'} e^{2Ks_1} \sum_{l\in\mathbb{Z}} |f_l^{(\mathtt{v},0)}|^2 \langle l\rangle^{2p} e^{2|l|s} = K^{2p'} e^{2Ks_1} \|\Pi_K f\|^2_{s,a,p}.\end{aligned} \tag{4.6}$$

The latter bounds holds also for the norm (2.21), hence (2.32) holds. Similarly the estimate (2.32) holds also for $\mathtt{v} = w$. Moreover for $\mathtt{v} = w$ we can write

$$(\mathbb{1} - \Pi_K) f^{(w,0)}(\theta) = \sum_{|l|>K} e^{\mathrm{i}\ell\cdot\theta} f_l^{(w,0)} + \sum_{|l|\le K} (\mathbb{1} - \Pi_{\ell_K}) f_l^{(w,0)} e^{\mathrm{i}\ell\cdot\theta},$$

hence one has for $p, p' \in \mathbb{N}$

$$\begin{aligned}\|(\mathbb{1} - \Pi_K) f^{(w,0)}(\theta)\|^2_{s,a,p} &\le \sum_{|l|>K} \langle l\rangle^{2p} \|f_l^{(w,0)}\|^2_{a,\mathtt{p}_0} e^{2s|l|} + \sum_{|l|>K} \langle l\rangle^{2p} e^{2s|l|} \|(\mathbb{1} - \Pi_{\ell_K}) f_l^{(w,0)}\|^2_{a,\mathtt{p}_0} \\ &+ \sum_{|l|>K} \langle l\rangle^{2\mathtt{p}_0} \|(\mathbb{1} - \Pi_{\ell_K}) f_l^{(w,0)}\|^2_{a,p} e^{2s|l|} + \sum_{l\in\mathbb{Z}} \langle l\rangle^{2\mathtt{p}_0} e^{2s|l|} \|(\mathbb{1} - \Pi_{\ell_K}) f_l^{(w,0)}\|^2_{a,p} \\ &+ \sum_{|l|>K} \langle l\rangle^{2\mathtt{p}_0} \|\Pi_{\ell_K} f_l^{(w,0)}\|^2_{a,p} e^{2s|l|} + \sum_{|l|\le K} \langle l\rangle^{2p} e^{2s|l|} \|(\mathbb{1} - \Pi_{\ell_K}) f_l^{(w,0)}\|^2_{a,\mathtt{p}_0} \\ &\le 2K^{-2p'} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2(p+p')} \|f_l^{(w,0)}\|^2_{a,\mathtt{p}_0} e^{2s|l|} \\ &+ 2K^{-2p'} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2\mathtt{p}_0} \|f_l^{(w,0)}\|^2_{a,p+p'} e^{2s|l|} \\ &+ \mathtt{c} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2(p+p')} K^{-2(p'+p)} K^{2\mathtt{p}_0} K^{2(p-\mathtt{p}_0)} \|f_l^{(w,0)}\|^2_{a,\mathtt{p}_0} e^{2s|l|} \\ &+ \mathtt{c} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2\mathtt{p}_0} e^{2s|l|} K^{2(p-\mathtt{p}_0)} K^{-2(p+p'-\mathtt{p}_0)} \|f_l^{(w,0)}\|^2_{a,p+p'} \\ &\le CK^{-2p'} \|f^{(w,0)}\|^2_{s,a,p+p'},\end{aligned} \tag{4.7}$$

and the latter bounds holds also for the norm (2.21). Similar bounds holds also for $\mathtt{v} = \theta, y$, hence (2.33) holds. Condition (2.34) is trivial. Finally, items 4 and 5 can be checked easily since the map generated by vector fields in $\mathcal{A}$ are simply translations. □

By looking at the homological equation it is clear that the minimal requirement for the vector field is that $F^{(y)}(\theta, y, 0) = -F^{(y)}(-\theta, y, 0)$, otherwise even when $\ell_{a,p} = \emptyset$ one can easily produce examples in which invariant tori do not exist[7].

Following Sevryuk (see for instance [7] and references therein) one expects to require that the vector field satisfies some appropriate symmetry: this can be stated by saying that the vector field is reversible w.r.t. some involution. Naturally one needs also the "unperturbed vector field" N_0 to be reversible w.r.t. the chosen involution, being the vector field that identifies the approximately invariant torus. In the applictaions to PDEs, one typically deals with N_0 of the form

$$N_0 = \omega^{(0)} \cdot \partial_\theta + \mathrm{i}\Lambda^{(0)} w \partial_w = \omega^{(0)} \cdot \partial_\theta + \mathrm{i}\Omega^{(0)} z^+ \partial_{z^+} - \mathrm{i}\Omega^{(0)} z^- \partial_{z^-} \tag{4.8}$$

with $\Omega^{(0)}$ a linear operator which is θ-independent and block-diagonal w.r.t. all the h_K. Thus N_0 is a diagonal operator as in Definition 2.20. Unfortunately such N_0 is not reversible w.r.t. the "simple" involution $(\theta, y, w) \to$

[7] Consider for instance $\dot{y} = 1$.

$(-\theta, y, w)$, but it is reversible w.r.t. the involution $S : (\theta, y, (z^+, z^-)) \to (-\theta, y, (z^-, z^+))$. Therefore, as for the subspace $\mathcal{E}$ we choose

$$\mathcal{E} = \mathcal{E}^{(0)} := \left\{ F \in \mathcal{V}_{\vec{v},p} : \quad \begin{pmatrix} F^{(\theta)}(-\theta, y, (z^-, z^+)) \\ F^{(y)}(-\theta, y, (z^-, z^+)) \\ F^{(z^+)}(-\theta, y, (z^-, z^+)) \\ F^{(z^-)}(-\theta, y, (z^-, z^+)) \end{pmatrix} = - \begin{pmatrix} -F^{(\theta)}(\theta, y, (z^+, z^-)) \\ F^{(y)}(\theta, y, (z^+, z^-)) \\ F^{(z^-)}(\theta, y, (z^+, z^-)) \\ F^{(z^+)}(\theta, y, (z^+, z^-)) \end{pmatrix} \right\}, \tag{4.9}$$

i.e. the vector fields which are reversible w.r.t. the involution S.

The conditions of Definitions 2.17 and 2.19 are trivially fulfilled with $\mathtt{n} = 1$ and

$$\mathcal{B}_{\mathcal{E}} := \{ g = (0, g^{(y)}(\theta), g^{(w)}(\theta)) : \qquad g^{(y)}(-\theta) = g^{(y)}(\theta), \quad g^{(z^+)}(-\theta) = g^{(z^-)}(\theta) \}. \tag{4.10}$$

Now consider a vector field of the form $\mathcal{E} \ni F = N_0 + G$ and our aim is to apply Theorem 2.25 to F provided that G is sufficiently small. The simplest possible choice of compatible change of variables is, $\mathcal{L}_n := \mathbb{1}$ for all n. With such choices, our scheme is the standard a Nash-Moser algorithm to find solutions of the torus embedding equation (1.2). Indeed each Φ_n is a traslation in the y, w direction

$$y \to y + g_n^{(y)}(\theta), \quad w \to w + g_n^{(w)}(\theta)$$

so that

$$\mathcal{H}_n : \ y \to y + h_n^{(y)}(\theta), \quad w \to w + h_n^{(w)}(\theta), \quad h_n = \sum_{j=0}^{n} g_j, \tag{4.11}$$

$$F_n = F(\theta, y + h_n^{(y)}, w + h_n^{(w)}) - \partial_\theta h_n \cdot F^{(\theta)}(\theta, y + h_n^{(y)}, w + h_n^{(w)}).$$

Note that h_n is simply an *approximate solution* for the torus embedding equation (1.2), indeed one has that $F_n^{(y)}(\theta, 0, 0), F_n^{(w)}(\theta, 0, 0) \to 0$ as $n \to \infty$.

The difference with the standard Nash-Moser algorithm is therefore only in the point of view: instead of looking for a torus embedding, we are looking for a translation in the y, w variables which puts the embedding to zero.

With the above assumptions and assuming also that the smallness conditions (2.54) are satisfied, then we can apply Theorem 2.25. Now we show that in this case the set of parameters satisfying the the Mel'nikov conditions of Definition 3.4 also satisfies the homological equation of Definition 2.23.

Proposition 4.3. *Let $\gamma > 0$, $K \geq K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider the vector field $F \in \mathcal{W}_{\vec{v}^o,p}$ with (see (4.8))*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}_s^d \to V_{a,p},$$

which is C^3-tame up to order $q = \mathfrak{p}_2 + 2$. Assume that $\gamma \sim \mathrm{diam}\mathcal{O}_0$ and $F \in \mathcal{E}$ defined in (4.9). If $\mathcal{O}$ satisfies the Mel'nikov conditions of Definition 3.4 for $(F, K, \vec{v}^o)$ then $\mathcal{O}$ satisfies the Homological equation of Definition 2.23 provided that we fix parameters $\mu = \mu_1$.

Proof. We note that (4.3) implies $\Pi_{\mathcal{X}} \mathrm{ad}(\Pi_{\mathcal{X}}^{\perp} G) = \Pi_{\mathcal{X}} \mathrm{ad}(\Pi_{\mathcal{N}} G)$ so we may set $g = \mathfrak{W}\, \Pi_K \Pi_{\mathcal{X}} G \in \mathcal{B}_{\mathcal{E}}$ for $\xi \in \mathcal{O}$. It is easily seen that the first of (2.46) follows from (3.3). As for the second equation we use the sharpness of $|\cdot|_{\vec{v},p}$. Indeed by Lemma B.1 we know that $C_{\vec{v},p+1}(G) |g|_{\vec{v},\mathfrak{p}_1} + C_{\vec{v},\mathfrak{p}_1}(G) |g|_{\vec{v},p+\nu+1}$ is a tameness constant for $[\Pi_{\mathcal{X}}^{\perp} G, g]$. Then the bound follows by Lemma 4.2. Regarding 2.c, one simply notes that formula (3.4) implies (2.47).
Finally in order to prove 2.d, we start by showing that the inequality (2.48) holds by substituting the l.h.s. with a tameness constant $C_{\vec{v},\mathfrak{p}_2-1}(\Pi_{\mathcal{X}} \Phi_* F)$; this follows from Lemmata B.3, B.1, B.6 and Remark B.2, see the proof of estimate (5.12) for more details. Therefore, the bound (2.48) follows from the sharpness of $|\cdot|_{\vec{v},p}$ for any p. $\square$

By Proposition 4.3 the set $\mathcal{O}_\infty$ of Theorem 2.25 contains the intersection over n of the sets in which the Mel'nikov conditions are satisfied for $(F_n, K_n, \vec{v}_n^0)$, therefore we now analyze the Mel'nikov conditions. The operator $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}}F_n)$ has the form

$$\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}}F_n) = \left(F_n^{(\theta)}(\theta,0,0)\cdot\partial_\theta\right)\mathbb{1} + \begin{pmatrix} F_n^{(y,y)}(\theta) & F_n^{(y,w)}(\theta) \\ F_n^{(w,y)}(\theta) & F_n^{(w,w)}(\theta)\end{pmatrix}. \tag{4.12}$$

Recall that $F^{(\mathtt{v}_i,\mathtt{v}_j)}$ are defined in (2.24). Note that this operator maps $\mathcal{B}_{\mathcal{E}}$ in $\mathcal{X}\cap\mathcal{E}$. Finding $\mathfrak{W}$ satisfying (3.3) and (3.4) is now equivalent to finding an approximate inverse for a K_n-truncation of (4.12), which unfortunately seems a quite delicate question.

A possible simplification occurs if instead of (4.2) we consider the decomposition (recall (2.27))

$$\mathcal{N} := \langle\mathcal{V}^{(\theta,0)}\rangle\oplus\mathcal{V}^{(w,w)}\oplus\mathcal{V}^{(y,y)}\oplus\mathcal{V}^{(y,w)}\oplus\mathcal{V}^{(w,y)} \qquad \mathcal{X}=\mathcal{A}:=\mathcal{V}^{(y,0)}\oplus\mathcal{V}^{(w,0)}\oplus\mathcal{V}_0^{(\theta,0)},$$

and leave $\mathcal{E}$ unchanged; it is easily seen that the equivalent of Lemma 4.2 holds, and that

$$\mathcal{B}_{\mathcal{E}} := \{g=(g^{(\theta)}(\theta),g^{(y)}(\theta),g^{(w)}(\theta))\ :\ g^{(\theta)}(-\theta)=-g^{(\theta)}(\theta)\,,\ g^{(y)}(-\theta)=g^{(y)}(\theta)\,,\ g^{(z^+)}(-\theta)=g^{(z^-)}(\theta)\}\,. \tag{4.13}$$

Note that in this case we would obtain a stronger result, since the dynamics on the model torus would be linear.

We divide $\mathcal{X}$ by degree decomposition as in (3.1), with $\mathtt{b}=2$ and $\mathcal{X}_1=\mathcal{V}^{(y,0)}\oplus\mathcal{V}^{(w,0)}$, $\mathcal{X}_2=\mathcal{V}_0^{(\theta,0)}$; this decomposition is triangular by Remark 3.2 and the equivalent of Proposition 4.3 holds.

As before we fix $\mathcal{L}_n=\mathbb{1}$ for all n. Now the maps Φ_n are a translation in the y,w direction composed with a torus diffeomorphism . They are hence of the form

$$\theta\to\theta+h_n^{(\theta)}(\theta),\quad y\to y+h_n^{(y)}(\theta)\,,\quad w\to w+h_n^{(w)}(\theta) \tag{4.14}$$

defined in such a way that $F_n^{(\theta,0)}(\theta)=\omega^{(n)}+O(|g_n|)$ (here $\omega^{(n)}$ is the average of $F_n^{(\theta,0)}(\theta)$ w.r.t. θ).

Regarding the Mel'nikov conditions we have that, by definition, $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}}F_n)$ is block-diagonal on $\mathcal{X}=\mathcal{X}_1\oplus\mathcal{X}_2$ and its action on $\mathcal{X}_1$ is of the form

$$\left(\omega^{(n)}\cdot\partial_\theta\right)\mathbb{1} + \begin{pmatrix} F_n^{(y,y)}(\theta) & F_n^{(y,w)}(\theta) \\ F_n^{(w,y)}(\theta) & F_n^{(w,w)}(\theta)\end{pmatrix}, \tag{4.15}$$

while the action on $\mathcal{X}_2$ is simply $\omega^{(n)}\cdot\partial_\theta$. Thus the Mel'nikov conditons (3.3),(3.4) on the component $\mathcal{X}_2$ amount to requiring that $\omega^{(n)}$ is γ,τ diophantine up to order K_n. All the difficulty is now reduced to inverting (4.15).

Note that, under the same Diophantine hypotheses on $\omega^{(n)}$, the operator (4.12) can be reduced to the form (4.15) by choosing at each step n, the change of variables $\mathcal{L}_n$ to be the torus diffeomorphism which reduces $F_n^{(\theta)}(\theta,0,0)$ to its mean value. Of course one needs to verify that the $\mathcal{L}_n$ are in fact a sequence of compatible changes of variables as in Defintion 2.24.

If we assume that the subspaces ℓ_K of Hypothesis 2.1 are finite dimensional then the invertibility of $\Pi_{K_n}\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}}F_n)\Pi_{K_n}$ can be imposed by requiring that its eigenvalues are non-zero (the so-called first Mel'nikov condition); however, unless ℓ_K is uniformly bounded (i.e. when $\ell_{a,p}$ is a finite dimensional space) it is not at all trivial to obtain from such condition the bounds (3.3) and (3.4).

To the best of our knowledge the only examples in which one has enough control on (4.15) as it is, are the forced cases, i.e. when $F^{(\theta)}=\omega_0$ and there are no y variables, that is $d_1=0$. In this case one can use the so-called multiscale approach; see for instance [49, 25, 28]. Otherwise one needs a more refined decomposition; see below.

4.2 Example 2: Hamiltonian KAM/Nash-Moser.

The following section is essentially a reformulation in our notations of the approach proposed in [45]. We start by remarking that when the vector field (2.53) is Hamiltonian, it is natural to apply to it only symplectic changes of variables: this amounts to completing the maps introduced in (4.14) to symplectic ones. We now describe our procedure and at the end of the subsection we state the Theorem with an application to the NLS equation.

Definition 4.4 (Symplectic structure). *Recall that we assumed $\ell_{a,p}$ to have the product structure of Definition 4.1. We endow the phase space with the symplectic structure $d\theta \wedge dy + \mathrm{i} dz^+ \wedge dz^-$.*

We consider the decomposition

$$\mathcal{N} := \langle \mathcal{V}^{(\theta,0)} \rangle \oplus \mathcal{V}^{(w,w)} \oplus \mathcal{V}^{(y,w,w)}, \quad \mathcal{X} := \mathcal{V}_0^{(\theta,0)} \oplus \mathcal{V}^{(y,0)} \oplus \mathcal{V}^{(y,y)} \oplus \mathcal{V}^{(y,w)} \oplus \mathcal{V}^{(w,0)}. \tag{4.16}$$

This decomposition satisfies (2.17) with $\mathtt{n} = 2$ and it is a degree decomposition with $\deg(y) = 2$. Using the notation of (3.1) we have $\mathtt{b} = 3$ and

$$\mathcal{X}_1 = \mathcal{V}^{(y,0)}, \quad \mathcal{X}_2 = \mathcal{V}^{(y,w)} \oplus \mathcal{V}^{(w,0)}, \quad \mathcal{X}_3 = \mathcal{V}_0^{(\theta,0)} \oplus \mathcal{V}^{(y,y)}.$$

We remark that if one wants to solve the torus embedding equation taking advantage of the Hamiltonian structure, then the decompostion (4.16) appears naturally since it is the minimal decomposition containing (4.2) and preserving the Hamiltonian structure. More precisely given a change of coordinates as in (4.14), completing it to a symplectic one produces an element of $\mathcal{X}$ in (4.16).

Note that by (i) of Definition 2.19, the bigger is the set $\mathcal{X}$, the more delicate is the choice of $\mathcal{A}$.

Definition 4.5 (Finite rank vector fields). *We consider vector fields $f : \mathbb{T}^d_s \times D_{a,p+\nu}(r) \to V_{a,p}$ of the form*

$$\begin{aligned} &f = \sum_{v \in \mathtt{V}} f^{(v,0)} \partial_v + (f^{(y,y)} y + f^{(y,w)} \cdot w) \cdot \partial_y, \\ &f^{(y_i,w)} \in H^p(\mathbb{T}^d_s; \ell_{-a,-\mathfrak{p}_0-\nu}) \cap H^{\mathfrak{p}_0}(\mathbb{T}^d_s; \ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu}), \quad \langle f^{(\theta,0)} \rangle = 0 \end{aligned} \tag{4.17}$$

and we set for $p \geq \mathfrak{p}_1$

$$\begin{aligned} |f|_{s,a,p} := &\sum_{u=\theta,y,w} \|f^{(u,0)}\|_{s,a,p} + \max_{i,j=1,\dots,d_1} \|f^{(y_i,y_j)}\|_{s,a,p} + \\ &+ \frac{1}{r_0^{\mathtt{s}}} \max_{i=1,\dots,d_1} \Big(\|f^{(y_i,w)}\|_{H^p(\mathbb{T}^d_s;\ell_{-a,-\mathfrak{p}_0-\nu})} + \|f^{(y_i,w)}\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu})} \Big) \end{aligned} \tag{4.18}$$

We say that f is of finite rank *if $|f|_{s,a,p} < \infty$. We denote by $\mathcal{A}_{s,a,p}$ the space of finite rank vector fields. Given a compact set $\mathcal{O} \subseteq \mathcal{O}_0$ we denote by $\mathcal{A}_{\vec{v},p}$ with $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ the set of Lipschitz families $\mathcal{O} \to \mathcal{A}_{s,a,p}$ with the corresponding γ-weighted Lipschitz norm which we denote by $|\cdot|_{\vec{v},p}$.*

Remark 4.6. *Note that $\mathcal{V}^{(y,w)}$ is not contained in the set of finite rank vector fields; indeed in general by the identification of $\ell^*_{a,p}$ with $\ell_{-a,-p}$ one has that and $g \in \mathcal{V}^{(y_i,w)}$ can be written as $g^{(y_i,w)}(\theta) \cdot w \, \partial_{y_i}$ where*

$$g^{(y_i,w)} \in H^p(\mathbb{T}^d_s, \ell_{-a,-\mathfrak{p}_0-\nu}) \cap H^{\mathfrak{p}_0}(\mathbb{T}^d_s, \ell_{-a,-p-\nu}).$$

On the other hand (4.17) *is a stronger condition. Our – notationally quite unpleasant – choice of $\ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu}$ is needed in order to verify condition* (2.31) *in Definition 2.18.*

Definition 4.7. *Given $K > 0$ and a vector field $f \in \mathcal{A}$ we define the projection $\Pi_K f$ as*

$$
\begin{aligned}
&(\Pi_K f^{(\mathtt{v},0)})(\theta) := \sum_{|\ell|\le K} f_\ell^{(\mathtt{v},0)} e^{\mathrm{i}\ell\cdot\theta}, \qquad \mathtt{v} = \theta, y\,,\\
&(\Pi_K f^{(w,0)})(\theta) := \sum_{|\ell|\le K} \Pi_{\ell_K} f_\ell^{(w,0)} e^{\mathrm{i}\ell\cdot\theta}, \quad (\Pi_K) f^{(y_i,y_j)}(\theta) := \sum_{|\ell|\le K} f_\ell^{(y_i,y_j)} e^{\mathrm{i}\ell\cdot\theta}, i,j = 1,\dots,d\,,\\
&(\Pi_K f^{(y_i,w)})(\theta) := \sum_{|\ell|\le K} \Pi_{\ell_K} f_\ell^{(y_i,w)} e^{\mathrm{i}\ell\cdot\theta}\,,
\end{aligned}
\tag{4.19}
$$

and we define $E^{(K)}$ as the subspace of $\mathcal{A}_{\vec{v},p}$ where Π_K acts as the identity.

Lemma 4.8. *The finite rank vector fields of Definition 4.5 satisfy all the conditions of Definition 2.18.*

Proof. The Hilbert structure comes from the fact that (4.18) is defined by using the norm of an Hilbert space on each component. Item 1 follows by the definition while item 2 formula (2.30) is proved in B.7. To prove bound (2.31) in item 2 we reason as follow. Let us study the (y, w) component since the other are trivial. For $\Phi = \mathbb{1}$ one has that

$$
\|f^{(y,w)} \circ \Phi\|_{s,a,\mathfrak{p}_1} = \max_{i=1,\dots,d} \frac{1}{r_0^{\mathtt{s}}} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2\mathfrak{p}_1} |f_l^{y_i,w} \cdot w|^2 \le \max_{i=1,\dots,d} \frac{1}{r_0^{\mathtt{s}}} \sum_{l\in\mathbb{Z}} \langle l\rangle^{2\mathfrak{p}_1} \|f_l^{(y_i,w)}\|^2_{-a,-\mathfrak{p}_0-\nu} \|w\|^2_{a,\mathfrak{p}_0+\nu} \tag{4.20}
$$

using the Cauchy-Schwartz inequality. By the sharpness of the latter inequality we deduce that any tameness constant must be larger that the right hand side of (4.20), which in turn is bounded from below by $\frac{1}{2}|f^{(y,w)}|_{s,a,\mathfrak{p}_1}$.

Items 4, 5 are proved in Lemmata B.8 and B.9 in the Appendix.

Finally we need to show that item 3 holds. Now given a vector field $f \in \mathcal{A}_{\vec{v},p}$, the components $f^{(\mathtt{v},0)}$ are discussed in (4.6) and (4.7), and the components $f^{(\mathtt{v},0)}$ and $f^{(y,y)}$ can be treated in the same way. By the definition of the projector 4.7 one has that the last component (y, w) behaves essentially as the component $(w, 0)$. Hence again the smoothing bounds hold by reasoning as done in (4.6) and (4.7). □

We choose (J is the standard symplectic matrix)

$$
\mathcal{E} = \mathcal{E}^{(0)}_{\mathrm{Ham}} := \Big\{F \in \mathcal{V}_{\vec{v},p} :\ F = (\partial_y H, -\partial_\theta H, \mathrm{i}J\partial_w H)\,, \quad H(\theta, y, z^+, \overline{z^+}) \in \mathbb{R}\Big\}, \tag{4.21}
$$

while the regular vector fields are given by Definition 4.5. Note that by construction $\mathcal{E}^{(0)}_{\mathrm{Ham}} \cap \mathcal{X} \equiv \mathcal{A}$, indeed the condition $J\partial_w H(\theta, y, w) = F^{(w)}(\theta, y, w) \in \ell_{a,p}$ implies that $\partial_w F^{(y)} := -\partial_w\partial_\theta H(\theta, y, w) \in \ell_{a,p}$ as well. Then $\mathcal{B}_{\mathcal{E}}$ is the space of regular Hamiltonian vector fields. The conditions of Definitions 2.17 and 2.19 are trivially fulfilled. Note that the degree decomposition preserves the Hamiltonian structure.

Lemma 4.9. *Consider a tame vector field $f \in \mathcal{A}_{\vec{v},p} \cap \mathcal{E}$ (i.e. regular vector field according to Definition 4.5 which is Hamiltonian). There exists a $\mathtt{c}$ (depending at most on $\mathfrak{p}_0$ and on the dimensions d, d_1) such that for any tameness constant*

$$
|f|_{\vec{v},p} \le \mathtt{c}\, C_{\vec{v},p+1}(f) \tag{4.22}
$$

for any $p \ge \mathfrak{p}_1$.

Proof. On the components $(\mathtt{v}, 0)$, $\mathtt{v} = \theta, y, w$ and (y, y) the bound (4.22) is proved in Lemma 4.2. Let us study the $(y, w)-$component. First recall that, since we are in a Hamiltonian setting, then one has $f^{(y,w)}(\theta) = -iJ\partial_\theta f^{(w,0)}(\theta)$. Hence for $\Phi \equiv \mathbb{1}$ and $y = 0 = w$ one has for any $p \ge \mathfrak{p}_0$ that

$$
|f^{(y,w)}(\theta)|_{\vec{v},p} \le |\partial_\theta f^{(w,0)}(\theta)|_{\vec{v},p} \le |f^{(w,0)}(\theta)|_{\vec{v},p+1} = \|f^{(w,0)} \circ \Phi\|_{\vec{v},p+1} \le \mathtt{c}C_{\vec{v},p+1}(f). \tag{4.23}
$$

Threfore the assertion follows. □

As in Subsection 4.1 we now relate the Mel'nikov conditions to the homological equation.

