Reducibility of first order linear operators on tori via Moser's theorem

R. Feola*, F. Giuliani† , R. Montalto**, M. Procesi†

* SISSA, Trieste, rfeola@sissa.it

† Roma Tre, fgiuliani@mat.uniroma3.it, procesi@mat.uniroma3.it

** University of Milan, riccardo.montalto@unimi.it, Via Saldini 50, 20133, Milan, Italy

Abstract

In this paper we prove reducibility of a class of first order, quasilinear, quasi-periodic time dependent PDEs on the torus

$$\partial_t u + \zeta \cdot \partial_x u + a(\omega t, x) \cdot \partial_x u = 0, \quad x \in \mathbb{T}^d,\, \zeta \in \mathbb{R}^d,\, \omega \in \mathbb{R}^\nu.$$

As a consequence we deduce a stability result on the associated Cauchy problem in Sobolev spaces. By the identification between first order operators and vector fields this problem can be formulated as the problem of finding a change of coordinates which conjugates a weakly perturbed constant vector field on $\mathbb{T}^{\nu+d}$ to a constant diophantine flow. For this purpose we generalize Moser's straightening theorem: considering smooth perturbations we prove that the corresponding *straightening* torus diffeomorphism is smooth, under the assumption that the perturbation is small only in some given Sobolev norm and that the initial frequency belongs to some Cantor-like set. In view of applications in KAM theory for PDEs we provide also *tame* estimates on the change of variables.

Contents

1 Introduction and main results

In the last years there has been a lot of advances in the study of KAM theory and almost global existence for classes of quasi-linear and fully nonlinear PDEs on the circle.

In these results, the main issue is to prove the reducibility of quasi-periodically time dependent linear operators. For instance these operators arise from linearizing such PDEs at small quasi-periodic approximate solutions, whose study is required by Nash-Moser type schemes, together with a careful quantitative analysis.

Given a linear PDE with coefficients which depend on time in a quasi-periodic way, we say that it is reducible if there exists a bounded change of variables depending quasi-periodically on time (say mapping $H^s \to H^s$ for all times), which makes constant its coefficients. This is a problem which is interesting on itself and has been studied for PDEs both on compact and non-compact domains. We mention among others [15], [8],[9],[24],[31],[10],[33] for the case of linear equations. Regarding reducible KAM theory for non linear PDEs the literature is quite vast, we mention the classical papers [27],[41],[29],[14] for PDEs on the circle, [21],[17],[22],[37],[16] for PDEs on $\mathbb{T}^n$. These works all deal with PDEs with bounded nonlinearities. Regarding unbounded cases we mention [28], [30], [11] for semilinear PDEs and [5],[6],[20],[23],[13],[4] for the quasilinear case.

In this paper we discuss the reducibility of a class of quasi-periodic time dependent transport equations on any multi-dimensional torus of the form

$$\partial_t u + \zeta \cdot \partial_x u + a(\omega t, x) \cdot \partial_x u = 0, \quad x \in \mathbb{T}^d, \zeta \in \mathbb{R}^d, \tag{1.1}$$

where $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$, (ω, ζ) varies in a bounded domain $\mathcal{O}_0 \subset \mathbb{R}^{\nu+d}$ and $a : \mathbb{T}^{\nu+d} \to \mathbb{R}^d$ is a C^∞ function. The main result of the paper is the following.

Theorem 1. *There exists $s_1 \in \mathbb{N}$ large enough and $\delta(s_1) \in (0,1)$ small enough such that if $\|a\|_{s_1} \equiv \|a\|_{H^{s_1}(\mathbb{T}^{\nu+d}, \mathbb{R}^d)} \le \delta(s_1)$ then there exists a Borel set $\mathcal{O}_\infty \subseteq \mathcal{O}_0$ such that the following holds. For any $(\omega, \zeta) \in \mathcal{O}_\infty$ there exists a quasi-periodic family of linear operators $\Psi(\omega t)$ which are bounded and invertible $H^s(\mathbb{T}^d, \mathbb{R}) \to H^s(\mathbb{T}^d, \mathbb{R})$, for any $s \ge 0$ such that $u(t,x)$ is a solution of the PDE (1.1) if and only if $v := \Psi(\omega t)^{-1}[u]$ solves the PDE with constant coefficients*

$$\partial_t v + m_\infty \cdot \partial_x v = 0$$

where m_∞ is a constant vector in $\mathbb{R}^d$ satisfying $|m_\infty - \zeta| = O(\|a\|_{s_1})$ uniformly with respect to $(\omega, \zeta) \in \mathcal{O}_0$. Moreover, the Lebesgue measure of the set $\mathcal{O}_0 \setminus \mathcal{O}_\infty$ converges to 0 when $\|a\|_{s_1}$ goes to zero.

As a direct consequence, the Sobolev norms of the solutions of the Cauchy problem associated to (1.1) are controlled uniformly in time. A more detailed and quantitative version of these results is stated in Theorem 4.1. From Theorem 1 we also deduce a series of corollaries, such as the fact that $(\omega, \zeta) \in \mathcal{O}_\infty$ there is no growth of the Sobolev norms (see Theorem 4.1).

Progress for the reducibility of PDEs on high dimensional compact manifolds is very recent, see [33],[10], and as far as we know there is still no general theory. In our case the key ingredient for proving the reducibility of the equation 1.1 is to use the identification between derivation operators and vector fields in order to change the PDE reduction problem into a corresponding *straightening* problem for a vector field. The idea of analyzing a first order linear differential operator through its associated vector field is the so called method of characteristics, which is the classical way in which a first order linear PDE is reduced to a (possibly non-linear) ODE. A similar approach is used in [10] for reducibility of a class of quadratic second order perturbations of the Harmonic oscillator on $\mathbb{R}^N$. The main idea is to exploit the identification between classical Hamiltonian functions (with N-degrees of freedom) and their quantum analogs so that the reducibility theorem for an operator corresponds to a finite dimensional KAM theorem.
Although our strategy is strongly taylored for the particular class of operators (1.1) we are confident that our result can find applications for different models. For instance Theorem 4.1

and Corollary 4.2 will be applied in [19] and in [32], in order to reduce to constant coefficients the highest order term of classes of linear operators. Note that in view of the application to [19], one needs also to study the dependence of the transformation $\Psi^{(\infty)}$ with respect to an additional parameter λ which is in a Banach space. This is done in Corollary 3.3.

The reducibility of the PDE (1.1) amounts to showing that there exists a linear time dependent operator $\Psi(\omega t)$ acting on $C^\infty(\mathbb{T}^d,\mathbb{R})$ and a map $\mathbb{R}^\nu\times\mathbb{R}^d\mapsto\mathbb{R}^d$, $(\omega,\zeta)\mapsto m(\omega,\zeta)$ such that

$$\Psi^{-1}\Psi_t+\Psi^{-1}((\zeta+a(\omega t,x))\cdot\partial_x\Psi)\cdot\partial_x=m(\omega,\zeta)\cdot\partial_x\,. \tag{1.2}$$

We look for the operator $\Psi(\omega t)$ of the special form

$$\Psi(\omega t)v:=v\circ\Psi(\omega t)\,,\quad \Psi(\omega t):x\mapsto x+\beta(\omega t,x)\,,$$

where $\beta\in C^\infty(\mathbb{T}^{d+\nu},\mathbb{T}^d)$. Here with an abuse of notation we are representing with the same symbol Ψ the time dependent torus diffeomorphism and its action on functions. In this way our reduction equation (1.2) becomes

$$\Psi^{-1}(\omega\cdot\partial_\varphi\beta+(\zeta+a(\varphi,x))\cdot(1+\partial_x\beta))=m_\infty(\omega,\zeta) \tag{1.3}$$

with unknowns $m_\infty\in\mathbb{R}^d$ and $\beta\in C^\infty(\mathbb{T}^\nu\times\mathbb{T}^d,\mathbb{R}^d)$.
One can directly see that solving the equation 1.3 is equivalent to finding a change of variables of the form $\Psi:(\varphi,x)\to(\varphi,x+\beta(\varphi,x))$ which straightens the vector field

$$X:=\omega\cdot\frac{\partial}{\partial\varphi}+(\zeta+a(\varphi,x))\cdot\frac{\partial}{\partial x}\,,\qquad\text{namely}\quad \Psi_*X=\omega\cdot\frac{\partial}{\partial\varphi}+m_\infty(\omega,\zeta)\cdot\frac{\partial}{\partial x}. \tag{1.4}$$

Under appropriate hypotheses on the size of a and the non-degereracy of (ω,ζ) the existence of such a change of variables follows by the classical Arnold-Moser theorem, which we now describe in the more general setting of vector fields on the torus $\mathbb{T}^N$. We set X_α to be the linear vector field with frequency α, namely the vector field which generates the flow $\theta\mapsto\theta+\alpha t$ with $\alpha\in\mathbb{R}^N$. Then the result can be stated as follows: for all diophantine frequencies α and for any appropriately small vector field f there exists a Lipschitz function $\alpha\to\lambda(\alpha)$ and a change of variables on $\mathbb{T}^N$ which conjugates $X_{\alpha+\lambda(\alpha)}+f$ to X_α and thus *straightens* the vector field[1] This result was first proved by Arnold in [2, 1] in the case of analytic vector fields, then by Moser in [34] for finitely differentiable vector fields. The problem was further investigated by Rüssmann [38, 39], Pöschel [35, 36], Herman [26, 25], Salamon [40] with the purpose of giving optimal bounds on the regularity hypotheses needed on the vector field. See also [42], [18],[31].
This classical results are not sufficient to prove Theorem 1, since we need a good control of *all* the Sobolev norms of the change of variables by requiring only smallness assumptions on fixed low norm of the perturbation. To this purpose we prove a *tame* version of Moser's theorem, which can be stated as follows.

Theorem 2 (Tame Moser Theorem). *Consider a vector field on $\mathbb{T}^N$ of the form*

$$X_0:=(\xi+f_0(\theta))\cdot\frac{\partial}{\partial\theta}\,,\quad f_0\in C^\infty(\mathbb{T}^N,\mathbb{R}^N).$$

There exist $s_1\in\mathbb{N}$ large enough, and $\eta_\star=\eta_\star(s_1)>0$ small enough such that if $\|f_0\|_{s_1}^{\gamma,\mathcal{O}_0}\le\eta_\star$ then there exists a Borel set $\mathcal{O}_\infty\subseteq\mathcal{O}_0$ such that the following holds. For any $\xi\in\mathcal{O}_\infty$ there exists $\beta\in C^\infty(\mathbb{T}^N,\mathbb{R}^N)$ such that $\mathbb{T}^N\to\mathbb{T}^N:\theta\mapsto\widehat\theta=\theta+\beta(\theta)$ is a diffeomorphism and

$$\Psi_*X_0:=(\Psi)^{-1}\big(\xi+f_0+(\xi+f_0)\cdot\partial_\theta\beta\big)\cdot\frac{\partial}{\partial\widehat\theta}=\alpha_\infty(\xi)\cdot\frac{\partial}{\partial\widehat\theta}\,.$$

[1] It is easily seen that $\alpha\to\alpha+\lambda(\alpha)$ can be extended to a Lipeomorphism on $\mathbb{R}^N$ so that the statement can be equivalently rephrased as the conjugation between $X_\xi+f$ and $X_{\alpha(\xi)}$, where the initial frequency $\xi\in\mathbb{R}^N$ is chosen so that the final frequency $\alpha(\xi)$ is diophantine.

where $\alpha_\infty(\xi)$ *is a constant vector in* $\mathbb{R}^N$ *satisfying* $|\alpha_\infty(\xi) - \mathrm{Id}_{\mathbb{R}^N}| = O(\|f_0\|_{s_1})$ *uniformly with respect to* $\xi \in \mathcal{O}_0$. *For any* $\xi \in \mathcal{O}_\infty$, *the function* β *satisfies the tame estimate* $\|\beta\|_s \leq C(s,N)\|f_0\|_{s+\sigma}$, *for some* $\sigma \in \mathbb{N}$, $C(s,N) > 0$ *large enough, for any* $s \geq 0$. *Moreover, the Lebesgue measure of the set* $\mathcal{O}_0 \setminus \mathcal{O}_\infty$ *converges to* 0 *when* $\|f_0\|_{s_1}$ *goes to zero.*

Once we have proved Theorem 2, then Theorem 1 follows directly by setting $N = \nu + d$, $\theta = (\varphi, x) \in \mathbb{T}^{\nu+d}$, $\xi = (\omega, \zeta)$ and $\mathbb{T}^{\nu+d} \to \mathbb{R}^{\nu+d}$, $(\varphi, x) \mapsto f_0(\varphi, x) := (0, a(\varphi, x))$. A more detailed and quantitative version of the statement above is given in Theorem 3.1.
The main point is that we consider C^∞ vector fields f_0 which are small in some *low* Sobolev norm with no further quantitative information on the higher Sobolev norms. Then we prove that the change of variables predicted by Moser's theorem is in fact C^∞ and we give a very good control of the Sobolev norms of this diffeomorphism in terms of the Sobolev norms of f_0. Note that in order to obtain this result we do not rely on approximation by analytic fuctions, as in most of the literature, instead we approach the problem in the spirit of the Nash-Moser theory where one uses interpolation and smoothing estimates in order to control the loss of regularity due to the presence of small divisors.
As explained before, our more general formulation of the Arnold-Moser theorem is well suited for applications to the reducibility of linear first-order operators on $H^s(\mathbb{T}^N, \mathbb{R})$, which are the main motivation of our work. In any case, as far as we know, the tame estimates on the change of variables were not previously known so that our result may have some interest also in the context of ODEs.

Reducibility of quasi-linear operators and lack of dispersion. The reducibility for quasi-linear, quasi-periodic time dependent operators has been recently developed in the framework of the KAM theory for PDEs.

In one-space dimension, Baldi-Berti-Montalto [5] proposed a method for proving the reducibility of a class of linear operators arising from forcing the Airy equation $u_t + u_{xxx} = 0$ by a quasi-linear, quasi-periodic in time perturbation. Actually their strategy finds applications for several other models in one spatial dimension. Such a procedure is split in two steps: a regularization procedure, in which the operator is reduced to a diagonal one plus a bounded remainder, and a KAM scheme which completely diagonalizes the linear operator by reducing *quadratically* the size of the remainders.
The reduction procedure that one has to perform depends strongly on the linear dispersion law of the PDE, namely on the function $j \mapsto \lambda(j)$ where $\mathrm{i}\lambda_j$ the eigenvalue of the linearized operator at the origin associated to the eigenfunction $e^{\mathrm{i}jx}$, $j \in \mathbb{Z}$. For instance when the dispersion law is superlinear, as in the KdV case where $\lambda_j = j^3$, in order to reduce to constant coefficients the leading term one applies a torus diffeomorphism of the form $x \to x + \beta(\omega t, x)$ where β solves

$$(1 + a(\varphi, x))(1 + \beta_x)^3 = m(\varphi)\,, \tag{1.5}$$

which can be solved directly by integration.
Comparing this equation with (1.3) one sees that here ∂_φ does not appear and the right hand side still depends on φ. This is due to the fact that the leading derivative is the spatial one, ∂_{xxx}.
When there is no dispersion, namely the dispersion law is linear as in the case we deal with, where $\lambda_j = j$, the equation that defines the correct change of variables is harder to solve. The time and the space play the same role and one has to deal with transport-like equations as (1.3). This was our first motivation for proving Theorem 1.

We remark that in [4] a transport equation similar to equation (1.3) appears in the reducibility of the linearized operator. In this case the corresponding vector field is more degenerate, in the sense that, in the basis $\partial/\partial\varphi$, $\partial/\partial x$, it has the form $(\omega, O(\varepsilon))$, instead of $(\omega, \zeta + O(\varepsilon))$ as in (1.4). In [4] the authors reduce the problem to the study of a nonlinear ODE and then they apply a Nash-Moser Hormander Theorem, see [7], to solve it. In our case such a strategy seems to fail.

Strategy of the proof. Let us now briefly discuss the strategy of proof of Theorem 2. We construct the transformation Ψ by means of an iterative KAM-type scheme in Sobolev class. Our method is based on constructing a sequence of transformations $(\Phi_n)_{n\in\mathbb{N}}$ of the form $\Phi_n(\theta) = \theta + g_n(\theta)$, $n \in \mathbb{N}$ which reduces quadratically the size of the vector field. More precisely, at the n-th step of our procedure, we deal with an operator

$$X_n := (\alpha_n(\xi) + f_n(\xi, \theta)) \cdot \frac{\partial}{\partial\theta},$$

with f_n in H^s for all s.

Then, we choose a function $g_n(\xi, \theta)$ so that

$$\alpha_n \cdot \partial_\theta g_n + \Pi_{N_n} f_n(\theta) = \langle f_n \rangle_{\mathbb{T}^N} \tag{1.6}$$

where $N_n = N_0^{(\frac{3}{2})^n}$, Π_{N_n} is the orthogonal projection on the Fourier modes $|k| \le N_n$ and $\langle f_n \rangle_{\mathbb{T}^N}$ is the average of the vector valued function f_n with respect to $\theta \in \mathbb{T}^N$. The equation (1.6) can be solved by imposing the *diophantine conditions*

$$|\alpha_n \cdot k| \ge \frac{\gamma}{\langle k \rangle^\tau}, \qquad \forall 0 < |k| \le N_n,$$

for some $\gamma \in (0, 1)$ and $\tau > \nu$. Then, setting $\Phi_n(\theta) = \theta + g_n(\theta)$, one gets

$$X_{n+1} = (\Phi_n)_* X_n = (\alpha_{n+1}(\xi) + f_{n+1}) \cdot \frac{\partial}{\partial\theta},$$

where $\alpha_{n+1} := \alpha_n + \langle f_n \rangle_{\mathbb{T}^N}$, $|f_{n+1}| \simeq |\Pi_{N_n}^\perp f_n| + |f_n|^2$.

