KAM theory for the reversible derivative wave equation

Massimiliano Berti, Luca Biasco, Michela Procesi

Abstract: We prove the existence of Cantor families of small amplitude, analytic, quasi-periodic solutions of derivative wave equations, with zero Lyapunov exponents and whose linearized equation is reducible to constant coefficients. This result is derived by an abstract KAM theorem for infinite dimensional reversible dynamical systems[1].

2000AMS SUBJECT CLASSIFICATION: 37K55, 35L05.

Contents

[1] *Massimiliano Berti*, Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi di Napoli Federico II, Via Cintia, Monte S. Angelo, I-80126, Napoli, Italy, `m.berti@unina.it`.

Luca Biasco, Dipartimento di Matematica, Università di Roma 3, Largo San Leonardo Murialdo, I-00146, Roma, Italy, `biasco@mat.uniroma3.it`.

Michela Procesi, Dipartimento di Matematica "G. Castelnuovo", Università degli Studi di Roma la Sapienza, P.le A. Moro 5, 00185 ROMA, Italy, `michela.procesi@mat.uniroma1.it`.

This research was supported by the European Research Council under FP7 and partially by the PRIN2009 grant "Critical point theory and perturbative methods for nonlinear differential equations".

1 Introduction

In the last years many progresses have been obtained concerning KAM theory for nonlinear PDEs, since the pioneering works of Kuksin [23] and Wayne [37] for 1-d semilinear wave (NLW) and Schrödinger (NLS) equations. A challenging frontier concerns PDEs with nonlinearities containing derivatives. In this direction the first existence result of quasi-periodic solutions has been proved by Kuksin [24]-[25], see also Kappeler-Pöschel [21], for perturbed KdV equations.

In this paper we develop KAM theory for derivative wave equations (DNLW) proving existence and stability of small amplitude analytic quasi-periodic solutions. The DNLW is not an Hamiltonian PDE, but may have a reversible structure, that we shall exploit.

Most of the existence results of quasi-periodic solutions proved so far concern *Hamiltonian* PDEs, see e.g. [25], [26], [28], [29], [14], [8], [9], [11], [12], [17], [16], [37], [5], [6], [36], [31], [30]. It was however remarked by Bourgain that the construction of periodic and quasi-periodic solutions using the Lyapunov-Schmidt decomposition and the Newton iteration method of Craig-Wayne [14] and [8]-[11] is a-priori not restricted to Hamiltonian systems. This approach appears as a general implicit function type result, in large part independent of the Hamiltonian character of the equations. For example in [10] Bourgain proved the existence (not the stability) of periodic solutions for the non-Hamiltonian derivative wave equation

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} + \mathrm{m}\mathtt{y} + \mathtt{y}_t^2 = 0\,, \quad \mathrm{m} > 0\,, \quad \mathtt{x} \in \mathbb{T}\,. \tag{1.1}$$

Actually also KAM theory is not only Hamiltonian in nature, but may be formulated for general vector fields, as realized in the seminal work of Moser [27]. This paper, in particular, started the analysis of reversible KAM theory for finite dimensional systems, later extended by Arnold [1] and Sevryuk [35]. The reversibility property implies that the averages over the fast angles of some components of the vector field are zero, thus removing the "secular drifts" of the actions which are incompatible with quasi-periodic solutions.

Recently, Zhang-Gao-Yuan [38] have proved existence and stability of C^∞-quasi periodic solutions for the derivative NLS equation $\mathrm{i}u_t + u_{\mathtt{xx}} + |u_{\mathtt{x}}|^2 u = 0$ with Dirichlet boundary conditions. Such equation is reversible, but not Hamiltonian. The result [38] is proved adapting the KAM scheme developed for the Hamiltonian DNLS in Liu-Yuan [20]. In turn [20] extends the approach of Kuksin [24]-[25], Kappeler-Pöschel [21], which is valid for more dispersive PDEs, like KdV. The derivative nonlinear wave equation (DNLW), which is not dispersive (the eigenvalues of the principal part of the differential operator grow linearly at infinity) is excluded by both these approaches.

In the recent paper [4] we have extended KAM theory to deal with Hamiltonian derivative wave equations like

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} + \mathrm{m}\mathtt{y} + f(D\mathtt{y}) = 0\,, \quad \mathrm{m} > 0\,, \quad D := \sqrt{-\partial_{\mathtt{xx}} + \mathrm{m}}\,, \quad \mathtt{x} \in \mathbb{T}\,.$$

This kind of Hamiltonian pseudo-differential equations has been introduced by Bourgain [8] and Craig [13] as models to study the effect of derivatives versus dispersive phenomena. The key of [4] is the proof of the first order asymptotic expansion of the perturbed normal frequencies, obtained using the notion of quasi-Töplitz function. This concept was introduced by Procesi-Xu [31] and it is connected to the Töplitz-Lipschitz property in Eliasson-Kuksin [16], see also [19]. Of course we could not deal in [4] with the derivative wave equation, which is not Hamiltonian.