Proposition 4.10. *Let $\gamma > 0$, $K \geq K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider a vector field $F \in \mathcal{W}_{\vec{v}^o,p} \cap \mathcal{E}^{(0)}_{\rm Ham}$ of the form*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}\,,$$

where $\mathcal{E}^{(0)}_{\rm Ham}$ is defined (4.21) and N_0 is defined in (4.8) with $\Omega^{(0)}$ self-adjoint. Assume that F is C^4-tame up to order $q = \mathfrak{p}_2 + 2$. Assume that $\gamma \sim {\rm diam}\mathcal{O}_0$ and set

$$\Gamma_p := \gamma^{-1} C_{\vec{v},p}(G), \quad \Theta_p := \gamma^{-1} C_{\vec{v},p}(\Pi^{\perp}_{\mathcal{N}} G). \tag{4.24}$$

If $\mathcal{O}$ satisfies the Mel'nikov conditions of Definition 3.4 for $(F, K, \vec{v}^o)$ then $\mathcal{O}$ satisfies the homological equation of Definition 2.23 provided that we fix parameters μ and $\mathtt{t}$ as in (3.7).

Proof. We wish to apply Proposition 3.5 in order to prove that items 1. and 2.a-b-c of Definition 2.23 are satisfied for $\mathcal{O}$ satisfying the Mel'nikov conditions. In order to do so we need to prove (3.6). The desired bounds follow from Lemma 4.9 and from the bounds (B.1) on tameness constants of commutators. We now prove item 2.d. We claim that there exists a choice of a tameness constant $C_{\vec{v},\mathfrak{p}_2-1}(\Pi_{\mathcal{X}}\Phi_* F)$ which satisfies (2.48). Indeed this follows from Lemmata B.3, B.1, B.6 and Remark B.2, see the proof of estimate (5.12) for more details. The bound (2.48) follows from Lemma 4.9.

In fact one may prove (2.48) directly (obtaining a slightly better bound). Let $\Phi = \mathbb{1} + f$ be the time one flow map of the field g defined by Proposition 3.5 and $\Phi^{-1} = \mathbb{1} + \tilde{f}$ its inverse. Note that f is of the form (4.17) and $\Pi_{\mathcal{X}}\Phi_* F = \Pi_{\mathcal{X}} F \circ \Phi^{-1} + df[F \circ \Phi^{-1}]$. The bound on the first summand follows by item 1. The only non trivial term in the second summand is given by

$$f^{(y,w)}(\theta)\big[d_w F^{(w)}(\Phi^{-1}(\theta,0,0))[w]\big] = \big((d_w F^{(w)}(\Phi^{-1}(\theta,0,0)))^* f^{(y,w)}(\theta)\big) \cdot w.$$

By the Hamiltonian structure the operator[8]

$$i\sigma_3 d_w F^{(w)}\,, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is self-adjoint, hence the bound (2.48) follows by the tame estimates on F and the fact that $f^{(y,w)} \in \ell_{-a,-\mathfrak{p}_0-\nu}$. Hence the assertion follows. ☐

Again, under the same assumptions as in Proposition 4.10 and of course assuming the smallness conditions (2.54), we can apply Theorem 2.25; by Proposition 4.10 the set $\mathcal{O}_\infty$ of Theorem 2.25 contains the intersection over n of the sets $\mathcal{C}_n$ in which the Mel'nikov conditions are satisfied for $(F_n, K_n, \vec{v}^{\rm o}_n)$, therefore we now analyze the Mel'nikov conditions.

The operator $\Pi_{\mathcal{X}}{\rm ad}(\Pi_{\mathcal{N}} F_n)$, restricted to the blocks $\mathcal{X}_1, \mathcal{X}_3$ coincide with the operator $(\omega^{(n)} \cdot \partial_\theta)\,\mathbb{1}$ while on the block $\mathcal{X}_2 = \mathcal{V}^{(w,0)} \oplus \mathcal{V}^{(y,w)}$, we get

$$\begin{pmatrix} (\omega^{(n)} \cdot \partial_\theta)\,\mathbb{1} - F_n^{(w,w)}(\theta) & 0 \\ -F_n^{(y,w,w)}(\theta) & (\omega^{(n)} \cdot \partial_\theta)\,\mathbb{1} + \big(F_n^{(w,w)}(\theta)\big)^* \end{pmatrix}. \tag{4.25}$$

Note that the operators appearing on the diagonal of (4.25) are $i\sigma_3$ times a self-adjoint operator, moreover the whole operator maps Hamiltonian vector fields into Hamiltonian vector fields.

As before, the Mel'nikov conditons (3.3),(3.4) on the component $\mathcal{X}_1, \mathcal{X}_3$ amount to requiring that $\omega^{(n)}$ is γ, τ diophantine up to order K_n.

In conclusion we have proved the following Theorem.

[8] We use the standard notation for the Pauli matrices.

Theorem 4.11. *Consider a vector field* $F \in \mathcal{W}_{\vec{v}^0,p} \cap \mathcal{E}^{(0)}_{\rm Ham}$ *of the form*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p}\,,$$

where $\mathcal{E}^{(0)}_{\rm Ham}$ *is defined* (4.21) *and* N_0 *is defined in* (4.8) *with* $\Omega^{(0)}$ *self-adjoint. Assume that* F *is* C^4*-tame up to order* $q = \mathfrak{p}_2 + 2$*. Fix* $\gamma > 0$ *such that* $\gamma \sim {\rm diam}\mathcal{O}_0$ *and assume that* G *satisfies the smallness conditions* (2.54) *of Theorem 2.25. Then there exists an invariant torus for* F *provided that* ξ *belong to the set* $\mathcal{O}_\infty$ *of Theorem 2.25. Finally* $\mathcal{O}_\infty$ *contains* $\bigcap_n \mathcal{C}_n$ *where* $\mathcal{C}_n$ *is the set of* ξ *such that* ω_n *is* $(\gamma,\tau)-$*diophantine and the matrix* $\mathfrak{N}_n$ *in* (4.25) *is approximatively invertible with tame bounds like* (3.3) *and* (3.4).

We claim that, in the applications, proving the approximate invertibility of (4.25) is significantly simpler than proving the invertibility of (4.15).

4.3 Example 3: Reversible KAM/Nash-Moser.

We now wish to obtain a triangular decomposition for the Melnikov conditions as in (4.25) but without restricting to Hamiltonian vector fields. To this purpose we set

$$\mathcal{N} := \mathcal{V}^{(\theta,0)} \oplus \mathcal{V}^{(w,w)}\,, \quad \mathcal{X} := \mathcal{V}^{(y,0)} \oplus \mathcal{V}^{(y,y)} \oplus \mathcal{V}^{(y,w)} \oplus \mathcal{V}^{(w,0)}. \tag{4.26}$$

or

$$\mathcal{N} := \langle \mathcal{V}^{(\theta,0)} \rangle \oplus \mathcal{V}^{(w,w)}\,, \quad \mathcal{X} := \mathcal{V}^{(\theta,0)}_0 \oplus \mathcal{V}^{(y,0)} \oplus \mathcal{V}^{(y,y)} \oplus \mathcal{V}^{(y,w)} \oplus \mathcal{V}^{(w,0)}. \tag{4.27}$$

such choices are compatible with Definition 2.17 with $\mathtt{n} = 1$. Note that both cases come from a degree decomposition provided that we fix $1 < \deg(y) < 2$, therefore they are trivially triangular. Now the degree decomposition of (3.1), say in case (4.27), reads $\mathtt{b} = 4$ and gives $\mathcal{X}_1 = \mathcal{V}^{(y,0)}$, $\mathcal{X}_2 = \mathcal{V}^{(w,0)}$, $\mathcal{X}_3 = \mathcal{V}^{(y,w)}$ and $\mathcal{X}_4 = \mathcal{V}^{(\theta,0)}_0 \oplus \mathcal{V}^{(y,y)}$.

We define the space of regular vector field $\mathcal{A}_{\vec{v},p}$ as the "finite rank vector field" of Definition 4.5. and introduce the smoothing operator Π_K as in Definition 4.7. By lemma 4.8 such vector fields satisfy all the conditions of Definition 2.18.

Regarding the choice of $\mathcal{E}$, we require the reversibility condition (4.9), moreover, in order to satisfy condition (i) and (iii) of Definition 2.19 we set

$$\mathcal{E} = \mathcal{E}^{(1)} := \Big\{ F \in \mathcal{E}^{(0)} :\ d_w F^{(y)}(\theta,y,w) \in H^p(\mathbb{T}^d_s;\ell_{-a,-\mathfrak{p}_0-\nu}) \cap H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu})\,, \Big\}. \tag{4.28}$$

If we set $\mathcal{L}_n = \mathbb{1}$ as before, we get changes of variables of the form

$$y \to y + h^{(y,0)}(\theta) + h^{(y,y)}(\theta) y + h^{(y,w)}(\theta)\cdot w\,, \quad w \to w + h^{(w,0)}(\theta)\,, \quad \theta \to \theta + h^{(\theta,0)}(\theta)$$

i.e. the changes of variables (4.14) of Example 1, composed with a (y,w)-linear change of variable of finite rank. Note that a regular $g \in \mathcal{X}$ is in $\mathcal{B}_\mathcal{E}$ if it satisfies (4.13).

On $\mathcal{E}^{(1)}$ we give a slightly stronger definition of tame vector field.

Definition 4.12. *We say that a* C^3-*tame vector field* $F \in \mathcal{E}^{(1)}$ *is "adjoint-tame" if there exists a choice of tameness constants* $C_{\vec{v},p}(F)$ *such that, for any* Φ *generated by* $g \in \mathcal{B}_\mathcal{E}$ *and for any* h *as in Definition 2.13, the adjoint*[9] *of* $d_{\mathtt{U}}F(\Phi)$ *is tame and satisfies the bounds. Setting*

$$X^p := H^{p+\nu}(\mathbb{T}^d_s;\mathbb{C}^{d_1} \times \ell_{-a,-\mathfrak{p}_0-\nu}) \cap H^{\mathfrak{p}_0+\nu}(\mathbb{T}^d_s;\mathbb{C}^{d_1} \times \ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu})$$

and

$$Y^p := H^p(\mathbb{T}^d_s;V_{-a,-\mathfrak{p}_0}) \cap H^{\mathfrak{p}_0}(\mathbb{T}^d_s;V_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0})$$

for $p \ge \mathfrak{p}_1$*, one has (see formula* (2.22)*)*

$$(T_1)^* \quad \|(d_{\mathtt{U}}F(\Phi))^*[h]\|_{\gamma,\mathcal{O},X^p} \le \big(C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)|g|_{\vec{v}',p+\nu}\big)\|h\|_{\gamma,\mathcal{O},Y^{\mathfrak{p}_0}} + C_{\vec{v},\mathfrak{p}_0}(F)\|h\|_{\gamma,\mathcal{O},Y^p}. \tag{4.29}$$

[9] We recall that, given a linear operator $A : X \to Y$ its adjoint is $A^* : Y^* \to X^*$. Our condition implies that $(d_{\mathtt{U}}F(\Phi))^*$ is bounded from $Y_1 \to X_1$, with $Y_1 \subset Y^*$ and $X_1 \subset X^*$ this is hence a much stronger condition.

We have the following Lemmata.

Lemma 4.13. *Consider a regular vector field $f \in \mathcal{A}_{\vec{v},p}$. Then f satisifes (4.29) with $C_{\vec{v},p}(f) = |f|_{\vec{v},p}$. Moreover there exists a $\mathtt{c}$ (depending at most on $\mathfrak{p}_0$ and on the dimensions d, d_1) such that for any tameness constant satisfying (4.29) one has*

$$|f|_{\vec{v},p} \le \mathtt{c}\, C_{\vec{v},p}(f) \tag{4.30}$$

for any $p \ge \mathfrak{p}_1$.

Proof. We only sketch the proof of the Lemma when Φ is the identity: the general case is essentially identical due to the simple structure of $\mathcal{A}_{\vec{v},p}$ and $\mathcal{B}$. The only non-trivial components are $f^{(y_i,w)} \cdot w$. The adjoint of the differential is then the map $\lambda \to f^{(y_i,w)}\lambda$ with $\lambda \in H^p(\mathbb{T}^d_s)$. The result follows by the definition of $|\cdot|_{\vec{v},p}$. □

Lemma 4.14. *The adjoint-tame vector fields are closed with respect to close to identity changes of variables $\Phi = \mathbb{1} + f$ generated by $\psi \in \mathcal{B}_{\mathcal{E}}$. In particular, setting $F_+ = \Phi_* F$, one has that the tameness constants in* (B.7) *satisfy condition* $(T1)^*$.

Proof. Fix a vector field $F : \mathbb{T}^d_s \times D_{a,p+\nu}(r) \times \mathcal{O} \to V_{a,p}$ which is C^3–tame, By Remark 2.7, we know that if $|f|_{\vec{v},\mathfrak{p}_1} = \mathtt{c}\rho$ with $\mathtt{c}$ small enough then there exists $\Phi^{-1} = \mathbb{1} + \tilde{f}$ with $|\tilde{f}|_{\vec{v},p} \sim |f|_{\vec{v},p} \sim |\psi|_{\vec{v},p}$ and one has, by Lemma B.3 $F_+ := \Phi_* F : \mathbb{T}^d_{s-2\rho s_0} \times D_{a,p+\nu}(r - 2\rho r_0) \times \mathcal{O} \to V_{a-2\rho a_0,p}$ is C^3–tame up to order $q - \nu - 1$, with scale of constants

$$C_{\vec{v}_2,p}(F_+) \le (1+\rho)\Big(C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F) C_{\vec{v}_1,p+\nu+1}(f)\Big), \tag{4.31}$$

where $\vec{v} := (\lambda, \mathcal{O}, s, a, r)$, $\vec{v}_1 := (\lambda, \mathcal{O}, s - \rho s_0, a - \varrho a_0, r - \rho r_0)$ and $\vec{v}_2 := (\lambda, \mathcal{O}, s - 2\rho s_0, a - 2\rho a_0, r - 2\rho r_0)$. Now by Lemma 4.13 since $f, \tilde{f}$ are "regular" vector fields, they are also "adjoint-tame".

Consider a transformation Γ generated by $g \in \mathcal{B}_{\mathcal{E}}$. We need to check that $(d_{\mathtt{U}} F_+(\Gamma))^*[h]$ satisfies (4.29) with $C_{\vec{v},p}(F) \rightsquigarrow C_{\vec{v}_2,p}(F_+)$.

One can write $F_+ = F \circ \Phi^{-1} + df(\Phi)[F \circ \Phi^{-1}]$ and study the two summands separately. First note that $\Phi \circ \Gamma = \Psi = \mathbb{1} + k$ with $k \in \mathcal{B}$ such that

$$|k|_{\vec{v}_1,p} \le |\psi|_{\vec{v},p} |g|_{\vec{v},\mathfrak{p}_1} + |\psi|_{\vec{v},\mathfrak{p}_1} |g|_{\vec{v},p}.$$

In particular k is "adjoint-tame" since it is regular. On the other hand if one has two linear (in y, w) vector fields A and B which are "adjoint-tame", then AB is clearly "adjoint-tame". Now one can write $(F_+ \circ \Gamma) = F(\Phi^{-1} \circ \Gamma) + df(\Phi \circ \Gamma)[F(\Phi^{-1} \circ \Gamma)]$. Let us study for instance the first summand. Essentially (4.29) follows by the chain rule, the property $(T1)^*$ on F, and the tame estimates on the differential of k and on its adjoint.

One has that $d_{\mathtt{U}} F(\Phi^{-1} \circ \Gamma) = d_{\mathtt{U}} F(\Phi^{-1} \circ \Gamma) d_{\mathtt{U}}(\Phi^{-1} \circ \Gamma)$. The estimate (4.29) follows by $(T1)^*$ on the field F and the estimates on k. Jus as an example consider the term from the differential of $df[F \circ \Phi^{-1}]$ is, for $i = 1, \dots, d_1$, the operator $f^{(y_i,w)} \cdot d_w F^{(w)}(\theta, y, w)[\cdot]$. For $h \in H^p(\mathbb{T}^d_s; \mathbb{C})$ we have the estimate

$$\begin{aligned}
&\|(d_w F^{(w)}(\Psi))^* f^{(y_i,w)} h\|_{\gamma,\mathcal{O},H^{p+\nu}(\mathbb{T}^d_s;\ell_{-a,-\mathfrak{p}_0-\nu}) \cap H^{\mathfrak{p}_0+\nu}(\mathbb{T}^d_s;\ell_{-a,p-\mathfrak{p}_1-\mathfrak{p}_0-\nu})} \overset{(T1)^*}{\le} \\
&\le (C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)|k|_{\vec{v},p+\nu}) \|f^{(y_i,w)} h\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\mathbb{C})} + C_{\vec{v},\mathfrak{p}_0}(F)(1 + |k|_{\vec{v},\mathfrak{p}_1}) \|f^{(y_i,w)} h\|_{H^p(\mathbb{T}^d_s;\mathbb{C})} \\
&\le (C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F)|k|_{\vec{v},p+\nu}) |\psi|_{\vec{v},\mathfrak{p}_0} \|h\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\mathbb{C})} + \\
&+ C_{\vec{v},\mathfrak{p}_0}(F)(1 + |k|_{\vec{v},\mathfrak{p}_1})(\|h\|_{H^p(\mathbb{T}^d_s;\mathbb{C})} |\psi|_{\vec{v},\mathfrak{p}_1} + \|h\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\mathbb{C})} |\psi|_{\vec{v},p}) \\
&\le C_{\vec{v},\mathfrak{p}_0}(F)(1 + 2|\psi|_{\vec{v},\mathfrak{p}_1} |g|_{\vec{v},\mathfrak{p}_1}) |\psi|_{\vec{v},\mathfrak{p}_1} \|h\|_{H^p(\mathbb{T}^d_s;\mathbb{C})} \\
&+ \|h\|_{H^{\mathfrak{p}_0}(\mathbb{T}^d_s;\mathbb{C})} \Big[C_{\vec{v},p}(F) |\psi|_{\vec{v},\mathfrak{p}_1} + C_{\vec{v},\mathfrak{p}_0}(F) |\psi|_{\vec{v},p+\nu} + C_{\vec{v},\mathfrak{p}_0}(F) |\psi|_{\vec{v},\mathfrak{p}_1} |g|_{\vec{v},p+\nu}\Big].
\end{aligned} \tag{4.32}$$

□

Proposition 4.15. *Let $\gamma > 0$, $K \geq K_0$, consider a compact set $\mathcal{O} \subset \mathcal{O}_0$ and set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. Consider a vector field $F \in \mathcal{W}_{\vec{v}^o,p} \cap \mathcal{E}^{(1)}$ of the form*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p} \,,$$

where $\mathcal{E}^{(1)}$ is defined (4.28) and N_0 is defined in (4.8) with $\Omega^{(0)}$ self-adjoint. Assume that F is C^3-tame up to order $q = \mathfrak{p}_2 + 2$ and adjoint-tame. Assume that $\gamma \sim \mathrm{diam}\mathcal{O}_0$ and set

$$\Gamma_p := \gamma^{-1} C_{\vec{v},p}(G), \quad \Theta_p := \gamma^{-1} C_{\vec{v},p}(\Pi^{\perp}_{\mathcal{N}} G). \tag{4.33}$$

If $\mathcal{O}$ satisfies the Mel'nikov conditions of Definition 3.4 for $(F, K, \vec{v}^o)$ then $\mathcal{O}$ satisfies the Homological equation of Definition 2.23 provided that we fix parameters μ and $\mathtt{t}$ as in (3.7). Moreover $\Phi_* F$ (defined in (2.48)) is adjoint-tame.

Proof. We wish to apply Proposition (3.5) in order to prove that items (1) and 2.a,b,c of Definition 2.23 are satisfied. In order to do so we need to prove (3.6). One has that (3.6) follows directly by using the adjoint-tameness estimates on $\Pi_{\mathcal{R}} G$ and $\Pi_{\mathcal{N}} G$ as done in the proof of Lemma 4.14. Regarding item 2.*d*. To prove estimate (2.48) one can use the adjoint-tameness of G to get the bound for the term $\Phi_* G$. The term $\Phi_* N_0$ must be treated as done in (5.31) of Proposition 5.1 by using the norm $|\cdot|_{\vec{v},p}$ instead of the tameness constant. This can be done using the (3.6) and the fact that g in Definition 2.23 satisfies the homological equation (2.47). The adjoint-tameness of $\Phi_* F$ follows by Lemma 4.14 since $g \in \mathcal{B}_{\mathcal{E}}$ by definition. □

As in the previous examples, under the same assumptions as in Proposition 4.15 and of course assuming (2.54), we can apply Theorem 2.25; by Proposition 4.15 the set $\mathcal{O}_\infty$ of Theorem 2.25 contains the intersection over n of the sets in which the Mel'nikov conditions are satisfied for $(F_n, K_n, \vec{v}^0_n)$, therefore we now analyze the Mel'nikov conditions (3.4) in this case.
The operator $\Pi_{\mathcal{X}} \mathrm{ad}(\Pi_{\mathcal{N}} F_n)$decomposes as follows: we get the operator $(\omega^{(n)} \cdot \partial_\theta)\, \mathbb{1}$ on the blocks $\mathcal{X}_1, \mathcal{X}_4$ while on the blocks $\mathcal{X}_2, \mathcal{X}_3$, we get

$$\left(\omega^{(n)} \cdot \partial_\theta\right) \mathbb{1} - F_n^{(w,w)}(\theta) \,, \quad \left(\omega^{(n)} \cdot \partial_\theta\right) \mathbb{1} + \left(F^{(w,w)}(\theta)\right)^* \tag{4.34}$$

respectively. As in the previous example the Diophantine condition on $\omega^{(n)}$ is used in order to solve the homological equations on $\mathcal{X}_1, \mathcal{X}_4$. In conclusion we have proved the following Theorem, which is the analogous of Theorem 4.11 in the reversible case, simply requiring less regularity for F, adjoint-tameness and of course the reversible structure instead of the Hamiltonian one.

Theorem 4.16. *Consider a vector field $F \in \mathcal{W}_{\vec{v}^o,p} \cap \mathcal{E}^{(1)}$ of the form*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p} \,,$$

where $\mathcal{E}^{(1)}$ is defined (4.28) and N_0 is defined in (4.8) with $\Omega^{(0)}$ self-adjoint. Assume that F is C^3-tame up to order $q = \mathfrak{p}_2 + 2$ and adjoint-tame. Fix $\gamma > 0$ such that $\gamma \sim \mathrm{diam}\mathcal{O}_0$ and assume that G satisfies the smallness conditions (2.54) of Theorem 2.25. Then there exists an invariant torus for F provided that ξ belong to the set $\mathcal{O}_\infty$ of Theorem 2.25. Finally $\mathcal{O}_\infty$ contains $\bigcap_n \mathcal{C}_n$ where $\mathcal{C}_n$ is the set of ξ such that ω_n is $(\gamma, \tau)-$diophantine and the matrices (4.34) are approximatively invertible with tame bounds like (3.3) and (3.4).

As in the previous example one could also apply our KAM scheme with the decomposition (4.2) but using as changes of variables operators $\mathcal{L}_n$ which block diagonalize (4.12) into (4.34).

4.4 Example 4: KAM with reducibility.

Up to now we have just reduced our problem to the inversion of (4.25) or (4.34). Inverting such operators is not trivial and requires some subtle multiscale arguments as discussed in Example 2, subsection 4.5 (see also

[45]).
Clearly a major simplification would appear if we were able to diagonalize (4.34)-(4.25). This is indeed the classical KAM approach (see [4], [15]) but it requires much stronger non resonance conditions, i.e. the second Mel'nikov conditions. How to use reducibility in order to prove bounds of the form (4.34)-(4.25) for nonlinear PDEs on a circle has been discussed in various papers, see [40], [39], [41]. Here we briefly recall the main point in the simplest possible context. We consider a Hamiltonian case and assume to work with the decomposition (4.16). Let suppose that in Definition 4.4 we have

$$h_{a,p} = \oplus_{j\in\mathbb{N}} h_j, \quad \ell_{a,p} = \oplus_{j\in\mathbb{N}} \ell_j\,, \quad \ell_j = h_j \times h_j\,, \quad w_j = (z_j^+, z_j^-)$$

with h_j finite dimensional subspaces, for simplicity suppose them one-dimensional. Then one may introduce finite dimensional monomial subspaces[10] $\mathcal{V}^{(\mathtt{v}, z_{j_1}^{\sigma_1}, \cdots, z_{j_k}^{\sigma_k})}$, with $\sigma_i = \pm 1$.

For a linear operator $A(\theta) \in \mathcal{L}(\ell_{a,p}, \ell_{a,p})$ one considers its block decomposition $\{A_j^i\}_{i,j\in\mathbb{Z}^r}$ and the off-diagonal decay norm

$$(|A|^{\rm dec}_{s,a,p})^2 := \sum_{h=(h_1,h_2)\in\mathbb{N}\times\mathbb{Z}^d} \langle h\rangle^{2p} e^{2(a|h_1|+s|h_2|)} \sup_{j\in\mathbb{N}} |A^j_{j+h_1}(h_2)|^2 \tag{4.35}$$

where $|A_j^i|$ is the operator norm on $\mathcal{L}(\ell_i, \ell_j)$. Then we consider the corresponding weighted Lipschitz norm which we denote $|\cdot|^{\rm dec}_{\vec{v},p}$. This gives a special role to diagonal θ-independent vector fields, so we define

$$\mathcal{N}_0 := \langle \mathcal{V}^{(\theta,0)}\rangle \bigoplus_{j,\sigma} \langle \mathcal{V}^{(z_j^\sigma, z_j^\sigma)}\rangle \cap \mathcal{V}^{(w,w)}\,,$$

Then one can choose $\mathcal{E}$ as

$$\mathcal{E}^{(2)}_{\rm Ham} := \{F \in \mathcal{E}^{(0)}_{\rm Ham} : \quad d_w F^{(w)}(\theta, y, u) = D + M\} \tag{4.36}$$

with D diagonal and M a bounded operator with finite $|\cdot|^{\rm dec}_{\vec{v},p}$ norm. We are in the framework of [15] or [48], but we are **not** requiring analiticity, therefore we need some tameness properties, which we ensure by choosing the norm $|\cdot|^{\rm dec}_{\vec{v},p}$. Let us briefly recall the notations. We consider the Hamiltonian vector field

$$F_0 = \omega_0 \cdot \partial_\theta + \mathrm{i} \sum_j \Omega_j^{(0)} z_j^+ \partial_{z_j^+} - \mathrm{i} \sum_j \Omega_j^{(0)} z_j^- \partial_{z_j^-} + G_0(\xi, y, \theta, w)$$

with the following assumptions.