We show that f_n converges to 0 in any Sobolev space H^s, $s \ge 0$ provided $\|f_0\|_{H^{s_1}}$ is small for some fixed Sobolev index s_1 (which has to be taken large enough). Then we define the change of variables $\Psi^{(n)} := \Phi_0 \circ \Phi_1 \circ \ldots \circ \Phi_n$ and show that it is of the form $\Psi^{(n)}(\theta) = \theta + h_n(\theta)$. Moreover the sequence $h_n \in C^\infty(\mathbb{T}^N, \mathbb{R}^N)$ converges to a function $h_\infty = h$ in any H^s, $s \ge 0$. Then we verify that $\Psi(\theta) = \theta + h(\theta)$ is a torus diffeomorphism which conjugates X_0 to $\alpha_\infty(\xi) \cdot \frac{\partial}{\partial\theta}$. The final step is to show that the condition $\alpha_\infty(\xi)$ diophantine is sufficient to ensure that all the Melnikov conditions are satisfied.

We remark that the sequence of transformations $\Psi^{(n)} := \Phi_0 \circ \Phi_1 \circ \ldots \circ \Phi_n$ does not converge with respect to the operator norm of bounded linear operators $H^s \to H^s$, since $\Phi_n - \mathrm{Id} = O(g_n \cdot \partial_\theta)$ is small in size but unbounded of order one. This is the reason for approaching the problem from the point of view of the straightening of the corresponding vector field.

The paper is organized as follows. In Section 2, we introduce the functional setting needed to state precisely Theorem 3.1. In Section 3, we state our main results at the level of vector fields. Then in Section 4 we state and prove the corresponding results on reducibility of PDEs. In Section 5 we prove Theorem 3.1, which is based on the iterative Lemma 5.2. The inductive step of such a Lemma is basically proved in Lemma 5.1. Finally, in the Appendix A, we collect some technical lemmata concerning Sobolev spaces and operators induced by diffeomorphisms of the torus which are of importance for our proofs.

Acknowledgements: Two of the authors were supported by the ERC grant *Hamiltonian PDEs and small divisor problems: a dynamical systems approach* n. 306414 under FP7 and

by PRIN 2015 *Variational methods, with applications to problems in mathematical physics and geometry.* Another one is supported by the Swiss National Science Foundation: *Hamiltonian systems of infinite dimension*, n. 200020-165537.
The authors wish to thank P. Baldi, M. Berti, L. Biasco, L. Chierchia, E. Haus, J. Massetti and C. Saffirio for helpful discussions.

2 Functional setting

Sobolev functions. Given $N, m \in \mathbb{N}$, we consider real valued functions $u \in L^2(\mathbb{T}^N, \mathbb{R}^m)$:

$$u(\theta) = \sum_{k \in \mathbb{Z}^N} u_k e^{\mathrm{i} k\cdot\theta}, \quad \overline{u}_k = u_{-k}.$$

We use the simplified notation L^2 to denote $L^2(\mathbb{T}^N, \mathbb{R}^m)$. We define the Sobolev space

$$H^s(\mathbb{T}^N, \mathbb{R}^m) := \left\{ u \in L^2 : \|u\|_s^2 := \sum_{k \in \mathbb{Z}^N} \langle k \rangle^{2s} |u_k|^2 < \infty \right\} \tag{2.1}$$

where $\langle k \rangle := \max(1, |k|)$.

If we separate the variables $\theta = (\varphi, x) \in \mathbb{T}^{\nu+d}$, we may consider a real valued function $u(\varphi, x) \in L^2$ as a φ-dependent family of functions $u(\varphi, \cdot) \in L^2(\mathbb{T}^d_x, \mathbb{R}^m)$ with the Fourier series expansion

$$u(\varphi, x) = \sum_{j \in \mathbb{Z}^d} u_j(\varphi)\, e^{\mathrm{i} j\cdot x} = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}^d} u_{\ell,j}\, e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)}.$$

In this case it may be more convenient to describe the Sobolev space

$$H^s(\mathbb{T}^{\nu+d}, \mathbb{R}^m) := \left\{ u \in L^2(\mathbb{T}^{\nu+d}, \mathbb{R}^m) : \|u\|_s^2 := \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}^d} \langle \ell, j \rangle^{2s} |u_{\ell j}|^2 < \infty \right\} \tag{2.2}$$

where $\langle \ell, j \rangle := \max\{1, |\ell|, |j|\}$, $|\ell| := \sum_{i=1}^\nu |\ell_i|$, $|j| := \sum_{i=1}^d |j_i|$. If $\mathfrak{s}_0 := [N/2] + 1$ then for $s \geq \mathfrak{s}_0$ the spaces $H^s := H^s(\mathbb{T}^N, \mathbb{R}^m)$ are embedded in $L^\infty(\mathbb{T}^N, \mathbb{R}^m)$ and they have the algebra and interpolation structure, namely $\forall u, v \in H^s$ with $s \geq \mathfrak{s}_0$:

$$\|uv\|_s \leq C(s, N) \|u\|_s \|v\|_s,$$
$$\|u\, v\|_s \leq C(N)\, \|u\|_s \|v\|_{\mathfrak{s}_0} + C(s, N) \|u\|_{\mathfrak{s}_0} \|v\|_s. \tag{2.3}$$

Here $C(N), C(s, N)$ are positive constants independent of u, v. The above estimates are classical and one can see for instance the Appendix of [12] for the proof.

Lipschitz norm. Fix $N \in \mathbb{N}$ and let $\mathcal{O}$ be a compact subset of $\mathbb{R}^N$. For a function $u \colon \mathcal{O} \to E$, where $(E, \|\cdot\|_E)$ is a Banach space, we define the sup-norm and the lip-seminorm of u as

$$\|u\|_E^{\sup} := \|u\|_E^{\sup,\mathcal{O}} := \sup_{\xi \in \mathcal{O}} \|u(\xi)\|_E,$$
$$\|u\|_E^{lip} := \|u\|_E^{lip,\mathcal{O}} := \sup_{\substack{\xi_1,\xi_2 \in \mathcal{O},\\ \xi_1 \neq \xi_2}} \frac{\|u(\xi_1) - u(\xi_2)\|_E}{|\xi_1 - \xi_2|}. \tag{2.4}$$

Fix $\gamma > 0$. We will use the following Lipschitz norms

$$\|u\|_s^{\gamma,\mathcal{O}} := \|u\|_s^{sup,\mathcal{O}} + \gamma \|u\|_{s-1}^{lip,\mathcal{O}}, \quad u \in H^s, \quad \forall s \geq [N/2] + 3, \tag{2.5}$$

$$|m|^{\gamma,\mathcal{O}} := |m|^{sup,\mathcal{O}} + \gamma |m|^{lip,\mathcal{O}}, \quad m \in \mathbb{R}. \tag{2.6}$$

Diffeomorphisms of the torus. We consider diffeomorphisms of the N-dimensional torus

$$\Phi\colon \mathbb{T}^N \to \mathbb{T}^N, \qquad \Phi\colon \theta \mapsto \theta + h(\theta) = \widehat{\theta} \tag{2.7}$$

where $h : \mathbb{T}^N \to \mathbb{R}^N$ is some C^∞ function with $\|\beta\|_{\mathfrak{s}_0+1} \le \frac{1}{2}$. We denote the inverse diffeomorphism as

$$\Phi^{-1}\colon \mathbb{T}^N \to \mathbb{T}^N, \qquad \Phi^{-1}\colon \theta \mapsto \theta + \tilde{h}(\theta)$$

with $\tilde{h}$ a C^∞ function. With an abuse of notation we identify transformations like (2.7) with the corresponding linear operators acting on $H^s(\mathbb{T}^N)$ as

$$\Phi\colon H^s \to H^s, \qquad u(\theta) \mapsto \Phi u(\theta) := u(\theta + h(\theta)). \tag{2.8}$$

Similarly we consider the action of Φ on the vector fields on $\mathbb{T}^N$ by the pushforward. Explicitly we denote by $\mathrm{T}\,(\mathbb{T}^N)$ the tangent space of $\mathbb{T}^N$.
Now given a vector field $X\colon \mathbb{T}^N \to \mathrm{T}\,(\mathbb{T}^N)$

$$X(\theta) = \sum_{j=1}^{N} X_j(\theta)\,\frac{\partial}{\partial\theta_j}\,, \quad X_1, \dots, X_N \in C^\infty(\mathbb{T}^N, \mathbb{R}) \tag{2.9}$$

its pushforward is

$$(\Phi_* X)(\theta) = d\Phi(\Phi^{-1}(\theta))[X(\Phi^{-1}(\theta))] = \sum_{i=1}^{N} \Phi^{-1}\left(X_i + \sum_{j=1}^{N} \frac{\partial X_i}{\partial\theta_j} X_j\right)\frac{\partial}{\partial\widehat{\theta}_i}.$$

We refer to the Appendix A for technical lemmata on the tameness properties of the Lipschitz, Sobolev norms and bounds for the diffeomorphisms of the torus.

Reversible vector fields. Let $S : \mathbb{T}^N \to \mathbb{T}^N$ be the involution $\theta \mapsto -\theta$. We say that a vector field $X : \mathbb{T}^N \to T(\mathbb{T}^N)$ is *reversible* if

$$X \circ S = -S \circ X\,.$$

This is equivalent to say that X is even with respect to the variable θ. A diffeomorphism of the torus $\Phi : \mathbb{T}^N \to \mathbb{T}^N$ is said to be *reversibility preserving* if

$$\Phi \circ S = S \circ \Phi\,.$$

The above definition is equivalent to require that Φ is odd with respect to the variable θ. It is a straightforward calculation to verify that if X is reversible and Φ is reversibility preserving, then the push-forward $\Phi_* X$ is still reversible.

Linear operators A C^∞ vector field $X(\theta) = \sum_{j=1}^N X_j(\theta)\frac{\partial}{\partial\theta_j}$ induces a linear operator acting on the space of functions $u : \mathbb{T}^N \to \mathbb{R}$, that we denote by $X(\theta)\cdot\partial_\theta = \sum_{j=1}^N X_j(\theta)\partial_{\theta_j}$. More precisely, the action of such a linear operator is given by

$$H^s(\mathbb{T}^N, \mathbb{R}) \to H^{s-1}(\mathbb{T}^N, \mathbb{R})\,, \quad u(\theta) \mapsto X(\theta)\cdot\partial_\theta u(\theta).$$

3 A tame version of Moser's theorem

In this section we state a detailed version of Theorem 2, which is a revisited version of Moser's theorem on weakly perturbed vector fields on the torus.

Let us fix $N \in \mathbb{N}$, $\mathcal{O}_0$ a compact subset of $\mathbb{R}^N$ with positive Lebesgue measure and consider

$$\tau := N + 2, \quad s_0 \ge [N/2] + 3, \quad \gamma \in (0, 1). \tag{3.1}$$

Notations. We denote for a vector valued function $u(\theta)$ its average as

$$\langle u\rangle := \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} u\, d\theta.$$

We denote any constant depending only on $N, \mathcal{O}_0$ as C and correspondingly we say $a \lesssim b$ if $a \le Cb$. A constant depending on parameters p is denoted by C_p and as above we say $a \lesssim_p b$ if $a \le C_p b$.

Theorem 3.1 (Tame Moser Theorem). *Consider for $\xi \in \mathcal{O}_0 \subseteq \mathbb{R}^N$ a Lipschitz family of vector fields on $\mathbb{T}^N$*

$$X_0 := (\xi + f_0(\theta;\xi)) \cdot \frac{\partial}{\partial\theta} \tag{3.2}$$

$$f_0(\cdot;\xi) \in H^s(\mathbb{T}^N, \mathbb{R}^N) \quad \forall s \ge s_0. \tag{3.3}$$

There exist $s_1 \ge s_0 + 2\tau + 4$, and $\eta_\star = \eta_\star(s_1) > 0$ such that if

$$\gamma^{-1}\|f_0\|_{s_1}^{\gamma,\mathcal{O}_0} := \delta \le \eta_\star \tag{3.4}$$

then there exists a Lipschitz function $\alpha_\infty : \mathcal{O}_0 \to \mathbb{R}^N, \xi \mapsto \alpha(\xi)$ with

$$|\alpha_\infty - \mathrm{Id}_{\mathbb{R}^N}|^{\gamma,\mathcal{O}_0} \le \gamma\delta, \tag{3.5}$$

such that in the set

$$\mathcal{O}_\infty^{2\gamma} := \Big\{\xi \in \mathcal{O}_0 : \ |\alpha_\infty(\xi)\cdot k| > \frac{2\gamma}{\langle k\rangle^\tau}\,,\ \forall k \in \mathbb{Z}^N \setminus \{0\}\Big\} \tag{3.6}$$

the following holds. There exists a map

$$\beta : \mathcal{O}_\infty^{2\gamma} \times \mathbb{T}^N \to \mathbb{R}^N\,, \quad \|\beta\|_s^{\gamma,\mathcal{O}_\infty^{2\gamma}} \lesssim_s \gamma^{-1}\|f\|_{s+2\tau+4}^{\gamma,\mathcal{O}_0}, \quad \forall s \ge s_0, \quad s \in \mathbb{N} \tag{3.7}$$

so that $\Psi : \theta \mapsto \theta + \beta(\theta) = \widehat{\theta}$ is a diffeomorphism of $\mathbb{T}^N$ and for all $\xi \in \mathcal{O}_\infty^{2\gamma}$

$$\Psi_* X_0 := (\Psi)^{-1}\big(\xi + f_0 + (\xi + f_0)\cdot\partial_\theta\beta\big)\cdot\frac{\partial}{\partial\widehat{\theta}} = \alpha_\infty(\xi)\cdot\frac{\partial}{\partial\widehat{\theta}}\,. \tag{3.8}$$

Furthermore, if f_0 is a reversible vector field (i.e. $f_0 = \mathrm{even}(\theta)$), then the diffeomorphism $\theta \mapsto \theta + \beta(\theta)$ is reversibility preserving (i.e. the function $\beta = \mathrm{odd}(\theta)$).
Finally, the Lebesgue measure of the set $\mathcal{O}_0 \setminus \mathcal{O}_\infty^{2\gamma}$ satisfies the bound

$$|\mathcal{O}_0 \setminus \mathcal{O}_\infty^{2\gamma}| \lesssim \gamma\,. \tag{3.9}$$

Remark 3.2. Note that condition (3.5) implies that the map $\xi \mapsto \alpha(\xi)$ is a lipeomorphism and $|\xi(\alpha)|^{lip} \le 2$. Hence the estimate of the measure of the complementary set $\mathcal{O}_0 \setminus \mathcal{O}_\infty^{2\gamma}$ is trivial.