The goal of this paper is to develop KAM theory for a class of reversible derivative wave equations

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} + \mathtt{m}\mathtt{y} = g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{y}_t) \,, \quad \mathtt{x} \in \mathbb{T} \,, \tag{1.2}$$

implying the *existence* and the *stability* of *analytic* quasi-periodic solutions, see Theorem 1.1. Note that the nonlinearity in (1.2) has an explicit $\mathtt{x}$-dependence (unlike [4]).

We can not expect the existence result for any nonlinearity. For example, (1.2) with the nonlinear friction term $g = \mathtt{y}_t^3$ has no non trivial smooth periodic/quasi-periodic solutions, see Proposition 1.2. This case is ruled out by assuming the reversibility condition

$$g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, -\mathtt{v}) = g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{v}) \tag{1.3}$$

satisfied for example by (1.1). Under condition (1.3) the equation (1.2) is *reversible*, namely the associated first order system

$$\mathtt{y}_t = \mathtt{v} \,, \quad \mathtt{v}_t = \mathtt{y}_{\mathtt{xx}} - \mathtt{m}\mathtt{y} + g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{v}) \,, \tag{1.4}$$

is reversible with respect to the involution

$$S(\mathtt{y}, \mathtt{v}) := (\mathtt{y}, -\mathtt{v}) \ , \quad S^2 = I \,. \tag{1.5}$$

Reversibility is an important property in order to allow the existence of periodic/quasi-periodic solutions, albeit not sufficient. For example, the reversible equation $\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_{\mathtt{x}}^3$, $\mathtt{x} \in \mathbb{T}$ (proposed in [13], page 89), has no smooth periodic/quasi-periodic solutions except the constants, see Proposition 1.1. In order to find quasi-periodic solutions we also require the parity assumption

$$g(-\mathtt{x}, \mathtt{y}, -\mathtt{y}_{\mathtt{x}}, \mathtt{v}) = g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{v}) \,, \tag{1.6}$$

which rules out nonlinearities like $\mathtt{y}_{\mathtt{x}}^3$. Actually, for the wave equation (1.2) the role of the time and space variables $(t, \mathtt{x})$ is highly symmetric. Then, considering $\mathtt{x}$ "as time" (spatial dynamics idea) the term $\mathtt{y}_{\mathtt{x}}^3$ is a friction and condition (1.6) is the corresponding reversibility condition.

After Theorem 1.1 we shall further comment on the assumptions.

Before stating our main results, we mention the classical bifurcation theorems of Rabinowitz [32] about periodic solutions (with rational periods) of dissipative forced derivative wave equations

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} + \alpha \mathtt{y}_t + \varepsilon F(\mathtt{x}, t, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{y}_t) = 0 \,, \quad \mathtt{x} \in [0, \pi]$$

with Dirichlet boundary conditions, and in [33] for fully-non-linear forced wave equations

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} + \alpha \mathtt{y}_t + \varepsilon F(\mathtt{x}, t, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{y}_t, \mathtt{y}_{tt}, \mathtt{y}_{t\mathtt{x}}, \mathtt{y}_{\mathtt{xx}}) = 0 \ , \quad \mathtt{x} \in [0, \pi] \,.$$

This latter result is quite subtle because, from the point of view of the initial value problem, it is uncertain whether a solution can exist for more than a finite time due to the formation of shocks. Here the presence of the dissipation $\alpha \neq 0$ allows the existence of a periodic solution. We mention also [34] for a third order singular perturbation problem of a second order ordinary differential equation.

Finally, concerning quasi-linear wave equations we mention the recent Birkhoff normal form results of Delort [15], which imply long time existence for solutions with small initial data. To our knowledge, these are the only results of this type on compact manifolds. For quasi-linear wave equations in $\mathbb{R}^d$ there is a wide literature since the nonlinear effects of derivatives may be controlled by dispersion.

1.1 Existence and stability of quasi-periodic solutions of DNLW

We consider derivative wave equations (1.2) where $\mathrm{m} > 0$, the nonlinearity

$$g : \mathbb{T} \times \mathcal{U} \to \mathbb{R}\,, \quad \mathcal{U} \subset \mathbb{R}^3 \text{ open neighborhood of } 0\,,$$

is real analytic and satisfies the "reversibility" and "parity" assumptions (1.3), (1.6). Moreover g vanishes at least quadratically at $(\mathtt{y}, \mathtt{y}_\mathtt{x}, \mathtt{v}) = (0, 0, 0)$, namely

$$g(\mathtt{x}, 0, 0, 0) = (\partial_\mathtt{y} g)(\mathtt{x}, 0, 0, 0) = (\partial_{\mathtt{y}_\mathtt{x}} g)(\mathtt{x}, 0, 0, 0) = (\partial_\mathtt{v} g)(\mathtt{x}, 0, 0, 0) = 0\,.$$