A. *Non-degeneracy.* We require that for all $j \neq k$, $\ell \in \mathbb{T}^d$

$$\big|\{\xi \in \mathcal{O}_0 : \omega_0 \cdot l + \Omega_j^{(0)} \pm \Omega_k^{(0)} = 0\}\big| = 0\,, \qquad \Omega_j^{(0)} \pm \Omega_k^{(0)} \neq 0 \quad \forall \xi \in \mathcal{O}_0 \tag{4.37}$$

B. *Frequency Asymptotics.* We assume that $\xi \to \omega^{(0)}(\xi)$ is a lipeomorphism and

$$|\omega(\xi)|, |\omega(\xi)|^{\rm Lip}, |\Omega_j^{(0)}(\xi) - j^\nu|, |\Omega_j^{(0)}(\xi)|^{\rm Lip} < M\,, \quad |\xi(\omega)|^{\rm Lip} \le L\,, \quad \forall \xi \in \mathcal{O}_0$$

for some $\nu > 1$.

C. *Regularity.* We require $G_0 \in \mathcal{E}^{(2)}_{\rm Ham}$ with $D = 0$, more precisely G is C^4 tame bounded Hamiltonian vector field with $\Pi_{\mathcal{N}_0} G = 0$ well defined and Lipschitz for $\xi \in \mathcal{O}_0$ a compact set of positive measure.

[10] Of course if $\ell_{a,p}$ is infinite dimensional then it is not true that for instance $\bigoplus_{i,j} \mathcal{V}^{(w_i,w_j)} = \mathcal{V}^{(w,w)}$.

D. *Smallness.* In Constraint 2.21 fix

$$\alpha = 0\,, \quad \chi = 3/2\,, \quad \kappa_1 = 2\kappa_0 + 1, \kappa_3 = 3\kappa_0 + 1\,, \quad \kappa_2 = 4\kappa_0 + 1\,,$$

$$\eta = \mu + 2\kappa_2 + 3\,, \quad \Delta\mathfrak{p} = 9\kappa_0 + 3, \quad \mathfrak{p}_1 = \frac{d+2}{2} + \nu + 1$$

this leaves as only parameter μ. Suppose that $\mathtt{G}_0 \sim \mathtt{R}_0 \sim 1$, fix ε_0 small and set $K_0 = \varepsilon_0^{-1/5\kappa_0}$ so that the smallness conditions (2.43),(2.44a)–(2.44c) are satisfied.

We assume finally that

$$\mathcal{P}_0 := \Pi_{\mathcal{N}} F_0 - \Pi_{\mathcal{N}_0} F_0 = \Pi_{\mathcal{N}} G_n - \Pi_{\mathcal{N}_0} G_0 = P^{(0)}(\theta) w \partial_w + \frac{\mathrm{i}}{2} Jw \cdot \sum_i \partial_{\theta_i} P^{(0)}(\theta) w \partial_{y_i}$$

is small i.e.

$$\gamma_0^{-1} |P^{(0)}|^{\mathrm{dec}}_{\vec{v}_0, \mathfrak{p}_1} \le \varepsilon_0$$

We are in the context of Subsection 4.2, so we have Proposition 4.10 with $\mathtt{b} = 3, \mathtt{t} = 1$, and for convenience let us set

$$\mu = 4(\mu_1 + \nu) + 5\,, \quad \mu_1 = 2\tau + 2\,. \tag{4.38}$$

Theorem 4.17. *Fix $\tau > d + 1 + \frac{2}{\nu - 1}$ and $\varepsilon_0 (LM)^{\tau+1} \ll 1$. Let F_0 be a C^4 tame vector field up to order $q = \mathfrak{p}_2 + 2$ satisfying assumptions A. to D. Then for γ_0 small enough there exists positive measure Cantor like set $\mathcal{O}_\infty(\gamma_0) \subset \mathcal{O}_0$ of asymptotically full Lebesgue measure as $\gamma \to 0$ and a symplectic close to identity change of variables $\mathcal{H}_\infty$ such that*

$$\Pi_{\mathcal{X}} (\mathcal{H}_\infty)_* F_0 = 0\,, \quad (\Pi_{\mathcal{N}} - \Pi_{\mathcal{N}_0})(\mathcal{H}_\infty)_* F_0 = 0, \,, \quad \forall \xi \in \mathcal{O}_\infty(\gamma_0)$$

so that F_0 has a reducible KAM torus.

We apply Theorem 2.25 to F_0 and we aim to show that we can produce a non-empty set $\mathcal{O}_\infty$ by choosing the $\mathcal{L}_n$ appropriately. By definition, at each step n, set

$$\Pi_{\mathcal{N}_0} F_n = \mathcal{D}_n = \omega_n \cdot \partial_\theta + \mathrm{i} \sum_j \Omega_j^{(n)} z_j^+ \partial_{z_j^+} - \mathrm{i} \sum_j \Omega_j^{(n)} z_j^- \partial_{z_j^-}\,, \tag{4.39}$$

$$\Pi_{\mathcal{N}} F_n - \Pi_{\mathcal{N}_0} F_n = \mathcal{P}_n = P^{(n)}(\theta) w \partial_w + \frac{\mathrm{i}}{2} Jw \cdot \sum_i \partial_{\theta_i} P^{(n)}(\theta) w \partial_{y_i}. \tag{4.40}$$

Lemma 4.18 (KAM reduction step)**.** *Fix $\vec{v}_n = (\gamma_n, \mathcal{O}_n, s_n, a_n, r_n)$, $\vec{v}_n^0 = (\gamma_n, \mathcal{O}_0, s_n, a_n, r_n)$ and $K_n \gg 1$ as in (2.52). Given a Hamiltonian vector field $\mathcal{D}_n + \mathcal{P}_n$ as in (4.39) with*

$$\rho_n (M\gamma_n)^{-1} K_n^{2\tau+1} |P^{(n)}|^{\mathrm{dec}}_{\vec{v}_n, \mathfrak{p}_1} \ll 1$$

there exists a symplectic change of variables $\mathcal{L}_{n+1}$, which conjugates

$$(\mathcal{L}_{n+1})_* (\mathcal{D}_n + \mathcal{P}_n) := \widehat{\mathcal{D}}_n + \widehat{\mathcal{P}}_n$$

(here $\widehat{\mathcal{D}}_n$ is the projection of the conjugate vector field onto $\mathcal{N}_0$) so that

1. *$\mathcal{L}_{n+1}$ is the time one flow of the Hamiltonian vector field*

$$\mathcal{S}_{n+1} := S^{(n+1)}(\theta) w \partial_w + \frac{\mathrm{i}}{2} Jw \cdot \sum_i \partial_{\theta_i} S^{(n+1)}(\theta) w \partial_{y_i}\,, \quad \text{with } |S^{(n+1)}|^{\mathrm{dec}}_{\vec{v}_n^0, p} \le M \gamma_n^{-1} K_n^{2\tau+1} |P^{(n)}|^{\mathrm{dec}}_{\vec{v}_n, p}.$$

2. *Set*

$$\widehat{\mathcal{O}}_n := \{\xi \in \mathcal{O}_n : \ |\omega^{(n)} \cdot l + \Omega_j^{(n)} \pm \Omega_k^{(n)}| \geq \frac{M\gamma_n}{K_n^\tau}, \quad |\ell| \leq K_n\}. \tag{4.41}$$

For all $\xi \in \widehat{\mathcal{O}}_n$*,* $\mathcal{S}_{n+1}$ *solves the Homological equation*

$$[\mathcal{S}_{n+1}, \mathcal{D}_n] + \Pi_{K_n}\mathcal{P}_n = 0\,, \quad (\Pi_K P)_{ij} = \begin{cases} P_{ij} & \text{if } |i-j| \leq K \\ 0 & \text{otherwise} \end{cases}$$

3. *Setting* $\widehat{v}_n := (\gamma_n, \widehat{\mathcal{O}}_n, s_n, a_n, r_n)$ *we have the bounds*

$$|\widehat{P}^{(n)}|^{\rm dec}_{\widehat{v}_n, \mathfrak{p}_1} \leq |S^{(n+1)}|^{\rm dec}_{\vec{v}, \mathfrak{p}_1} |P^{(n)}|^{\rm dec}_{\vec{v}, \mathfrak{p}_1} + K_n^{-(\mathfrak{p}_2 - \mathfrak{p}_1)} |P^{(n)}|^{\rm dec}_{\vec{v}_n, \mathfrak{p}_2}\,,$$
$$|\widehat{P}^{(n)}|^{\rm dec}_{\widehat{v}_n, \mathfrak{p}_2} \leq |P^{(n)}|^{\rm dec}_{\vec{v}_n, \mathfrak{p}_2} + {\rm const} |S^{(n+1)}|^{\rm dec}_{\vec{v}, \mathfrak{p}_2} |P^{(n)}|^{\rm dec}_{\vec{v}_n, \mathfrak{p}_2}\,,$$

Now we may apply this reduction step at each step in our iteration in Theorem 2.25, provided that we show that the $\mathcal{L}_n$ are compatible changes of variables. This we prove by induction. In fact we recursively obtain that

$$\gamma_n^{-1} |P^{(n)}|^{\rm dec}_{\vec{v}, \mathfrak{p}_1} \leq K_0^{\kappa_1} \varepsilon_0 K_n^{-\kappa_1}\,, \quad \gamma_n^{-1} |P^{(n)}|^{\rm dec}_{\vec{v}, \mathfrak{p}_2} \leq 2\mathtt{G}_0 K_n^{\kappa_1}$$

so that the $\mathcal{L}_n$ are compatible since $\kappa_3 > \kappa_1 + 2\tau + 1$. Now we can choose a set $\mathcal{O}_{n+1}$ which satisfies the Melnikov conditions (4.25) at step n for $\widehat{F}_n = (\mathcal{L}_n)_* F_n$. Since $\Pi_{\mathcal{N}} \widehat{F}_n = (\mathcal{L}_n)_* \Pi_{\mathcal{N}} F_n$, this amounts to finding an approximate inverse for ${\rm ad}(\widehat{\mathcal{D}}_n + \widehat{\mathcal{P}}_n)$ which satisfies (3.3) with μ_1 fixed in (4.38). Now, finding a partial inverse of (4.25) is equivalent to finding an inverse to ${\rm ad}\, \widehat{\mathcal{D}}_n$, since ${\rm ad}\, \widehat{\mathcal{P}}_n$ can be put inside the remainder, see formula (2.47). In turn the invertibility of ${\rm ad}\, \widehat{\mathcal{D}}_n$ on $\mathcal{X}_2$ is ensured by requiring lower bounds on the eigenvalues, i.e. choosing at each step

$$\mathcal{O}_{n+1} := \{\xi \in \widehat{\mathcal{O}}_n : \ |\widehat{\omega}_n \cdot l + \widehat{\Omega}_j^{(n)}| \geq \frac{\gamma_n M}{K_n^\tau}\,, \quad |\ell| \leq K_n\} \tag{4.42}$$

in Theorem 2.25. One can easily check that, troughout the algorithm $\xi \to \omega^{(n)}(\xi)$ and $\xi \to \widehat{\omega}^{(n)}(\xi)$ are lipeomorphisms and

$$|\omega^{(n)} - \omega_0| + \gamma_0 |\omega^{(n)} - \omega_0|^{\rm Lip}, |\Omega_j^{(n)} - \Omega_j^{(0)}| + \gamma_0 |\Omega_j^{(n)} - \Omega_j^{(0)}|^{\rm Lip} < \gamma_0 \varepsilon_0\,, \quad |\xi_n(\omega)|^{\rm Lip} \leq 2L\,, \quad \text{in } \mathcal{O}_n,$$

the same for the corresponding $\widehat{\cdot}$ quantities.

Lemma 4.19. *For* $\tau > d + 1 + \frac{2}{\nu - 1}$ *and* $\varepsilon_0 (LM)^{\tau+1} \ll 1$ *we have that* $|\mathcal{O}_0 \setminus \cap_n \mathcal{O}_n| \to 0$ *as* $\gamma_0 \to 0$*.*

Proof. This is proved in [15] Corollary C. □

The main point in this approach is that at each step, while constructing our approximate solutions by inverting (4.25), we apply a change of variables which approximately diagonalizes the linearized operator up to a negligible remainder. Condition (4.41), ensures that the sequence of linear changes of variables converges. In fact this last approach could be made slightly more flexible, indeed in our construction we have used the norm (4.35) on the component $\mathcal{V}^{(w,w)} \oplus \mathcal{V}^{(y,w,w)}$ in order to perform the reduction. This imposes some unnecessary conditions on the changes of variables, since the only unavoidable request on the $\mathcal{L}_n$ is that they preserve $\mathcal{E}$ and do not disrupt the bounds. A typical example is a change of variables of the form $\mathcal{L}w(x) = w(x + a(\theta, x))$. One does not expect such change of variables to have finite (4.35) norm but, if a is chosen appropriately, it can be a compatible change of variables.

4.5 Application to the NLS.

Let us specify to a PDE context; typical examples are $\ell_{a,p} = H^p(\mathbb{T}^d_a)$ or $\ell_p = \ell_{a=0,p} = H^p(\mathbb{G})$ with $\mathbb{G}$ a compact Lie group or homogeneous manifold. Then we may suppose that $\Omega^{(0)}$ in (4.8) is diagonal w.r.t. the Fourier basis (and self-adjoint) while G is a composition operator w.r.t. the w variable. For simplicity of the exposition, let us restrict ourself to semi-linear PDEs with no derivative in the nonlinearity, so that $G(w(x)) = f(w(x))$ with f a C^k map $\mathbb{C}^2 \to \mathbb{C}^2$. More precisely we choose

$$\mathcal{E} = \mathcal{E}^{(1)}_{\rm Ham} := \Big\{ F \in \mathcal{E}^{(0)}_{\rm Ham} : \quad d_w F^{(w)}(\theta, y, u) = D + M + R\,, \quad u \in D_{a,p}(r) \Big\} \tag{4.43}$$

where D is a diagonal operator M is a multiplication operator while R is finite rank. Since we apply symplectic changes of variables which are the identity plus a traslation plus a finite rank operator, $\mathcal{E}^{(1)}_{\rm Ham}$ is preserved throughout our algorithm.

As an example consider the NLS equation on a simple *compact* Lie group $\mathbb{G}$:

$$\mathrm{i}\partial_t u + \Delta u + M_\xi u = \epsilon g(|u|^2) u\,, \tag{4.44}$$

here $g(y) \in C^q(\mathbb{R}, \mathbb{R})$ with q large, while M_ξ is a Fourier multiplier

$$M_\xi \phi_i(x) = \xi_i \phi_i(x)\,, \quad i = 1, \cdots, d$$

where $\phi_n(x)$ are distinct eigenfunctions of the Laplace-Beltrami operator. We introduce the variables $\theta, y, w = (z, \bar z)$ by writing (for $I_n > 0$)

$$u(x) = \sum_{j=1}^d \sqrt{I_j + y_j} e^{\mathrm{i}\theta_j} \phi_j(x) + z(x)\,, \tag{4.45}$$

where $z(x)$ belongs to the orthogonal complement of $\mathrm{Span}(\phi_i(x))_{i=1}^d$ which by definition is $\ell_p = \ell_{0,p}$. As norm we choose the one induced on ℓ_p by $H^p(\mathbb{G})$. This change of variables is symplectic and one obtains a Hamiltonian vector field $F = N_0 + G$ of the form (1.3), where $\omega^{(0)}_i(\xi) = \lambda_i + \xi_i$, with λ_i being the eigenvalue of the Laplace-Beltrami operator associated to ϕ_i. Correspondingly $\Lambda^{(0)}$ is the Laplace-Beltrami operator restricted to the complement of $\mathrm{Span}(\phi_n(x))_{n=1}^d$. Since the non-linearity is a composition operator on H^p, then classical results ensure the C^k tameness of the vector field for all k. Moreover the restriction of a multiplication operator to ℓ_p is a multiplication operator up to a finite rank operator (ℓ_p is the complement of a finite dimensional subspace), then $F \in \mathcal{E}^{(1)}_{\rm Ham}$. Now we fix $\gamma_0 > 1/2$ and take $\mathtt{G}_0, \mathtt{R}_0, \varepsilon_0 \sim \epsilon$ in (2.21). In this way the NLS equations satisfies all the hypotheses of Theorem 4.11 and hence we deduce the existence of an invariant torus in the set $\mathcal{O}_\infty$.

Let us now discuss the measure estimates for the sets $\mathcal{C}_n$ in this setting. It is easily seen that $\omega^{(n)}_j = \lambda_j + \xi_j + O(\epsilon)$ so imposing the diophantine conditions on such sequence is simple. As one could expect the key point is the inversion of the operator in (4.25). In turn, since the operator is triangular, this amounts to inverting the diagonal, i.e. the operator

$$\mathfrak{L}_n = (\omega^{(n)} \cdot \partial_\theta)\, \mathbb{1} - F^{(w,w)}_n(\theta). \tag{4.46}$$

acting on $\mathcal{V}^{(w,0)}$. We first remark that $F^{(w,w)}_n(\theta)$ is the linearized of $F^{(w)}$ at an approximate solution up to a finite rank term. This follows from the fact that our changes of variables are traslations plus finite rank. From this we deduce that $F^{(w,w)}_n(\theta)$ is a multiplication operator plus a finite rank one. Now in order to prove that the estimates (3.3) and (3.4) hold for (4.46) for a large set of ξ one can use a multiscale theorem, such as the one in [19],[24] or [28]. Indeed one may verify that $\mathfrak{L}_n$ in (4.46) fits all the hypotheses of Proposition 5.8 in [28] so that the tame estimates on the inverse follow by conditions on the eigenvalues. The measure estimates follow by eigenvalue variation (again just as in [28]).

If $\mathbb{G} = \mathbb{T}^k$, this general strategy is carried on in full details in [45], in the more complicated case of a multiplication potential; see also the application to the wave equation [50]. Since the authors follow the Nash-Moser approach they never apply the symplectic change of variables which block-diagonalize, but only deduce its existence by the Hamiltonian structure.

If $\mathbb{G} = \mathbb{T}$ and we look for odd solutions, then the hypoteses A. to D. of subsection 4.4 are satisfied and using the same change of variables as in (4.45) and reasoning as above, one can see that $F \in \mathcal{E}^{(2)}_{\rm Ham}$, see (4.36), and therefore one can apply Theorem 4.17 and obtain that the invariant torus is also reducible; note that in this case the procedure is complete in the sense that the set $\mathcal{O}_\infty$ has positive measure.

Theorem. *The NLS equation* (4.44) *admits reducible and linearly stable quasi-periodic solutions for all* ε *small enough and for all* ξ *in a positive measure set.*

Note that one could avoid adding the Fourier multiplier in (4.44) and obtain the parameters by using Birkhoff Normal form (this is done for instance in [51]).
In our application, we have not considered a reversible case: this is due to the fact that the natural structure for the semilinear NLS is the Hamiltonian one. On the other hand, if one considers DNLS (Derivative NLS) there are interesting examples which are not Hamiltonian but instead they are reversible; see for instance [32, 41]. We believe (see [52, 46]) that our result can be fruitfully applied also to the fully nonlinear autonomous case.

4.6 Some comments

The examples 1-4 show that our KAM approach interpolates from the Nash-Moser scheme to the KAM. Each different choice of decomposition depends on the specific application one is studying. Note that we could always choose the simplest decomposition (4.2) (or (4.16) in the Hamiltonian setting) and achieve the block-decoupling of the linearized operators through the compatible change of variables $\mathcal{L}_n$.

A final remark is in order. In the PDEs setting of subsection 4.5, applying the changes of variables of Lemma 4.18 implies some loss of information. Indeed it is easily to see that changes of variables do not preserve $\mathcal{E}^{(1)}_{\rm Ham}$ unless we can show that the matrices in $S^{(n)}$ are Töplitz (up to finite rank).

Unfortunately in most applications this is not the case: indeed in classic KAM scheme the PDEs structure is essentially ignored and one works in $\mathcal{E}^{(2)}_{\rm Ham}$. This has been an obstacle in extending the KAM theory to higher spatial dimensions. In the latter case it is convenient to choose as $\mathcal{E}$ a slightly more involved class of vector field, the so called *quasi-Töplitz* or *Töplitz-Lipschitz* vector fields. The idea is to retain some information on the original PDE structure by showing that the linearized operator in the $w-$component can be "approximated" by piecewise multiplication operators.

A very good idea is to follow the approach of [39],[40] where the Authors take advantage of the Second Mel'nikov conditions (4.41) but do not apply the changes of variables which diagonalize the linearized operator. In this way at each step they preserve the PDE structure. The key observation is that, in a Sobolev regularity setting, the bounds (3.4) and (3.3) follow from corresponding bounds in some special coordinate system (i.e. the one in which the operator is diagonal) provided that the change of variables to the new coordinates is well-defined as an operator from the phase space to itself. Then there is no need to actually apply the change of variables. In the analytic context however this approach presents some difficulties, as one can see easily already in the case of the torus diffeomorphism, i.e. in studying the conjugate of a vector field F under the map

$$\mathbb{T}^d_s \ni \theta \mapsto \theta + g(\theta), \tag{4.47}$$

for some $s \geq 0$, and $g(\theta)$ small. Note that this change of variables is necessary in order to pass from (4.12) to (4.15). The map (4.47) induces an operator on the functions $f \in H^p(\mathbb{T}^d_s)$ defined as $(\mathcal{T}f)(\theta) = f(\theta + g(\theta))$. It is easy to check that in the Sobolev context, i.e. $s = 0$, the map (4.47) is a diffeomorphism of $\mathbb{T}^d$ into itself and hence $\mathcal{T}$ is well-defined from $H^p(\mathbb{T}^d)$ into itself. On the contrary, if $s > 0$ one has that the map (4.47) maps $\mathbb{T}^d_s$ into $\mathbb{T}^d_{s'}$ with $s' < s$, in other words there is a loss of analyticity tied to the size of g. In this case one cannot follow the strategy used in [40]. By following directly such strategy one loses all the analyticity after a finite number of steps. This is due to the fact that, even if the iterative Nash-Moser scheme is coordinate independent, some of the estimates we perform actually depend on the system of coordinates. The basic idea of our approach is in fact to use at each step the system of coordinates which is more adapted to the problem one is studying. This point explains also the rôle of the compatible transformations $\mathcal{L}_n$. Indeed such transformation are not fundamental in proving the convergence of the scheme, but are introduced as a degree of freedom in order to study the set of good parameters in 3.4. In particular such transformations are the key point in order to study problems in the analytic setting.

Up to now we have only considered degree decompositions, however one can certainly consider cases in which $\mathcal{N}, \mathcal{X}, \mathcal{R}$ are not triangular. For instance we may consider $\mathcal{E}$ as in Example 4.2 with the same $\mathcal{X}$ as in (4.16) (i.e. it contains only terms of degree ≤ 0 w.r.t. the decomposition with $\deg(y) = 2$) but where $\mathcal{R}$ contains only (and all) functions of degree ≥ 3. Now this choice respects Definitions 2.17 and 2.19 but clearly the decomposition is not triangular since $\Pi_{\mathcal{X}}\mathrm{ad}N$ is not block diagonal. However it is easily seen that $\Pi_{\mathcal{X}}\mathrm{ad}(R) = 0$ for any $R \in \mathcal{R}$ so that $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{X}}^{\perp}F) = \Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}}F)$. In fact if one divides

$$\mathcal{N} = \bigoplus_{j=0}^{2} \mathcal{N}_j$$

in terms of increasing degree then $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}_0}F)$ is block diagonal on $\mathcal{X} = \oplus \mathcal{X}_j$ while $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}_0}^{\perp}F)$ is upper triangular, so that solving the homological equation only depends on inverting $\Pi_{\mathcal{X}}\mathrm{ad}(\Pi_{\mathcal{N}_0}F)$ as in the previous example. This means that we can apply Theorem2.25 having a pretty explicit description of the set $\mathcal{O}_\infty$.

On the smallness condition Let us comment smallness conditions in Constraint 2.21. First of all one can note that conditions (2.42) are trivially non empty. Indeed given any $\mu, \nu, \kappa_3, \alpha, \mathfrak{p}_1, \chi$, (which typically are fixed by the problem) then (2.42a),(2.42b) and (2.42c) give lower bounds one κ_1, κ_2, η, while (2.42d) and (2.42e) constrain $\mathfrak{p}_2$ in a non-empty interval.

The conditions on $\varepsilon_0, \mathtt{G}_0, \mathtt{R}_0$ are more subtle. Indeed (2.43) gives a lower bound on the size of K_0, but then it is not trivial to show that (2.44) can be fulfilled. Clearly if $\mathtt{G}_0 \sim \mathtt{R}_0 \sim \varepsilon_0$ all the conditions reduce to a smallness condition on ε_0 in terms of K_0. If $\mathtt{G}_0$ or $\mathtt{R}_0$ are large the problem gets more complicated, see also Remark 2.22. Unfortunately this situation appears in many applications, for this reason we have not made any simplifying assumption on the relative sizes.

As one sees in (2.54), the parameters $\mathtt{G}_0, \mathtt{R}_0, \varepsilon_0$ give upper bounds on the size of $G_0, \Pi_{\mathcal{N}^\perp}G_0, \Pi_{\mathcal{X}}G_0$ w.r.t. the parameter γ_0. In applications G_0, γ_0 are essentially given so that the only hope of modulating the bounds comes from modifying the domain $\mathbb{T}^d_s \times D_{a,p}(r) \times \mathcal{O}_0$. To this purpose we first make the trivial remark that if one shrinks to a suitably small neighborhood of zero then polynomials of high degree become very small. In order to formalize this fact we define a scaling degree as follows.

$$\mathtt{s}(\theta) = 0\,, \quad \mathtt{s}(y) = \mathtt{s}\,, \quad \mathtt{s}(w) = 1.$$

Then given a monomial vector field this fixes the scaling as

$$\mathtt{s}(y^j e^{\mathrm{i}\theta\cdot\ell} w^\alpha \partial_{\mathtt{v}}) = \mathtt{s}j + |\alpha| - \mathtt{s}(\mathtt{v}).$$

By construction the scaling is additive w.r.t. commutators and behaves just as the degree in Remark 3.2. We have the following result.

Lemma 4.20. *Consider a tame vector field F as in Definition 2.13 of minimal scaling $\bar{\mathtt{s}}$. Consider the rescaling $r_0 \rightsquigarrow \delta r_0$, $r \rightsquigarrow \delta r$. Then one has*

$$C_{\vec{v}_1,p}(F) \leq \delta^{\bar{\mathtt{s}}} C_{\vec{v},p}(F), \tag{4.48}$$

with $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ and $\vec{v}_1 = (\gamma, \mathcal{O}, s, a, \delta r)$.