Corollary 3.3. *(Parameter dependence) Under the assumptions of Theorem 3.1, suppose that $f_0 = f_0(\theta, \lambda; \xi)$ depends on some parameter $\lambda \in B_E$, where B_E is a ball centered at the origin of a Banach space E. Consequently the frequency map $\alpha_\infty = \alpha_\infty(\xi, \lambda)$, given by Theorem 3.1, is defined for all $\xi \in \mathcal{O}_0$ and $\lambda \in B_E$. This clearly implies that the set (3.6) depends on λ. Denoting with $\Delta_{12} g(\lambda) := g(\lambda_1) - g(\lambda_2)$ for some $\lambda_1, \lambda_2 \in B_E$, if*

$$K^{\tau+1}\|\Delta_{12} f_0\|_{s_1} \le \rho\,, \tag{3.10}$$

for some $\rho \in (0,1)$ small enough and $K > 0$ large enough, then the following holds. There exists $\mu > \tau$ such that for all $\xi \in \mathcal{O}_\infty^{2\gamma}(\lambda_1)$ we have

$$|\Delta_{12}\alpha_\infty| \lesssim \|\Delta_{12} f_0\|_{s_1} + K^{-\mu}\gamma^{-1}\max\{\|f_0(\lambda_1)\|_{s_1}\,,\ \|f_0(\lambda_2)\|_{s_1}\}. \tag{3.11}$$

As explained in the introduction we now divide the variables θ in time and space, $\theta = (\varphi, x) \in \mathbb{T}^{\nu+d}$. Similarly we write $\xi = (\omega, \zeta) \in \mathbb{R}^{\nu+d}$. We have the following result

Proposition 3.4. *Consider for $(\omega, \zeta) \in \mathcal{O}_0 \subseteq \mathbb{R}^{\nu+d}$ a Lipschitz family of vector fields on $\mathbb{T}^{\nu+d}$*

$$X_0 := \omega \cdot \frac{\partial}{\partial \varphi} + (\zeta + a_0(\varphi, x; \omega, \zeta)) \cdot \frac{\partial}{\partial x} \tag{3.12}$$

$$a_0(\cdot; \omega, \zeta) \in H^s(\mathbb{T}^{\nu+d}, \mathbb{R}^d) \quad \forall s \geq s_0. \tag{3.13}$$

Fix $s_1, \eta_\star$ as in Theorem 3.1, if

$$\gamma^{-1} \|a_0\|_{s_1}^{\gamma, \mathcal{O}_0} := \delta \leq \eta_\star \tag{3.14}$$

then there exists a Lipschitz function $m_\infty : \mathcal{O}_0 \to \mathbb{R}^d, (\omega, \zeta) \mapsto m_\infty(\omega, \zeta)$ such that denoting $\alpha_\infty(\omega, \zeta) = (\omega, m_\infty(\omega, \zeta))$ we have

$$|\alpha_\infty - \mathrm{Id}_{\mathbb{R}^{\nu+d}}|^\gamma \leq \gamma \delta, \tag{3.15}$$

such that in the set

$$\mathcal{O}_\infty^{2\gamma} := \Big\{ (\omega, \zeta) \in \mathcal{O}_0 : \; |\omega \cdot \ell + m_\infty(\omega, \zeta) \cdot j| > \frac{2\gamma}{\langle \ell, j \rangle^\tau}, \; \forall (\ell, j) \in \mathbb{Z}^{\nu+d} \setminus \{0\} \Big\} \tag{3.16}$$

the following holds. There exists a map

$$\beta : \mathbb{T}^{\nu+d} \times \mathcal{O}_\infty^{2\gamma} \to \mathbb{R}^d, \quad \|\beta\|_s^{\gamma, \mathcal{O}_\infty^{2\gamma}} \lesssim_s \gamma^{-1} \|a_0\|_{s+2\tau+4}^{\gamma, \mathcal{O}_0}, \quad \forall s \geq s_0, \; s \in \mathbb{N} \tag{3.17}$$

so that $\Psi : (\varphi, x) \mapsto (\varphi, x + \beta(\varphi, x; \omega, \zeta)) = (\varphi, \widehat{x})$ is a diffeomorphism of $\mathbb{T}^{\nu+d}$ and for all $(\omega, \zeta) \in \mathcal{O}_\infty^{2\gamma}$

$$\begin{aligned} \Psi_* X_0 &:= \omega \cdot \frac{\partial}{\partial \varphi} + \Psi^{-1}\big(\omega \cdot \partial_\varphi \beta + \zeta + a_0 + (\zeta + a_0) \cdot \partial_x \beta\big) \cdot \frac{\partial}{\partial \widehat{x}} \\ &= \omega \cdot \frac{\partial}{\partial \varphi} + m_\infty(\omega, \zeta) \cdot \frac{\partial}{\partial \widehat{x}}. \end{aligned} \tag{3.18}$$

Furthermore if a_0 is a reversible vector field (i.e. $a_0 = \mathrm{even}(\varphi, x)$) then the diffeomorphism $(\varphi, x) \mapsto \big(\varphi, x + \beta(\varphi, x)\big)$ is reversibility preserving (i.e. $\beta = \mathrm{odd}(\varphi, x)$).

Remark 3.5. The vector field in (3.12) can be written in the form (3.2) by setting $f_0 = (0, \ldots, 0, a_0)$. Then condition (3.14) implies (3.4) and Theorem 3.1 holds. In the previous proposition we are simply stating that the change of variables Ψ, predicted by Theorem 3.1, is the identity on the φ variables (and so is the frequency map α_∞).

We now wish to consider the case where ζ is not an independent parameter but a given function of ω. We have the following

Corollary 3.6. *Consider for $\omega \in \Omega_0 \subseteq \mathbb{R}^\nu$ a Lipschitz family of vector fields on $\mathbb{T}^{\nu+d}$*

$$X_0 := \omega \cdot \frac{\partial}{\partial \varphi} + (m_0(\omega) + a_0(\varphi, x; \omega)) \cdot \frac{\partial}{\partial x}. \tag{3.19}$$

Here $m_0(\omega)$ is a Lipschitz function and

$$a_0(\cdot; \omega) \in H^s(\mathbb{T}^{\nu+d}, \mathbb{R}^d) \quad \forall s \geq s_0. \tag{3.20}$$

Fix $s_1, \eta_\star$ as in Theorem 3.1, if

$$\gamma^{-1} \|a_0\|_{s_1}^{\gamma, \Omega_0} := \delta \leq \eta_\star \tag{3.21}$$

then in the set

$$\Omega_\infty^{2\gamma} := \Big\{\omega \in \Omega_0 : \ |\omega\cdot\ell + m_\infty(\omega, m_0(\omega))\cdot j| > \frac{2\gamma}{\langle \ell, j\rangle^\tau}\,,\ \forall(\ell,j)\in\mathbb{Z}^{\nu+d}\setminus\{0\}\Big\} \tag{3.22}$$

the map β of Proposition 3.4, restricted to $\zeta = m_0(\beta)$ diagonalizes X_0 as in formula (3.18)*. Moreover*

$$\begin{aligned} &\|\beta\|_s^{\gamma,\Omega_\infty^{2\gamma}} \lesssim_s (1+|m_0|^{lip,\Omega_0})\gamma^{-1}\|a_0\|_{s+2\tau+4}^{\gamma,\Omega_0}\,,\\ &|m_0-m_\infty|^\gamma \lesssim \|a_0\|_{s_1}^{\gamma,\Omega_0} \end{aligned} \tag{3.23}$$

Furthermore if a_0 is a reversible vector field (i.e. $a_0 = \mathrm{even}(\varphi,x)$) then the diffeomorphism $(\varphi,x)\mapsto\big(\varphi, x+\beta(\varphi,x)\big)$ is reversibility preserving (i.e. $\beta = \mathrm{odd}(\varphi,x)$).

Proof. We wish to apply Proposition (3.4), we consider the map $M:\omega\mapsto(\omega,m_0(\omega))$ and denote by $\mathcal{O}_0$ the image of Ω_0 through this map. Then we consider $\widetilde{a}_0(\varphi,x;\omega,\zeta) = a_0(\varphi,x;\omega)$ in this way the dependence on ζ is trivial and we have

$$\gamma^{-1}\|\widetilde{a}_0\|_{s_1}^{\gamma,\mathcal{O}_0} = \gamma^{-1}\|a_0\|_{s_1}^{\gamma,\Omega_0} := \delta \le \eta_\star .$$

We thus apply Proposition (3.4) to the vector field

$$\widetilde{X}_0 := \omega\cdot\frac{\partial}{\partial\varphi} + (\zeta + \widetilde{a}_0(\varphi,x;\omega,\zeta))\cdot\frac{\partial}{\partial x}.$$

We produce a change of variables $x\mapsto x+\widetilde{\beta}(\varphi,x;\omega,\zeta)$ which diagonalizes $\widetilde{X}_0$ in the set $\mathcal{O}_\infty^{2\gamma}$. We now restrict our parameter set to $\zeta = m_0(\omega)$. By definition $\omega\in\Omega_0 \Leftrightarrow (\omega,m_0(\omega))\in\mathcal{O}_0$, so the restriction of $\mathcal{O}_\infty^{2\gamma}$ to $\zeta = m_0(\omega)$ is $\Omega_\infty^{2\gamma}$. It remains to prove the bound (3.23). Setting $\beta(\varphi,x,\omega) = \widetilde{\beta}(\varphi,x;\omega,m_0(\omega))$ we have

$$\|\beta\|_s^{sup,\Omega_\infty^{2\gamma}} = \|\widetilde{\beta}\|_s^{sup,\mathcal{O}_\infty^{2\gamma}}\,,\quad \|\beta\|_s^{lip,\Omega_\infty^{2\gamma}} \le \|\widetilde{\beta}\|_s^{lip,\mathcal{O}_\infty^{2\gamma}}(1+|m_0|^{lip,\Omega_0})$$

hence the result follows. □

4 Reducibility

In this section we state a more detailed version of Theorem 1 regarding reducibility of quasi-periodic transport type equations and the control of the Sobolev norms for their solutions. They are obtained as a consequence of Theorem 3.1.

Theorem 4.1. *Consider the transport equation*

$$\partial_t u + \big(\zeta + a_0(\omega t, x;\omega,\zeta)\big)\cdot\partial_x u = 0\,. \tag{4.1}$$

Then if (3.13), (3.14) *are fullfilled, for $(\omega,\zeta)\in\mathcal{O}_\infty^{2\gamma}$ (see* (3.16)*), under the change of variable $u = \Psi(\omega t)[v] = v(x+\beta(\omega t,x))$ defined in* (3.17)*, the PDE* (4.1) *transforms into the equation with constant coefficients*

$$\partial_t v + m_\infty(\omega,\zeta)\cdot\partial_x v = 0\,. \tag{4.2}$$

As a consequence, for any $s\ge 0$, $u_0\in H^s(\mathbb{T}^d)$ the only solutions of the Cauchy problem

$$\begin{cases}\partial_t u + \big(\zeta + a_0(\omega t,x;\omega,\zeta)\big)\cdot\partial_x u = 0\\ u(0,x) = u_0(x)\end{cases}$$

satisfies $\|u(t)\|_{H^s_x}\lesssim_s \|u_0\|_{H^s_x}$ for any $t\in\mathbb{R}$. Furthermore, if $a_0 = \mathrm{even}(\varphi,x)$, then $\beta = \mathrm{odd}(\varphi,x)$.

Proof. Let $(\omega,\zeta)\in\mathcal{O}_\infty^{2\gamma}$. By a direct calculation, under the change of coordinates $u=\Psi(\omega t)[v]=v(x+\beta(\omega t,x))$, the equation (4.1) transforms into the PDE

$$\partial_t v+\Psi(\omega t)^{-1}\Big(\omega\cdot\partial_\varphi\beta+\zeta+a_0+(a_0+\zeta)\cdot\partial_x\beta\Big)\cdot\partial_x v=0\,.$$

By applying Proposition 3.4, one gets that

$$\Psi(\omega t)^{-1}\Big(\omega\cdot\partial_\varphi\beta+\zeta+a_0+(a_0+\zeta)\cdot\partial_x\beta\Big)=m_\infty(\zeta,\omega)$$

implying that v solves the equation (4.2). Such a PDE with constant coefficients can be integrated explicitly, implying that for any $s\geq 0$, $\|v(t)\|_{H^s_x}=\|v(0)\|_{H^s_x}$. Note that, since $\beta\in C^\infty$, by Lemma A.3, $\Psi(\omega t)^{\pm1}$ is a bounded linear operator $H^k(\mathbb{T}^d)\to H^k(\mathbb{T}^d)$ for any $k\in\mathbb{N}$ with $\sup_{\varphi\in\mathbb{T}^\nu}\|\Psi(\varphi)^{\pm1}\|_{\mathcal{B}(H^k_x)}<+\infty$. By using the classical Riesz-Thorin interpolation Theorem for linear operators one gets that $\Psi(\varphi)\in\mathcal{B}(H^s_x)$ with $\sup_{\varphi\in\mathbb{T}^\nu}\|\Psi(\varphi)^{\pm1}\|_{\mathcal{B}(H^s_x)}<+\infty$ for any $s\geq 0$. Then given $u_0\in H^s(\mathbb{T}^d)$, one gets that

$$\|u(t)\|_{H^s_x}=\|\Psi(\omega t)[v(t)]\|_{H^s_x}\lesssim_s\|v(t)\|_{H^s_x}\lesssim_s\|v(0)\|_{H^s_x}\lesssim_s\|\Psi(\omega t)^{-1}[u_0]\|_{H^s_x}\lesssim_s\|u_0\|_{H^s_x}$$

and this concludes the proof. □

We remark that, as explained in Proposition 3.4, the measure of $\mathcal{O}_0\setminus\mathcal{O}_\infty^{2\gamma}$ is of order γ, which guarantees that the reducibility result holds for a positive measure set of parameters (ω,ζ). On the other hand we are not able to give a bound on $\Omega_0\setminus\Omega_\infty^{2\gamma}$ unless we impose some further conditions. We give an example which we believe is interesting for applications.

Corollary 4.2. *Let $d=1$ and consider the one-dimensional transport equation*

$$\partial_t u+\big(m_0(\omega)+a_0(\omega t,x;\omega)\big)\partial_x u=0,\quad x\in\mathbb{T}\,.$$

Assume that $m_0(\omega)$ satisfies

$$\inf_{\omega\in\Omega_0}|m_0|\geq c\,,\quad |m_0(\omega)|^{lip,\Omega_0}\leq L$$

for some appropriate c,L which depend on the set Ω_0. Assume finally that (3.20), (3.21) hold. Then the analog of Theorem 4.1 holds for any value of the frequency ω in the set $\Omega_\infty^{2\gamma}$ defined in (3.22). Moreover if $\eta_\star,L$ are small enough, depending on the set Ω_0, then the set $\Omega_\infty^{2\gamma}$ satisfies the measure estimate $|\Omega_0\setminus\Omega_\infty^{2\gamma}|\lesssim\gamma$.

Note that this includes the case $m_0\neq 0$ and constant in ω.

Proof. The fact that the analog of Theorem 4.1 holds is just a repetition of the proof of that statement. The only non trivial thing is to show the measure estimate. We first note that

$$\inf_{\omega\in\Omega_0}|m_\infty|\geq c-\eta_\star\,,\quad |m_\infty|^{lip,\Omega_0}\leq L+\eta_\star$$

Fix ℓ,j and compute the measure of a resonant set

$$\mathcal{R}_{\ell j}:=\Big\{\omega\in\Omega_0\ :\ |\omega\cdot\ell-m_\infty j|\leq\frac{2\gamma}{\langle\ell\rangle^\tau}\Big\}.$$

We claim that if $\mathcal{R}_{\ell,j}\neq\emptyset$ then $|j|\leq\mathtt{k}|\ell|$, where

$$\mathtt{k}:=\sup_{\omega\in\Omega_0}\frac{|\omega|+2\gamma}{|m_\infty|}\leq(c-\eta_\star)^{-1}(\sup_{\omega\in\Omega_0}|\omega|+2).$$

Note that, since Ω_0 is compact, $\mathtt{k}$ is finite provided $\eta_\star$ is small. To prove our claim we note that

$$|m_\infty j| \le |\omega\cdot\ell| + |\omega\cdot\ell - m_\infty j| \le |\omega||\ell| + \frac{2\gamma}{\langle\ell\rangle^\tau} \le (|\omega| + 2\gamma)|\ell|.$$

We claim that $|\mathcal{R}_{\ell,j}| \le C\gamma/\langle\ell\rangle^{\tau+1}$, for some $C>0$, for any $(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{Z}\setminus\{(0,0)\}$. We write

$$\omega = \frac{\ell}{|\ell|}s + v\,, \quad v\cdot\ell = 0\,,$$

so that setting

$$\phi(s) := \omega\cdot\ell - m_\infty(\omega)j = |\ell|s + m_\infty\Big(\frac{\ell}{|\ell|}s + v\Big)j\,.$$

we have (recall $|j| \le \mathtt{k}|\ell|$)

$$\frac{|\phi(s_1)-\phi(s_2)|}{|s_1-s_2|} \ge |\ell|(1-\mathtt{k}|m_\infty|^{lip,\Omega_0}) \ge \frac{|\ell|}{2}\,,$$

provided that

$$\mathtt{k}|m_\infty|^{lip,\Omega_0} = \sup_{\omega\in\Omega_0}(|\omega|+2\gamma)\frac{|m_\infty|^{lip,\Omega_0}}{\inf_{\omega\in\Omega_0}|m_\infty|} \le \frac{L+\eta_\star}{c-\eta_\star}(\sup_{\omega\in\Omega_0}|\omega|+2) < \frac{1}{2}. \tag{4.3}$$

Note that these equations give an upper bound on $L,\eta_\star$. We assume that the desired smallness conditions hold and hence we have our claim.

$$|\Omega_0\setminus\Omega_\infty^{2\gamma}| \le \sum_{\ell\in\mathbb{Z}^\nu\setminus\{0\},|j|\le\mathtt{k}|\ell|}|\mathcal{R}_{\ell,j}| \le \mathtt{k}\gamma\sum_{\ell\in\mathbb{Z}^\nu\setminus\{0\}}\frac{1}{|\ell|^\tau} \lesssim \gamma.$$

since $\tau > \nu+1$. □

We also state the following corollary, concerning the solvability of a forced quasi-periodic transport equation. Such a corollary will be applied in [32].