In addition we need a "non-degeneracy" assumption on the leading order term of the nonlinearity (in order to verify the usual "twist" condition required in KAM theory). For definiteness, we have developed all the calculations for

$$g = \mathtt{y}\mathtt{y}_\mathtt{x}^2 + O_5(\mathtt{x}, \mathtt{y}, \mathtt{y}_\mathtt{x}, \mathtt{y}_t) \tag{1.7}$$

where O_5 contains terms of order at least five in $(\mathtt{y}, \mathtt{y}_\mathtt{x}, \mathtt{y}_t)$ (terms of order four could also be considered, see Remark 7.1).

Because of (1.3), it is natural to look for "reversible" quasi-periodic solutions, namely such that $\mathtt{y}(t, \mathtt{x})$ is even and $\mathtt{v}(t, \mathtt{x})$ is odd in time. Moreover, because of (1.6) it is natural to restrict to solutions which are even is $\mathtt{x}$ (standing waves), namely with

$$(\mathtt{y}, \mathtt{v})(-\mathtt{x}) = (\mathtt{y}, \mathtt{v})(\mathtt{x})\,, \ \ \forall \mathtt{x} \in \mathbb{T}\,. \tag{1.8}$$

Note, in particular, that $\mathtt{y}$ satisfies the *Neumann* boundary conditions $\mathtt{y}_\mathtt{x}(t, 0) = \mathtt{y}_\mathtt{x}(t, \pi) = 0$.

Summarizing we look for reversible quasi-periodic *standing wave* solutions of (1.2), satisfying

$$\mathtt{y}(t, \mathtt{x}) = \mathtt{y}(t, -\mathtt{x})\,, \ \forall t\,, \quad \mathtt{y}(-t, \mathtt{x}) = \mathtt{y}(t, \mathtt{x})\,, \ \forall \mathtt{x} \in \mathbb{T}\,. \tag{1.9}$$

For every finite choice of the *tangential* sites $\mathcal{I}^+ \subset \mathbb{N} \setminus \{0\}$, the linear Klein-Gordon equation

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{x}\mathtt{x}} + \mathrm{m}\mathtt{y} = 0\,, \quad \mathtt{x} \in \mathbb{T}\,, \tag{1.10}$$

possesses the family of quasi-periodic standing wave solutions

$$\mathtt{y} = \sum_{j \in \mathcal{I}^+} \frac{\sqrt{8\xi_j}}{\lambda_j} \cos(\lambda_j\, t) \cos(j\mathtt{x})\,, \quad \lambda_j := \sqrt{j^2 + \mathrm{m}}\,, \tag{1.11}$$

parametrized by the amplitudes $\xi_j \in \mathbb{R}_+$.

Theorem 1.1. *For every finite choice of the tangential sites* $\mathcal{I}^+ \subset \mathbb{N} \setminus \{0\}$ *and for all* $\mathrm{m} > 0$*, except finitely many (depending on* $\mathcal{I}^+$*), the DNLW equation* (1.2) *with a real analytic nonlinearity satisfying* (1.3)*,* (1.6)*,* (1.7) *admits small-amplitude, analytic (both in* t *and* $\mathtt{x}$*), quasi-periodic solutions*

$$\mathtt{y} = \sum_{j \in \mathcal{I}^+} \frac{\sqrt{8\xi_j}}{\lambda_j} \cos(\omega_j^\infty(\xi)\, t) \cos(j\mathtt{x}) + o(\sqrt{\xi}), \quad \omega_j^\infty(\xi) \overset{\xi \to 0}{\approx} \sqrt{j^2 + \mathrm{m}} \tag{1.12}$$

satisfying (1.9)*, for a Cantor-like set of parameters with asymptotical density* 1 *at* $\xi = 0$*. These quasi-periodic solutions have zero Lyapunov exponents and the linearized equations can be reduced to constant coefficients (in a phase space of functions even in* $\mathtt{x}$*). The term* $o(\sqrt{\xi})$ *in* (1.12) *is small in some analytic norm.*

This theorem answers the question, posed by Craig in [13], of developing a general theory for quasi-periodic solutions for reversible derivative wave equations. With respect to Bourgain [10], we prove existence of quasi-periodic solutions (not only periodic) as well as their stability.

Let us comment on the hypothesis of Theorem 1.1.

1. **Reversibility and Parity.** As already said, the "reversibility" and "parity" assumptions (1.3), (1.6), rule out nonlinearities like $\mathtt{y}_t^3$ and $\mathtt{y}_{\mathtt{x}}^3$ for which periodic/quasi-periodic solutions of (1.2) do not exist. We generalize these non-existence results in Propositions 1.1, 1.2.