This definition of scaling induces a natural scaling decomposition of a vector field. In particular we remark that by construction $\mathcal{N}$ contains necessarily some terms of scaling zero, while $\mathcal{X}$ contains terms of negative scaling. However one can fix the scaling so that $\mathcal{R}$ contains only terms of positive scaling. Hence by Lemma 4.20 terms of positive scaling can be made small. In conclusion given the constants $\bar{\mathtt{G}}_0, \bar{\varepsilon}_0, \bar{\mathtt{R}}_0$ in a ball $\bar{r}_0$, one can try to fulfill 2.44 by rescaling $r_0 = \delta \bar{r_0}$. in this way $\mathtt{R}_0$ becomes smaller, $\mathtt{G}_0$ at best does not grow, while ε_0 necessarily grows. This procedure produces upper and lower bounds on δ.

Consider the decomposition in (4.16) where the degree is equal to the scaling. Then $\mathcal{N}$ contains only terms of degree zero and hence $\mathtt{G}_0$ is scaling invariant. $\mathcal{R}$ contains only terms of positive degree so that rescaling $\mathtt{R}_0 \rightsquigarrow \delta \bar{\mathtt{R}}_0$ at least. $\mathcal{X}$ has negative degree ≥ -2, hence $\varepsilon_0 \rightsquigarrow \delta^{-2} \bar{\varepsilon}_0$. As explained in Remark 2.22 constants $\bar{\mathtt{G}_0}, \bar{\varepsilon}_0, \bar{\mathtt{R}}_0$ are expected to depend only on γ. Using the rescaling we have introduced the extra parameter δ, which should be chosen in terms of γ in order to fulfill the conditions (2.44). Actually it can be useful to make a finer analysis for $\bar{\varepsilon}_0$. Indeed one can bound separately the terms of degree $-2, -1, 0$ in $\bar{\varepsilon}_0$. Let us denote them by $\bar{\varepsilon}_0^{(i)}$, $i = -2, -1, 0$. Then $\varepsilon_0 \rightsquigarrow \sum_{i=-2}^{0} \delta^i \bar{\varepsilon}_0^{(i)}$, and hence the smallness conditions can be taken asymetrically. The same holds for $\mathtt{R}_0$. This has been discussed with slightly different notation in [48].

5 Proof of the result

We divide the proof of Theorem 2.25 into two pieces: we first show one step in full details (this is the same in both cases) and then we prove that it is indeed possible to perform infinitely many steps and that the procedure converges.

5.1 The KAM step

Proposition 5.1. *Let γ_0, a_0, r_0, s_0, ε_0 and K_0 be the constants appearing in Theorem 2.25. Fix $\gamma, a, r, s \geq 0$ so that*

$$\frac{\gamma_0}{2} \leq \gamma \leq \gamma_0, \ \frac{a_0}{2} \leq a \leq a_0, \ \frac{s_0}{2} \leq s \leq s_0, \ \frac{r_0}{2} \leq r \leq r_0 \tag{5.1}$$

and $0 < \rho < 1$ such that

$$r - 8\rho r_0 > \frac{r_0}{2}, \ \text{ if } s_0 \neq 0 \text{ then } s - 8\rho s_0 > \frac{s_0}{2}, \ \text{if } a_0 \neq 0 \text{ then } \ a - 8\rho a_0 > \frac{a_0}{2}. \tag{5.2}$$

Consider a vector field:

$$F : \mathbb{T}^d_s \times D_{a,p+\nu}(r) \times \mathcal{O}_0 \to V_{a,p}, \tag{5.3}$$

which is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. Let N_0 be the diagonal vector field appearing in Theorem 2.25 and write $F = N_0 + (F - N_0) = N_0 + G$. Set $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$, denote by $\mathcal{O} \subseteq \mathcal{O}_0$ some set of parameters ξ for which $F(\xi) \in \mathcal{E}$ and $|\Pi_{\mathcal{X}} G|_{\vec{v}, \mathfrak{p}_2 - 1} \leq \mathtt{C} C_{\vec{v}, \mathfrak{p}_2}(\Pi^{\perp}_{\mathcal{N}} G)$ and define also

$$\Gamma_p := \gamma^{-1} C_{\vec{v},p}(G), \quad \Theta_p := \gamma^{-1} C_{\vec{v},p}(\Pi^{\perp}_{\mathcal{N}} G), \quad \delta := \gamma^{-1} |\Pi_{\mathcal{X}} G|_{\vec{v}, \mathfrak{p}_1}. \tag{5.4}$$

Fix $K > K_0$ and assume that

$$\rho^{-1} K^{\mu+\nu+3} \Gamma_{\mathfrak{p}_1} \delta \leq \epsilon \tag{5.5}$$

with $\epsilon = \epsilon(\mathfrak{p}_1, d)$ small enough. Consider a map $\mathcal{L}$ compatible with $(F, K, \vec{v}, \rho)$ (see Definition 2.24) and set

$$\hat{F} = N_0 + \hat{G} := (\mathcal{L})_* F \, .$$

Consider any $\mathcal{O}_+ \subseteq \mathcal{O}$ solving the homological equation for $(\hat{F}, K, \vec{v}^o_2, \rho)$ and set

$$\begin{aligned} \vec{w}_i &= (\gamma, \mathcal{O}_+, s - i\rho s_0, a - i\rho a_0, r - i\rho r_0), \\ \vec{v}^o_i &= (\gamma, \mathcal{O}_0, s - i\rho s_0, a - i\rho a_0, r - i\rho r_0), \end{aligned} \tag{5.6}$$

for $i = 1, \ldots, 8$. The following holds:

(i) there exists an invertible (see Def. 2.8) change of variables

$$\Phi_+ := \mathbb{1} + f_+ : \mathbb{T}^d_{s-4\rho s_0} \times D_{a-4\rho a_0,p}(r - 4\rho r_0) \times \mathcal{O}_0 \longrightarrow \mathbb{T}^d_{s-2\rho s_0} \times D_{a-2\rho a_0,p}(r - 2\rho r_0), \tag{5.7}$$

with f_+ a regular vector field (see Def. 4.5) such that Φ_+ is $\mathcal{E}$ preserving for all $\xi \in \mathcal{O}_+$ and is generated by a vector field g_+ which satisfies the bound

$$\begin{aligned} &|g_+|_{\vec{w}_2,\mathfrak{p}_1} \le \mathtt{C}\Gamma_{\mathfrak{p}_1} K^{\mu}\delta \\ &|g_+|_{\vec{w}_2,\mathfrak{p}_2} \le \mathtt{C}K^{\mu+1}(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1}))\,; \end{aligned} \tag{5.8}$$

(ii) fix s_+, a_+, r_+ as

$$s_+ = s - 8\rho s_0, \quad a_+ = a - 8\rho a_0, \quad r_+ = r - 8\rho r_0 \tag{5.9}$$

then

$$F_+ := (\Phi_+)_*\hat{F} = N_0 + G_+ : \mathbb{T}^d_{s_+} \times D_{a_+,p+\nu}(r_+) \times \mathcal{O}_0 \to V_{a,p}\,; \tag{5.10}$$

(iii) setting $\gamma_0/2 \le \gamma_+ \le \gamma$, $\vec{v}_+ := (\gamma_+, \mathcal{O}_+, a_+, s_+) = \vec{w}_8$, F_+ is tame and denoting the tameness constants as

$$\Gamma_{+,p} := \gamma_+^{-1} C_{\vec{v}_+,p}(G_+), \quad \Theta_{+,p} := \gamma_+^{-1} C_{\vec{v}_+,p}(\Pi^{\perp}_{\mathcal{N}} G_+), \quad \delta_+ := \gamma_+^{-1}|\Pi_{\mathcal{X}} G_+|_{\vec{v}_+,\mathfrak{p}_1}$$

one can fix

$$\begin{aligned} \Gamma_{+,\mathfrak{p}_1} &= \frac{\gamma}{\gamma_+}(1 + \varepsilon_0 K^{-1})\Gamma_{\mathfrak{p}_1} + \mathtt{C}K^{\mu}\Gamma_{\mathfrak{p}_1}\delta(K^{\nu+1}\Gamma_{\mathfrak{p}_1} + \varepsilon_0 K^{-\eta})\,, \\ \Gamma_{+,\mathfrak{p}_2} &= \mathtt{C}\Big(\Gamma_{\mathfrak{p}_2} + \varepsilon_0\Gamma_{\mathfrak{p}_1}K^{\kappa_3} + \Gamma_{\mathfrak{p}_1}K^{\mu+\nu+2}(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\alpha(p-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1}))\Big)\,, \end{aligned} \tag{5.11}$$

$$\begin{aligned} \Theta_{+,\mathfrak{p}_1} &= \frac{\gamma}{\gamma_+}(1 + \varepsilon_0 K^{-1})\Theta_{\mathfrak{p}_1} + \mathtt{C}K^{\mu}\Gamma_{\mathfrak{p}_1}\delta(K^{\nu+1}\Gamma_{\mathfrak{p}_1} + \varepsilon_0 K^{-\eta})\,, \\ \Theta_{+,\mathfrak{p}_2} &= \mathtt{C}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\mu+\nu+2}\Gamma_{\mathfrak{p}_1}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\big)\big)\,, \\ &\gamma_+^{-1}|\Pi_{\mathcal{X}}(G_+)|_{\vec{v}_+,\mathfrak{p}_2-1} \le \Theta_{+,\mathfrak{p}_2} \end{aligned} \tag{5.12}$$

and

$$\begin{aligned} \delta_+ \le \mathtt{C}\Gamma_{\mathfrak{p}_1}&\big(\delta^2\Gamma^2_{\mathfrak{p}_1}K^{2\mu+2\nu+4} + \delta\varepsilon_0 K^{\mu-\eta}\big) + K^{\mu+\nu+2-(\mathfrak{p}_2-\mathfrak{p}_1)}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}\big) \\ &+ \Gamma_{\mathfrak{p}_1}K^{\mu+\nu+2-(\mathfrak{p}_2-\mathfrak{p}_1)}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\big)\,. \end{aligned} \tag{5.13}$$

Proof. First of all we note that by the definition of $\mathcal{L}$ one has

$$\hat{F} : \mathbb{T}^d_{s-2\rho s_0} \times D_{a-2\rho a_0,p+\nu}(r - 2\rho r_0) \times \mathcal{O}_0 \to V_{a,p}$$

and $\hat{F} \in \mathcal{E}$ for each ξ in $\mathcal{O}$. By (2.51) we have

$$\begin{aligned} \Pi_{\mathcal{X}}\hat{G} &= \Pi_{\mathcal{X}}\hat{F} = \Pi_{\mathcal{X}}(\mathcal{L})_*F = \Pi_{\mathcal{X}}(\mathcal{L})_*\Pi_{\mathcal{X}}F = \Pi_{\mathcal{X}}(\mathcal{L})_*\Pi_{\mathcal{X}}G\,, \\ \Pi^{\perp}_{\mathcal{N}}\hat{G} &= \Pi^{\perp}_{\mathcal{N}}\hat{F} = \Pi^{\perp}_{\mathcal{N}}(\mathcal{L})_*F = \Pi^{\perp}_{\mathcal{N}}(\mathcal{L})_*\Pi^{\perp}_{\mathcal{N}}F = \Pi^{\perp}_{\mathcal{N}}(\mathcal{L})_*\Pi^{\perp}_{\mathcal{N}}G\,. \end{aligned}$$

Now since G is $C^{\mathtt{n}}$-tame and $\mathcal{N}, \mathcal{X}$ have maximal degree $\le \mathtt{n}$ we have that the tameness constants of $\Pi_{\mathcal{N}}G, \Pi_{\mathcal{R}}G$ as well as $|\Pi_{\mathcal{X}}G|$ are controlled by the tameness constant of G .

By (2.50) we have the bounds

$$\begin{aligned} C_{\vec{w}_2,\mathfrak{p}_1}(\hat{G}) &\le C_{\vec{v},\mathfrak{p}_1}(G)(1 + \varepsilon_0 K^{-1}) \le \gamma\Gamma_{\mathfrak{p}_1}(1 + \varepsilon_0 K^{-1})\,, \\ C_{\vec{w}_2,\mathfrak{p}_2}(\hat{G}) &\le C_{\vec{v},\mathfrak{p}_2}(G) + \varepsilon_0 K^{\kappa_3} C_{\vec{v},\mathfrak{p}_1}(G) \le \gamma(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\,, \end{aligned} \tag{5.14}$$

$$
\begin{aligned}
&C_{\vec{w}_2,\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}\hat{G}) \le C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}G)(1+\varepsilon_0 K^{-1}) \le \gamma\Theta_{\mathfrak{p}_1}(1+\varepsilon_0 K^{-1})\,,\\
&C_{\vec{w}_2,\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp}\hat{G}) \le C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp}G) + \varepsilon_0 K^{\kappa_3} C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}G) \le \gamma(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1})\,,
\end{aligned}
\tag{5.15}
$$

and

$$
\begin{aligned}
&C_{\vec{w}_2,\mathfrak{p}_1}(\Pi_{\mathcal{X}}\hat{G}) \le C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{X}}G)(1+\varepsilon_0 K^{-1}) \le 2\gamma\delta\,,\\
&C_{\vec{w}_2,\mathfrak{p}_2}(\Pi_{\mathcal{X}}\hat{G}) \le C_{\vec{v},\mathfrak{p}_2}(\Pi_{\mathcal{X}}G) + \varepsilon_0 K^{\kappa_3} C_{\vec{v},\mathfrak{p}_1}(\Pi_{\mathcal{X}}G) \le \gamma(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1})\,,\\
&|\Pi_{\mathcal{X}}\hat{G}|_{\vec{w}_2,\mathfrak{p}_2-1} \le \gamma(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}).
\end{aligned}
\tag{5.16}
$$

Our aim is to define for $\xi \in \mathcal{O}_+$ a vector field g_+ as the "approximate" solution of the equation

$$
\Pi_K \Pi_{\mathcal{X}}[g_+, \Pi_{\mathcal{X}}^{\perp}\hat{F}] = \Pi_K\Pi_{\mathcal{X}}\hat{F}. \tag{5.17}
$$

By Definition, if $\xi \in \mathcal{O}_+$ then we can find g_+ satisfying properties (a),(b),(c),(d) of Definition 2.23.

By (2.46) one gets

$$
|g_+|_{\vec{w}_2,p} \le \mathtt{C}\gamma^{-1}K^{\mu}(|\Pi_K\Pi_{\mathcal{X}}\hat{G}|_{\vec{w}_2,p} + K^{\alpha(p-\mathfrak{p}_1)}|\Pi_K\Pi_{\mathcal{X}}\hat{G}|_{\vec{w}_2,\mathfrak{p}_1}\gamma^{-1}C_{\vec{w}_2,p}(\hat{G})) \tag{5.18}
$$

and hence, using (5.14),(5.16) and (2.32) we have (5.8) Moreover, by condition (a) of Definition 2.23, one has $g_+ \in \mathcal{B}_{\mathcal{E}}$ for $\xi \in \mathcal{O}_+$ and $|g_+|_{\vec{v}_2^0,\mathfrak{p}_1} \le \mathtt{C}|g_+|_{\vec{w}_2,\mathfrak{p}_1}$.

Now, if ϵ in (5.5) is small enough, by Definition 2.18 item 5, g_+ generates a change of variables $\Phi_+ = \mathbb{1} + f_+$and $|f_+|_{\vec{v}_3^0,p} \le 2|\tilde{g}_+|_{\vec{v}_2^0,p}$. Finally, for possibly smaller ϵ one has (5.7), by using Definition 2.18 item 4. Note that the smallness conditions on ϵ come only from this two conditions.

First we note that, since N_0 is diagonal (recall Definition 2.20), one has

$$
\begin{aligned}
G_+ &:= (\Phi_+)_* N_0 + (\Phi_+)_*\hat{G} - N_0 = \int_0^1 dt\,(\Phi_+)_*^t[g_+, N_0] + (\Phi_+)_*\hat{G}\\
&= \int_0^1 dt(\Phi_+)_*^t \Pi_K\Pi_{\mathcal{X}}[g_+, N_0 + \Pi_{\mathcal{X}}^{\perp}\hat{G}] - \int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}] + (\Phi_+)_*\hat{G}\\
&= \int_0^1 dt(\Phi_+)_*^t(\Pi_K\Pi_{\mathcal{X}}\hat{G} + u) - \int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g,\Pi_{\mathcal{X}}^{\perp}\hat{G}] + (\Phi_+)_*\hat{G}
\end{aligned}
\tag{5.19}
$$

where

$$
u := \Pi_K\Pi_{\mathcal{X}}[g_+, N_0 + \Pi_{\mathcal{X}}^{\perp}\hat{G}] - \Pi_K\Pi_{\mathcal{X}}\hat{G} \tag{5.20}
$$

and u in (5.20) satisfies (2.47), so by applying (5.14) and (5.16), we get

$$
\begin{aligned}
&|u|_{\vec{w}_2,\mathfrak{p}_1} \le \mathtt{C}\gamma\varepsilon_0\Gamma_{\mathfrak{p}_1}K^{-\eta+\mu}\delta\\
&|u|_{\vec{w}_2,\mathfrak{p}_2} \le \mathtt{C}\gamma K^{\mu+1}\Big((\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1})\Gamma_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\Big)\,;
\end{aligned}
\tag{5.21}
$$

Regarding the first summand of (5.19), using Lemma B.3, we have

$$
C_{\vec{w}_8,\mathfrak{p}_1}\left(\int_0^1 dt\Phi_*^t(\Pi_K\Pi_{\mathcal{X}}\hat{G} + u)\right) \overset{(5.21),(5.5)}{\le} \gamma\mathtt{C}\delta(1+\gamma\varepsilon_0\Gamma_{\mathfrak{p}_1}K^{-\eta+\mu}) \tag{5.22}
$$

Moreover

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_2}\left(\int_0^1 dt\Phi_*^t(\Pi_K\Pi_{\mathcal{X}}\hat{G} + u)\right) &\le (1+\rho)(C_{\vec{w}_6,\mathfrak{p}_2}(\Pi_{\mathcal{X}}\hat{G} + u) + C_{\vec{w}_6,\mathfrak{p}_1}(\Pi_{\mathcal{X}}\hat{G}+u)|f_+|_{\vec{w}_7,\mathfrak{p}_2+\nu+1})\\
&\overset{(2.47),(2.42),(5.5)}{\le} \mathtt{C}\gamma K^{\mu+1}\Big((\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1})\Gamma_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\Big)\,.
\end{aligned}
\tag{5.23}
$$

The second term of (5.19) can be estimated as follows. First we note that, since $\hat{G}$ is $C^{\mathtt{n}+2}$-tame up to order[11] $q+1$, then the vector field $\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]$ is $C^{\mathtt{n}+1}$-tame up to order q. This implies, due to the presence of the projection onto $\mathcal{X}$, that it is indeed C^k-tame for all k. The tameness constant is given by

$$\begin{aligned}C_{\vec{w}_8,p}(\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]) &\overset{(2.34)}{\leq} KC_{\vec{w}_8,p-1}(\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]) \\ &\leq \mathtt{C}K^{\nu+1}(|g_+|_{\vec{w}_8,p}C_{\vec{w}_8,\mathtt{p}_0}(\Pi_{\mathcal{X}}^{\perp}\hat{G})+|g_+|_{\vec{w}_8,\mathtt{p}_0}C_{\vec{w}_8,p}(\Pi_{\mathcal{X}}^{\perp}\hat{G})) \\ &\overset{rmkB.2}{\leq} \mathtt{C}K^{\nu+1}(|g_+|_{\vec{w}_8,p}C_{\vec{w}_8,\mathtt{p}_0}(\hat{G})+|g_+|_{\vec{w}_8,\mathtt{p}_0}C_{\vec{w}_8,p}(\hat{G})),\end{aligned} \tag{5.24}$$

hence, by (5.14) and (5.8), we have

$$C_{\vec{v}_+,\mathtt{p}_1}(\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]) \leq \gamma\mathtt{C}\Gamma_{\mathtt{p}_1}^2K^{\mu+\nu+1}\delta\,, \tag{5.25}$$

$$C_{\vec{v}_+,\mathtt{p}_2}(\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}])\leq\mathtt{C}\gamma K^{\mu+\nu+2}\Gamma_{\mathtt{p}_1}\Big((\Theta_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Theta_{\mathtt{p}_1}+K^{\alpha(\mathtt{p}_2-\mathtt{p}_1)}\delta(\Gamma_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Gamma_{\mathtt{p}_1})\Big)\,. \tag{5.26}$$

Therefore using Lemma B.3 we have

$$C_{\vec{w}_8,\mathtt{p}_1}(\int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]) \leq \gamma\mathtt{C}\Gamma_{\mathtt{p}_1}^2K^{\mu+\nu+1}\delta \tag{5.27}$$

and

$$\begin{aligned}C_{\vec{w}_8,\mathtt{p}_2}(\int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]) \leq \gamma\mathtt{C}K^{\mu+\nu+2}\Gamma_{\mathtt{p}_1}\Big(&\Theta_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Theta_{\mathtt{p}_1} \\ &+K^{\alpha(\mathtt{p}_2-\mathtt{p}_1)}\delta(\Gamma_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Gamma_{\mathtt{p}_1})\Big)\end{aligned} \tag{5.28}$$

by (5.5).

Finally, again by Lemma B.3, we estimate the third summand in (5.19) as

$$C_{\vec{w}_8,\mathtt{p}_1}(\Phi_*(\hat{G})) \overset{(5.14),(5.8)}{\leq} \gamma\Gamma_{\mathtt{p}_1}(1+\mathtt{C}\delta\Gamma_{\mathtt{p}_1}K^{\mu})(1+\varepsilon_0K^{-1}) \tag{5.29}$$

and

$$\begin{aligned}C_{\vec{w}_8,\mathtt{p}_2}((\Phi_+)_*\hat{G}) &\leq (1+\mathtt{c}^{-1}|f_+|_{\vec{w}_7,\mathtt{p}_1})\Big(C_{\vec{w}_6,\mathtt{p}_2}(\hat{G})+C_{\vec{w}_6,\mathtt{p}_1}(\hat{G})|f_+|_{\vec{w}_7,\mathtt{p}_2+\nu+1}\Big) \\ &\overset{(5.14),(5.8)}{\leq} \gamma\mathtt{C}\Big(\Gamma_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Gamma_{\mathtt{p}_1} \\ &\qquad +\Gamma_{\mathtt{p}_1}K^{\mu+\nu+2}(\Theta_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Theta_{\mathtt{p}_1}+K^{\alpha(p-\mathtt{p}_1)}\delta(\Gamma_{\mathtt{p}_2}+\varepsilon_0K^{\kappa_3}\Gamma_{\mathtt{p}_1}))\Big)\,.\end{aligned} \tag{5.30}$$

The bounds (5.11) follow by collecting together (5.22), (5.23), (5.27),(5.29) and (5.30).

Let us study $\Theta_{+,p}$. First of all we see that

$$\Pi_{\mathcal{N}}^{\perp}G_+=\Pi_{\mathcal{N}}^{\perp}F_+=\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*\hat{F}=\Pi_{\mathcal{N}}^{\perp}\Big((\Phi_+)_*N_0+(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})+(\Phi_+)_*(\Pi_{\mathcal{N}}^{\perp}\hat{G})\Big)\,. \tag{5.31}$$

In order to estimate the first term $\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*N_0$, we first note that

$$\begin{aligned}\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*N_0 &= \Pi_{\mathcal{N}}^{\perp}((\Phi_+)_*N_0-N_0)=\Pi_{\mathcal{N}}^{\perp}\int_0^1 dt\,(\Phi_+^t)_*[g_+,N_0]=\Pi_{\mathcal{N}}^{\perp}\int_0^1 dt\,(\Phi_+^t)_*(\Pi_K\Pi_{\mathcal{X}}[g_+,N_0]) \\ &= \Pi_{\mathcal{N}}^{\perp}\Big(\int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g_+,N_0+\Pi_{\mathcal{X}}^{\perp}\hat{G}]-\int_0^1 dt(\Phi_+)_*^t\Pi_K\Pi_{\mathcal{X}}[g_+,\Pi_{\mathcal{X}}^{\perp}\hat{G}]\Big)\,,\end{aligned} \tag{5.32}$$

[11] due to the truncation Π_K

substituting (5.20) and using Remark B.2 in order to remove the projection we have

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_* N_0) &\overset{(5.22),(5.27)}{\le} \mathtt{C}\gamma K^{\mu}\Gamma_{\mathfrak{p}_1}\delta(K^{\nu+1}\Gamma_{\mathfrak{p}_1}+\varepsilon_0 K^{-\eta}),\\
C_{\vec{w}_8,\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_* N_0) &\overset{(5.23),(5.27)}{\le} \mathtt{C}\gamma K^{\mu+\nu+2}\Gamma_{\mathfrak{p}_1}\Big(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}+K^{\alpha(p-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\Big).
\end{aligned}
\tag{5.33}
$$

In order to bound the second summand we use Lemma B.6 with $\mathcal{U}=\mathcal{N}$ and obtain

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})) &\le \mathtt{C}|f_+|_{\vec{w}_6,\mathfrak{p}_1+\nu+1}C_{\vec{w}_6,\mathfrak{p}_1}(\Pi_{\mathcal{N}}\hat{G}) \overset{rmkB.2}{\le} \mathtt{C}|f_+|_{\vec{w}_6,\mathfrak{p}_1+\nu+1}C_{\vec{w}_6,\mathfrak{p}_1}(\hat{G})\\
&\overset{(5.14)}{\le} \gamma\mathtt{C}K^{\mu+\nu+1}\delta\Gamma_{\mathfrak{p}_1}^2,
\end{aligned}
\tag{5.34}
$$

and

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}\hat{G})) &\le \mathtt{C}\Big(C_{\vec{w}_2,\mathfrak{p}_2}(\hat{G})|f_+|_{\vec{w}_6,\mathfrak{p}_1}+C_{\vec{w}_2,\mathfrak{p}_1}(\hat{G})|f_+|_{\vec{w}_6,\mathfrak{p}_2}\Big)\\
&\overset{(5.14)}{\le} \gamma\mathtt{C}\Gamma_{\mathfrak{p}_1}K^{\mu+1}(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}+K^{\alpha(p-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1}))
\end{aligned}
\tag{5.35}
$$

Regarding the third summand, using Remark B.2, Lemma B.3 and (5.15) we obtain

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_1}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}^{\perp}\hat{G})) &\overset{(5.8)}{\le} \gamma(1+\mathtt{C}\Gamma_{\mathfrak{p}_1}K^{\mu}\delta)(1+\varepsilon_0 K^{-1})\Theta_{\mathfrak{p}_1},\\
C_{\vec{w}_8,\mathfrak{p}_2}(\Pi_{\mathcal{N}}^{\perp}(\Phi_+)_*(\Pi_{\mathcal{N}}^{\perp}\hat{G})) &\overset{(5.8)}{\le} \mathtt{C}\gamma(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1})+\\
&\quad+\mathtt{C}\gamma K^{\mu+\nu+2}\Theta_{\mathfrak{p}_1}\big(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}+K^{\alpha(p-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\big).
\end{aligned}
\tag{5.36}
$$

By collecting together (5.33), (5.34), (5.35) and (5.36) we obtain the first two lines of (5.12). In order to prove the last of (5.12) we use item (d) of Definition 2.23. Indeed we substitute (5.8) ,(5.14) and (5.15) in (2.48) and we obtain the desired bound.