Corollary 4.3. *(Forced case) Assume the hypotheses of Corollary 3.6 and let $f := f(\varphi,x)$ be some $C^\infty(\mathbb{T}^{\nu+d})$ function.*
For every $\omega\in\Omega_\infty^{2\gamma}$ (see (3.22)), there exist a C^∞ function $b(\varphi,x;\omega)$ and a constant $\mathtt{c}=\mathtt{c}(\omega)$ (depending in a Lipschitz way on the parameter ω) such that

$$\omega\cdot\partial_\varphi b(\varphi,x) + (m_0 + a_0(\varphi,x))\partial_x b(\varphi,x) + f(\varphi,x) = \mathtt{c}, \quad (\varphi,x)\in\mathbb{T}^{\nu+1}. \tag{4.4}$$

Furthermore, there exists a constant $\sigma=\sigma(\tau,\nu)>0$ such that the following estimates hold:

$$\begin{aligned} &\|b\|_s^{\gamma,\Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-1}\Big(\|f\|_{s+\sigma} + \|a_0\|_{s+\sigma}^{\gamma,\Omega_0}\|f\|_{s_0+\sigma}\Big)\,, \quad \forall s\ge s_0\,,\\ &|\mathtt{c}|^{\gamma,\Omega_0} \lesssim \|f\|_{s_0}^{\gamma,\Omega_0}\,. \end{aligned} \tag{4.5}$$

Proof. The equation (4.4) can be written as

$$\mathcal{L}b + f = \mathtt{c} \tag{4.6}$$

where

$$\mathcal{L} := \omega\cdot\partial_\varphi - (m_0 + a_0(\varphi,x))\partial_x, \tag{4.7}$$

By Corollary 3.6, in particular by (5.82), we have that $\mathcal{L} = \Psi^{-1}\mathcal{L}_\infty\Psi$, where

$$\Psi h(\varphi,x) := h(\varphi,x+\beta(\varphi,x)), \quad \mathcal{L}_\infty := \omega\cdot\partial_\varphi - m_\infty\partial_x. \tag{4.8}$$

Then

$$\mathcal{L}_\infty \Psi b = \Psi\big(\mathsf{c} - f\big). \tag{4.9}$$

Using that $\Psi(1) = 1$ we get

$$\mathcal{L}_\infty \Psi b = \mathsf{c} - \Psi\big(f\big). \tag{4.10}$$

and we choose c such that

$$\mathsf{c} = \langle \Psi\big(f\big)\rangle_{\mathbb{T}^{\nu+1}} \tag{4.11}$$

so that the r. h. s. of the equation (4.10) has zero average. By the fact that Ψ is bounded from H^s to itself for any $s \geq s_0$ and $f \in C^\infty$ then $g := \mathsf{c} - \Psi\big(f\big) \in C^\infty$.
Since $g \in C^\infty$ and has zero average then the equation $\mathcal{L}_\infty[h] = g$, for any $\omega \in \mathcal{O}^{2\gamma}_\infty$, has a C^∞-solution which is given by

$$h(\varphi, x) := \mathcal{L}_\infty^{-1}[g](\varphi, x) = \sum_{(\ell,j)\neq(0,0)} \frac{g_{\ell j}}{\mathrm{i}(\omega\cdot\ell - m_\infty j)}\, e^{\mathrm{i}(\ell\cdot\varphi + jx)}. \tag{4.12}$$

Furthermore, using the estimate on m_∞ given in (3.23), the following standard estimate holds:

$$\|h\|_s^{\gamma,\Omega^{2\gamma}_\infty} \lesssim_s \gamma^{-1}\|g\|_{s+2\tau+1}^{\gamma,\Omega^{2\gamma}_\infty}, \quad \forall s \geq 0\,. \tag{4.13}$$

Therefore the function

$$b := \Psi^{-1}\mathcal{L}_\infty^{-1}[\mathsf{c} - \Psi\big(f\big)] \tag{4.14}$$

is a C^∞-solution of the equation (4.6).
Finally, the estimates (4.5) follow by (4.11), (4.14) and by applying the estimates (A.7), (A.8), the smallness condition (3.21) and the estimates (3.23), (4.13). □

5 An iterative KAM scheme

We prove Theorem 3.1 and Corollary 3.3. This is done by applying recursively a KAM step which we now describe.

5.1 KAM step

Consider for $\xi \in \mathcal{O} \subseteq \mathcal{O}_0$ a Lipschitz family of vector fields on $\mathbb{T}^N$

$$\begin{aligned} X &:= (\alpha + f(\theta;\xi))\cdot\frac{\partial}{\partial\theta} \\ |\alpha|^{lip,\mathcal{O}_0} &\leq M < 2\,, \quad f(\cdot;\xi) \in H^s(\mathbb{T}^N, \mathbb{R}^d) \quad \forall s \geq s_0. \end{aligned} \tag{5.1}$$

Given $K \gg 1$ and $\gamma > 0$ assume that for some domain $\mathcal{O} \subseteq \mathcal{O}_0$ we have (recall (3.1))

$$\gamma^{-1}K^{2\tau+2s_0+1}\|f\|_{s_0}^{\gamma,\mathcal{O}} \leq \delta := \delta(s_0) \tag{5.2}$$

for some δ small enough. Let

$$\mathcal{O}_+ \equiv \mathcal{C}_{K,\mathcal{O}} := \{\xi \in \mathcal{O} : \ |\alpha\cdot k| > \frac{\gamma}{\langle k\rangle^\tau}\,,\ \forall k \in \mathbb{Z}^N\setminus\{0\},\ \ |k| \leq K\}\,, \tag{5.3}$$

and for all $\xi \in \mathcal{O}_+$ set $g(\theta;\xi)$ to be

$$g(\theta;\xi) := \sum_{|k|\leq K} g_k e^{\mathrm{i}k\cdot\theta}\,, \tag{5.4}$$

where

$$g_k = -\frac{f_k}{\mathrm{i}\alpha\cdot k}\,, \quad \forall k \in \mathbb{Z}^N\,, \quad 0 < |k| \leq K\,. \tag{5.5}$$

Lemma 5.1. *The function g defined in (5.4) satisfies*

$$\|g\|_s^{\gamma,\mathcal{O}_+} \lesssim \gamma^{-1}\|\Pi_K f\|_{s+2\tau+1}^{\gamma,\mathcal{O}}, \quad \forall\, s \geq s_0. \tag{5.6}$$

The map

$$\Phi : \theta \mapsto \theta + g(\theta) \tag{5.7}$$

is a diffeomorphism of $\mathbb{T}^N$. We have that the pushforward of the vector field X in (5.1) under the map Φ in (5.7) has the form

$$\Phi_* X := (\alpha_+ + f_+(\theta;\xi)) \cdot \frac{\partial}{\partial\theta} \tag{5.8}$$

where $\alpha_+ \in \mathbb{R}^N$ is defined and Lipschitz for $\xi \in \mathcal{O}_0$, the function f_+ is defined and Lipschitz for all $\xi \in \mathcal{O}_+$ (see (5.3)) and the following bounds hold:

$$|\alpha - \alpha_+|^{\gamma,\mathcal{O}_0} \lesssim \|f\|_{s_0}^{\gamma,\mathcal{O}}, \tag{5.9}$$

$$\begin{aligned} \|f_+\|_{s_0}^{\gamma,\mathcal{O}_+} &\lesssim K^{s_0-s_1}\|f\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}\gamma^{-1}K^{2\tau+2}(\|f\|_{s_0}^{\gamma,\mathcal{O}})^2, \\ \|f_+\|_s^{\gamma,\mathcal{O}_+} &\leq \|f\|_s^{\gamma,\mathcal{O}} + C_s\gamma^{-1}K^{2\tau+2s_0+1}\|f\|_{s_0}^{\gamma,\mathcal{O}}\|f\|_s^{\gamma,\mathcal{O}}. \end{aligned} \tag{5.10}$$

Moreover if f is a reversible vector field, then $\theta \mapsto \theta + g(\theta)$ is a reversibility preserving map, implying that f_+ is a reversible vector field.
Let $\lambda_1, \lambda_2 \in B_E$, $\xi \in \mathcal{O}_+(\lambda_1) \cap \mathcal{O}_+(\lambda_2)$, $s_1, \mathtt{b} > 0$ satisfying

$$s_0 < \mathtt{b} + 2\tau + 3s_0 + 2 < s_1\,. \tag{5.11}$$

There exists $\delta' := \delta'(s_1)$ such that if

$$\gamma^{-1}\big(\|f(\lambda_1)\|_{s_1} + \|f(\lambda_2)\|_{s_1}\big) \leq \delta' \tag{5.12}$$

then for $\omega \in \mathcal{O}_+(\lambda_1) \cap \mathcal{O}_+(\lambda_2)$

$$\|\Delta_{12} g\|_s \lesssim_s \gamma^{-1}\Big(\|\Pi_K \Delta_{12} f\|_{s+\tau} + \gamma^{-1}|\Delta_{12}\alpha|\|\Pi_K f\|_{s+2\tau+1}\Big), \quad \forall s \geq 0 \tag{5.13}$$

$$|\Delta_{12}(\alpha_+ - \alpha)| \leq \|\Delta_{12} f\|_{s_0}\,, \tag{5.14}$$

$$\begin{aligned} \|\Delta_{12} f_+\|_{s_0-1} &\lesssim_{s_0,\mathtt{b}} K^{-1-\mathtt{b}}\|\Delta_{12} f\|_{s_0+\mathtt{b}} + K^{\tau+s_0}\gamma^{-1}\|\Delta_{12} f\|_{s_0-1}\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2) \\ &\quad + K^{2\tau+s_0}\gamma^{-2}|\Delta_{12}\alpha|\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2)^2, \\ \|\Delta_{12} f_+\|_{s_0+\mathtt{b}} &\lesssim_{s_0,\mathtt{b}} K^{2\tau+s_0}\Big(\|\Delta_{12} f\|_{s_0+\mathtt{b}} + |\Delta_{12}\alpha|\Big)\,. \end{aligned} \tag{5.15}$$

Proof. By definition of $\|\cdot\|_s$ (see (2.1)), (5.4) and (5.3) we have $\|g\|_s \lesssim \gamma^{-1}K^\tau\|a\|_s$ for all $\xi \in \mathcal{O}_+$. By (5.4) we have, for $\xi, \xi' \in \mathcal{O}_+$,

$$|\Delta_{\xi,\xi'} g_k| \leq \frac{|\Delta_{\xi,\xi'} f_k|}{|\alpha(\xi)\cdot k|} + \frac{|f_k(\xi')||\Delta_{\xi,\xi'}\alpha||k|}{|\alpha(\xi)\cdot k||\alpha(\xi')\cdot k|} \tag{5.16}$$

hence by (5.3) we get (5.6) and, by using smoothing properties of the projector Π_K (see Lemma A.1, item (iv)) we have

$$\|g\|_s^{\gamma,\mathcal{O}_+} \lesssim \gamma^{-1}K^{2\tau+1}\|f\|_s^{\gamma,\mathcal{O}}. \tag{5.17}$$

We claim that g satisfies the hypotheses of Lemma A.3, hence Φ is a diffeomorphism. Indeed, since $s_0 \geq [N/2] + 3 > \mathfrak{s}_0 + 1$, by (5.17) and (5.2) we have

$$|g|_{1,\infty}^{\gamma,\mathcal{O}_+} \lesssim_{s_0} \|g\|_{s_0}^{\gamma,\mathcal{O}_+} \lesssim_{s_0} \gamma^{-1}K^{2\tau+1}\|f\|_{s_0}^{\gamma,\mathcal{O}} \leq \frac{1}{2}$$

provided that δ in (5.2) is sufficiently small. By applying Lemma A.3, and by the Sobolev embedding, one gets that the inverse diffeomorphism $y \mapsto y + \widetilde{g}(\varphi, y)$ satisfies the estimate

$$\|\widetilde{g}\|_s^{\gamma,\mathcal{O}_+} \lesssim_s \|g\|_{s+s_0}^{\gamma,\mathcal{O}_+}, \quad s \geq 0. \tag{5.18}$$

By definition of pushforward (we rename the new variables as θ)

$$\Phi_* X := (\Phi)^{-1}\big(\alpha + f + (\alpha + f)\cdot \partial_\theta g\big) \cdot \frac{\partial}{\partial \theta}$$

and by the definition of g in (5.4)

$$\begin{aligned}&\alpha + f + (\alpha + f)\cdot \partial_\theta g\\ &= \alpha + \langle f \rangle + \Pi_K^\perp f + f \cdot \partial_\theta g.\end{aligned}$$

Now we extend $\langle f \rangle$ from $\mathcal{O}$ to the whole $\mathcal{O}_0$ by Kirtzbraun theorem, preserving the Lipschitz norm.
We denote the extension by $\langle f \rangle^{\text{Ext}}$ and set

$$\begin{aligned}\alpha_+ &:= \alpha + \langle f \rangle^{\text{Ext}}, \quad \xi \in \mathcal{O}_0,\\ f_+(\theta) &:= \Phi^{-1}\Big(\Pi_K^\perp f + f \cdot \partial_\theta g\Big), \quad \xi \in \mathcal{O}_+.\end{aligned} \tag{5.19}$$

The bounds (5.9) follow since

$$|\langle f \rangle^{\text{Ext}}|^{\gamma,\mathcal{O}_0} \leq |\langle f \rangle|^{\gamma,\mathcal{O}} \leq \|f\|_{s_0}^{\gamma,\mathcal{O}}.$$

As for (5.10) we repeatedly use Lemma A.3, indeed setting

$$F := \Pi_K^\perp f + f \cdot \partial_\theta g \tag{5.20}$$

we have by (A.8) for $s \in \mathbb{N},\ s \geq s_0$

$$\begin{aligned}\|f_+\|_s^{\gamma,\mathcal{O}_+} &\leq \|F\|_s^{\gamma,\mathcal{O}_+} + C_s\big(\|F\|_s^{\gamma,\mathcal{O}_+}\|\widetilde{g}\|_{s_0+1}^{\gamma,\mathcal{O}_+} + \|\widetilde{g}\|_{s+s_0}^{\gamma,\mathcal{O}_+}\|F\|_{s_0}^{\gamma,\mathcal{O}_+}\big)\\ &\overset{(5.18)}{\leq} \|F\|_s^{\gamma,\mathcal{O}_+} + C_s\big(\|F\|_s^{\gamma,\mathcal{O}_+}\|g\|_{2s_0+1}^{\gamma,\mathcal{O}_+} + \|g\|_{s+2s_0}^{\gamma,\mathcal{O}_+}\|F\|_{s_0}^{\gamma,\mathcal{O}_+}\big)\\ \|F\|_s^{\gamma,\mathcal{O}_+} &\leq \|\Pi_K^\perp f(\varphi,x)\|_s^{\gamma,\mathcal{O}} + C_s(\|f\|_{s_0}^{\gamma,\mathcal{O}}\|g\|_{s+1}^{\gamma,\mathcal{O}_+} + \|f\|_s^{\gamma,\mathcal{O}}\|g\|_{s_0+1}^{\gamma,\mathcal{O}_+}).\end{aligned} \tag{5.21}$$

Then if $s = s_0$ by applying the smoothing estimates (A.4) in the second inequality in (5.21) we get

$$\begin{aligned}\|F\|_{s_0}^{\gamma,\mathcal{O}_+} &\leq K^{s_0-s_1}\|f\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}K\|f\|_{s_0}^{\gamma,\mathcal{O}}\|g\|_{s_0}^{\gamma,\mathcal{O}_+}\\ &\overset{(5.17)}{\leq} K^{s_0-s_1}\|f\|_{s_1}^{\gamma,\mathcal{O}} + \gamma^{-1}C_{s_0}K^{2\tau+2}(\|f\|_{s_0}^{\gamma,\mathcal{O}})^2,\\ \|f_+\|_{s_0}^{\gamma,\mathcal{O}_+} &\leq (K^{s_0-s_1}\|f\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}\gamma^{-1}K^{2\tau+2}(\|f\|_{s_0}^{\gamma,\mathcal{O}})^2)(1 + C_{s_0}\gamma^{-1}K^{2\tau+2s_0+1}\|f\|_{s_0}^{\gamma,\mathcal{O}}).\end{aligned} \tag{5.22}$$

If $s > s_0$ by (5.21) and (5.17) we just get

$$\|f_+\|_s^{\gamma,\mathcal{O}_+} \leq \|F\|_s^{\gamma,\mathcal{O}_+}(1 + C_s\gamma^{-1}K^{2\tau+s_0+2}\|f\|_{s_0}^{\gamma,\mathcal{O}}) + C_s\gamma^{-1}K^{2\tau+2s_0+1}\|f\|_s^{\gamma,\mathcal{O}}\|F\|_{s_0}^{\gamma,\mathcal{O}_+} \tag{5.23}$$

with

$$\|F\|_s^{\gamma,\mathcal{O}_+} \leq \|\Pi_K^\perp f\|_s^{\gamma,\mathcal{O}} + 2C_s\gamma^{-1}K^{2\tau+2}\|f\|_{s_0}^{\gamma,\mathcal{O}}\|f\|_s^{\gamma,\mathcal{O}}, \tag{5.24}$$

by using (5.2) we get (5.10).

If the vector field f is reversible, then $f = \mathrm{even}(\theta)$, therefore by the formulae (5.4), (5.5), one has that $g = \mathrm{odd}(\theta)$, implying that the diffeomorphism $\theta \mapsto \theta + g(\theta)$ is reversibility preserving. It then follows that the function $F = \mathrm{even}(\theta)$, by (5.20). By (5.19), one has that $f_+ = \Phi^{-1}F$, hence

$$f_+(-\theta) = F(-\theta + g(-\theta)) = F\Big(- \big(\theta + g(\theta)\big)\Big) = F(\theta + g(\theta)) = f_+(\theta)$$

implying that $f_+ = \mathrm{even}(\theta)$. This proves that f_+ is a reversible vector field.