In order to prove the bound for δ_+ we first write

$$
\Pi_{\mathcal{X}}G_+ = \Pi_{\mathcal{X}}(\hat{F}+[\hat{F},g_+]+Q), \qquad Q:=(\Phi_+)_*\hat{F}-(\hat{F}+[\hat{F},g_+])
\tag{5.37}
$$

and hence

$$
\begin{aligned}
\Pi_{\mathcal{X}}G_+ &= \Pi_{\mathcal{X}}\hat{F}+\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}^{\perp}\hat{F},g_+]+\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{F},g_+]+\Pi_{\mathcal{X}}Q\\
&= u+\Pi_K^{\perp}\Big(\Pi_{\mathcal{X}}\hat{G}+\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}^{\perp}\hat{G},g_+]+\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{G},g_+]\Big)+\Pi_K\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{G},g_+]+\Pi_{\mathcal{X}}Q\\
&= u+\Pi_K^{\perp}\Big(\Pi_{\mathcal{X}}\hat{G}+\Pi_{\mathcal{X}}[\hat{G},g_+]\Big)+\Pi_K\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{G},g_+]+\Pi_{\mathcal{X}}Q
\end{aligned}
\tag{5.38}
$$

where u is defined in (5.20) and bounded in (5.21).

The second summand in (5.38) can be bounded as

$$
\begin{aligned}
|\Pi_K^{\perp}\Big(\Pi_{\mathcal{X}}\hat{G}+\Pi_{\mathcal{X}}[\hat{G},g_+]\Big)|_{\vec{w}_8,\mathfrak{p}_1} &\overset{(2.33)}{\le} K^{-(\mathfrak{p}_2-\mathfrak{p}_1)+2}|\Pi_{\mathcal{X}}\hat{G}+\Pi_{\mathcal{X}}[\hat{G},g_+])|_{\vec{w}_6,\mathfrak{p}_2-2}\\
&\overset{(5.16),(2.46)}{\le} \mathtt{C}\gamma K^{-(\mathfrak{p}_2-\mathfrak{p}_1)+\nu+\mu+2}\Gamma_{\mathfrak{p}_1}(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}+K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1}))\\
&\quad+\mathtt{C}\gamma K^{-(\mathfrak{p}_2-\mathfrak{p}_1)+2}(\Theta_{\mathfrak{p}_2}+\varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}).
\end{aligned}
\tag{5.39}
$$

We can choose the tameness constant of the third summand in (5.38) as follows

$$
\begin{aligned}
C_{\vec{w}_8,\mathfrak{p}_1}(\Pi_K\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{F},g_+]) &\overset{(2.34)}{\le} KC_{\vec{w}_8,\mathfrak{p}_1-1}(\Pi_K\Pi_{\mathcal{X}}[\Pi_{\mathcal{X}}\hat{F},g_+]) \overset{(B.1)}{\le} \mathtt{C}K^{\nu+1}C_{\vec{w}_8,\mathfrak{p}_1}(\Pi_{\mathcal{X}}\hat{F})|g_+|_{\vec{w}_8,\mathfrak{p}_1}\\
&\overset{(5.16),(5.8)}{\le} \gamma\mathtt{C}K^{\mu+\nu+1}\Gamma_{\mathfrak{p}_1}\delta^2.
\end{aligned}
\tag{5.40}
$$

Finally we deal with the last summand in (5.38) as follows. Using Remark B.5 and Definition 2.20 one can reason as in (5.19) and write

$$\Pi_{\mathcal{X}} Q = \Pi_{\mathcal{X}} \left(\int_0^1 dt \int_0^t ds (\Phi_+)_*^s \left(\left[g_+, [g_+, \hat{G}] \right] + \left[g_+, \Pi_K \Pi_{\mathcal{X}} (\hat{F} + u - [g_+, \Pi_{\mathcal{X}}^{\perp} \hat{G}]) \right] \right) \right) . \tag{5.41}$$

Regarding the first summand in (5.41) one uses (B.12b) and obtains

$$\begin{aligned} C_{\vec{w}_8, \mathfrak{p}_1} \left(\Pi_{\mathcal{X}} \int_0^1 ds \int_0^t ds (\Phi_+)_*^s \left(\left[g_+, [g_+, \hat{G}] \right] \right) \right) &\leq C_{\vec{w}_6, \mathfrak{p}_1+2}(\hat{G}) |g_+|_{\vec{w}_6, \mathfrak{p}_1} |g_+|_{\vec{w}_6, \mathfrak{p}_1+\nu+2} \\ &\overset{(5.8)}{\leq} \mathtt{C} K^{\nu+2+2\mu} \Gamma_{\mathfrak{p}_1}^2 \delta^2 \left(C_{\vec{w}_6, \mathfrak{p}_1+2}(\Pi_K \hat{G}) + C_{\vec{w}_6, \mathfrak{p}_1+2}(\Pi_K^{\perp} \hat{G}) \right) \\ &\overset{(5.14)}{\leq} \mathtt{C} \gamma \delta^2 \Gamma_{\mathfrak{p}_1}^2 K^{\nu+2\mu+4} \Big(\Gamma_{\mathfrak{p}_1} + K^{-(\mathfrak{p}_2-\mathfrak{p}_1)} (\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3} \Gamma_{\mathfrak{p}_1}) \Big) . \end{aligned} \tag{5.42}$$

Again we are using the fact that G is $C^{\mathtt{n}+2}$-tame to infer that the double commutator is $C^{\mathtt{n}}$-tame, and then the projection onto $\mathcal{X}$ in order to recover the C^k-tameness for all k. Regarding the second summand in (5.41), since each term is a polynomial in $E^{(K)}$, using Lemmata B.3, B.1-(iii) and the bounds (5.22), (5.25) we obtain

$$\begin{aligned} C_{\vec{w}_8, \mathfrak{p}_1} \Big(\int_0^1 \int_0^t (\Phi_+)_*^s [g_+, \Pi_K \Pi_{\mathcal{X}} (\hat{F} + u - [g_+, \Pi_{\mathcal{X}}^{\perp} \hat{G}])] \Big) &\leq \\ \leq \mathtt{C} |g_+|_{\vec{w}_6, \mathfrak{p}_1+1} C_{\vec{w}_6, \mathfrak{p}_1+\nu+1} (\Pi_K \Pi_{\mathcal{X}} (\hat{F} + u - [g_+, \Pi_{\mathcal{X}}^{\perp} \hat{G}])) & \\ \leq \mathtt{C} \gamma K^{2\mu+2\nu+4} \Gamma_{\mathfrak{p}_1}^3 \delta^2 . & \end{aligned} \tag{5.43}$$

Therefore, collecting together (5.43) and (5.42) we obtain

$$C_{\vec{v}_+, \mathfrak{p}_1} (\Pi_{\mathcal{X}} Q) \leq \mathtt{C} \gamma \delta^2 \Gamma_{\mathfrak{p}_1}^2 K^{2\mu+2\nu+4} \Big(\Gamma_{\mathfrak{p}_1} + K^{-(\mathfrak{p}_2-\mathfrak{p}_1)} (\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3} \Gamma_{\mathfrak{p}_1}) \Big). \tag{5.44}$$

In conclusion one has

$$\begin{aligned} C_{\vec{v}_+, \mathfrak{p}_1} (\Pi_{\mathcal{X}} F_+) \leq \mathtt{C} \gamma \Gamma_{\mathfrak{p}_1} K^{\mu+\nu+1} \Big(& \delta^2 \Gamma_{\mathfrak{p}_1} K^{\mu+\nu+3} \big(\Gamma_{\mathfrak{p}_1} + K^{-(\mathfrak{p}_2-\mathfrak{p}_1)} (\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3} \Gamma_{\mathfrak{p}_1}) \big) + \\ & K^{-(\mathfrak{p}_2-\mathfrak{p}_1)} (\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3} \Theta_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)} \delta (\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3} \Gamma_{\mathfrak{p}_1}) \Big) + \Gamma_{\mathfrak{p}_1} \delta \varepsilon_0 K^{-\eta+\mu} . \end{aligned} \tag{5.45}$$

Recalling that the norm $|\cdot|_{\vec{v}_+, \mathfrak{p}_1}$ is the sharp tameness constant (see (2.31)), then (5.45) implies the bound (5.13) since $\delta \Gamma_{\mathfrak{p}_1} K^{\mu+\nu+3} \leq 1$ by (5.5). ☐

5.2 Proof of Theorem 2.25: iterative scheme.

We now prove Theorem 2.25 by induction on n. The induction basis is trivial with $g_0 = 0$. Assuming (2.56) up to n we prove the inductive step using the "KAM step" of Proposition 5.1. First of all we ensure that

$$\rho_n^{-1} K_n^{\mu+\nu+3} \Gamma_{n, \mathfrak{p}_1} \delta_n \leq \mathtt{c}, \tag{5.46}$$

which, by the inductive hypothesis and (2.52) reads

$$2^{n+9} K_0^{(\mu+\nu+3-\kappa_2)\chi^n} \mathtt{G}_0 \varepsilon_0 K_0^{\kappa_2} \leq \mathtt{c}; \tag{5.47}$$

this is true because by (2.42b) and the fact that K_0 is large enough depending on $\chi, d, \mathfrak{p}_0$, the left hand side (5.47) is decreasing in n so that (5.46) follows form

$$K_0^{\mu+\nu+4} \mathtt{G}_0 \varepsilon_0 < 1$$

which is indeed implied by (2.44a) because $\mathtt{G}_0 \geq \mathtt{R}_0$.

Hence we can apply the "KAM step" to $F_n := (\Phi_n \circ \mathcal{L}_n)_* F_{n-1} \in \mathcal{W}_{\vec{v}^0_n, \mathfrak{p}_2}$ which is a $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We fix $(K_n, \gamma_n, a_n, s_n, r_n, \rho_n, \mathcal{O}_n) \rightsquigarrow (K, \gamma, a, s, r, \rho, \mathcal{O})$, $\Gamma_{n,p} \rightsquigarrow \Gamma_p$, $\Theta_{n,p} \rightsquigarrow \Theta_p$, $\delta_n \rightsquigarrow \delta$, $(\gamma_{n+1}, a_{n+1}, s_{n+1}, r_{n+1}, \rho_{n+1}, \mathcal{O}_{n+1}) \rightsquigarrow (\gamma_+, a_+, s_+, r_+, \rho_+, \mathcal{O}_+)$. The KAM steps produces a bounded regular vector field g_{n+1} and a left invertible change of variables $\Phi_{n+1} = \mathbb{1} + f_{n+1}$ such that $F_{n+1} := (\Phi_{n+1} \circ \mathcal{L}_n)_* F_n \in \mathcal{W}_{\vec{v}^0_{n+1}, \mathfrak{p}_2}$ is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We now verify that the bounds (2.56) hold with $\Gamma_{n+1,p} \rightsquigarrow \Gamma_{+,p}$, $\Theta_{n+1,p} \rightsquigarrow \Theta_{+,p}$, , $\delta_{n+1} \rightsquigarrow \delta_+$.

Let us prove **(i)**. By substituting into (5.8) we immediately obtain the bounds for g_{n+1} of (2.56).

Now we recall that, by definition

$$\frac{\gamma_n}{\gamma_{n+1}} = 1 + \frac{1}{2^{n+3} - 1}.$$

We use (5.11) together with the inductive hypotheses to obtain

$$\Gamma_{n+1,\mathfrak{p}_1} \leq \Big(1 + \frac{1}{2^{n+3} - 1}\Big)\mathtt{G}_n + 2\varepsilon_0 K_n^{-1}\mathtt{G}_0 + \mathtt{C} K_n^{\mu-\kappa_2}\mathtt{G}_0 \varepsilon_0 K_0^{\kappa_2}(K_n^{\nu+1}\mathtt{G}_0 + \varepsilon_0 K_n^{-\eta}) \leq \mathtt{G}_{n+1}\,,$$

which follow by requiring

$$\max(2^n K_n^{-1}\varepsilon_0, 2^n K_n^{\mu+\nu+1-\kappa_2} K_0^{\kappa_2}\mathtt{G}_0\varepsilon_0, 2^n K_n^{-\eta-\kappa_2+\mu} K_0^{\kappa_2}\varepsilon_0) \leq \mathtt{c}\,,$$

and as before this follows by (2.42c) and (2.44a).

Regarding $\Theta_{n+1,\mathfrak{p}_1}$, using (5.12) we get

$$\Theta_{n+1,\mathfrak{p}_1} \leq \Big(1 + \frac{1}{2^{n+3} - 1}\Big)\mathtt{R}_n + 2\varepsilon_0 K_n^{-1}\mathtt{R}_0 + \mathtt{C} K_n^{\mu+\nu+1-\kappa_2}\mathtt{G}_0^2\varepsilon_0 K_0^{\kappa_2} + K_n^{-\eta-\kappa_2+\mu}\mathtt{G}_0\varepsilon_0^2 K_0^{\kappa_2} \leq \mathtt{R}_{n+1}.$$

which again follows from (2.42c) and (2.44a).

For $\delta_{n+1} \rightsquigarrow \delta_+$, we apply (5.13) and get

$$\begin{aligned}\delta_{n+1} \leq & \mathtt{C}\mathtt{G}_0\Big(\varepsilon_0^2 K_0^{\kappa_2}(\mathtt{G}_0^2 K_0^{\kappa_2} K_n^{2\mu+2\nu+4-2\kappa_2} + K_n^{\mu-\eta-\kappa_2}) + \\ (K_n^{\kappa_1} + \varepsilon_0 K_n^{\kappa_3})&(\mathtt{R}_0 K_n^{\mu+\nu+1-\Delta\mathfrak{p}} + \varepsilon_0 K_0^{\kappa_2}\mathtt{G}_0 K_n^{\mu+\nu+1-(1-\alpha)\Delta\mathfrak{p}-\kappa_2})\Big) + (K_n^{\kappa_1} + \varepsilon_0 K_n^{\kappa_3})\mathtt{R}_0 K_n^{-\Delta\mathfrak{p}+1} \leq \varepsilon_0 K_0^{\kappa_2} K_n^{-\chi\kappa_2}\end{aligned}$$

which follows by (2.42b), (2.42c), (2.42d), (2.44a) and (2.44b).

Regarding $\Gamma_{n+1,\mathfrak{p}_2}$, by (5.11) we get

$$\Gamma_{n+1,\mathfrak{p}_2} \leq \mathtt{C}\mathtt{G}_0(K_n^{\kappa_1} + \varepsilon_0 K_n^{\kappa_3})\Big(1 + K_n^{\mu+\nu+2}(\mathtt{R}_0 + K_n^{\alpha\Delta\mathfrak{p}-\kappa_2}\varepsilon_0 K_0^{\kappa_2}\mathtt{G}_0)\Big) \leq \mathtt{G}_0 K_n^{\chi\kappa_1}$$

which follows by (2.42a), (2.42e) and (2.44c).

Finally, by (5.12)

$$\Theta_{n+1,\mathfrak{p}_2} \leq \mathtt{C}\mathtt{R}_0(K_n^{\kappa_1} + \varepsilon_0 K_n^{\kappa_3}) + \mathtt{C} K_n^{\mu+\nu+2}(K_n^{\kappa_1} + \varepsilon_0 K_n^{\kappa_3})\mathtt{G}_0\big(\mathtt{R}_0 + K_n^{\alpha\Delta\mathfrak{p}-\kappa_2}\varepsilon_0 K_0^{\kappa_2}\mathtt{G}_0\big) \leq \mathtt{R}_0 K_n^{\chi\kappa_1}$$

which follows again by (2.42a), (2.42e) and (2.44c).

We now prove **(ii)**. Setting $\vec{w}_{4,n} = (\gamma_n, \mathcal{O}_{n+1}, s_n - 4\rho_n s_0, a_n - 4\rho_n a_0, r_n - 4\rho_n r_0)$, we prove inductively

$$\|(\mathcal{H}_{n+1} - \mathbb{1})(u)\|_{\vec{w}_{4,n},\mathfrak{p}_1} \leq \varepsilon_0 \sum_{k=0}^{n+1} \frac{1}{2^k}. \tag{5.48}$$

Note that the choice of $\vec{w}_{4,n}$ (5.48) is consistent with the fact that $F_n := (\mathcal{H}_n)_* F_0$ is defined on the domain

$$\mathbb{T}^d_{s_n} \times D_{a_n,p}(r_n)\,.$$

For $n=-1$ this is obvious. Then by induction we have

$$\begin{aligned}
\|(\mathcal{H}_{n+1}-\mathbb{1})(u)\|_{\vec{w}_{4,n},\mathtt{p}_1} &\le \|(\mathcal{H}_n-\mathbb{1})(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}+\|f_{n+1}(\mathcal{L}_{n+1}\circ\mathcal{H}_n(u))\|_{\vec{w}_{4,n},\mathtt{p}_1}+\|(\mathcal{L}_{n+1}-\mathbb{1})\mathcal{H}_n(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}\\
&\overset{(5.48),(2.49)}{\le} \varepsilon_0\sum_{k=0}^{n}\frac{1}{2^k}+2|g_{n+1}|_{\vec{w}_{4,n},\mathtt{p}_1}\|\mathcal{L}_{n+1}\circ\mathcal{H}_n(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}+\mathtt{C}\varepsilon_0 K_n^{-1}\|\mathcal{H}_n(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}\\
&\overset{(5.8)}{\le} \varepsilon_0\sum_{k=0}^{n}\frac{1}{2^k}+K_0^{\kappa_2}\varepsilon_0\mathtt{G}_0K_n^{-\kappa_2+\mu+1}+2\mathtt{C}\varepsilon_0K_n^{-1}\\
&\overset{(2.52)}{\le} \varepsilon_0\Big(\sum_{k=0}^{n}\frac{1}{2^k}+\frac{1}{2^{n+1}}\Big)
\end{aligned} \tag{5.49}$$

for K_0 large enough. Moreover as before

$$\|(\mathcal{H}_{n+1}-\mathcal{H}_n)u\|_{\vec{w}_{4,n},\mathtt{p}_1}\le\|(\mathcal{L}_{n+1}-\mathbb{1})\mathcal{H}_n(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}+\|f_{n+1}(\mathcal{L}_{n+1}\circ\mathcal{H}_n)(u)\|_{\vec{w}_{4,n},\mathtt{p}_1}\le\varepsilon_0 2^{-(n+1)} \tag{5.50}$$

which implies that the sequence $\mathcal{H}_n$ is Cauchy and therefore there exists a limit map $\mathcal{H}_\infty=\lim_{n\to\infty}\mathcal{H}_n$. Moreover

$$\begin{aligned}
\|(\mathcal{H}_\infty-\mathbb{1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathtt{p}_1} &\le \|(\mathcal{H}_1-\mathbb{1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathtt{p}_1}+\sum_{n\ge2}\|(\mathcal{H}_n-\mathcal{H}_{n-1})(u)\|_{\gamma_\infty,\mathcal{O}_\infty,\frac{s_0}{2},\frac{a_0}{2},\mathtt{p}_1}\\
&\le 2\varepsilon_0\,,
\end{aligned} \tag{5.51}$$

so that for ε_0 small enough also (2.57) holds. We are left with the proof of (2.58). By definition the limit vector field is

$$F_\infty:=\lim_{n\to\infty}F_n,\quad F_n:=(\Phi_n\circ\mathcal{L}_n)_*F_{n-1}. \tag{5.52}$$

On the other hand we have

$$\mathcal{F}_\infty:=(\mathcal{H}_\infty)_*F_0$$

We want to prove that $\mathcal{F}_\infty=F_\infty$; setting $\mathcal{F}_n:=(\mathcal{H}_n)_*F_0$ we show inductively that $\mathcal{F}_n=F_n$ for any $n\ge1$. For $n=1$ one has $\mathcal{H}_1=\Phi_1\circ\mathcal{L}_1$ and hence $F_1=\mathcal{F}_1$. Now assume that $\mathcal{F}_{n-1}=F_{n-1}$. By definition one has

$$\mathcal{H}_n=\mathcal{K}_n\circ\mathcal{H}_{n-1}:=(\Phi_n\circ\mathcal{L}_n)\circ\mathcal{H}_{n-1},\qquad \mathcal{H}_n^{-1}=\mathcal{H}_{n-1}^{-1}\circ\mathcal{K}_n^{-1}.$$

Hence we have

$$\begin{aligned}
F_n-\mathcal{F}_n&=(\mathcal{K}_n)_*F_{n-1}-(\mathcal{H}_n)_*F_0=d\mathcal{K}_nF_{n-1}\circ\mathcal{K}_n^{-1}-d\mathcal{H}_nF_0\circ\mathcal{H}_n^{-1}\\
&=d\mathcal{K}_nF_{n-1}\circ\mathcal{K}_n^{-1}-d\mathcal{K}_nd\mathcal{H}_{n-1}F_0\circ\mathcal{H}_{n-1}^{-1}\circ\mathcal{K}_n^{-1}\\
&=d\mathcal{K}_n(F_{n-1}-d\mathcal{H}_{n-1}F_0\circ\mathcal{H}_{n-1}^{-1})\circ\mathcal{K}_n^{-1}=0.
\end{aligned} \tag{5.53}$$

This concludes the proof of Theorem 2.25. ∎

A Smooth functions and vector fields on the torus

Here we provide some technical results.

The following one is a general result about smooth maps on the torus. First of all, for any $p\ge0$ and $\zeta\ge0$ we denote as usual

$$H^p(\mathbb{T}^b_s;\mathbb{C}):=\Big\{u=\sum_{l\in\mathbb{Z}^b}u_le^{\mathrm{i}l\cdot\theta}:\|u\|^2_{s,p}:=\sum_{l\in\mathbb{Z}^b}\langle l\rangle^{2p}|u_l|^2e^{2s|l|}<\infty\Big\}\,, \tag{A.1}$$

the space of functions which are analytic on the strip $\mathbb{T}^b_s$, Sobolev on its boundary, and have Fourier coefficients u_l. By Cauchy formula for analytic complex functions we have that this u is uniquely determined by the values that assume on the edge of the domain i.e $z = x \pm i\sigma s$ where $\sigma \in \{\pm 1\}^b$. We can define a natural norm using the Sobolev norm of the function on the boundary

$$|u|_{s,p}^2 := \sum_{\sigma \in \{\pm 1\}^b} \int_{\mathbb{T}^b} \langle \nabla \rangle^{2p} |u(x + i\sigma s)|^2 \tag{A.2}$$

Using the Fourier basis it reads

$$|u|_{s,p}^2 := \sum_{\sigma \in \{\pm 1\}^b} \sum_{l \in \mathbb{Z}^b} \langle l \rangle^{2p} |u_l|^2 e^{-2s\sigma \cdot l} .$$

Lemma A.1. *The norm* $|\cdot|_{s,p}$ *and*

$$\|u\|_{s,p}^2 := \sum_{l \in \mathbb{Z}^b} \langle l \rangle^{2p} |u_l|^2 e^{2s|l|} \tag{A.3}$$

are equivalent.

The following Lemma lists some important properties of Sobolev spaces $H^s := H^s(\mathbb{T}^b; \mathbb{C})$ with norm

$$\|u\|_s^2 := \sum_{l \in \mathbb{Z}^b} \langle l \rangle^{2p} |u_l|^2 .$$

The same results holds also for our analytic norm in (A.1). The proof of the Lemma is classical.

Lemma A.2. *Let* $s_0 > d/2$. *Then*

(i) **Embedding.** $\|u\|_{L^\infty} \le C(s_0) \|u\|_{s_0}$, $\forall\, u \in H^{s_0}$.

(ii) **Algebra.** $\|uv\|_{s_0} \le C(s_0) \|u\|_{s_0} \|v\|_{s_0}$, $\forall\, u, v \in H^{s_0}$.

(iii) **Interpolation.** *For* $0 \le s_1 \le s \le s_2$, $s = \lambda s_1 + (1 - \lambda) s_2$,

$$\|u\|_s \le \|u\|_{s_1}^{\lambda} \|u\|_{s_2}^{1-\lambda}, \quad \forall\, u \in H^{s_2}. \tag{A.4}$$

(iv) **Asymmetric tame product.** *For* $s \ge s_0$ *one has*

$$\|uv\|_s \le C(s_0) \|u\|_s \|v\|_{s_0} + C(s) \|u\|_{s_0} \|v\|_s, \quad \forall\, u, v \in H^s. \tag{A.5}$$

(vi) **Mixed norms asymmetric tame product.** *For* $s \ge 0$, $s \in \mathbb{N}$ *setting* $|u|_s^\infty := \sum_{|\alpha| \le s} ||D^\alpha u||_{L^\infty}$ *the norm in* $W^{s,\infty}$ *one has*

$$\|uv\|_s \le \frac{3}{2} \|u\|_{L^\infty} \|v\|_s + C(s) |u|_{s,\infty} \|v\|_0, \forall\, u \in W^{s,\infty}, v \in H^s. \tag{A.6}$$

If $u := u(\lambda)$ *and* $v := v(\lambda)$ *depend in a Lipschitz way on* $\lambda \in \Lambda \subset \mathbb{R}^b$, *all the previous statements hold if one replace the norms* $\|\cdot\|_s$, $|\cdot|_s^\infty$ *with* $\|\cdot\|_{s,\lambda}$, $|\cdot|_{s,\lambda}^\infty$ *defined as in* (2.21).

We now introduce the space

$$W^{p,\infty}(\mathbb{T}_\zeta^b) := \left\{ \beta : \mathbb{T}_\zeta^b \to \mathbb{T}_\zeta^b \ : \ |\beta|_{p,\zeta,\infty} := \sum_{k=0}^{p} \|d^k \beta\|_{L^\infty(\mathbb{T}_\zeta^b)} < \infty \right\}, \tag{A.7}$$

and note that one has $H^{\zeta, p + \mathfrak{p}_0}(\mathbb{T}_\zeta^b) \subset W^{p,\infty}(\mathbb{T}_\zeta^b)$.

Lemma A.3 (Diffeo). *Let $\beta \in W^{p,\infty}(\mathbb{T}^b_\zeta)$ for some $p, \zeta \geq 0$ such that*

$$\|\beta\|_{\zeta,\mathfrak{p}_0} \leq \frac{\delta}{2C_1}, \quad \|\beta\|_{\zeta,\mathfrak{p}_0} \leq \frac{1}{2C_2}, \quad 0 < \delta < \frac{\zeta}{2}, \;\; C_1, C_2 > 0, \tag{A.8}$$

and let us consider $\Phi : \mathbb{T}^b_\zeta \to \mathbb{T}^b_{2\zeta}$ of the form

$$x \mapsto x + \beta(x) = \Phi(x). \tag{A.9}$$

Then the following is true.
(i) There exists $\Psi : \mathbb{T}^b_{\zeta-\delta} \to \mathbb{T}^b_\zeta$ of the form $\Psi(y) = y + \tilde{\beta}(y)$ with $\tilde{\beta} \in W^{p,\infty}(\mathbb{T}^b_{\zeta-\delta})$ satisfying

$$\|\tilde{\beta}\|_{\zeta-\delta,\mathfrak{p}_0} \leq \frac{\delta}{2}, \quad \|\tilde{\beta}\|_{\zeta-\delta,p} \leq 2\|\beta\|_{\zeta,p}\,, \tag{A.10}$$

such that for all $x \in \mathbb{T}^b_{\zeta-2\delta}$ one has $\Psi \circ \Phi(x) = x$.
(ii) For all $u \in H^{\zeta,p}(\mathbb{T}^b_\zeta)$, the composition $(u \circ \Phi)(x) = u(x + \beta(x))$ satisfies

$$\|u \circ \Phi\|_{\zeta-\delta,p} \leq C(\|u\|_{\zeta,p} + |d\beta|_{p-1,\zeta,\infty}\|u\|_{\zeta,\mathfrak{p}_0}). \tag{A.11}$$

Proof. For $\zeta = 0$ the result is proved in [38] thus in the following we assume $\zeta > 0$.
(i) First of all recall that, if $\mathfrak{p}_0 \geq b/2$ then $\|u\|_{L^\infty} \leq \|u\|_{\zeta,\mathfrak{p}_0}$. We look for $\tilde{\beta}$ such that

$$\tilde{\beta}(y) = -\beta(y + \tilde{\beta}(y)). \tag{A.12}$$

The idea is to rewrite the problem as a fixed point equation. We define the operator $\mathcal{G} : H^{\zeta,p} \to H^{\zeta,p}$ as $\mathcal{G}(\tilde{\beta}) = -\beta(y + \tilde{\beta})$. First of all we need to show that $\mathcal{G}$ maps the ball $B_{\delta/2} := \{\|u\|_{\zeta-\delta,p} < \delta/2\}$ into itself.

One has

$$\|\mathcal{G}(\tilde{\beta})\|_{\zeta-\delta,\mathfrak{p}_0} = \left\| \sum_{n\geq 0} \frac{1}{n!}(\partial^n\beta)\tilde{\beta}^n \right\|_{\zeta-\delta,\mathfrak{p}_0} \leq \sum_{n\geq 0} \frac{1}{n!}\|\beta\|_{\zeta-\delta,\mathfrak{p}_0+n}\|\tilde{\beta}\|^n_{\zeta-\delta,\mathfrak{p}_0}\,, \tag{A.13}$$

where $\partial\beta$ denotes the derivative of β w.r.t. its argument. Note that for any $u \in H^{\zeta+\delta,s}$ and $\tau > 0$ one has

$$\|u\|_{\zeta,s+\tau} \leq \left(\frac{\tau}{e}\right)^\tau \frac{1}{\delta^\tau}\|u\|_{\zeta+\delta,s}\,; \tag{A.14}$$

indeed

$$\|u\|^2_{\zeta,p+\tau} = \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2(p+\tau)} e^{2\zeta|l|}|u_l|^2 \leq \sum_{l\in\mathbb{Z}^b} \langle l\rangle^{2p}|l|^{2\tau}e^{-2\delta(|l|)}e^{2(\zeta+\delta)|l|}|b_l|^2\,,$$

and the function $f(x) := x^{2\tau}e^{-2\delta x}$ reach its maximum at $x = \tau/\delta$ and $f(\tau/\delta) = (\tau/\delta e)^{2\tau}$, so that (A.14) follows. Then using (A.14) and the fact that $n! = (1/\sqrt{2\pi n})(n/e)^n(1 + O(1/n))$ as $n \to \infty$, we obtain

$$\begin{aligned} \|\mathcal{G}(\tilde{\beta})\|_{\zeta-\delta,\mathfrak{p}_0} &\leq \sum_{n\geq 0} \frac{1}{n!}\left(\frac{n}{e}\right)^n \frac{1}{\delta^n}\|\beta\|_{\zeta,\mathfrak{p}_0}\|\tilde{\beta}\|^n_{\zeta-\delta,\mathfrak{p}_0} \leq \|\beta\|_{\zeta,\mathfrak{p}_0}\sum_{n\geq 0} C\left(\frac{\|\tilde{\beta}\|_{\zeta-\delta,\mathfrak{p}_0}}{\delta}\right)^n \\ &\leq 2C\|\beta\|_{\zeta,\mathfrak{p}_0} \overset{(A.8)}{\leq} \frac{\delta}{2}\,. \end{aligned} \tag{A.15}$$

Finally we show that $\mathcal{G}$ is a contraction. One has

$$\begin{aligned}
\|\mathcal{G}(\tilde\beta_1)-\mathcal{G}(\tilde\beta_2)\|_{\zeta-\delta,p} &= \left\|\sum_{n\geq1}\frac{1}{n!}(\partial^n\beta)\tilde\beta_1^n-\sum_{n\geq1}\frac{1}{n!}(\partial^n\beta)\tilde\beta_2^n\right\|_{\zeta-\delta,p}\\
&= \left\|\sum_{n\geq1}\frac{1}{n!}(\partial^n\beta)(\tilde\beta_1-\tilde\beta_2)\left(\sum_{k=0}^{n-1}\tilde\beta_1^k\tilde\beta_2^{n-1-k}\right)\right\|_{\zeta-\delta,p}\\
&\leq \|\tilde\beta_1-\tilde\beta_2\|_{\zeta-\delta,p}\sum_{n\geq1}\frac{1}{n!}\left(\frac{n}{e}\right)^n\frac{1}{\delta^n}\|\beta\|_{\zeta,p}\|\sum_{k=0}^{n-1}\|\tilde\beta_1\|^k_{\zeta-\delta,p}\|\tilde\beta_2\|^{n-1-k}_{\zeta-\delta,s}\\
&\leq \|\tilde\beta_1-\tilde\beta_2\|_{\zeta-\delta,p}C_2\|\beta\|_{\zeta-\delta,p}\overset{(A.8)}{\leq}\frac{1}{2}\|\tilde\beta_1-\tilde\beta_2\|_{\zeta-\delta,p}\,.
\end{aligned}\tag{A.16}$$

Then we deduce that there exists a unique fixed point in $B_{\delta/2}$, hence a solution of the equation (A.11).
(ii) One can follow almost word by word the proof of Lemma 11.4 in [38] using the norm (A.2) instead of (A.3) and the interpolation properties of the $W^{p,\infty}(\mathbb{T}^b_\zeta)$-norms. ■

Remark A.4. *Note that by Lemma A.1, one has*

$$\begin{aligned}
\|f^{(\theta)}(\theta,y,w)\|_{s,a,p} &\approx \frac{1}{s_0}\max_{1\leq i\leq d}\sum_{\sigma\in\{\pm1\}^d}\|f^{(\theta_i)}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^p}\,,\\
\|f^{(y)}(\theta,y,w)\|_{s,a,p} &\approx \frac{1}{r_0^{\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{\sigma\in\{\pm1\}^d}\|f^{(y_i)}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^p}\,,\\
\|f^{(w)}(\theta,y,w)\|_{s,a,p} &\approx \frac{1}{r_0}\sum_{\sigma\in\{\pm1\}^d}\left(\|\mathtt{f}_{\mathtt{p}_0}(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^p(\mathbb{T}^d_s)}+\|\mathtt{f}_p(\mathrm{Re}(\theta)+\mathrm{i}\sigma s)\|_{H^{\mathtt{p}_0}(\mathbb{T}^d_s)}\right)
\end{aligned}$$

where $\mathtt{f}_p(\theta)$ is defined as in (2.12)*. In particular this means that for all $s\geq0$, $a\geq0$ and $p\geq\overline{\mathtt{p}}>n/2$ one has the standard algebra, interpolation and tame properties w.r.t. composition with functions in $H^p(\mathbb{T}^d_s)$; see for instance [25, 40, 41, 28] just to mention a few.*

From Lemma A.3 and Remark A.4 above we deduce the following result.

Lemma A.5. *Given a tame vector field $f\in\mathcal{V}_{\vec v,p}$ with scale of constants $C_p(f)$ of the form* (2.14) *and given a map $\Phi(\theta)=\theta+\beta(\theta):\mathbb{T}^d_{s'}\to\mathbb{T}^d_s$ as in* (A.9) *with $b=d$ and $\zeta=s$, then the composition $f\circ\Phi$ is a tame vector field with constant*

$$C_p(f\circ\Phi)\leq C_p(f)+C_{\mathtt{p}_0}(f)\|\beta\|_{s,p+\nu+3}. \tag{A.17}$$

Moreover if f is a regular vector field, i.e. it satisfies (4.17)*, then*

$$|f\circ\Phi|_{\vec v_1,p}\leq|f|_{\vec v,p}+|f|_{\vec v,\mathtt{p}_0}\|\beta\|_{s,p+\nu+3}. \tag{A.18}$$

where $\vec v_1=(\lambda,\mathcal{O},s',a)$.

Proof. By Lemma A.3 one has that if $\|\beta\|_{s,\mathtt{p}_1}$ is sufficiently small, then the vector field $f\circ\Phi$ is defined on $\mathbb{T}^d_{s-\rho s_0}\times D_{a,p}(r-\rho r_0)$. Lemma A.3 guarantees that for a function $u(\theta)\in\mathbb{C}$ the estimate (A.11) holds. Hence also the components $f^{(v)}(\theta,y,w)$ for $v=\theta,y$ satisfy the same bounds (recall that for the norm (2.7) y,w are parameters). Let us study the composition of $f^{(w)}(\theta+\beta(\theta),y,w)$. By Remark 2.4 one has

$$\|f^{(w)}\|_{s,a,p}=\frac{1}{r_0}(\|\mathtt{f}_{\mathtt{p}_0}\|_{s,p}+\|\mathtt{f}_p\|_{s,\mathtt{p}_0}).$$

Now $\mathtt{f}_p:\mathbb{T}^d_s\to\mathbb{C}$ and hence we can apply Lemma A.3 to obtain the result. The bounds on the derivatives follow in the same way. □

B Properties of Tame and regular vector fields

We now discuss the main properties of C^k-tame and regular vector fields; in particular we need to control the changes in the tameness constants when conjugating via changes of variables generated by regular bounded vector fields.

Lemma B.1. *Consider any two* C^k-tame *vector fields* $F, G \in \mathcal{W}_{\vec{v},p}$*, then the following holds.*

(i) For $l = 1, \dots, d$ *one has that* $\partial_{\theta_l} F$ *is a* C^k*-tame vector field up to order* $q-1$ *with tameness constants* $C_{\vec{v},p+1}(F)$*.*

(ii) For $l = 1, \dots, d$ *one has that* $\partial_{y_l} F, d_w F[w]$ *are* C^{k-1}*-tame vector fields up to order* q *with tameness constants* $C_{\vec{v},p}(F)$*. for any* $h \geq 0$*.*

(iii) The commutator $[G, F]$ *is a* C^{k-1}*-tame vector field up to order* $q-1$ *with scale of constants*

$$C_{\vec{v},p}([G,F]) \leq \mathtt{C}(C_{\vec{v},p+\nu_G+1}(F) C_{\vec{v},\mathfrak{p}_0+\nu_F+1}(G) + C_{\vec{v},\mathfrak{p}_0+\nu_G+1}(F) C_{\vec{v},p+\nu_F+1}(G)), \tag{B.1}$$

where ν_F*,* ν_G *are the loss of regularity of* F*,* G *respectively.*

(iv) If F *is a polynomial of maximal degree* k *in* y, w *then it is* C^∞*-tame up to order* q*.*

Proof. Let us check item (i). We consider a map $\Phi := \mathbb{1} + f$ as in Definition 2.13. Recall that $\|f\|_{s,a,\mathfrak{p}_1} < 1/2$. One has that

$$\begin{aligned} \|(\partial_\theta F) \circ \Phi\|_{s,a,p} &\leq \|\partial_\theta (F \circ \Phi)\|_{s,a,p} + \|(\partial_\theta F) \circ \Phi \cdot \partial_\theta f\|_{s,a,p} \\ &\leq \|\partial_\theta (F \circ \Phi)\|_{s,a,p} + \|(\partial_\theta F) \circ \Phi\|_{s,a,p} \|\partial_\theta f\|_{s,a,\mathfrak{p}_0} + \|(\partial_\theta F) \circ \Phi\|_{s,a,\mathfrak{p}_0} \|\partial_\theta f\|_{s,a,p}. \end{aligned} \tag{B.2}$$

Now, for $p = \mathfrak{p}_0$, by (B.2) one gets

$$\big(1 - 2\|\partial_\theta f\|_{s,a,\mathfrak{p}_0}\big) \|(\partial_\theta F) \circ \Phi\|_{s,a,\mathfrak{p}_0} \leq \|\partial_\theta (F \circ \Phi)\|_{s,a,\mathfrak{p}_0} \overset{(T1)}{\leq} C_{s,a,p+1}(F)(1 + \|\Phi\|_{s,a,\mathfrak{p}_0}), \tag{B.3}$$

hence, for $p > \mathfrak{p}_0$ one has

$$\|(\partial_\theta F) \circ \Phi\|_{s,a,p} \leq \mathtt{c}(C_{s,a,p+1}(F) + C_{s,a,\mathfrak{p}_0}(F) \|\Phi\|_{s,a,p+1}), \tag{B.4}$$

for some $\mathtt{c}$ independent of p. Equation (B.4) implies property $(T1)$ for the vector field $(\partial_\theta F)(\theta, y, w)$. Clearly it holds for $p \leq q-1$. The other bounds follows similarly. Finally items $(ii), (iii), (iv)$ follow by the definitions. □

Remark B.2. *For* $k \geq 0$ *and any* $\mathtt{v} \in \mathtt{V}, \mathtt{v}_1, \dots, \mathtt{v}_k \in \mathtt{U}$ *consider any monomial subspace* $\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)}$ *as in* (2.23)*. Then for all* C^k*-tame vector fields* F *one has that* $\Pi_{\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)}} F$ *is* C^∞*-tame (up to the same order as* F*) and one can choose the constant as* $C_{\vec{v},p}(\Pi_{\mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_k)}} F) = C_{\vec{v},p}(F)$*. The same holds for the direct sum* $\mathcal{U}$ *of a finite number of monomial spaces and their orthogonal, namely one can chose*

$$C_{\vec{v},p}(\Pi_{\mathcal{U}} F) = C_{\vec{v},p}(F), \qquad C_{\vec{v},p}(\Pi_{\mathcal{U}}^\perp F) = C_{\vec{v},p}(F)\,.$$

Lemma B.3 (Conjugation). *Consider a tame left invertible map* $\Phi = \mathbb{1} + f$ *with tame inverse* $\Psi = \mathbb{1} + h$ *as in Definition 2.8 such that* (2.18) *holds. Assume that* $\mathfrak{p}_1 \geq \mathfrak{p}_0 + \nu + 1$ *and the fields* f, h *are such that* $C_{\vec{v},\mathfrak{p}_1}(f) = C_{\vec{v},\mathfrak{p}_1}(h) \leq \mathtt{c}\rho$ *for* $\rho > 0$ *and* $\mathtt{c}$ *the same appearing in Remark 2.7* [12]*. For any vector field*

$$F : \mathbb{T}^d_s \times D_{a,p+\nu}(r) \times \mathcal{O} \to V_{a,p}, \tag{B.5}$$

[12] By Remark 2.7 the smallness of the constants $C_{\vec{v},\mathfrak{p}_1}(f), C_{\vec{v},\mathfrak{p}_1}(h)$ automatically implies that Φ, Ψ satisfy (2.17).

which is C^k–tame up to order q, one has that the push–forward

$$F_+ := \Phi_* F : \mathbb{T}^d_{s-2\rho s_0} \times D_{a,p+\nu}(r - 2\rho r_0) \times \mathcal{O} \to V_{a-2\rho a_0,p} \tag{B.6}$$

is C^k–tame up to order $q - \nu - 1$, with scale of constants

$$C_{\vec{v}_2,p}(F_+) \le (1+\rho)\Big(C_{\vec{v},p}(F) + C_{\vec{v},\mathfrak{p}_0}(F) C_{\vec{v}_1,p+\nu+1}(f)\Big), \tag{B.7}$$

where $\vec{v} := (\lambda, \mathcal{O}, s, a, r)$, $\vec{v}_1 := (\lambda, \mathcal{O}, s - \rho s_0, a - \varrho a_0, r - \rho r_0)$ and $\vec{v}_2 := (\lambda, \mathcal{O}, s - 2\rho s_0, a - 2\varrho a_0, r - 2\rho r_0)$.

Proof. By (2.19) the vector field F_+ is defined in $\mathbb{T}_{s-2\rho s_0} \times D_{a,p}(r - 2\rho r_0) \times \mathcal{O}$. Then, given a change of coordinates $\Gamma : \mathbb{T}^d_{s_1} \times D_{a',p'}(r_1) \times \mathcal{O} \to \mathbb{T}_{s-2\rho s_0} \times D_{a,p}(r - 2\rho r_0)$ we can consider the composition of F_+ with Γ; in particular

$$\Psi \circ \Gamma : \mathbb{T}^d_{s_1} \times D_{a',p'}(r_1) \times \mathcal{O} \longrightarrow \mathbb{T}_s \times D_{a,p+\nu}(r),$$

namely the domain of F. Let us check the property (T_0) for the vector field F_+. In the following we will keep track only of the index p. One has

$$\begin{aligned} &\|\Psi(\Gamma)\|_p \le C_p(f) + (1 + C_{\mathfrak{p}_0}(f))\|\Gamma\|_p, \\ &\|\Psi(\Gamma)\|_{\mathfrak{p}_0+\nu} \le 1 + 2C_{\mathfrak{p}_0+\nu}(f). \end{aligned} \tag{B.8}$$

so that we get

$$\begin{aligned} \|F_+(\Gamma)\|_p &\le \|F(\Psi(\Gamma))\|_p + \|df(\Psi(\Gamma))[F(\Psi(\Gamma))]\|_p \\ &\overset{(\mathrm{T}_1)}{\le} (1 + C_{\mathfrak{p}_0+1}(f))\|F(\Psi(\Gamma))\|_p + (C_{p+1}(f) + C_{\mathfrak{p}_0+1}(f)\|\Psi(\Gamma)\|_p)\,\|F(\Psi(\Gamma))\|_{\mathfrak{p}_0} \\ &\overset{(\mathrm{T}_0)}{\le} (1 + C_{\mathfrak{p}_0+1}(f))\,[C_p(F) + C_{\mathfrak{p}_0}(F)\|\Psi(\Gamma)\|_{p+\nu}] \\ &\qquad + (C_{p+1}(f) + C_{\mathfrak{p}_0+1}(f)\|\Psi(\Gamma)\|_p)\,[C_{\mathfrak{p}_0}(F) + C_{\mathfrak{p}_0}(F)\|\Psi(\Gamma)\|_{\mathfrak{p}_0+\nu}]\ , \end{aligned} \tag{B.9}$$

and therefore

$$\begin{aligned} \|F_+(\Gamma)\|_p \le\ & C_p(F)(1 + C_{\mathfrak{p}_0+\nu+1}(f)) + 5C_{\mathfrak{p}_0}(F)(1 + C_{\mathfrak{p}_0+1}(f))C_{p+\nu+1}(f) \\ &+ \|\Gamma\|_{p+\nu}\left[C_{\mathfrak{p}_0}(F)(1 + 3C_{\mathfrak{p}_0+\nu+1}(f))^2\right], \end{aligned} \tag{B.10}$$

that is (T_0). The other properties are obtained with similar calculations using also the fact that the vector field f is linear in the variables y, w. Hence F_+ is tame with scale of constants in (B.7). □

Remark B.4. *In Lemma B.3, if $f, h \in E^{(K)}$, then the smoothing estimate* (2.32) *applied to $|f|_{\vec{v},p+\nu+1}$ implies that F_+ is C^k–tame up to order q. The same holds if Φ is generated by a vector field $g \in E^{(K)}$.*

Remark B.5. *Consider a vector field g which generates a well defined flow Φ^t for $t \le 1$ and set $\Phi := \Phi^1$. Then, for all $p \ge \mathfrak{p}_0$ for all vector fields F such that the push-forward with Φ^t is well defined, one has*

$$\begin{aligned} L &:= \Phi_* F - F = \int_0^1 \Phi^t_*[g, F]dt, \\ Q &:= \Phi_* F - [g, F] - F = \int_0^1 \int_0^t \Phi^s_*[g, [g, F]]dsdt. \end{aligned} \tag{B.11}$$

If moreover $g \in \mathcal{B}$ satisfies Definition 2.18 item 5, and F is as in formula (B.5) *then L is C^{k-1} tame and Q is C^{k-2} tame up to order $q - \nu - 1$ with constants*

$$C_{\vec{v}_2,p}(L) \le \mathtt{C}\Big(C_{\vec{v},p+1}(F)|g|_{\vec{v},\mathfrak{p}_0+\nu+1} + C_{\vec{v},\mathfrak{p}_0+1}(F)|g|_{\vec{v},p+\nu+1}\Big) \tag{B.12a}$$

$$C_{\vec{v}_2,p}(Q) \le \mathtt{C}\Big(C_{\vec{v},p+2}(F)|g|^2_{\vec{v},\mathfrak{p}_0+\nu+1} + C_{\vec{v},\mathfrak{p}_0+2}(F)|g|_{\vec{v},\mathfrak{p}_0+\nu+2}|g|_{\vec{v},p+\nu+2}\Big) \tag{B.12b}$$

$\vec{v} := (\lambda, \mathcal{O}, s, a, r)$, $\vec{v}_2 := (\lambda, \mathcal{O}, s - 2\rho s_0, a, r - 2\rho r_0)$. Finally if $g \in E^{(K)}$ then L, Q are tame up to order q.