Now let $\mathtt{b}$ satisfy (5.11) and $\lambda_1, \lambda_2 \subseteq B_E$ where $B_E \subseteq E$ is a ball in the Banach space E. We introduce the notation

$$\mathcal{M}_s(f, \lambda_1, \lambda_2) := \max\{\|f(\lambda_1)\|_s\,,\, \|f(\lambda_2)\|_s\}\,, \quad \forall s \geq 0\,.$$

The bound in (5.13) follows since

$$|\Delta_{12} g_k| \leq \frac{|\Delta_{12} f_k|}{|\alpha(\lambda_1)\cdot k|} + \frac{|f_k(\lambda_2)||\Delta_{12}\alpha||k|}{|\alpha(\lambda_1)\cdot k||\alpha(\lambda_2)\cdot k|} \leq \gamma^{-1}\langle k\rangle^{\tau}|\Delta_{12} f_k| + \gamma^{-2}\langle k\rangle^{2\tau+1}|f_k(\lambda_2)||\Delta_{12}\alpha|$$

for all $\xi \in \mathcal{O}_+(\lambda_1)\cap\mathcal{O}_+(\lambda_2)$. Now under the same hypotheses

$$\begin{aligned} \Delta_{12}\alpha_+ &= \Delta_{12}\alpha + \langle \Delta_{12} f\rangle, \\ \Delta_{12} f_+ &= (\Delta_{12}\Phi^{-1})\big[\Pi_K^{\perp} f(\lambda_1) + f(\lambda_1)\cdot\partial_\theta g(\lambda_1)\big] \\ &\quad + \Phi^{-1}(\lambda_2)\big[\Pi_K^{\perp}\Delta_{12} f + (\Delta_{12} f)\cdot\partial_\theta g(\lambda_1) + f(\lambda_2)\cdot\partial_\theta(\Delta_{12} g)\big]\,. \end{aligned} \tag{5.25}$$

By using the mean value Theorem, by applying Lemma A.3 and using the estimate (5.18), for any $s \in [s_0 - 1, s_0 + \mathtt{b}]$, one gets

$$\begin{aligned} \|\Delta_{12}\Phi^{-1}[u]\|_s &\lesssim_s \|u\|_{s+1}\Big(1 + \|g(\lambda_1)\|_{s+2s_0} + \|g(\lambda_2)\|_{s+2s_0}\Big)\|\Delta_{12} g\|_{s+s_0} \\ &\overset{(5.6)}{\lesssim_s} \|u\|_{s+1}\Big(1 + C_{s_0}\gamma^{-1}\mathcal{M}_{s+2s_0+\tau}(f,\lambda_1,\lambda_2)\Big)\|\Delta_{12} g\|_{s+s_0} \\ &\overset{s_2+\mathtt{b}+2s_0\leq s_1}{\lesssim_s} \|u\|_{s+1}\Big(1 + C_{s_0}\gamma^{-1}\mathcal{M}_{s_1}(f,\lambda_1,\lambda_2)\Big)\|\Delta_{12} g\|_{s+s_0} \\ &\overset{(5.12)}{\lesssim_s} \|u\|_{s+1}\|\Delta_{12} g\|_{s+s_0}\,. \end{aligned} \tag{5.26}$$

Using (5.12), applying the estimate (5.13) and Lemma A.1-(iv), one gets

$$\|\Delta_{12}\Phi^{-1}[u]\|_s \lesssim_s \gamma^{-1}\|u\|_{s+1}\Big(\|\Pi_K\Delta_{12} f\|_{s+s_0+\tau} + \gamma^{-1}|\Delta_{12}\alpha|\|\Pi_K f\|_{s+s_0+2\tau+1}\Big)\,, \tag{5.27}$$

for $s \in [s_0 - 1, s_0 + \mathtt{b}]$. Then we have

$$\begin{aligned} \|u\|_{s+1} = \|\Pi_K^{\perp} f(\lambda_1) + f(\lambda_1)\cdot\partial_\theta g(\lambda_1)\|_{s+1} &\leq \|f(\lambda_1)\|_{s+1}\big(1 + C\|g(\lambda_1)\|_{s+2}\big) \\ &\overset{(5.17)}{\leq} \|f(\lambda_1)\|_{s+1}\Big(1 + C\gamma^{-1}\|\Pi_K f(\lambda_1)\|_{s+\tau+2}\Big) \\ &\overset{(5.12)}{\lesssim_s} \|f(\lambda_1)\|_{s+1} \end{aligned}$$

taking δ' in (5.12) small enough. The above estimates imply that

$$\begin{aligned} &\|(\Delta_{12}\Phi^{-1})\big[\Pi_K^{\perp} f(\lambda_1) + f(\lambda_1)\cdot\partial_\theta g(\lambda_1)\big]\|_s \\ &\lesssim_s \gamma^{-1}\|f\|_{s+1}\Big(\|\Pi_K\Delta_{12} f\|_{s+s_0+\tau} + \gamma^{-1}|\Delta_{12}\alpha|\|\Pi_K f\|_{s+s_0+2\tau+1}\Big). \end{aligned}$$

Specializing the above estimate for $s = s_0 - 1$ and $s = s_0 + \mathtt{b}$, one gets

$$\begin{aligned}&\|(\Delta_{12}\Phi^{-1})\big[\Pi_K^\perp f(\lambda_1) + f(\lambda_1)\cdot\partial_\theta g(\lambda_1)\big]\|_{s_0-1}\\ &\lesssim_{s_0} \gamma^{-1}K^{\tau+s_0}\|\Delta_{12}f\|_{s_0-1}\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2) + \gamma^{-2}K^{2\tau+s_0}|\Delta_{12}\alpha|\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2)^2\,,\end{aligned} \tag{5.28}$$

$$\begin{aligned}&\|(\Delta_{12}\Phi^{-1})\big[\Pi_K^\perp f(\lambda_1) + f(\lambda_1)\cdot\partial_\theta g(\lambda_1)\big]\|_{s_0+\mathtt{b}}\\ &\lesssim_{s_0,\mathtt{b}} \gamma^{-1}\|f\|_{s_0+\mathtt{b}+1}\Big(\|\Pi_K\Delta_{12}f\|_{2s_0+\mathtt{b}+\tau} + \gamma^{-1}|\Delta_{12}\alpha|\|\Pi_K f\|_{2s_0+\mathtt{b}+2\tau+1}\Big)\\ &\overset{(5.6),3s_0+\mathtt{b}+2\tau+2\le s_1,(5.12)}{\lesssim_{s_0,\mathtt{b}}} K^{\tau+s_0}\Big(\|\Delta_{12}f\|_{s_0+\mathtt{b}} + |\Delta_{12}\alpha|\Big)\,.\end{aligned} \tag{5.29}$$

Furthermore, by Lemma A.3, for any $s\in[s_0, s_0+\mathtt{b}]$

$$\begin{aligned}\|\widetilde{g}(\lambda_2)\|_{s+s_0} &\overset{(5.18)}{\lesssim_s} \|g(\lambda_2)\|_{s+2s_0} \overset{(5.6)}{\lesssim_s} \gamma^{-1}\|\Pi_K f(\lambda_2)\|_{s+2s_0+\tau}\\ &\overset{\mathtt{b}+3s_0+\tau\le s_1}{\lesssim_s} \gamma^{-1}\|f(\lambda_2)\|_{s_1} \overset{(5.12)}{\lesssim_s} 1\end{aligned} \tag{5.30}$$

provided that δ' in (5.12) is small enough. In this way we have, for any $s\in[s_0,s_0+\mathtt{b}]$,

$$\|\Phi^{-1}[u]\|_s \lesssim_s \|u\|_s(1+\|\tilde{g}\|_{s+s_0}) \lesssim_s \|u\|_s,$$

and consequently

$$\begin{aligned}&\|\Phi^{-1}(\lambda_2)\big[\Pi_K^\perp\Delta_{12}f + (\Delta_{12}f)\cdot\partial_\theta g(\lambda_1) + f(\lambda_2)\cdot\partial_\theta(\Delta_{12}g)\big]\|_s\\ &\lesssim_s \|\Pi_K^\perp\Delta_{12}f\|_s + \|\Delta_{12}f\|_s\|g(\lambda_1)\|_{s+1} + \|f(\lambda_2)\|_s\|\Delta_{12}g\|_{s+1}\\ &\lesssim_s \|\Pi_K^\perp\Delta_{12}f\|_s + \gamma^{-1}\|\Delta_{12}f\|_s\|\Pi_K f(\lambda_1)\|_{s+1+\tau}\\ &+\gamma^{-1}\|f(\lambda_2)\|_s\Big(\|\Pi_K\Delta_{12}f\|_{s+\tau+1} + \gamma^{-1}|\Delta_{12}\alpha|\|\Pi_K f\|_{s+2\tau+2}\Big).\end{aligned}$$

Then, using also Lemma A.1-(iv), we have for $s = s_0 - 1$

$$\begin{aligned}&\|\Phi^{-1}(\lambda_2)\big[\Pi_K^\perp\Delta_{12}f + (\Delta_{12}f)\cdot\partial_\theta g(\lambda_1) + f(\lambda_2)\cdot\partial_\theta(\Delta_{12}g)\big]\|_{s_0-1}\\ &\overset{(A.1)-(iv)}{\lesssim_{s_0,\mathtt{b}}} K^{-1-\mathtt{b}}\|\Delta_{12}f\|_{s_0+\mathtt{b}} + K^\tau\gamma^{-1}\|\Delta_{12}f\|_{s_0-1}\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2)\\ &\quad+K^{2\tau}\gamma^{-2}|\Delta_{12}\alpha|\mathcal{M}_{s_0}(f,\lambda_1,\lambda_2)^2,\end{aligned} \tag{5.31}$$

similarly for $s = s_0 + \mathtt{b}$.

$$\begin{aligned}&\|\Phi^{-1}(\lambda_2)\big[\Pi_K^\perp\Delta_{12}f + (\Delta_{12}f)\cdot\partial_\theta g(\lambda_1) + f(\lambda_2)\cdot\partial_\theta(\Delta_{12}g)\big]\|_{s_0+\mathtt{b}}\\ &\overset{(5.6),(5.13),s_0+\mathtt{b}+2\tau+2\le s_1}{\lesssim_{s_0,\mathtt{b}}} \|\Pi_K^\perp\Delta_{12}f\|_{s_0+\mathtt{b}} + K^\tau\gamma^{-1}\|\Delta_{12}f\|_{s_0+\mathtt{b}}\mathcal{M}_{s_1}(f,\lambda_1,\lambda_2)\\ &+K^{2\tau+1}\gamma^{-2}|\Delta_{12}\alpha|\mathcal{M}_{s_1}(f,\lambda_1,\lambda_2)^2 \overset{(5.12)}{\lesssim_{s_0,\mathtt{b}}} K^{2\tau+1}\big(\|\Delta_{12}f\|_{s_0+\mathtt{b}} + |\Delta_{12}\alpha|\big)\,.\end{aligned} \tag{5.32}$$

The estimate (5.15) then follows by recalling (5.25) and by applying the estimates (5.29), (5.32). □

5.2 KAM iteration

Now we describe the iteration of the KAM step.

Lemma 5.2. *Consider the vector field* X_0 *in* (3.2). *Recall* (3.1) *and set*

$$\chi := \frac{3}{2}\,, \quad \mu > 4\tau + 2s_0 + 4, \quad \varrho > 2\tau + 2s_0 + 1, \quad s_1 > \chi\mu + s_0,$$
$$\kappa > 8\tau + 2s_0 + 4\,, \quad \mathtt{b} > \mu\chi + \kappa + 1. \tag{5.33}$$

There exist K_0 *depending on* s_0, ν *and* $\delta_* := \delta_*(s_1)$ *small such that if*

$$\delta_0(s_1) K_0^{\varrho} \leq \delta_*, \qquad \text{where} \qquad \delta_0(s_1) := \gamma^{-1}\|f_0\|_{s_1}^{\gamma,\mathcal{O}_0} \tag{5.34}$$

then, setting $K_n := K_0^{\chi^n}$, $\chi := 3/2$,

$$\mathcal{O}_{n+1} = \mathcal{C}_{K_n,\mathcal{O}_n} := \left\{ \xi \in \mathcal{O}_n : |\alpha_n(\xi)\cdot k| \geq \frac{\gamma}{\langle k\rangle^{\tau}}, \quad \forall k \in \mathbb{Z}^N \setminus \{0\}, \quad |k| \leq K_n \right\}, \tag{5.35}$$

$$g_{n+1}(\theta;\xi) := \sum_{|k|\leq K_n} g_k^{(n+1)} e^{\mathrm{i}k\cdot\theta}\,, \quad \xi \in \mathcal{O}_{n+1}$$
$$g_k^{(n+1)} := \frac{a_k^{(n)}}{\mathrm{i}\alpha_n \cdot k}\,, \quad k \in \mathbb{Z}^N \setminus \{0\}\,, \quad |k| \leq K_n, \tag{5.36}$$

$$\delta_n(s) := \gamma^{-1}\|f_n\|_s^{\gamma,\mathcal{O}_n}, \tag{5.37}$$

and

$$\lambda := \lambda(s) := 1/(s - s_0 + 1), \quad \mathtt{M}(s) := \max\{\delta_0(s_1), \delta_0(s)\}, \tag{5.38}$$

the following holds for any $n \geq 0$.

$(\mathcal{P}_1)_n$. *Set* $g_0 = 0$. *The torus diffeomorphism* $\Phi_n : \theta \mapsto \theta + g_n(\theta)$ *is well defined and the induced operator* (2.8) *acts on* H^s *to itself* $\forall s \geq s_0$. *Setting*

$$X_n := (\Phi_n)_* X_{n-1} := \big(\alpha_n(\xi) + f_n(\theta;\xi)\big) \cdot \frac{\partial}{\partial\theta} \tag{5.39}$$

we have the bounds

$$|\alpha_n - \alpha_{n-1}|^{\gamma} \lesssim \gamma\delta_0(s_1) K_n^{-\mu} K_0^{\mu}\,, \quad |\alpha_n|^{lip} \leq M_0 + C\gamma^{-1}\delta_0(s_1) \tag{5.40}$$

and there exists a positive constant $C_1(s)$ *such that*

$$\delta_n(s_0) \leq \delta_0(s_1)\, K_0^{\mu}\, K_n^{-\mu}, \quad \delta_n(s) \leq C_1(s)\, \delta_0(s)(1 + \sum_{j=1}^{n} 2^{-j})\,, \quad s \geq s_0. \tag{5.41}$$

As a consequence

$$\delta_n(s) \leq C(s)\, K_n^{-\lambda\mu} K_0^{\lambda\mu} \mathtt{M}(s+1)\,, \tag{5.42}$$
$$\|g_n\|_s^{\gamma,\mathcal{O}_n} \leq \delta_n(s + 2\tau + 1) \leq C_2(s)\, K_n^{-\lambda\mu} K_0^{\lambda\mu}\, \mathtt{M}(s + 2\tau + 2), \quad s \geq s_0 \tag{5.43}$$

for some $C_2(s) > 0$.

$(\mathcal{P}_2)_n$. *The torus diffeomorphism defined by*

$$\begin{cases} \Psi_0 = \mathrm{Id}, \\ \Psi_n = \Phi_n \circ \Psi_{n-1} \end{cases} \tag{5.44}$$

is of the form $\Psi_n : \theta \mapsto \theta + h_n(\theta)$ *with, for all* $s \geq s_0$, *(recall* (5.43) *for the definition of* $\mathtt{M}(s)$*)*

$$\|h_n\|_{s_0}^{\gamma,\mathcal{O}_n} \leq C(s_0)\delta_0(s_1) \sum_{j=0}^{n} 2^{-j}\,, \quad \|h_n\|_s^{\gamma,\mathcal{O}_n} \leq C_3(s)\, \mathtt{M}(s + 2\tau + s_0 + 2) \sum_{j=0}^{n} 2^{-j}\,, \tag{5.45}$$

$$\|h_{n-1}-h_n\|_{s_0}^{\gamma,\mathcal{O}_n} \le C(s_0)\delta_0(s_1)2^{-n}\,, \quad \|h_{n-1}-h_n\|_{s}^{\gamma,\mathcal{O}_n} \le C_4(s)\,\mathtt{M}(s+2\tau+s_0+3)2^{-n}. \tag{5.46}$$

Moreover, if f_0 *is a reversible vector fields, then* $\theta \mapsto \theta + g_n(\theta)$, $\theta \mapsto \theta + h_n(\theta)$ *are reversibility preserving maps and* f_n *is a reversible vector field.*

$(\mathcal{P}_3)_n$. *Let* $\lambda_1, \lambda_2 \in B_E$. *There exists a constant* $C_*(s_1) > 0$ *and* $\tilde{\delta} := \tilde{\delta}(s_1)$ *such that if*

$$K_0^{2\tau+s_0+\chi\mu}\gamma^{-1}\big(\|f_0(\lambda_1)\|_{s_1} + \|f_0(\lambda_2)\|_{s_1}\big) \le \tilde{\delta} \tag{5.47}$$

then for any $n \ge 0$, *for any* $\xi \in \mathcal{O}_n(\lambda_1)\cap\mathcal{O}_n(\lambda_2)$, *the following estimates hold:*

$$\|\Delta_{12} f_n\|_{s_0-1} \le C_*(s_1)K_n^{-\mu}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\,, \tag{5.48}$$

$$\|\Delta_{12} f_n\|_{s_0+\mathtt{b}} \le C_*(s_1)K_n^{\kappa}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\,, \tag{5.49}$$

$$|\Delta_{12}(\alpha_{n+1}-\alpha_n)| \le \|\Delta_{12}f_n\|_{s_0-1}, \tag{5.50}$$

$$|\Delta_{12}\alpha_n| \lesssim \|\Delta_{12}f_0\|_{s_0+\mathtt{b}}, \tag{5.51}$$

$$\|\Delta_{12}h_n\|_{s_0-1} \le C_*(s_1)\gamma^{-1}\sum_{j=0}^{n}2^{-j}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\,. \tag{5.52}$$

$(\mathcal{P}_4)_n$. *Let* $\lambda_1, \lambda_2 \in B_E$, $0 < \gamma - \rho < \gamma < 1$ *satisfy*

$$K_{n-1}^{\tau+1}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}} \le \rho\,. \tag{5.53}$$

Then $\mathcal{O}_n^{\gamma}(\lambda_1) \subseteq \mathcal{O}_n^{\gamma-\rho}(\lambda_2)$.

Proof. The statements $(\mathcal{P}_{1,2})_0$ are trivial. $(\mathcal{P}_3)_0$ follows taking $C_*(s_1)$ large enough, for instance $C_*(s_1) > K_0^{\mu}$. The statement $(\mathcal{P}_3)_0$ holds by setting $\mathcal{O}_0^{\gamma}(\lambda_1) = \mathcal{O}_0 = \mathcal{O}_0^{\gamma-\rho}(\lambda_2)$.
Now suppose that $(\mathcal{P}_{1,2})_n$ hold and we prove that $(\mathcal{P}_{1,2})_{n+1}$ also hold.

Proof of $(\mathcal{P}_1)_{n+1}$. We have to prove that the $(n+1)$-th diffeomorphism of the torus is well defined from H^s to itself for all $s \ge s_0$. In particular, we show that (5.2) holds with $K \rightsquigarrow K_n$ and $f \rightsquigarrow f_n$.
We have

$$\delta_n(s_0)K_n^{2\tau+2s_0+1} \le \delta_0(s_1)\,K_n^{2\tau+2s_0+1-\mu}K_0^{\mu}. \tag{5.54}$$

By (5.33) $\mu > 2\tau+2s_0+1$. Hence $K_n^{2\tau+2s_0+1-\mu}$ is a decreasing sequence and by (5.34), (5.33) $(\rho > 2\tau+2s_0+1)$

$$\delta_0(s_1)\,K_n^{2\tau+2s_0+1-\mu}K_0^{\mu} \le \delta_0(s_1)K_0^{2\tau+2s_0+1} \le \delta_*. \tag{5.55}$$

Then by (5.54), (5.55) and taking $\delta_* \le \delta$ (recall (5.34) and (5.2)) we get our first claim

$$\delta_n(s_0)K_n^{2\tau+2s_0+1} \le \delta.$$

In order to prove (5.41) we apply the KAM step with $f_+ \rightsquigarrow f_{n+1}$.
We start by estimating the low norm. By (5.22)

$$\begin{aligned}\delta_{n+1}(s_0) &\le \gamma^{-1}(K_n^{s_0-s_1}\|f_n\|_{s_1}^{\gamma,\mathcal{O}_n} + C_{s_0}\gamma^{-1}K_n^{2\tau+2}(\|f_n\|_{s_0}^{\gamma,\mathcal{O}_n})^2)(1+C_{s_0}\gamma^{-1}K_n^{2\tau+2s_0+1}\|f_n\|_{s_0}^{\gamma,\mathcal{O}_n})\\ &\le (K_n^{s_0-s_1}\delta_n(s_1) + C_{s_0}K_n^{2\tau+2}(\delta_n(s_0)^2)(1+C_{s_0}K_n^{2\tau+2s_0+1}\delta_n(s_0)\;).\end{aligned} \tag{5.56}$$

We first note that $C_{s_0}K_n^{2\tau+2s_0+1-\mu}K_0^{\mu}\,\delta_0(s_1) < 1$, indeed, since $\mu > 2\tau+2s_0+1$, this is a decreasing sequence and by (5.34) (taking δ_* small enough) and $\varrho > 2\tau+2s_0+1$ (see (5.33))

$$C_{s_0}K_0^{2\tau+2s_0}\,\delta_0(s_1) \le C_{s_0}\delta_* < 1.$$

Hence (recall (5.38))

$$\delta_{n+1}(s_0) \le 2K_n^{s_0-s_1}\delta_n(s_1) + 2C_{s_0}K_n^{2\tau+2}\delta_n(s_0)^2 \le \delta_0(s_1)\, K_{n+1}^{-\lambda\mu}K_0^{\lambda\mu}$$

provided that

$$\begin{cases} 2\delta_n(s_1)K_n^{-(s_1-s_0)} < \frac{1}{2}\delta_0(s_1)\, K_0^{\mu}\, K_{n+1}^{-\mu}, \\ 2\, C_{s_0}\, \delta_n(s_0)^2 K_n^{2\tau+2} < \frac{1}{2}\delta_0(s_1)\, K_0^{\mu}\, K_{n+1}^{-\mu}. \end{cases} \tag{5.57}$$

Thus, by the inductive hypotesis (5.41), we have to prove

$$8C_1(s_1)K_n^{-(s_1-s_0)+\chi\mu}K_0^{-\mu} < 1\,, \quad 4C_{s_0}\delta_0(s_1)K_0^{\mu}K_n^{2\tau+2-(2-\chi)\mu} < 1. \tag{5.58}$$

By (5.33) we have

$$s_1 - s_0 > \chi\mu\,, \quad \mu > \frac{2\tau+2}{2-\chi} = 4\tau+4 \overset{(3.1)}{=} 4\nu+12, \tag{5.59}$$

then the sequences in (5.58) are decreasing and we just need

$$8C_1(s_1)K_0^{-(s_1-s_0)+\mu(\chi-1)} < 1\,, \quad 4C_{s_0}\delta_0(s_1)K_0^{2\tau+2+\mu(\chi-1)} < 1,$$

which follows by taking K_0 sufficiently large (depending on C_{s_0} and $C_1(s_1)$ in (5.41)) and by (5.34), since

$$\varrho > 2\tau+2+\mu(\chi-1) = 2\tau+2+\frac{\mu}{2}.$$

Regarding the estimates in high norm, by (5.10) we have for all $s > s_0$

$$\|f_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} \le \|f_n\|_s^{\gamma,\mathcal{O}_n} + C_s\gamma^{-1}K_n^{2\tau+1+2s_0}\|f_n\|_{s_0}^{\gamma,\mathcal{O}_n}\|f_n\|_s^{\gamma,\mathcal{O}_n}. \tag{5.60}$$

First we prove the following bounds for $s > s_0$

$$\delta_n(s) \le (\mathtt{C}(s))^n\delta_0(s)(1+\sum_{j=1}^{n}2^{-j}), \quad 0 \le n \le n_0(s), \tag{5.61}$$

$$\delta_n(s) \le (\mathtt{C}(s))^{n_0(s)}\delta_0(s)(1+\sum_{j=1}^{n}2^{-j}), \quad n \ge n_0(s), \tag{5.62}$$

for some constant $\mathtt{C}(s)$ and for a suitable $n_0(s)$. For $n=0$ (5.61) holds by taking $\mathtt{C}(s) \ge 1$. For $n \le n_0(s)-1$ we have, by (5.60),

$$\begin{aligned} \delta_{n+1}(s) &\le \delta_n(s)(1+C_s\gamma K_n^{2\tau+2s_0+1}\delta_n(s_0)) \\ &\le (\mathtt{C}(s))^n\delta_0(s)(1+\sum_{j=1}^{n}2^{-j})(1+C_s\gamma K_n^{2\tau+2s_0+1}\delta_n(s_0)) \\ &\le (\mathtt{C}(s))^{n+1}\delta_0(s)(1+\sum_{j=1}^{n+1}2^{-j}) \end{aligned}$$

provide that

$$\frac{C_s}{\mathtt{C}(s)}\gamma K_n^{2\tau+2s_0+1-\mu}K_0^{\mu}\delta_0(s_1) \le \frac{2^{-(n+1)}}{1+\sum_{j=1}^{n}2^{-j}}. \tag{5.63}$$

Considering that $n \le n_0(s)-1$, by (5.33) and (5.34) with δ_* small enough, we get (5.63).
Now consider $n \ge n_0(s)$. By (5.60) we have

$$\begin{aligned} \delta_{n+1}(s) &\le \delta_n(s)(1+C_s\gamma K_n^{2\tau+2s_0+1}\delta_n(s_0)) \\ &\le (\mathtt{C}(s))^{n_0(s)}\delta_0(s)(1+\sum_{j=1}^{n}2^{-j})(1+C_s\gamma K_n^{2\tau+2s_0+1}\delta_n(s_0)) \\ &\le (\mathtt{C}(s))^{n_0(s)}\delta_0(s)(1+\sum_{j=1}^{n+1}2^{-j}) \end{aligned}$$

provide that

$$C_s\gamma K_n^{2\tau+2s_0+1-\mu}K_0^{\mu}\delta_0(s_1) \le \frac{2^{-(n+1)}}{1+\sum_{j=1}^{n}2^{-j}}. \tag{5.64}$$

The bound (5.64) follows by (5.33), (5.34) and by choosing $n_0(s)$ large enough. Hence we proved the second estimate in (5.41) by setting

$$C_1(s) := (\mathtt{C}(s))^{n_0(s)}.$$

Now we prove (5.43). For $s \ge s_0$, setting $\lambda = 1/(s-s_0+1)$, we have

$$\|f_n\|_s^{\gamma,\mathcal{O}_n} \le (\|f_n\|_{s_0}^{\gamma,\mathcal{O}_n})^{\lambda}(\|f_n\|_{s+1}^{\gamma,\mathcal{O}_n})^{1-\lambda}, \tag{5.65}$$

from which we may deduce that (recall (5.41))

$$\delta_n(s) \le (K_n^{-\mu}K_0^{\mu}\delta_0(s_1))^{\lambda}(\delta_n(s+1))^{1-\lambda} \le 2\,C_1(s+1)\,K_n^{-\lambda\mu}K_0^{\lambda\mu}\mathtt{M}(s+1). \tag{5.66}$$

By (5.66)

$$\|g_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} \le \delta_n(s+2\tau+1) \le 2\,C_1(s+2\tau+2)\,K_n^{-\lambda\mu}K_0^{\lambda\mu}\mathtt{M}(s+2\tau+2)\,, \quad s \ge s_0 \tag{5.67}$$

which is (5.43) taking $C_2(s) \ge 2\,C_1(s+2\tau+2)$. The bounds (5.41) trivially implies (5.40).

Proof of $(\mathcal{P}_2)_{n+1}$. By construction

$$h_{n+1}(\theta) = g_{n+1}(\theta) + h_n(\theta + g_{n+1}(\theta)) \tag{5.68}$$

thus, by (A.8), for $s \in \mathbb{N},\ s \ge s_0$

$$\begin{aligned}\|h_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} \le \|g_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} + \|h_n\|_s^{\gamma,\mathcal{O}_n}(1 + C_s\|g_{n+1}\|_{s_0+1}^{\gamma,\mathcal{O}_{n+1}}) \\ + C_s\|h_n\|_{s_0}^{\gamma,\mathcal{O}_n}\|g_{n+1}\|_{s+s_0}^{\gamma,\mathcal{O}_{n+1}}.\end{aligned} \tag{5.69}$$

First we show the following. By fixing an opportune $n_0(s) \in \mathbb{N}$, we have the bounds

$$\|h_n\|_s \le (\mathtt{C}(s))^n\mathtt{M}(s+2\tau+s_0+2)\sum_{j=0}^{n}2^{-j} \quad 0 \le n \le n_0(s), \tag{5.70}$$

$$\|h_n\|_s \le (\mathtt{C}(s))^{n_0(s)}\mathtt{M}(s+2\tau+s_0+2)\sum_{j=0}^{n}2^{-j} \quad n \ge n_0(s). \tag{5.71}$$

We recall that $h_0 := \alpha_0 = 0$. By (5.67), (5.45), (5.69) we have for $n \le n_0(s)-1$

$$\begin{aligned}&\|h_{n+1}\|_s^{\gamma,\mathcal{O}_n} \le C_2(s)\,K_n^{-\lambda\mu}K_0^{\lambda\mu}\,\mathtt{M}(s+2\tau+2+s_0)\,(1+C_sC(s_0)\delta_0(s_1)\sum_{j=0}^{n}2^{-j}) \\ &+ (\mathtt{C}(s))^n\,\mathtt{M}(s+2\tau+s_0+2)\,\sum_{j=0}^{n}2^{-j}(1+C_sK_n^{2\tau+2-\mu}K_0^{\mu}\delta_0(s_1)) \le \\ &\le (\mathtt{C}(s))^{n+1}\,\mathtt{M}(s+2\tau+s_0+2)\sum_{j=0}^{n+1}2^{-j}\end{aligned}$$

provided that we choose $\mathtt{C}(s)$ such that (recall (5.34))

$$(\mathtt{C}(s))^n \ge 2K_n^{-\lambda\mu}C_sC(s_0)\delta_0(s_1), \quad \mathtt{C}(s) \ge \max\{2(1+C_s), C_2(s)K_0^{\lambda\mu}\}. \tag{5.72}$$

Hence we proved (5.70). Now, by (5.67), (5.45), (5.69), we have for $n \geq n_0(s)$

$$\begin{aligned}\|h_{n+1}\|_s^{\gamma,\mathcal{O}_n} &\leq C_2(s)\, K_n^{-\lambda\mu} K_0^{\lambda\mu}\, \mathtt{M}(s+2\tau+2+s_0)\,(1+C_s C(s_0)\delta_0(s_1)\sum_{j=0}^{n} 2^{-j})\\ &+ (\mathtt{C}(s))^{n_0(s)}\, \mathtt{M}(s+2\tau+s_0+2)\sum_{j=0}^{n} 2^{-j}(1+C_s K_n^{2\tau+2-\mu} K_0^{\mu}\delta_0(s_1)) \leq\\ &\leq (\mathtt{C}(s))^{n_0(s)}\, \mathtt{M}(s+2\tau+s_0+2)\sum_{j=0}^{n+1} 2^{-j}\end{aligned}$$

provide that we choose $n_0(s)$ large enough so that (recall $\mu > 2\tau+2$)

$$C_s K_n^{2\tau+2-\mu} K_0^{\mu}\delta_0(s_1) \leq \frac{2^{-(n+1)}}{1+\sum_{j=1}^{n} 2^{-j}}$$

and by using (5.72) with $n = n_0(s)$.
Hence we proved (5.45) with

$$C_3(s) := \max\{(\mathtt{C}(s))^{n_0(s)}, C(s_0)\}.$$

In order to prove the first bound in (5.46) we use (5.67), (5.45), (A.6b) and we have

$$\|h_{n+1}-h_n\|_{s_0}^{\gamma,\mathcal{O}_{n+1}} \leq \|g_{n+1}\|_{2s_0}(1+C_{s_0}\|h_n\|_{s_0+1}^{\gamma,\mathcal{O}_n}). \tag{5.73}$$

By interpolation we get

$$\|h_n\|_{s_0+1}^{\gamma,\mathcal{O}_n} \leq (\|h_n\|_{s_0}^{\gamma,\mathcal{O}_n})^{1/2}(\|h_n\|_{s_0+2}^{\gamma,\mathcal{O}_n})^{1/2} \overset{(5.45)}{\leq} \tilde{C}(s_0)\delta_0(s_1)\sum_{j=0}^{n} 2^{-j} \leq 2\tilde{C}(s_0)\delta_0(s_1) \overset{(5.34)}{\leq} 1$$

where

$$\tilde{C}(s_0) := (C(s_0))^{1/2}(C_3(s_0+2))^{1/2}.$$

Hence we have by (5.73)

$$\begin{aligned}\|h_{n+1}-h_n\|_{s_0}^{\gamma,\mathcal{O}_{n+1}} \leq 2\|g_{n+1}\|_{2s_0}^{\gamma,\mathcal{O}_{n+1}} \overset{(5.6)}{\leq} K_n^{2\tau+1+s_0}\delta_n(s_0) &\overset{(5.41)}{\leq} \delta_0(s_1) K_n^{2\tau+1+s_0-\mu} K_0^{\mu}\\ &\leq C(s_0)\delta_0(s_1)2^{-(n+1)}\end{aligned}$$

provided that $C(s_0) > K_0^{\mu}$, $\mu > 2\tau+1+s_0$ and $K_0 > 1$ is large enough. Now we prove the second bound in (5.46).
By (5.67), (5.45), (A.6b), (5.43) we have

$$\begin{aligned}\|h_{n+1}-h_n\|_s^{\gamma,\mathcal{O}_{n+1}} &\leq \|g_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} + C(s)(\|h_n\|_{s+1}^{\gamma,\mathcal{O}_n}\|g_{n+1}\|_{s_0}^{\gamma,\mathcal{O}_{n+1}} + \|h_n\|_{s_0}^{\gamma,\mathcal{O}_n}\|g_{n+1}\|_{s+s_0}^{\gamma,\mathcal{O}_{n+1}})\\ &\leq C_4(s)\,\mathtt{M}(s+2\tau+s_0+3)\,2^{-(n+1)}\end{aligned}$$

provided that $C_4(s)$ is large enough and

$$K_n^{-\lambda\mu}K_0^{\lambda\mu} \leq 2^{-(n+1)}, \quad K_n^{2\tau+1}\delta_n(s_0) \leq 2^{-(n+3)}, \quad 2\delta_0(s_1)K_n^{-\lambda\mu}K_0^{\lambda\mu} \leq 1$$

which hold by taking $K_0 > 1$ large enough, by (5.41) and (5.34).
If f_n is a reversible vector field, by Lemma 5.1 (recall also the definitions (5.36))one has that $\theta \mapsto \theta + g_{n+1}(\theta)$ is a reversibility preserving map and f_{n+1} is a reversible vector field. Furthermore, since by the inductive hypotheses, $\theta \mapsto \theta + h_n(\theta)$ is a reversibility preserving map, by the formula (5.68) one immediately gets that $\theta \mapsto \theta + h_{n+1}(\theta)$ is reversibility preserving too.