Lemma B.6. *Consider any subspace $\mathcal{U}$ which is a finite direct sum of* monomial *subspaces as in formula* (2.23)*, having degree at most k, or their average in θ. Under the hypotheses of Lemma B.3, assume also that*[13] *$F \in \mathcal{U}$ and $f \in \mathcal{B}$. Then for all $p \geq \mathfrak{p}_0$ one has*

$$C_{\vec{v}_2,p}(\Pi_{\mathcal{U}}^{\perp}\Phi_* F) \leq \mathtt{C}\Big(C_{\vec{v},p}(F)\rho + C_{\vec{v},\mathfrak{p}_1}(F)|f|_{\vec{v}_1,p+\nu+1}\Big) \tag{B.13}$$

where $\vec{v} := (\lambda, \mathcal{O}, s, a)$, $\vec{v}_1 := (\lambda, \mathcal{O}, s - \rho s_0, a)$ and $\vec{v}_2 := (\lambda, \mathcal{O}, s - 2\rho s_0, a)$. Moreover if $f \in E^{(K)}$ then (B.13) *holds up to order q.*

Proof. One has

$$\Pi_{\mathcal{U}}^{\perp}\Big(\Phi_* F\Big) = \Pi_{\mathcal{U}}^{\perp}\Big(F \circ \Psi - F + d_u f(\Psi)[F \circ \Psi]\Big). \tag{B.14}$$

The last summand clearly satisfies the estimates (B.13). Regarding the first terms we write

$$\Pi_{\mathcal{U}}^{\perp}(F \circ \Psi - F) = \Pi_{\mathcal{U}}^{\perp}\big(d_\theta F[f^{(\theta)}] + \sum_{\mathtt{u}=y,w} d_{\mathtt{u}} F[f^{(\mathtt{u})}]\big)\,;$$

the second term satisfies (B.13) by property (T_1), while we claim that $d_\theta F[f^{(\theta)}] \in \mathcal{U}$. Indeed if $F \in \mathcal{V}^{(\mathtt{v},\mathtt{v}_1,\dots,\mathtt{v}_h)}$, it has the form

$$F = \frac{1}{\alpha(\mathtt{v}_1,\dots,\mathtt{v}_h)!}\Big(\prod_{i=1}^{h} d_{\mathtt{v}_i}\Big)F^{(\mathtt{v})}(\theta,0,0)[\mathtt{v}_1,\dots,\mathtt{v}_h],$$

so that deriving w.r.t. θ on both sides and computing at $g^{(\theta)}$ commutes with the derivation in $\mathtt{v}_i \in \mathtt{U}$ (this follows from the fact that if $f \in \mathcal{B}$ then $f^{(\theta)}(\theta, y, w) = f^{(\theta)}(\theta, 0, 0)$). If $\mathcal{U}$ is only the average of some monomial space, then clearly its θ-derivative is zero. □

B.1 Proof of Lemma 4.8

Lemma B.7. *All regular vector fields f as in Definition 4.5 are C^∞-tame up to order q with tameness constant*

$$C_{\vec{v},p}(f) = C_{d,q}|f|_{\vec{v},p}. \tag{B.15}$$

Proof. In view of Lemma B.1–(iv), we only need to prove that a regular vector field is C^1-tame. Consider a regular vector field f (see Definition 4.5) and a map $\Phi = \mathbb{1} + g$ as in Definition 2.13. For simplicity we drop the indices $\vec{v}, \vec{v}_1, \vec{v}_2$. Without loss of generality we can also assume that $g^{(\theta)}$ depends only on θ, since in (T_m) we first perform the y, w-derivatives and then compute at $\Phi = \mathbb{1} + g$. Let us check (T_0) for f. One has

$$\begin{aligned}
&(f \circ \Phi)^{(\theta)} := h^{(\theta,0)}(\theta), \qquad (f \circ \Phi)^{(w)} := h^{(w,0)}(\theta),\\
&(f \circ \Phi)^{(y)} := h^{(y,0)}(\theta) + h^{(y,y)}(\theta)\Phi^{(y)}(\theta, y, w) + h^{(y,w)}(\theta) \cdot \Phi^{(w)}(\theta, y, w),
\end{aligned}$$

where

$$h^{(v,v')}(\theta) := f^{(v,v')}(\theta + g^{(\theta,0)}(\theta)), \quad v, v' = 0, \theta, y, w.$$

We first give bounds on the norm of $f \circ \Phi$ in terms of the norms of h and Φ. The terms depending only on θ are trivially bounded by the norm of h. In the y-component one has

$$\begin{aligned}
\|h^{(y,y)}\Phi^{(y)}\|_{s,a,p}^2 &\leq C(d)\frac{1}{r_0^{2\mathtt{s}}}\sum_{\ell \in \mathbb{Z}^d}\sum_{i=1}^{d_1}\sum_{k=1}^{d_1}|(h^{(y_i,y_k)}g^{(y_k)})(\ell)|^2 e^{2s|\ell|}\langle \ell \rangle^{2p}\\
&= C(d)\frac{1}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{k=1}^{d_1}\|h^{(y_i,y_k)}(\theta)\Phi^{(y_k)}(\theta)\|_{s,p}^2 \qquad\qquad (B.16)\\
&\overset{(A.5)}{\leq} C(d)\frac{1}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{k=1}^{d_1}(\|h^{(y_i,y_k)}\|_{s,p}\|\Phi^{(y_k)}\|_{s,\mathfrak{p}_0} + \|h^{(y_i,y_k)}\|_{s,\mathfrak{p}_0}\|\Phi^{(y_k)}\|_{s,p})^2,
\end{aligned}$$

[13] hence they are both C^∞-tame.

hence one obtains

$$\|h^{(y,y)}\Phi^{(y)}\|_{s,a,p} \leq K\left(r_0^{\mathtt{s}}\|h^{(y,y)}\|_{s,a,p}\|\Phi^{(y)}\|_{s,a,\mathfrak{p}_0} + r_0^{\mathtt{s}}\|h^{(y,y)}\|_{s,a,\mathfrak{p}_0}\|\Phi^{(y)}\|_{s,a,p}\right). \tag{B.17}$$

Finally one has

$$\begin{aligned}
\|h^{(y,w)}\Phi^{(w)}\|_{s,a,p}^2 &\leq C\frac{1}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\|h^{(y_i,w)}\Phi^{(w)}\|_{s,p}^2 = C\frac{1}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l\in\mathbb{Z}^d}\langle l\rangle^{2p}e^{2s|l|}|(h^{(y_i,w)}\cdot\Phi^{(w)})(l)|^2 \\
&\leq C\frac{1}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l\in\mathbb{Z}^d}\langle l\rangle^{2p}e^{2s|l|}(\sum_{k\in\mathbb{Z}^d}|h^{(y_i,w)}(l-k)\cdot\Phi^{(w)}(k)|)^2 \\
&\leq \frac{C}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2p}e^{2s|l-k|}\langle k\rangle^{2\mathfrak{p}_0}e^{2s|k|}|h^{(y_i,w)}(l-k)\cdot\Phi^{(w)}(k)|^2 \\
&\quad + \frac{C}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2\mathfrak{p}_0}e^{2s|l-k|}\langle k\rangle^{2p}e^{2s|k|}|h^{(y_i,w)}(l-k)\cdot\Phi^{(w)}(k)|^2 \\
&\overset{(2.1)}{\leq} \frac{C}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2p}e^{2s|l-k|}\langle k\rangle^{2\mathfrak{p}_0}e^{2s|k|}\|h^{(y_i,w)}(l-k)\|_{-a,-\mathfrak{p}_0-\nu}^2\|\Phi^{(w)}(k)\|_{a,\mathfrak{p}_0+\nu}^2 \\
&\quad + \frac{C}{r_0^{2\mathtt{s}}}\sum_{i=1}^{d_1}\sum_{l,k\in\mathbb{Z}^d}\langle l-k\rangle^{2\mathfrak{p}_0}e^{2s|l-k|}\langle k\rangle^{2p}e^{2s|k|}\|h^{(y_i,w)}(l-k)\|_{-a,-\mathfrak{p}_0-\nu}^2\|\Phi^{(w)}(k)\|_{a,\mathfrak{p}_0+\nu}^2,
\end{aligned} \tag{B.18}$$

where in the third line we used the standard interpolation estimates and the fact that $\mathfrak{p}_0 > d/2$. By (B.18), since $p \leq q$, it follows that

$$\|h^{(y,w)}\Phi^{(w)}\|_{s,a,p} \leq C\left(r_0\|h^{(y,w)}\|_{H^p(\mathbb{T}_s^d;\ell_{-a,-\mathfrak{p}_0-\nu})}\|\Phi^{(w)}\|_{s,a,\mathfrak{p}_0+\nu} + r_0\|h^{(y,w)}\|_{H^{\mathfrak{p}_0}(\mathbb{T}_s^d;\ell_{-a,-\mathfrak{p}_0-\nu})}\|\Phi^{(w)}\|_{s,a,p}\right).$$

Now each component $h^{(v,v')}$ is a function $\mathbb{T}_s^d \to V_{a,p}$ composed with a diffeomorphism of the torus given by $\theta \mapsto \theta + g^{(\theta,0)}(\theta)$. Hence we obtain, by using Lemma A.3(ii) and Lemma A.5,

$$\|f\circ\Phi\|_{s,a,p} \leq C(d,q)(|f|_{s,a,p} + |f|_{s,a,\mathfrak{p}_0}\|\Phi\|_{s,a,p+\nu}), \tag{B.19}$$

The property (T$_1$) follows in the same way. □

Lemma B.8. *Consider a vector field $f \in \mathcal{B}$ such that*

$$f : \mathbb{T}_s^d \times D_{a,p}(r) \times \mathcal{O} \to V_{a,p} \tag{B.20}$$

and

$$|f|_{\vec{v},\mathfrak{p}_1} \leq \mathtt{c}\rho, \tag{B.21}$$

for some $\rho > 0$. If ρ is small enough, then for all $\xi \in \mathcal{O}$ the following holds.
(i) The map $\Phi := \mathbb{1} + f$ is such that

$$\Phi : \mathbb{T}_s^d \times D_{a,p}(r) \times \mathcal{O} \longrightarrow \mathbb{T}_{s+\rho s_0}^d \times D_{a,p}(r+\rho r_0). \tag{B.22}$$

(ii) There exists a vector field $h \in \mathcal{B}$ such that

- $|h|_{\vec{v}_1,p} \leq 2|f|_{\vec{v},p}$, *the map $\Psi := \mathbb{1} + h$ is such that*

$$\Psi : \mathbb{T}_{s-\rho s_0}^d \times D_{a,p}(r-\rho r_0) \times \mathcal{O} \to \mathbb{T}_s^d \times D_{a,p}(r). \tag{B.23}$$

- *for all* $(\theta, y, w) \in \mathbb{T}^d_{s-2\rho s_0} \times D_{a,\mathfrak{p}_1}(r-2\rho r_0)$ *one has*

$$\Psi \circ \Phi(\theta, y, w) = (\theta, y, w). \tag{B.24}$$

Proof. *(i)* We want to bound the components of $\Phi = \mathbb{1} + f$. First of all we see that for $\theta \in \mathbb{T}^d_s$ one has

$$|\Phi^{(\theta)}|_\infty \le s + |f^{(\theta)}|_\infty \le s + \|f^{(\theta)} \cdot \partial_\theta\|_{s,a,\mathfrak{p}_0} \overset{(B.21)}{\le} s + \rho s_0 \,, \tag{B.25}$$

where we used the standard Sobolev embedding Theorem. The bound on $\|\Phi^{(w)}\|_{a,\mathfrak{p}_0} \le r + \rho r_0$ follows directly by hypothesis (B.21). In order to obtain the estimates on the $y-$components we need to check that

$$\begin{gathered} |f^{(y,0)}(\theta)|_1 \le \mathtt{c}^{-1}\|f^{(y,0)}(\theta)\cdot\partial_y\|_{s,a,\mathfrak{p}_0}, \qquad |f^{(y,y)}(\theta)y|_1 \le \mathtt{c}^{-1}\|f^{(y,y)}(\theta)y\cdot\partial_y\|_{s,a,\mathfrak{p}_0}, \\ |f^{(y,w)}(\theta)w|_1 \le \mathtt{c}^{-1}\|f^{(y,w)}(\theta)w\cdot\partial_y\|_{s,a,\mathfrak{p}_0}\,. \end{gathered} \tag{B.26}$$

Since for a d-dimensional vector $\mathtt{v}$ one has $|\mathtt{v}|_1 \le d|\mathtt{v}|_\infty$ we get

$$|f^{(y,w)}(\theta)\cdot w|_1 \le d_1 \max_{v=y_1,\dots,y_{d_1}} \|f^{(v,w)}(\theta)\cdot w\|_\infty \le K(n,\mathfrak{p}_0)\|f^{(y,w)}(\theta)\cdot w\|_{s,\mathfrak{p}_0}\,. \tag{B.27}$$

The other bounds in (B.26) follow in the same way. The extension of the bounds for the Lipschitz norm is standard; see for instance [41]. Thus we obtain $|\Phi^{(y)}|_1 \le (r+\rho r_0)^2$ so that (B.22) follows.
(ii) The first d components of the map $(\theta_+, y_+, w_+) = \Phi(\theta, y, w)$ are $\theta_+ = \theta + f^{(\theta)}(\theta)$. If ρ is small enough we can apply Lemma A.3 in order to define an inverse map $h^{(\theta)}(\theta_+) \in W^{p,\infty}(\mathbb{T}^d_{s-\rho s_0})$ with $\|h^{(\theta)}\|_{s-\rho s_0,p} \le 2\|f^{(\theta)}\|_{s,p}$. Hence we set

$$\Psi^{(\theta)}(\theta_+) := \theta_+ + h^{(\theta)}(\theta_+)\,, \qquad \theta_+ \in \mathbb{T}^d_{s-\rho s_0}. \tag{B.28}$$

Regarding the other components we first solve y, w as functions of y_+, w_+, θ and then substitute (B.28). We have

$$\begin{aligned} w &= w_+ - f^{(w,0)}(\Psi^{(\theta)}(\theta_+)) \\ y &= (\mathbb{1} - f^{(y,y)}(\Psi^{(\theta)}(\theta_+)))^{-1}(y_+ - f^{(y,0)}(\Psi^{(\theta)}(\theta_+)) - f^{(y,w)}(\Psi^{(\theta)}(\theta_+))\cdot(w_+ - f^{(w,0)}(\Psi^{(\theta)}(\theta_+)))) \end{aligned}$$

which fixes the remaining components of h. The estimates on the norm of h follow by Lemma A.3 (*ii*) and by Lemma A.5. □

Lemma B.9. *Given any regular bounded vector field* $g \in \mathcal{B}$, $p \ge \mathfrak{p}_1$ *with* $|g|_{\vec{v},\mathfrak{p}_1} \le \mathtt{c}\rho$ *then for* $0 \le t \le 1$ *there exists* $f_t \in \mathcal{B}$ *such that the time$-t$ map of the flow of* g *is of the form* $\mathbb{1} + f_t$ *moreover we have* $|f_t|_{\vec{v},p} \le 2|g|_{\vec{v}_1,p}$ *where* $\vec{v}_1 = (\lambda, \mathcal{O}, s-\rho s_0, a, r)$.

Proof. The dynamical system associated with g is

$$\dot\theta = g^{(\theta,0)}(\theta), \tag{B.29a}$$
$$\dot y = g^{(y,0)}(\theta) + g^{(y,y)}(\theta)y + g^{(y,w)}(\theta)\cdot w, \tag{B.29b}$$
$$\dot w = g^{(w,0)}(\theta). \tag{B.29c}$$

We solve first (B.29a), then substitute into (B.29c) and finally substitute both into (B.29b) and hence the result follows by proving that the solution of (B.29a), with initial datum φ, has the form

$$\theta(t) = \varphi + h(t,\varphi),$$

with $h \in H^p(\mathbb{T}^d_{s-\rho s_0})$ a zero-average function. This latter statement follows by the standard theory of existence, uniqueness and smoothness w.r.t. the initial data. □

C Proof of Proposition 3.5

Proof. Given a tame vector field $F \in \mathcal{V}_{\vec{v},p}$ such that $F \in \mathcal{E}$ for all $\xi \in \mathcal{O}$, let us define

$$\mathfrak{A} := \mathfrak{N} + \mathfrak{R} := \Pi_K \Pi_{\mathcal{X}}([\Pi_{\mathcal{N}} F, \cdot]) + \Pi_K \Pi_{\mathcal{X}}([\Pi_{\mathcal{R}} F, \cdot]).$$

We note that $\mathfrak{N}, \mathfrak{R} : E^{(K)} \cap \mathcal{B}_{\mathcal{E}} \to E^{(K)} \cap \mathcal{X} \cap \mathcal{E}$.

Then the "approximate invertibility" of $\mathfrak{N}$ implies the "approximate invertibility" of $\mathfrak{A}$. Indeed let $\mathfrak{W} : E^{(K)} \cap \mathcal{X} \cap \mathcal{E} \to E^{(K)} \cap \mathcal{B}_{\mathcal{E}}$ be the "approximate right inverse" of $\mathfrak{N}$ defined in (3.3) and denote

$$\mathfrak{U} := \mathfrak{R}\mathfrak{W} : E^{(K)} \cap \mathcal{X} \cap \mathcal{E} \to E^{(K)} \cap \mathcal{X} \cap \mathcal{E}.$$

By (3.1) and (3.2) we have that $\mathfrak{U}$ is strictly upper triangular so $\mathfrak{U}^{\mathtt{b}} = 0$. Now we set $\mathfrak{B} = \mathfrak{N}\mathfrak{W} - \mathbb{1}$ which is "small" in the sense of (3.4). Then we have

$$(\mathfrak{N} + \mathfrak{R})\mathfrak{W}(\mathbb{1} + \mathfrak{U})^{-1} = (\mathbb{1} + \mathfrak{U} + \mathfrak{B})(\mathbb{1} + \mathfrak{U})^{-1} = \mathbb{1} + \mathfrak{B}(\mathbb{1} + \mathfrak{U})^{-1}, \quad (\mathbb{1} + \mathfrak{U})^{-1} = \sum_{j=0}^{\mathtt{b}-1} (-1)^j \mathfrak{U}^j . \tag{C.1}$$

Thus $\mathfrak{W}(\mathbb{1} + \mathfrak{U})^{-1}$ is an approximate inverse for $\mathfrak{A}$ in the sense that it is a true inverse for $\mathfrak{B} = 0$. Then for all $\xi \in \mathcal{O}$ let us set

$$\tilde{g} := \mathfrak{W}(\mathbb{1} + \mathfrak{U})^{-1} \Pi_K \Pi_X F. \tag{C.2}$$

As for the bounds we first notice that by (3.3) and (3.6) one has

$$|\mathfrak{U}X|_{\vec{v},p} \le K^{\mu_1+\nu+1}\big[\Theta_{\mathfrak{p}_1}|X|_{\vec{v},p} + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1} K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)|X|_{\vec{v},\mathfrak{p}_1}\big]. \tag{C.3}$$

Now we can prove inductively that

$$|\mathfrak{U}^j X|_{\vec{v},p} \le K^{j(\mu+\nu+1)}\Big[\Theta_{\mathfrak{p}_1}^j |X|_{\vec{v},p} + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1} K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)\mathtt{P}_j(\Theta_{\mathfrak{p}_1}, \Gamma_{\mathfrak{p}_1})|X|_{\vec{v},\mathfrak{p}_1}\Big], \tag{C.4}$$

where $\mathtt{P}_j(\Theta_{\mathfrak{p}_1}, \Gamma_{\mathfrak{p}_1})$ is a polynomial of degree $2(j-1)$ defined recursively as

$$\begin{aligned} &\mathtt{P}_1 := 1\,, \\ &\mathtt{P}_j(\Theta_{\mathfrak{p}_1}, \Gamma_{\mathfrak{p}_1}) := \Theta_{\mathfrak{p}_1}^{j-1} + 2\Theta_{\mathfrak{p}_1}(1+\Gamma_{\mathfrak{p}_1})\mathtt{P}_{j-1}(\Theta_{\mathfrak{p}_1}, \Gamma_{\mathfrak{p}_1}). \end{aligned} \tag{C.5}$$

Indeed for $j = 1$ this is exactly the bound (C.3); then assuming (C.4) to hold up to j we have

$$\begin{aligned} |\mathfrak{U}^{j+1}X|_{\vec{v},p} &= |\mathfrak{U}(\mathfrak{U}^j X)|_{\vec{v},p} \le K^{\mu+\nu+1}\big[\Theta_{\mathfrak{p}_1}|\mathfrak{U}^j X|_{\vec{v},p} + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1}K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)|\mathfrak{U}^j X|_{\vec{v},\mathfrak{p}_1}\big] \\ &\le K^{\mu+\nu+1}\Big(\Theta_{\mathfrak{p}_1}K^{j(\mu+\nu+1)}\Big[\Theta_{\mathfrak{p}_1}^j|X|_{\vec{v},p} + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1}K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)\mathtt{P}_j(\Theta_{\mathfrak{p}_1},\Gamma_{\mathfrak{p}_1})|X|_{\vec{v},\mathfrak{p}_1}\Big] \\ &\quad + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1}K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)K^{j(\mu+\nu+1)}\Big[\Theta_{\mathfrak{p}_1}^j + \Theta_{\mathfrak{p}_1}(1+2\Gamma_{\mathfrak{p}_1})\mathtt{P}_j(\Theta_{\mathfrak{p}_1},\Gamma_{\mathfrak{p}_1})\Big]|X|_{\vec{v},\mathfrak{p}_1}\Big) \\ &= K^{(j+1)(\mu+\nu+1)}\Big(\Theta^{j+1}|X|_{\vec{v},p} \\ &\quad + \big(\Theta_p(1+\Gamma_{\mathfrak{p}_1}) + \Theta_{\mathfrak{p}_1}K^{\alpha(p-\mathfrak{p}_1)}\Gamma_p\big)\big(\Theta_{\mathfrak{p}_1}^j + 2\Theta_{\mathfrak{p}_1}(1+\Gamma_{\mathfrak{p}_1})\mathtt{P}_j(\Theta_{\mathfrak{p}_1},\Gamma_{\mathfrak{p}_1})\big)|X|_{\vec{v},\mathfrak{p}_1}\Big) \end{aligned} \tag{C.6}$$

which is (C.4) for $j+1$ taking into account (C.5). Moreover, again by induction, the polynomials $\mathtt{P}_j$ satisfy the bound

$$|\mathtt{P}_j| \le 3^j \Theta_{\mathfrak{p}_1}^{j-1}(1+\Gamma_{\mathfrak{p}_1})^{j-1}\,, \tag{C.7}$$

uniformly in $\Theta_{\mathfrak{p}_1}, \Gamma_{\mathfrak{p}_1}$. Indeed for $j = 1$ this is trivial while assuming (C.7) up to j we have

$$\begin{aligned} |\mathtt{P}_{j+1}| &\overset{(C.5)}{\le} \Theta_{\mathfrak{p}_1}^j + 2\Theta_{\mathfrak{p}_1}(1+\Gamma_{\mathfrak{p}_1})|\mathtt{P}_j| \\ &\le \Theta_{\mathfrak{p}_1}^j + 2\Theta_{\mathfrak{p}_1}(1+\Gamma_{\mathfrak{p}_1})3^j\Theta_{\mathfrak{p}_1}^{j-1}(1+\Gamma_{\mathfrak{p}_1})^{j-1} = \Theta_{\mathfrak{p}_1}^j(1+2\cdot 3^j(1+\Gamma_{\mathfrak{p}_1})^j) \\ &\le (1+2\cdot 3^j)\Theta_{\mathfrak{p}_1}^j(1+\Gamma_{\mathfrak{p}_1})^j \le 3^{j+1}\Theta_{\mathfrak{p}_1}^j(1+\Gamma_{\mathfrak{p}_1})^j\,, \end{aligned} \tag{C.8}$$

where in the last inequality we used the fact that $1+2C^j \le C^{j+1}$ for $C \ge 3$. Summarizing we obtained

$$|\mathfrak{U}^j X|_{\vec{v},p} \le K^{j(\mu_1+\nu+1)} \Theta_{\mathtt{p}_1}^{j-1} \Big[\Theta_{\mathtt{p}_1} |X|_{\vec{v},p} + \big(\Theta_p (1+\Gamma_{\mathtt{p}_1}) + \Theta_{\mathtt{p}_1} K^{\alpha(p-\mathtt{p}_1)} \Gamma_p \big) 3^j (1+\Gamma_{\mathtt{p}_1})^{j-1} |X|_{\vec{v},\mathtt{p}_1} \Big], \tag{C.9}$$

so that (the second summand is zero for $j=0$)

$$|\mathfrak{W}\mathfrak{U}^j X|_{\vec{v},p} \le 4^j K^{j(\mu_1+\nu+1)+\mu_1} \Big[\Theta_{\mathtt{p}_1}^{j} |X|_{\vec{v},p} + \Theta_p \Theta_{\mathtt{p}_1}^{j-1} (1+\Gamma_{\mathtt{p}_1})^j |X|_{\vec{v},\mathtt{p}_1} + \Theta_{\mathtt{p}_1}^{j} K^{\alpha(p-\mathtt{p}_1)} (1+\Gamma_{\mathtt{p}_1})^j \Gamma_p |X|_{\vec{v},\mathtt{p}_1} \Big] \tag{C.10}$$

and finally

$$\begin{aligned} |\mathfrak{W}(1+\mathfrak{U})^{-1} X|_{\vec{v},p} \le\ & K^{\mu_1} (1+4K^{\mu_1+\nu+1}\Theta_{\mathtt{p}_1})^b |X|_{\vec{v},p} \\ & + K^{\mu_1}(1+\Gamma_{\mathtt{p}_1})(1+4K^{\mu_1+\nu+1}\Theta_{\mathtt{p}_1}(1+\Gamma_{\mathtt{p}_1}))^{b-1} \Theta_p |X|_{\vec{v},\mathtt{p}_1} \\ & + K^{\mu_1+\alpha(p-\mathtt{p}_1)} (1+4K^{\mu_1+\nu+1}\Theta_{\mathtt{p}_1}(1+\Gamma_{\mathtt{p}_1}))^b \Gamma_p |X|_{\vec{v},\mathtt{p}_1} \end{aligned} \tag{C.11}$$

where again the second summand is in fact zero if $b=0$. Therefore, since $\Theta_p \le \Gamma_p$, we obtain

$$|\mathfrak{W}(1+\mathfrak{U})^{-1} X|_{\vec{v},p} \le K^{(b+1)(\mu_1+\nu+1)} (1+\Theta_{\mathtt{p}_1}(1+\Gamma_{\mathtt{p}_1}))^b (1+\Gamma_{\mathtt{p}_1}) \Big[|X|_{\vec{v},p} + K^{\alpha(p-\mathtt{p}_1)} \Gamma_p |X|_{\vec{v},\mathtt{p}_1} \Big] \tag{C.12}$$

this concludes the proof of (2.46). The proof of (2.47) follows the same lines. Now we have defined a function $\tilde{g}$ on the set $\mathcal{O}$. In order to conclude the proof we need to extend this function to the whole $\mathcal{O}_0$. We know that regular vector fields in $E^{(K)}$ have a structure of Hilbert space w.r.t. the norm $|\cdot|_{s,a,\mathtt{p}_1}$ so we may apply Kirtzbraun Theorem in order to extend $\tilde{g}$ to a regular vector field in $E^{(K)}$ with the same Lipschitz norm, i.e. $|g|^{\mathrm{lip}}_{s,a,\mathtt{p}_1} \le \gamma^{-1} |\tilde{g}|_{\vec{v},p}$. As for the sup norm one clearly has

$$\sup_{\xi\in\mathcal{O}_0} |g|_{s,a,\mathtt{p}_1} \le \sup_{\xi\in\mathcal{O}_0} |\tilde{g}|_{s,a,\mathtt{p}_1} + \mathrm{diam}(\mathcal{O}_0) |\mathrm{g}|^{\mathrm{lip}}_{\mathrm{s,a,}\mathtt{p}_1}.$$

□

D Time analytic case

Theorem 2.25 does not make any assumptions on the analiticity parameters a_0, s_0 and relies on tame estimates in order to control the *high* Sobolev norm $\mathtt{p}_2 \ge \mathtt{p}_1 + \kappa_0 + \chi\kappa_2$. However if one makes the Ansatz that $s_0 > 0$ then we may take $\mathtt{p}_2 = \mathtt{p}_1$ and consequently have a simplified scheme, since we do not have to control the tameness constants in high norm but only the norm $|\cdot|_{\vec{v},\mathtt{p}_1}$. In order to do so we need to modify Definition 2.18 by substituting item *3.* with the following:

3′. For $K > 1$ there exists smoothing projection operators $\Pi_K : \mathcal{A}_{\vec{v},p} \to \mathcal{A}_{\vec{v},p}$ such that $\Pi_K^2 = \Pi_K$ and setting $\vec{v}_1 = (\gamma, \mathcal{O}, s+s_1, a, r)$, for $p_1 \ge 0$, one has

$$|\Pi_K F|_{\vec{v}_1,p+p_1} \le \mathtt{C} K^{p_1} e^{s_1 K} |F|_{\vec{v},p} \tag{D.13}$$

$$|F - \Pi_K F|_{\vec{v},p} \le \mathtt{C} K^{-p_1} e^{-s_1 K} |F|_{\vec{v}_1,p+p_1} \tag{D.14}$$

finally if $C_{\vec{v},p}(F)$ is any tameness constant for F then we may choose a tameness constant such that

$$C_{\vec{v},p+p_1}(\Pi_K F) \le \mathtt{C} K^{p_1} e^{s_1 K} C_{\vec{v}_1,p}(F) \tag{D.15}$$

We denote by $E^{(K)}$ the subspace where $\Pi_K E^{(K)} = E^{(K)}$.

Constraint D.1 (The exponents: analytic case). *We fix parameters $\varepsilon_0, \mathtt{R}_0, \mathtt{G}_0, \mu, \nu, \eta, \chi, \kappa_2$ such that the following holds.*

- $0 < \varepsilon_0 \le \mathtt{R}_0 \le \mathtt{G}_0$ *with* $\varepsilon_0 \mathtt{G}_0^3, \varepsilon_0 \mathtt{G}_0^2 \mathtt{R}_0^{-1} < 1$.
- *We have* $\mu, \nu \ge 0$, $1 < \chi < 2$, *finally setting* $\kappa_0 := \mu + \nu + 4$

$$\kappa_2 > \frac{2\kappa_0}{2-\chi}, \qquad \eta > \mu + (\chi - 1)\kappa_2 + 1, \tag{D.16}$$

- *there exists* $K_0 > 1$ *such that*

$$\log K_0 \ge \frac{1}{\log \chi} C, \tag{D.17}$$

with C *a given function of* $\mu, \nu, \eta, \kappa_2, s_0$ *(which goes to* ∞ *as* $s_0 \to 0$*) and moreover*

$$\mathtt{G}_0^2 \mathtt{R}_0^{-1} \varepsilon_0 K_0^{\kappa_0} \max(1, \mathtt{R}_0 \mathtt{G}_0 K_0^{\kappa_0 + (\chi-1)\kappa_2}) < 1, \tag{D.18a}$$

$$K_0^{\kappa_0 + (\chi-1)\kappa_2} e^{-\frac{s_0 K_0}{32}} \mathtt{G}_0 \varepsilon_0^{-1} \max\Big(1, \mathtt{R}_0\Big) \le 1, \tag{D.18b}$$

Now in order to state our result we define the good parameters and the changes of variables as in the general case but with $\mathfrak{p}_2 = \mathfrak{p}_1$, $\kappa_3 = \kappa_1 = \alpha = 0$. For clarity we restate our definition in this simpler case.

Definition D.2 (Homological equation). *Let* $\gamma > 0$, $K \ge K_0$, *consider a compact set* $\mathcal{O} \subset \mathcal{O}_0$ *and set* $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$ *and* $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$. *Consider a vector field* $F \in \mathcal{W}_{\vec{v}^o, p}$ *i.e.*

$$F = N_0 + G : \mathcal{O}_0 \times D_{a,p+\nu}(r) \times \mathbb{T}^d_s \to V_{a,p},$$

which is $C^{\mathtt{n}+2}$*-tame up to order* $q = \mathfrak{p}_1 + 2$. *We say* $\mathcal{O}$ *satisfies the* homological equation, *for* $(F, K, \vec{v}^o, \rho)$ *if the following holds.*

1. For all $\xi \in \mathcal{O}$ *one has* $F(\xi) \in \mathcal{E}$.

2. there exist a bounded regular vector field $g \in \mathcal{W}_{\vec{v}^o,p} \cap E^{(K)}$ *such that*

(a) $g \in \mathcal{B}_{\mathcal{E}}$ *for all* $\xi \in \mathcal{O}$,

(b) one has $|g|_{\vec{v}^o, \mathfrak{p}_1} \le \mathtt{C} |g|_{\vec{v}, \mathfrak{p}_1} \le \mathtt{c}\rho$ *and*

$$|g|_{\vec{v}, \mathfrak{p}_1} \le \gamma^{-1} K^{\mu} |\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v}, \mathfrak{p}_1} (1 + \gamma^{-1} C_{\vec{v}, p}(G)), \tag{D.19}$$

(c) setting $u := \Pi_K \Pi_{\mathcal{X}} (\mathrm{ad}(\Pi_{\mathcal{X}}^{\perp} F)[g] - F)$, *one has*

$$|u|_{\vec{v}, \mathfrak{p}_1} \le \varepsilon_0 \gamma^{-1} K^{-\eta + \mu} C_{\vec{v}, \mathfrak{p}_1}(G) |\Pi_K \Pi_{\mathcal{X}} G|_{\vec{v}, \mathfrak{p}_1}, \tag{D.20}$$

Remark D.3. *Note that if we take* $\mathfrak{p}_2 = \mathfrak{p}_1$ *then the second inequality in* (2.46) *as well as item 2(d) of Definition 2.23 follow from* (2.30) *and* (2.31).

Definition D.4 (Compatible changes of variables: analytic case). *Let the parameters in Constraint D.1 be fixed. Fix also* $\vec{v} = (\gamma, \mathcal{O}, s, a, r)$, $\vec{v}^o = (\gamma, \mathcal{O}_0, s, a, r)$ *with* $\mathcal{O} \subseteq \mathcal{O}_0$ *a compact set, parameters* $K \ge K_0, \rho < 1$. *Consider a vector field* $F = N_0 + G \in \mathcal{W}_{\vec{v}^o, p}$ *which is* $C^{\mathtt{n}+2}$*-tame up to order* $q = \mathfrak{p}_1 + 2$ *and such that* $F \in \mathcal{E} \quad \forall \xi \in \mathcal{O}$. *We say that a left invertible* $\mathcal{E}$*-preserving change of variables*

$$\mathcal{L}, \mathcal{L}^{-1} : \mathbb{T}^d_s \times D_{a, \mathfrak{p}_1}(r) \times \mathcal{O}_0 \to \mathbb{T}^d_{s+\rho s_0} \times D_{a - \rho a_0, \mathfrak{p}_1}(r + \rho r_0)$$

is compatible *with* $(F, K, \vec{v}, \rho)$ *if the following holds:*

(i) $\mathcal{L}$ *is "close to identity", i.e. denoting* $\vec{v}^o_1 := (\gamma, \mathcal{O}_0, s - \rho s_0, a - \rho a_0, r - \rho r_0)$ *one has*

$$\|(\mathcal{L} - \mathbb{1}) h\|_{\vec{v}^o_1, \mathfrak{p}_1} \le \mathtt{C} \varepsilon_0 K^{-1} \|h\|_{\vec{v}^o, \mathfrak{p}_1}. \tag{D.21}$$

(ii) $\mathcal{L}_*$ *conjugates the* $C^{\mathtt{n}+2}$*-tame vector field* F *to the vector field* $\hat{F} := \mathcal{L}_* F = N_0 + \hat{G}$ *which is* $C^{\mathtt{n}+2}$*-tame; moreover denoting* $\vec{v}_2 := (\gamma, \mathcal{O}, s - 2\rho s_0, a - 2\rho a_0, r - 2\rho r_0)$ *one may choose the tameness constants of* $\hat{G}$ *so that*

$$C_{\vec{v}_2,\mathtt{p}_1}(\hat{G}) \le C_{\vec{v},\mathtt{p}_1}(G)(1 + \varepsilon_0 K^{-1}), \tag{D.22}$$

(iii) $\mathcal{L}_*$ *"preserves the* $(\mathcal{N}, \mathcal{X}, \mathcal{R})$*-decomposition", namely one has*

$$\Pi_{\mathcal{N}}^{\perp}(\mathcal{L}_* \Pi_{\mathcal{N}} F) = 0, \qquad \Pi_{\mathcal{X}}(\mathcal{L}_* \Pi_{\mathcal{X}}^{\perp} F) = 0. \tag{D.23}$$

Then the result is the following.

Theorem D.5 (Abstract KAM: analytic case). *Let* N_0 *be a diagonal vector field as in Definition 2.20 and consider a vector field*

$$F_0 := N_0 + G_0 \in \mathcal{E} \cap \mathcal{W}_{\vec{v}_0,p} \tag{D.24}$$

which is $C^{\mathtt{n}+2}$*-tame up to order* $q = \mathtt{p}_1 + 2$*. Fix parameters* $\varepsilon_0, \mathtt{R}_0, \mathtt{G}_0, \mu, \nu, \eta, \chi, \kappa_2$ *satisfying Constraint D.1 and assume that*

$$\gamma_0^{-1} C_{\vec{v}_0,\mathtt{p}_1}(G_0) \le \mathtt{G}_0, \quad \gamma_0^{-1} C_{\vec{v}_0,\mathtt{p}_1}(\Pi_{\mathcal{N}}^{\perp} G_0) \le \mathtt{R}_0, \quad \gamma_0^{-1} |\Pi_{\mathcal{X}} G_0|_{\vec{v}_0,\mathtt{p}_1} \le \varepsilon_0. \tag{D.25}$$

For all $n \ge 0$ *we define recursively changes of variables* $\mathcal{L}_n, \Phi_n$ *and compact sets* $\mathcal{O}_n$ *as follows.*

Set $\mathcal{H}_{-1} = \mathcal{H}_0 = \Phi_0 = \mathcal{L}_0 = \mathbb{1}$*, and for* $0 \le j \le n-1$ *set recursively* $\mathcal{H}_j = \Phi_j \circ \mathcal{L}_j \circ \mathcal{H}_{j-1}$ *and* $F_j := (\mathcal{H}_j)_* F_0 := N_0 + G_j$*. Let* $\mathcal{L}_n$ *be any change of variables compatible with* $(F_{n-1}, K_{n-1}, \vec{v}_{n-1}, \rho_{n-1})$ *following Definition D.4, and* $\mathcal{O}_n$ *be any compact set*

$$\mathcal{O}_n \subseteq \mathcal{O}_{n-1}, \tag{D.26}$$

which satisfies the Homological equation for $((\mathcal{L}_n)_* F_{n-1}, K_{n-1}, \vec{v}_{n-1}^o, \rho_{n-1})$*. For* $n > 0$ *let* g_n *be the regular vector field defined in item (2) of Definition D.2 and set* Φ_n *the time-1 flow map generated by* g_n*.*

Then Φ_n *is left invertible and* $F_n := (\Phi_n \circ \mathcal{L}_n)_* F_{n-1} \in \mathcal{W}_{\vec{v}_n^o,p}$ *is* $C^{\mathtt{n}+2}$*-tame up to order* $q = \mathtt{p}_1 + 2$*. Moreover the following holds.*

(i) *Setting* $G_n = F_n - N_0$ *then*

$$\begin{aligned} \Gamma_{n,\mathtt{p}_1} &:= \gamma_n^{-1} C_{\vec{v}_n,\mathtt{p}_1}(G_n) \le \mathtt{G}_n, & \Theta_{n,\mathtt{p}_1} &:= \gamma_n^{-1} C_{\vec{v}_n,\mathtt{p}_1}(\Pi_{\mathcal{N}}^{\perp} G_n) \le \mathtt{R}_n, \\ \delta_n &:= \gamma_n^{-1} |\Pi_{\mathcal{X}} G_n|_{\vec{v}_n,\mathtt{p}_1} \le K_0^{\kappa_2} \varepsilon_0 K_n^{-\kappa_2}, & |g_n|_{\vec{u}_n,\mathtt{p}_1} &\le K_0^{\kappa_2} \varepsilon_0 \mathtt{G}_0 K_{n-1}^{-\kappa_2+\mu+1}, \end{aligned} \tag{D.27}$$

where $\vec{u}_n = (\gamma_n, \mathcal{O}_n, s_n + 12\rho_n s_0, a_n + 12\rho_n a_0, r_n + 12\rho_n r_0)$*.*

(ii) *The sequence* $\mathcal{H}_n$ *converges for all* $\xi \in \mathcal{O}_0$ *to some change of variables*

$$\mathcal{H}_\infty = \mathcal{H}_\infty(\xi) : D_{a_0,p}(s_0/2, r_0/2) \longrightarrow D_{\frac{a_0}{2},p}(s_0, r_0). \tag{D.28}$$

(iii) *Defining* $F_\infty := (\mathcal{H}_\infty)_* F_0$ *one has*

$$\Pi_{\mathcal{X}} F_\infty = 0 \quad \forall \xi \in \mathcal{O}_\infty := \bigcap_{n \ge 0} \mathcal{O}_n \tag{D.29}$$

and

$$\gamma_0^{-1} C_{\vec{v}_\infty,\mathtt{p}_1}(\Pi_{\mathcal{N}} F_\infty - N_0) \le 2\mathtt{G}_0, \quad \gamma_0^{-1} C_{\vec{v}_\infty,\mathtt{p}_1}(\Pi_{\mathcal{R}} F_\infty) \le 2\mathtt{R}_0$$

with $\vec{v}_\infty := (\gamma_0/2, \mathcal{O}_\infty, s_0/2, a_0/2)$*.*

Proof. The proof of Theorem D.5 is essentially identical to the one of Theorem 2.25. We give a sketch for completeness. The induction basis is trivial with $g_0 = 0$. Assuming (D.27) up to n we prove the inductive step using the "KAM step" of Proposition 5.1 with $\mathfrak{p}_1 = \mathfrak{p}_2$ and $\alpha = \kappa_3 = 0$ and the bound (5.13) substituted by

$$\begin{aligned}\delta_+ \le \mathtt{C}\Gamma_{\mathfrak{p}_1}\big(\delta^2\Gamma_{\mathfrak{p}_1}^2 K^{2\mu+2\nu+4} + \delta\varepsilon_0 K^{\mu-\eta}\big) + e^{-2\rho s_0 K} K^{\mu+\nu+1-(\mathfrak{p}_2-\mathfrak{p}_1)}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1}\big) \\ + \Gamma_{\mathfrak{p}_1} K^{\mu+\nu+1-(\mathfrak{p}_2-\mathfrak{p}_1)} e^{-2\rho s_0 K}\big(\Theta_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Theta_{\mathfrak{p}_1} + K^{\alpha(\mathfrak{p}_2-\mathfrak{p}_1)}\delta(\Gamma_{\mathfrak{p}_2} + \varepsilon_0 K^{\kappa_3}\Gamma_{\mathfrak{p}_1})\big)\,.\end{aligned} \tag{D.30}$$

Bound (D.30) follows using the smoothing properties (D.13), (D.14) in item (3′), in the equation (5.39).

First of all we note that

$$\rho_n^{-1} K_n^{\mu+\nu+3}\Gamma_{n,\mathfrak{p}_1}\delta_n \le \mathtt{c}, \tag{D.31}$$

which, by the inductive hypothesis and (2.52) reads

$$2^{n+9} K_0^{(\mu+\nu+3-\kappa_2)\chi^n}\mathtt{G}_0\varepsilon_0 K_0^{\kappa_2} \le \mathtt{c}; \tag{D.32}$$

this is true since by (D.16) and (D.17) the left hand side (D.32) is decreasing in n so that (D.31) follows form

$$K_0^{\mu+\nu+4}\mathtt{G}_0\varepsilon_0 < 1$$

which is indeed implied by (D.18a) because $\mathtt{G}_0 \ge \mathtt{R}_0$.

Hence we can apply the "KAM step" to $F_n := (\Phi_n \circ \mathcal{L}_n)_* F_{n-1} \in \mathcal{W}_{\vec{v}_n^0,\mathfrak{p}_2}$ which is a $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We fix $(K_n, \gamma_n, a_n, s_n, r_n, \rho_n, \mathcal{O}_n) \rightsquigarrow (K, \gamma, a, s, r, \rho, \mathcal{O})$, $\Gamma_{n,p} \rightsquigarrow \Gamma_p$, $\Theta_{n,p} \rightsquigarrow \Theta_p$, $\delta_n \rightsquigarrow \delta$, $(\gamma_{n+1}, a_{n+1}, s_{n+1}, r_{n+1}, \rho_{n+1}, \mathcal{O}_{n+1}) \rightsquigarrow (\gamma_+, a_+, s_+, r_+, \rho_+, \mathcal{O}_+)$. The KAM steps produces a bounded regular vector field g_{n+1} and a left invertible change of variables $\Phi_{n+1} = \mathbb{1} + f_{n+1}$ such that $F_{n+1} := (\Phi_{n+1} \circ \mathcal{L}_n)_* F_n \in \mathcal{W}_{\vec{v}_{n+1}^0,\mathfrak{p}_2}$ is $C^{\mathtt{n}+2}$-tame up to order $q = \mathfrak{p}_2 + 2$. We now verify that the bounds (D.27) hold with $\Gamma_{n+1,\mathfrak{p}_1} \rightsquigarrow \Gamma_{+,\mathfrak{p}_1}$, $\Theta_{n+1,\mathfrak{p}_1} \rightsquigarrow \Theta_{+,\mathfrak{p}_1}$, , $\delta_{n+1} \rightsquigarrow \delta_+$.

Let us prove **(i)**, the others follow exactly as in Theorem (2.25).

By substituting into (5.8) we immediately obtain the bounds for g_{n+1} of (D.27).

Now we recall that, by definition

$$\frac{\gamma_n}{\gamma_{n+1}} = 1 + \frac{1}{2^{n+3}-1}.$$

We use (5.11) together with the inductive hypotheses to obtain

$$\Gamma_{n+1,\mathfrak{p}_1} \le \Big(1 + \frac{1}{2^{n+3}-1}\Big)\mathtt{G}_n + 2\varepsilon_0 K_n^{-1}\mathtt{G}_0 + \mathtt{C}K_n^{\mu-\kappa_2}\mathtt{G}_0\varepsilon_0 K_0^{\kappa_2}(K_n^{\nu+1}\mathtt{G}_0 + \varepsilon_0 K_n^{-\eta}) \le \mathtt{G}_{n+1}\,,$$

which follow by requiring

$$\max(2^n K_n^{-1}\varepsilon_0, 2^n K_n^{\mu+\nu+1-\kappa_2} K_0^{\kappa_2}\mathtt{G}_0\varepsilon_0, 2^n K_n^{-\eta-\kappa_2+\mu} K_0^{\kappa_2}\varepsilon_0) \le \mathtt{c}\,,$$

and as before this follows by (D.16) and (D.18a).

Regarding $\Theta_{n+1,\mathfrak{p}_1}$, using (5.12) we get

$$\Theta_{n+1,\mathfrak{p}_1} \le \Big(1 + \frac{1}{2^{n+3}-1}\Big)\mathtt{R}_n + 2\varepsilon_0 K_n^{-1}\mathtt{R}_0 + \mathtt{C}K_n^{\mu+\nu+1-\kappa_2}\mathtt{G}_0^2\varepsilon_0 K_0^{\kappa_2} + K_n^{-\eta-\kappa_2+\mu}\mathtt{G}_0\varepsilon_0^2 K_0^{\kappa_2} \le \mathtt{R}_{n+1}.$$

which again follows from (D.16) and (D.18a).

For $\delta_{n+1} \rightsquigarrow \delta_+$, we apply (D.30) with $\mathfrak{p}_2 = \mathfrak{p}_1$, $\alpha = \kappa_3 = 0$ and get

$$\begin{aligned}\delta_{n+1} \le & \mathtt{C}\mathtt{G}_0\Big(\varepsilon_0^2 K_0^{\kappa_2}(\mathtt{G}_0^2 K_0^{\kappa_2} K_n^{2\mu+2\nu+4-2\kappa_2} + K_n^{\mu-\eta-\kappa_2}) + \\ & e^{-2\rho_n s_0 K_n}(\mathtt{R}_0 K_n^{\mu+\nu+1} + \varepsilon_0 K_0^{\kappa_2}\mathtt{G}_0 K_n^{\mu+\nu+1-\kappa_2})\Big) + e^{-2\rho_n s_0 K_n}\mathtt{R}_0 K_n \le \varepsilon_0 K_0^{\kappa_2} K_n^{-\chi\kappa_2}\end{aligned}$$

Now

$$\mathtt{CG}_0\Big(\varepsilon_0^2 K_0^{\kappa_2}(\mathtt{G}_0^2 K_0^{\kappa_2} K_n^{2\mu+2\nu+4-2\kappa_2} + K_n^{\mu-\eta-\kappa_2}) \le \frac{1}{2}\varepsilon_0 K_0^{\kappa_2} K_n^{-\chi\kappa_2}$$

by (D.16) and (D.18a). As for the second term, since $s_0 > 0$ all the summands are decreasing in n provided that K_0 is large enough.

□

References

[1] J. Moser, Convergent series expansions for quasi–periodic motions, Math. Ann. 169 (1967) 136–176.

[2] A. Kolmogorov, On conservation of conditionally periodic motions for a small change in hamil- ton's function, Dokl. Akad. Nauk SSSR 98 (1954) 527–530.

[3] V. I. Arnold, Small denominators and problems of stability of motion in classical celestial mechanics, Russian Math. Surveys 18 (1963) 85–193.

[4] J. Moser, Rapidly convergent iteration method and non-linear partial differential equations - i, Ann. Sc. Norm. Sup. Pisa 20 (2) (1966) 265–315.

[5] J. Pöschel, A lecture on the classical KAM theorem, in: Proc. Symp. Pure Math., Vol. 69, 2001, pp. 707–732.

[6] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, in: Number theory and dynamical systems (York, 1987), Vol. 134, Cambridge Univ. Press, Cambridge, 1989, pp. 5–8.

[7] M. Sevryuk, The reversible context 2 in KAM theory: the first steps, Regul. Chaotic Dyn 16 (1-2) (2011) 24–38.

[8] E. Zehnder, Generalized implicit function theorems with applications to some small divisors problems i-ii, Comm. Pure Appl. Math. (29) (1976) 49–113.

[9] J. Fejoz, Periodic and quasi-periodic motions in the many-body problem. mémoire d'habilitation à diriger des recherches. (2010).

[10] S. Kuksin, Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum, Funktsional Anal. i Prilozhen. 21 (3) (1987) 22–37.

[11] C. E. Wayne, Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory, Comm. Math. Phys. 127 (3) (1990) 479–528.

[12] W. Craig, C. E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure Appl. Math. 46 (11) (1993) 1409–1498.

[13] J. Bourgain, Construction of quasi-periodic solutions for hamiltonian perturbations of linear equations and applications to nonlinear pde, Internat. Math. Res. Notices.

[14] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comment. Math. Helv. 71 (2) (1996) 269–296. `doi:10.1007/BF02566420`.
URL `http://dx.doi.org/10.1007/BF02566420`

[15] J. Pöschel, A KAM-theorem for some nonlinear partial differential equations, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 23 (1) (1996) 119–148.
URL `http://www.numdam.org/item?id=ASNSP_1996_4_23_1_119_0`

[16] S. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 (1) (1996) 149–179. doi:10.2307/2118656.

[17] J. Bourgain, Periodic solutions of nonlinear wave equations, in: Harmonic Analysis and Partial Differential Equations (Chicago, IL, 1996), Chicago Lectures in Math, Univ. Chicago Press, 1999, pp. 69–97.

[18] S. Kuksin, A KAM theorem for equations of the Korteweg-de Vries type, Rev. Math. Phys. 10 (3) (1998) 1–64.

[19] J. Bourgain, Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. of Math. (2) 148 (2) (1998) 363–439. doi:10.2307/121001.

[20] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Comm. Math. Phys. 211 (2000) 497–525.

[21] J. Bourgain, Green's function estimates for lattice Schrödinger operators and applications, Vol. 158 of Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 2005.

[22] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2) (2006) 343–372. doi:10.1007/s00220-005-1497-0.

[23] L. H. Eliasson, S. B. Kuksin, KAM for the nonlinear Schrödinger equation, Ann. of Math. (2) 172 (1) (2010) 371–435. doi:10.4007/annals.2010.172.371.

[24] M. Berti, P. Bolle, Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential, Nonlinearity 25 (2012) 2579–2613.

[25] M. Berti, P. Bolle, Quasi-periodic solutions for Schrödinger equations with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential, J. European Math. Society 15 (2013) 229–286.

[26] J. Geng, J. You, X. Xu, KAM tori for cubic NLS with constant potentials, preprint.

[27] M. Procesi, C. Procesi, A KAM algorithm for the non–linear Schrödinger equation, Advances in Math. 272 (2015) 399–470.

[28] M. Berti, L. Corsi, M. Procesi, An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds, Comm. Math. Phys. 334 (3) (2015) 1413–1454.

[29] L. Corsi, E. Haus, M. Procesi, A KAM result on compact Lie groups, Acta Applicandae Mathematicae 137 (2015) 41–59.

[30] B. Grébert, E. Paturel, KAM for the Klein-Gordon equation on $\mathbb{S}^d$, Boll. Unione Mat Ital. 9 (2) (2016) 237–288.

[31] T. Kappeler, J. Pöschel, KAM and KdV, Vol. 45, Springer-Verlag, Berlin, 2003.

[32] J. Zhang, M. Gao, X. Yuan, KAM tori for reversible partial differential equations, Nonlinearity 24 (2011) 1189–1228.

[33] J. Liu, X. Yuan, A KAM Theorem for Hamiltonian partial differential equations with unbounded perturbations, Comm. Math. Phys, 307 (2011) 629–673.

[34] M. Berti, L. Biasco, M. Procesi, KAM theory for the Hamiltonian derivative wave equation, Annales Scientifiques de l'ENS 46 (2) (2013) 299–371.

[35] M. Berti, P. Biasco, M. Procesi, KAM theory for reversible derivative wave equations, Archive for Rational Mechanics and Analysis 212 (3) (2014) 905–955.

[36] G. Iooss, P. Plotnikov, J. Toland, Standing waves on an infinitely deep perfect fluid under gravity, Arch. Ration. Mech. Anal. 177 (3) (2005) 367–478.

[37] P. Baldi, Periodic solutions of forced Kirchhoff equations, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. 8 (5) (2009) 117–141.

[38] P. Baldi, Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type, Ann. I. H. Poincaré (C) Anal. Non Linéaire 30 (2013) 33–77.

[39] P. Baldi, M. Berti, R. Montalto, KAM for quasi-linear and fully nonlinear forced KdV, Math. Ann. 359 (2014) 471–536.

[40] P. Baldi, M. Berti, R. Montalto, KAM for autonomous quasilinear perturbations of KdV, arXiv:1404-3125 (2014).

[41] R. Feola, M. Procesi, Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations, Journal of Differential Equations.

[42] R. Montalto, Quasi-periodic solutions of forced Kirchoff equation, preprint, arXiv:1602.05093.

[43] M. Berti, R. Montalto, Quasi-periodic standing wave solutions of gravity-capillary water waves, preprint (2016).

[44] F. Giuliani, Quasi-periodic solutions for quasi-linear generalized KdV equations, preprint 2016, arXiv:1607.02583.

[45] M. Berti, P. Bolle, A Nash-Moser approach to KAM theory, to appear on Field Institute Communications.

[46] R. Feola, M. Procesi, KAM for quasi-linear autonomous NLS, in preparation.

[47] T. Kappeler, R. Montalto, Canonical coordinates with tame estimates for the defocusing NLS equation on the circle, preprint, arXiv:1607.04454.

[48] M. Berti, L. Biasco, Branching of Cantor manifolds of elliptic tori and applications to PDEs, Comm. Math. Phys. 305 (3) (2011) 741–796.

[49] J. Bourgain, Global solutions of nonlinear Schrödinger equations, Vol. 46 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 1999.

[50] M. Berti, P. Bolle, Quasi-periodic solutions for autonomous NLW on $\mathbb{T}^d$ with a multiplicative potential, in preparation.

[51] M. Procesi, C. Procesi, Reducible quasi-periodic solutions for the non–linear Schrödinger equation, BUMI 2 (2016) 189–236.

[52] R. Feola, Quasi-periodic solutions for fully nonlinear NLS, Ph.D. thesis, University ”La Sapienza” (2016).