Proof of $(\mathcal{P}_3)_{n+1}$. If we take (recall δ_* in (5.34), (5.33))

$$\tilde{\delta} \leq K_0^{-\varrho+2\tau+1}\delta_*,$$

since $f_0(\lambda_1)$ and $f_0(\lambda_2)$ satisfy the smallness assumption (5.47), then condition (5.34) holds for both $f_0(\lambda_1)$ and $f_0(\lambda_2)$ and we can apply the estimates proved in the steps $(\mathcal{P}_1)_n$, $(\mathcal{P}_2)_n$ obtaining that

$$\|f_n(\lambda_1)\|_{s_1}, \|f_n(\lambda_2)\|_{s_1} \overset{(5.41)}{\lesssim_{s_1}} \mathcal{M}_{s_1}(f_0,\lambda_1,\lambda_2). \tag{5.74}$$

This estimate implies that (5.12) is verified by (5.47). Then by applying (5.15), one gets that

$$\begin{aligned}
\|\Delta_{12}f_{n+1}\|_{s_0-1} &\lesssim_{s_0} K_n^{-1-\mathtt{b}}\|\Delta_{12}f_n\|_{s_0+\mathtt{b}} + K_n^{\tau+s_0}\|\Delta_{12}f_n\|_{s_0-1}\mathcal{M}_{s_0}(f_n,\lambda_1,\lambda_2) \qquad (5.75)\\
&\quad + K_n^{2\tau+s_0}\gamma^{-2}|\Delta_{12}\alpha_n|\mathcal{M}_{s_0}(f_n,\lambda_1,\lambda_2)^2\\
&\overset{(5.41),(5.50),(5.47),(5.48),(5.49)}{\lesssim_{s_0,\mathtt{b}}} K_n^{\kappa-1-\mathtt{b}}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\\
&\quad + C_*(s_1)K_n^{2\tau+s_0-2\mu}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}K_0^{2\mu}\delta_0(s_1)\\
&\leq C_*(s_1)K_{n+1}^{-\mu}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}} \qquad (5.76)
\end{aligned}$$

provided for any $n \geq 0$,

$$C(s_0,\mathtt{b})K_{n+1}^{\mu}K_n^{-\mathtt{b}-1+\kappa} \leq \frac{C_*(s_1)}{2}, \quad C(s_0,\mathtt{b})K_{n+1}^{\mu}K_n^{2\tau+s_0-2\mu}K_0^{2\mu}\delta_0(s_1) \leq \frac{C_*(s_1)}{2}.$$

As in the previous items the left hand side of these inequalities is decreasing in n, since by (5.33) we have $\mu > \frac{2\tau+s_0}{(2-\chi)}$, $\mathtt{b} > \mu\chi+\kappa-1$. Then our claim follows by taking $K_0, C_*(s_1) > 0$ large enough. Moreover

$$\begin{aligned}
\|\Delta_{12}f_{n+1}\|_{s_0+\mathtt{b}} &\lesssim_{s_0,\mathtt{b}} K_n^{2\tau+s_0}\Big(\|\Delta_{12}f_n\|_{s_0+\mathtt{b}} + |\Delta_{12}\alpha_n|\Big)\\
&\overset{(5.49),(5.51)}{\lesssim_{s_0,\mathtt{b}}} K_n^{\kappa+2\tau+s_0}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}} \leq K_{n+1}^{\kappa}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}
\end{aligned}$$

provided $C(s_0,\mathtt{b})K_n^{2\tau+s_0+\kappa} \leq K_{n+1}^{\kappa}$ for any $n \geq 0$. Such a condition is fullfilled, by taking $K_0 > 1$ large enough, since by (5.33) one has that $(\chi-1)\kappa > 2\tau+s_0$. Therefore, the estimates (5.48), (5.49) have been proved at the step $n+1$. The estimates (5.50), (5.51) follow by applying (5.14) by using a telescoping argument.
The estimates (5.6), (5.45), using that $2\tau+2s_0+3 \leq s_1$ imply that

$$\begin{aligned}
\|g_n(\lambda_1)\|_{2s_0}, \|g_n(\lambda_2)\|_{2s_0} &\lesssim_{s_0} K_n^{\tau+s_0}\gamma^{-1}\mathcal{M}_{s_0}(f_n,\lambda_1,\lambda_2)\\
&\lesssim_{s_0} K_n^{\tau+s_0-\mu}K_0^{\mu}\delta_0(s_1) \overset{(5.47)}{\leq} 1, \qquad (5.77)\\
\|h_n(\lambda_1)\|_{s_0}, \|h_n(\lambda_2)\|_{s_0} &\lesssim_{s_0} \mathcal{M}_{s_1}(a_0,\lambda_1,\lambda_2).
\end{aligned}$$

By (5.13), (5.41), (5.51) and (5.48), (5.47), one gets the estimate

$$\|\Delta_{12}g_n\|_{s_0-1} \lesssim_{s_0} \gamma^{-1}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}K_n^{2\tau+1-\mu}. \tag{5.78}$$

By the formula (5.68) one gets

$$\begin{aligned}
h_{n+1}(\theta;\lambda_1) - h_{n+1}(\theta;\lambda_2) &= g_{n+1}(\theta;\lambda_1) - g_{n+1}(\theta;\lambda_2)\\
&\quad + h_n(\theta+g_{n+1}(\theta;\lambda_1);\lambda_1) - h_n(\theta+g_{n+1}(\theta;\lambda_2);\lambda_2)\\
&= \Delta_{12}g_{n+1}(\theta) + (\Delta_{12}h_n)(\theta+g_{n+1}(\theta;\lambda_1))\\
&\quad + h_n(\theta+g_{n+1}(\theta;\lambda_1);\lambda_2) - h_n(\theta+g_{n+1}(\theta;\lambda_2);\lambda_2).
\end{aligned}$$

Using the triangular inequality, the mean value theorem, the estimates (5.77), (5.78), Lemma A.3 and the smallness condition (5.47), one gets the estimate

$$\begin{aligned}\|\Delta_{12}h_{n+1}\|_{s_0-1} &\leq C_{s_0}\gamma^{-1}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}K_n^{2\tau+1-\mu} \\ &\quad + \|\Delta_{12}h_n\|_{s_0-1}\Big(1+C_{s_0}K_n^{2\tau+1+s_0-\mu}\gamma^{-1}\mathcal{M}_{s_1}(f_0,\lambda_1,\lambda_2)\Big)\,.\end{aligned} \tag{5.79}$$

Then using the induction hypothesis (5.52), one gets

$$\begin{aligned}\|\Delta_{12}h_{n+1}\|_{s_0-1} &\leq C_{s_0}\gamma^{-1}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}K_n^{2\tau+1-\mu} + C_*(s_1)\sum_{j=0}^{n}2^{-j}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\gamma^{-1} \\ &\quad + C_*(s_1)C_{s_0}K_n^{2\tau+1+s_0-\mu}\gamma^{-1}\mathcal{M}_{s_1}(f_0,\lambda_1,\lambda_2)\sum_{j=0}^{n}2^{-j}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\gamma^{-1} \\ &\leq C_*(s_1)\sum_{j=0}^{n+1}2^{-j}\|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\gamma^{-1}\end{aligned} \tag{5.80}$$

provided

$$C_{s_0}K_n^{2\tau+1-\mu} \leq C_*(s_1)\sum_{j=0}^{n+1}2^{-j}, \qquad C_{s_0}\gamma^{-1}K_n^{2\tau+1+s_0-\mu}\mathcal{M}_{s_1}(f_0,\lambda_1,\lambda_2) \leq 1.$$

This condition is fullfilled, by (5.47), taking K_0 and $C_*(s_1)$ large enough, recalling that $K_n = K_0^{\chi^n}$ for any $n \geq 0$ and since $2\tau+1+s_0-\mu < 0$.
Finally, we prove the statement $(\mathcal{P}4)_{n+1}$. Let $\xi \in \mathcal{O}^{\gamma}_{n+1}(\lambda_1)$. By the definition (5.35), $\xi \in \mathcal{O}^{\gamma}_{n}(\lambda_1)$ and the induction hypothesis implies that $\xi \in \mathcal{O}^{\gamma-\rho}_{n}(\lambda_2)$. Since, trivially $\mathcal{O}^{\gamma}_{n}(\lambda_1) \subseteq \mathcal{O}^{\gamma-\rho}_{n}(\lambda_1)$, one has that

$$\xi \in \mathcal{O}^{\gamma}_{n+1}(\lambda_1) \subseteq \mathcal{O}^{\gamma-\rho}_{n}(\lambda_1) \cap \mathcal{O}^{\gamma-\rho}_{n}(\lambda_2)\,.$$

We can then apply the estimate (5.51) implying that for any $\omega \in \mathcal{O}^{\gamma}_{n+1}(\lambda_1) \subseteq \mathcal{O}^{\gamma-\rho}_{n}(\lambda_1) \cap \mathcal{O}^{\gamma-\rho}_{n}(\lambda_2)$ one has that

$$|\Delta_{12}\alpha_n| \lesssim \|\Delta_{12}f_0\|_{s_0+\mathtt{b}}\,.$$

Therefore, for any $k \in \mathbb{Z}^N \setminus \{0\}$, $|k| \leq K_n$, one has that

$$\begin{aligned}|\alpha_n(\xi;\lambda_2)\cdot k| &\geq |\alpha_n(\xi;\lambda_1)\cdot k| - |\Delta_{12}\alpha_n||k| \\ &\geq \frac{\gamma}{\langle k\rangle^{\tau}} - K_n\|\Delta_{12}a_0\|_{s_0+\mathtt{b}} \geq \frac{\gamma-\rho}{\langle k\rangle^{\tau}}\end{aligned} \tag{5.81}$$

By the condition (5.53). Then $\xi \in \mathcal{O}^{\gamma-\rho}_{n+1}(\lambda_2)$, which is the claimed statement. □

5.3 Proof of Theorem 3.1

Now we can prove the Theorem 3.1.

Proof. We fix s_1 as in (5.33) and choose $\eta_\star$ so that (3.4) implies (5.34), namely (recall (5.33))

$$K_0^{\varrho}\eta_* \leq \delta_*.$$

Consider now the sequence h_n defined in Lemma 5.2-($\mathcal{P}_2$). By formula (5.46) this is a Cauchy sequence in $H^s(\mathbb{T}^N)$ for all $s \geq s_0$. Let us denote by $h^{(\infty)}$ its limit. We note that $h^{(\infty)}$ belongs to $\cap_{s\geq s_0}H^s(\mathbb{T}^N)$, hence it is a C^∞ function in θ. As a consequence $\Psi^{(\infty)}$ is C^∞ torus diffeomorphism.

In the same way, by (5.40) the sequence α_n is a Cauchy sequence and we denote by α_∞ its limit. We claim that

$$(\Psi^{(\infty)})^{-1}\Big(\xi + f_0 + (\xi + f_0)\cdot\partial_\theta h^{(\infty)}\Big) = \alpha_\infty. \tag{5.82}$$

First we prove by induction that (recall (5.39))

$$(\Psi_n)_* X_0 = X_n. \tag{5.83}$$

For $n = 0$ this is trivially true. Now prove the $(n+1)$-th step. Recalling the definition (5.44), by the composition of pushforwards

$$(\Psi_{n+1})_* X_0 = (\Phi_{n+1})_* (\Psi_n)_* X_0 = (\Phi_{n+1})_* X_n = X_{n+1}.$$

Now by (5.83) we have that

$$(\Psi_n)^{-1}\Big(\xi + f_0 + (\xi + f_0)\cdot\partial_\theta h_n\Big) = \alpha_n + f_n. \tag{5.84}$$

By (5.41) the r. h. s. of (5.84) converges in H^{s_0} to α_∞. By the fact that h_n converges to $h^{(\infty)}$ in H^s, for every $s \geq s_0$, then

$$(\Psi^{(\infty)})^{-1}\Big(\xi + f_0 + (\xi + f_0)\cdot\partial_\theta h^{(\infty)}\Big) - (\Psi_n)^{-1}\Big(\xi + f_0 + (\xi + f_0)\cdot\partial_\theta h_n\Big)$$

converges to 0 in H^{s_0} by using triangle inequalities, the mean value theorem and the bounds given in Lemma A.3. Then we proved our claim.
By (5.82), setting $\Psi^{(\infty)} : \theta \mapsto (\theta + h^{(\infty)}(\theta))$, we have

$$\Psi_*^{(\infty)} X_0 = \alpha_\infty(\xi)\cdot\frac{\partial}{\partial\theta}, \quad \forall \xi \in \cap_n \mathcal{O}_n.$$

The bound (3.5) follows by (5.40). In order to complete the proof we need to show that

$$\mathcal{O}_\infty^{2\gamma} \subset \bigcap_n \mathcal{O}_n.$$

We prove this by induction. By definition $\mathcal{O}_\infty^{2\gamma} \subset \mathcal{O}_0$. Suppose that $\mathcal{O}_\infty^{2\gamma} \subset \mathcal{O}_n$ and we claim that $\mathcal{O}_\infty^{2\gamma}$ is included in $\mathcal{O}_{n+1}$.
Fix $\xi \in \mathcal{O}_\infty^{2\gamma}$ and $|k| \leq K_n$. Then by (5.40), (5.34) and recalling μ in (5.33) we have

$$|\alpha_n(\xi)\cdot k| \geq |\alpha_\infty(\xi)\cdot k| - |\alpha_\infty - \alpha_n||k| \geq \frac{2\gamma}{\langle k\rangle^\tau} - \delta_0(s_1) K_0^\mu K_{n-1}^{-\mu} K_n^2 K_n^\tau \geq \frac{\gamma}{\langle k\rangle^\tau}.$$

Finally, note that if f_0 is a reversible vector field, all the diffeomorphisms $\theta \mapsto \theta + h_n(\theta)$ are reversibility preserving, namely $h_n = \mathrm{odd}(\theta)$ for any $n \in \mathbb{N}$. Hence the limit function $h^{(\infty)} = \lim_{n\to+\infty} h_n$ is $\mathrm{odd}(\theta)$ implying that the map $\theta \mapsto \theta + h^{(\infty)}(\theta)$ is reversibility preserving. The proof of the theorem is then concluded. □

Proof of Corollary 3.3

Lemma (5.2)-$(\mathcal{P}4)_n$ implies that if $0 < \gamma - \rho < \gamma < 1$ and $\lambda_1, \lambda_2 \in B_E$ satisfies

$$K_{n-1}^{\tau+1} \|\Delta_{12} f_0\|_{s_0+\mathtt{b}} \leq \rho \tag{5.85}$$

then $\mathcal{O}_\infty^{2\gamma}(\lambda_1) \subseteq \cap_{m=0}^n \mathcal{O}_m^{\gamma-\rho}(\lambda_2)$. By choosing $K_n = K$, the smallness condition (3.10) implies (5.85). This is a standard argument in KAM iterations, see for instance [5]. By the estimate (5.40) and using a telescoping argument, one has that

$$\begin{aligned} |\alpha_\infty(\lambda_1) - \alpha_n(\lambda_1)|\,, |\alpha_\infty(\lambda_2) - \alpha_n(\lambda_2)| &\lesssim \gamma^{-1}\max\{\|f_0(\lambda_1)\|_{s_1}, \|f_0(\lambda_2)\|_{s_1}\} K_{n+1}^{-\mu} \\ &\overset{K_{n+1}>K_n=K}{\lesssim} \gamma^{-1}\max\{\|f_0(\lambda_1)\|_{s_1}, \|f_0(\lambda_2)\|_{s_1}\} K^{-\mu}\,. \end{aligned} \tag{5.86}$$

Finally by triangular inequality and using the estimate (5.51) together with the previous bound (5.86) one gets the claimed inequality (3.11).

A Technical Lemmata

In this Section we present standard tame and Lipschitz estimates for composition of functions and changes of variables.

Let us denote $L^\infty := L^\infty(\mathbb{T}^d, \mathbb{C})$ and $W^{s,\infty} := W^{s,\infty}(\mathbb{T}^d, \mathbb{C})$ with $d \geq 1$. The norms of these spaces are respectively indicated with $|\cdot|_{L^\infty} := |\cdot|_{0,\infty}$, $|\cdot|_{s,\infty}$ and are defined by

$$|u|_{L^\infty} := \sup_{x\in\mathbb{T}^d} |u(x)|, \quad |u|_{s,\infty} := \sum_{s_1 \leq s} |D^{s_1} u|_{L^\infty}, \quad |D^{s_1} u|_{L^\infty} := \sup_{|\vec{s}_1| = s_1} |\partial_x^{\vec{s}_1} u|_{L^\infty}, \tag{A.1}$$

here D^s is the s-th Fréchet derivative with respect to x, hence D^s is a symmetric multi-linear operator.

Let us denote with $H^s := H^s(\mathbb{T}^d, \mathbb{C})$ the space of Sobolev functions on $\mathbb{T}^d$ defined by

$$H^s(\mathbb{T}^d, \mathbb{C}) := \left\{ u \in L^2(\mathbb{T}^d) : \|u\|_s^2 := \sum_{j\in\mathbb{Z}^d} |u_j|^2 \langle j \rangle^{2s} < \infty \right\}. \tag{A.2}$$

We shall actually use the equivalent norm

$$\|u\|_s := \|u\|_{H^s(\mathbb{T}^d)} := \|u\|_{L^2(\mathbb{T}^d)} + \|D^s u\|_{L^2(\mathbb{T}^d)}, \quad \|D^s u\|_{L^2(\mathbb{T}^d)} := \sup_{|\vec{s}|=s} \|\partial_x^{\vec{s}} u\|_{L^2(\mathbb{T}^d)}. \tag{A.3}$$

Lemma A.1. *Let $s_0 > d/2$. Then the following holds.*

(i) **Embedding**. $|u|_{L^\infty} \leq \|u\|_{s_0}$ *for all* $u \in H^{s_0}$.

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

(iii) **Interpolation**. *For* $0 \leq s_1 \leq s \leq s_2$, $s = \lambda s_1 + (1-\lambda)s_2$, $\lambda \in [0,1]$,

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

Let $a_0, b_0 \geq 0$ *and* $p, q > 0$. *For all* $u \in H^{a_0+p+q}$, $v \in H^{b_0+p+q}$

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \leq \|u\|_{a_0+p+q}\|v\|_{b_0} + \|u\|_{a_0}\|v\|_{b_0+p+q}.$$

Similarly

$$|u|_{s,\infty} \leq C(s_1, s_2)|u|_{s_1,\infty}^\lambda |v|_{s_2,\infty}^{1-\lambda} \quad \forall u \in W^{s_2,\infty}$$

and for all $u \in W^{a_0+p+q}$, $v \in W^{b_0+p+q}$

$$|u|_{a_0+p,\infty}|v|_{b_0+q,\infty} \leq C(a_0, b_0, p, q)|u|_{a_0+p+q,\infty}|v|_{b_0,\infty} + |u|_{a_0,\infty}|v|_{b_0+p+q,\infty}.$$

(iv) *For any* $s, \alpha \geq 0$,

$$\|\Pi_N u\|_{s+\alpha} \leq N^\alpha \|u\|_s, \quad \|\Pi_N^\perp u\|_s \leq N^{-\alpha}\|u\|_{s+\alpha} \tag{A.4}$$

where

$$\Pi_N u(\varphi, x) := \sum_{|(\ell,j)|\leq N} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + jx)}, \quad \Pi_N^\perp := \mathrm{Id} - \Pi_N.$$

Remark A.2. If $u = u(\omega)$ and $v = v(\omega)$ depend in a Lipschitz way on a parameter $\mathcal{O} \subset \mathbb{R}^\nu$, all the previous statements hold by replacing $|\cdot|_{s,\infty}$, $\|\cdot\|_s$ with the Lipschitz norms $|\cdot|^{\gamma,\mathcal{O}}_{s,\infty}$, $\|\cdot\|^{\gamma,\mathcal{O}}_s$, provided that we take $s_0 > d/2+1$ (i.e. all the relations hold with s_0+1, for $s_0 > d/2$ and then we rename s_0). Indeed we first apply the formulas above to the variation $(u(\omega)-u(\omega'))/(\omega-\omega')$, this implies the desired bounds for the norm $\max\{\|u\|_s^{sup}, \gamma\|u\|^{lip}_{s-1}\}$. Since this norm is equivalent to $\|\cdot\|^{\gamma,\mathcal{O}}_s$, our claim follows.

Lemma A.3. (Change of variable) *Consider $p \in W^{s,\infty}(\mathbb{T}^d;\mathbb{R}^d)$, $s \geq 1$, with $|p|_{1,\infty} \leq 1/2$. Let $f(x) = x + p(x)$. Then:*
(i) $f : \mathbb{T}^d \to \mathbb{T}^d$ is a diffeomorphism, its inverse is $f^{-1}(y) = g(y) = y + q(y)$ with $q \in W^{s,\infty}(\mathbb{T}^d;\mathbb{R}^d)$ and $|q|_{s,\infty} \leq C|p|_{s,\infty}$. More precisely,

$$|q|_{L^\infty} = |p|_{L^\infty},\ |Dq|_{L^\infty} \leq 2|Dp|_{L^\infty},\ |Dq|_{s-1,\infty} \leq C|Dp|_{s-1,\infty}, \tag{A.5}$$

where the constant C depends on d, s

(ii) If $u \in H^s(\mathbb{T}^d;\mathbb{C})$, then $u \circ f(x) = u(x+p(x)) \in H^s(\mathbb{T}^d;\mathbb{C})$, and, with the same C as in (i) one has

$$\|u \circ f\|_s \leq \|u\|_s + C(\|u\|_s |p|_{1,\infty} + |Dp|_{s-1,\infty}\|u\|_1), \tag{A.6a}$$

$$\|u \circ f - u\|_s \leq C(|p|_{L^\infty}\|u\|_{s+1} + |p|_{s,\infty}\|u\|_2), \tag{A.6b}$$

(iii) Assume that $p = p_\omega$ depends in a Lipschitz way by a parameter $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$, and suppose, as above, that $|p_\omega|_{1,\infty} \leq 1/2$ for all ω. Then $q = q_\omega$ is also Lipschitz in ω, and

$$|q|^{\gamma,\mathcal{O}}_{s,\infty} \leq C\left(|p|^{\gamma,\mathcal{O}}_{s,\infty} + \{\sup_{\omega\in\mathcal{O}} |p_\omega|_{s,\infty}\}|p|^{\gamma,\mathcal{O}}_{1,\infty}\right) \leq C|p|^{\gamma,\mathcal{O}}_{s,\infty}, \tag{A.7}$$

$$\|u \circ f\|^{\gamma,\mathcal{O}}_s \leq \|u\|^{\gamma,\mathcal{O}}_s + C(\|u\|^{\gamma,\mathcal{O}}_s |p|^{\gamma,\mathcal{O}}_{1,\infty} + |p|^{\gamma,\mathcal{O}}_{s,\infty}\|u\|^{\gamma,\mathcal{O}}_2), \quad s \in \mathbb{N} \tag{A.8}$$

the constant C depends on d, s (it is independent on γ).

Proof. The estimate (A.5) is proved in [3], (A.6b) is proved in the Appendix of [5]. The bounds (A.6a) are slightly different from the corresponding ones of [3], [5]. This and the different choiche of wheighted Lipschitz norm reflect on the Lipschitz bounds (A.7) and (A.8). Let us prove (A.6a). We follow the proof of Lemma $B.4$-(ii) in [3] by treating in a different way some terms arising from the Faa di Bruno's formula. First we note that $\|u \circ f\|_0 \leq \|u\|_0(1 + 2|Dp|_\infty)$. Then we consider the expression

$$D^s(u \circ f) = \sum_{k=1}^{s} \sum_{j_1+\dots+j_k=s,\, j_i\geq 1} C_k\ (D^k u)[D^{j_1} f, \dots, D^{j_k} f]\,.$$

Here the coefficients C_k are integer numbers which take into account the combinatorics, it is easily seen that $C_s = 1$. We and note that $Df = \mathrm{Id} + Dp$, while $D^j f = D^j p$ for $j \geq 1$. Then we split the sum above in the following way

$$
\begin{aligned}
D^s(u\circ f) = & \ (D^s u)\circ f[Df,\dots,Df] + \sum_{k=1}^{s-1} \sum_{\substack{j_1+\dots+j_k=s,\\ \prod_i j_i>1}} C_k\ (D^k u)[D^{j_1}f,\dots,D^{j_k}f] \\
& = (D^s u)\circ f + \sum_{r=1}^{s}\binom{s}{r}\ (D^s u)\circ f[\underbrace{Dp,\dots,Dp}_{\times r},\underbrace{\mathrm{Id},\dots,\mathrm{Id}}_{\times s-r}] \\
& + \sum_{k=1}^{s-1} \sum_{\substack{j_1+\dots+j_k=s,\\ \prod_i j_i>1}} C_k\ (D^k u)\circ f[D^{j_1}f,\dots,D^{j_k}f] \qquad \text{(A.9)}\\
& = (D^s u)\circ f \\
& + \sum_{k=1}^{s} \sum_{\substack{r_1,r_2\\ 0<r_1+r_2\le k}} \sum_{\substack{j_1+\dots+j_{r_1}=s+r_1-k\\ j_i>1}} C_{k,r_1,r_2}\ (D^k u)\circ f[D^{j_1}p,\dots,D^{j_{r_1}}p,\underbrace{Dp,\dots,Dp}_{\times r_2},\underbrace{\mathrm{Id},\dots,\mathrm{Id}}_{\times k-r_1-r_2}]
\end{aligned}
$$

The first summand is estimated by noting that

$$\|(D^s u)\circ f\|_0 \le (1+2|Dp|_\infty)\|u\|_s.$$

Now we rename $j_i = h_i + 1$ for $i = 1, r_1$ in the second summand and set $G = Dp$, we get

$$\sum_{h_1+\dots+h_{r_1}=s-k} C_{k,r_1,r_2}\ (D^k u)\circ f[D^{h_1}G,\dots,D^{h_{r_1}}G,\underbrace{G,\dots,G}_{\times r_2},\underbrace{\mathrm{Id},\dots,\mathrm{Id}}_{\times k-r_1-r_2}].$$

The L^2 norms of the summands above are bounded by

$$2C_{k,r_1,r_2}\|u\|_k |G|_{h_1,\infty}\dots|G|_{h_{r_1},\infty}|G|^{r_2}_\infty.$$

Then one can follow exactly the same proof of Lemma $B.4$-(ii) in [3].

In order to prove (A.7) we use formula (6.15) of [5] which reads in terms of the Lipschitz seminorm

$$\gamma|q|^{lip,\mathcal{O}}_{s-1,\infty} \le C\left(|p|^{sup,\mathcal{O}}_{s-1,\infty} + \gamma|p|^{lip,\mathcal{O}}_{s-1,\infty} + |p|^{sup,\mathcal{O}}_{s,\infty}(|p|^{sup,\mathcal{O}}_{0,\infty} + \gamma|p|^{lip,\mathcal{O}}_{0,\infty})\right).$$

In order to prove (A.8) we compute the Lipschitz variation. We have

$$\|u(\omega,x+p_\omega(x)) - u(\omega',x+p_{\omega'}(x))\|_s = \|\Delta_\omega u\circ f_\omega + (u\circ\widehat{f}-u)\circ f_{\omega'}\Delta_\omega p\|_s\,,\quad \widehat{f} = f_\omega\circ g_{\omega'}\,,$$

now in the r.h.s. the first summand is bounded by using (A.6a); regarding the second summand we use the interpolation estimates for products and then (A.6a), (A.6b). Note that (A.6b) *loses one derivative*, this is why in the norm $\|\cdot\|^{\gamma,\mathcal{O}}$ we require the estimate of the Lipschitz variation of ω only for the norm $\|\cdot\|_{s-1}$. □

References

[1] V. I. Arnol′d. On the classical perturbation theory and the stability problem of planetary systems. *Dokl. Akad. Nauk SSSR*, 145:487–490, 1962.

[2] V. I. Arnol′d. Small denominators and problems of stability of motion in classical and celestial mechanics. *Uspehi Mat. Nauk*, 18(6 (114)):91–192, 1963.

[3] P. Baldi. Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type. *Ann. I. H. Poincaré (C) Anal. Non Linéaire*, 30:33–77, 2013.

[4] P. Baldi, M. Berti, E. Haus, and R. Montalto. Time quasiperiodic gravity water waves in finite depth. preprint 2017, arXiv:1708.01517.

[5] P. Baldi, M. Berti, and R. Montalto. KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation. *Math. Ann.*, 359(1-2):471–536, 2014.

[6] P. Baldi, M. Berti, and R. Montalto. KAM for autonomous quasi-linear perturbations of KdV. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 33(6):1589–1638, 2016.

[7] P. Baldi and E. Haus. A Nash-Moser-Hörmander implicit function theorem with applications to control and Cauchy problems for PDEs. *J. Funct. Anal.*, 273(12):3875–3900, 2017.

[8] D. Bambusi. Reducibility of 1-d Schrödinger equation with time quasiperiodic unbounded perturbations, II. *Comm. Math. Phys.*, 353(1):353–378, 2017.

[9] D. Bambusi and S. Graffi. Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods. *Comm. Math. Phys.*, 219(2):465–480, 2001.

[10] D. Bambusi, B. Grébert, A. Maspero, and D. Robert. Reducibility of the quantum harmonic oscillator in d-dimensions with polynomial time-dependent perturbation. *Anal. PDE*, 11(3):775–799, 2018.

[11] M. Berti, P. Biasco, and M. Procesi. KAM theory for reversible derivative wave equations. *Archive for Rational Mechanics and Analysis*, 212(3):905–955, 2014.

[12] M. Berti, P. Bolle, and M. Procesi. An abstract Nash-Moser theorem with parameters and applications to PDEs. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 27(1):377 – 399, 2010.

[13] M. Berti and R. Montalto. Quasi-periodic standing wave solutions of gravity-capillary water waves. Preprint 2016 arXiv:1602.02411, To appear on Memoirs Am. Math. Soc. MEMO-891, 2016.

[14] L. Chierchia and J. You. Kam tori for 1d nonlinear wave equations with periodic boundary conditions. *Comm. Math. Phys.*, 211:497–525, 2000.

[15] H. L. Eliasson and S. B. Kuksin. On reducibility of Schrödinger equations with quasiperiodic in time potentials. *Comm. Math. Phys.*, 286(1):125–135, 2009.

[16] L. H. Eliasson, B. Grébert, and S.B. Kuksin. KAM for the nonlinear beam equation. *Geom. Funct. Anal.*, 26(6):1588–1715, 2016.

[17] L. H. Eliasson and S. B. Kuksin. KAM for the nonlinear Schrödinger equation. *Ann. of Math. (2)*, 172(1):371–435, 2010.

[18] B. Fayad and R. Krikorian. Herman's last geometric theorem. *Annales scientifiques de l'École Normale Supérieure*, 42(2):193–219, 2009.

[19] R. Feola, F. Giuliani, and M. Procesi. Quasi-periodic solutions for Hamiltonian perturbation of Degasperis-Procesi equation. in preparation.

[20] R. Feola and M. Procesi. Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations. *J. Differential Equations*, 259(7):3389–3447, 2015.

[21] J. Geng and J. You. A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces. *Comm. Math. Phys.*, 262(2):343–372, 2006.

[22] J. Geng, J. You, and X. Xu. KAM tori for cubic NLS with constant potentials. Preprint.

[23] F. Giuliani. Quasi-periodic solutions for quasi-linear generalized KdV equations. *Journal of Differential Equations*, 262(10):5052 – 5132, 2017.

[24] B. Grébert and E. Paturel. KAM for the Klein-Gordon equation on $\mathbb{S}^d$. *Boll. Unione Mat Ital.*, 9(2):237–288, 2016.

[25] M. R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[26] M. R. Herman. Différentiabilité optimale et contre-exemples à la fermeture en topologie C^∞ des orbites récurrentes de flots hamiltoniens. *C. R. Acad. Sci. Paris Sér. I Math.*, 313(1):49–51, 1991.

[27] S. B. Kuksin. *Nearly integrable infinite-dimensional Hamiltonian systems*, volume 1556 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993.

[28] S. B. Kuksin. A KAM theorem for equations of the Korteweg-de Vries type. *Rev. Math. Phys.*, 10(3):1–64, 1998.

[29] S. B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math.*, 143(1):149–179, 1996.

[30] J. Liu and X. Yuan. A KAM Theorem for Hamiltonian partial differential equations with unbounded perturbations. *Comm. Math. Phys,*, 307:629–673, 2011.

[31] A. Maspero and D. Robert. On time dependent Schrödinger equations: Global well-posedness and growth of Sobolev norms. *Journal of Functional Analysis*, 273(2):721–781, 2017.

[32] R. Montalto. Growth of Sobolev norms for a class of Schrödinger equations with sublinear dispersion. in preparation.

[33] R. Montalto. A Reducibility result for a class of Linear Wave Equations on $\mathbb{T}^d$. *International Mathematics Research Notices*, page rnx167, 2017.

[34] J. Moser. A rapidly convergent iteration method and nonlinear differential equations. *Uspehi Mat. Nauk*, 23(4 (142)):179–238, 1968.

[35] J. Pöschel. *Über invariante Tori in differenzierbaren Hamiltonschen Systemen.* Bonner Mathematische Schriften [Bonn Mathematical Publications], 120. Universität Bonn, Mathematisches Institut, Bonn, 1980. Dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1979, Beiträge zur Differentialgeometrie [Contributions to differential geometry], 3.

[36] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Commun. Pure Appl. Math.*, 35:653–696, 1982.

[37] C. Procesi and M. Procesi. A normal form for the Schrödinger equation with analytic nonlinearities. *Communications in Mathematical Physics*, 312(2):501–557, 2012.

[38] H. Rüssmann. Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1970:67–105, 1970.

[39] H. Rüssmann. *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, pages 598–624. Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[40] D. A. Salamon. The Kolmogorov-Arnold-Moser theorem. *Math. Phys. Electron. J.*, 10:Paper 3, 37, 2004.

[41] C. Eugene Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Comm. Math. Phys.*, 127(3):479–528, 1990.

[42] J. Ch. Yoccoz. Travaux de Herman sur les tores invariants. *Séminaire Bourbaki*, 34:311–344, 1991-1992.