2. **Mass** $\mathrm{m} > 0$**.** The assumption on the mass $\mathrm{m} \neq 0$ is, in general, necessary. When $\mathrm{m} = 0$, a well known example of Fritz John (that we reproduce in Proposition 1.3) proves that (1.1) has no smooth solutions for all times except the constants. In Proposition 1.4 we prove non-existence of periodic/quasi-periodic solutions for DNLW equations satisfying both (1.3), (1.6), but with mass $\mathrm{m} = 0$. For the KAM construction, the mass $\mathrm{m} > 0$ is used in the Birkhoff normal form step (see section 7 and, in particular, Lemma 7.3). If the mass $\mathrm{m} < 0$ then the Sturm-Liouville operator $-\partial_{\mathtt{xx}} + \mathrm{m}$ may possess finitely many negative eigenvalues and one should expect the existence of partially hyperbolic tori.

3. $\mathtt{x}$**-dependence.** The nonlinearity g in (1.2) may explicitly depend on the space variable $\mathtt{x}$, i.e. this equations is NOT invariant under $\mathtt{x}$ translations. This is an important novelty with respect to the KAM theorem in [4] which used the conservation of momentum, as [17], [18]. The key idea is the introduction of the $\mathtt{a}$-weighted majorant norm for vector fields (Definition 2.4) which penalizes the high-momentum monomials (see (2.38)), see comments at the end of the section.

4. **Twist.** We have developed all the calculations for the cubic leading term $g = \mathtt{y}\mathtt{y}_{\mathtt{x}}^2 + \text{h.o.t.}$. In this case the third order Birkhoff normal form of the PDE (1.2) turns out to be (partially) integrable and the frequency-to-action map is invertible. This is the so called "twist-condition" in KAM theory. It could be interesting to classify the allowed nonlinearities. For example, among the cubic nonlinearities, we already know that for $\mathtt{y}_{\mathtt{x}}^3$, $\mathtt{y}^2\mathtt{y}_{\mathtt{x}}$ (and $\mathtt{v}^3$) there are no non-trivial periodic/quasi-periodic solutions, see Propositions 1.1-1.2. On the other hand, for $\mathtt{y}^3$ the Birkhoff normal form is (partially) integrable by [29] (for Dirichlet boundary conditions).

5. **Boundary conditions.** The solutions of Theorem 1.1 satisfy the Neumann boundary conditions $\mathtt{y}_{\mathtt{x}}(t, 0) = \mathtt{y}_{\mathtt{x}}(t, \pi) = 0$. For proving the existence of solutions under Dirichlet boundary conditions it would seem natural to substitute (1.6) with the oddness assumption

$$g(-\mathtt{x}, -\mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{v}) = -g(\mathtt{x}, \mathtt{y}, \mathtt{y}_{\mathtt{x}}, \mathtt{v}), \tag{1.13}$$

so that the subspace of functions $(\mathtt{y}, \mathtt{v})(\mathtt{x})$ odd in $\mathtt{x}$ is invariant under the flow evolution of (1.4). However, in order to find quasi-periodic solutions of (1.2), we need the *real-coefficients property* (1.31) which follows from (1.3) and (1.6), but not from (1.3) and (1.13). It is easy to check that (1.3), (1.13) and (1.31) imply the parity assumption (1.6). Of course, if a nonlinearity satisfies (1.3), (1.6) and also (1.13) we could look for also quasi-periodic solutions satisfying Dirichlet boundary conditions.

6. **Derivative vs quasi-linear NLW.** It has been proved by Klainermann-Majda [22] that all classical solutions of Hamiltonian quasi-linear wave equations like

$$\mathtt{y}_{tt} = (1 + \sigma(\mathtt{y}_{\mathtt{x}}))\mathtt{y}_{\mathtt{xx}} \tag{1.14}$$

with $\sigma^{(j)}(0) = 0$, $j = 1, \dots, p-1$, $\sigma^{(p)}(0) \neq 0$, do not admit smooth, small amplitude, periodic (a fortiori quasi-periodic) solutions except the constants. Actually, any non constant solution of (1.14), with sufficiently small initial data, develops a singularity in finite time in the second derivative $\mathtt{y}_{\mathtt{xx}}$. In this respect [22] may suggest that Theorem 1.1 is optimal regarding the order of (integer) derivatives in the nonlinearity. Interestingly, the solutions of the derivative wave equation (which is a semilinear PDE) found in Theorem 1.1 are analytic in both time t and space $\mathtt{x}$. Clearly the KAM approach developed in this paper fails for quasi-linear equations like (1.14) because the auxiliary vector field (whose flow defines the KAM transformations) is unbounded (of order 1). One could still ask for a KAM result for quasi-linear Klein Gordon equations (for which Delort [15] proved some steps of Birkhoff normal form). Note that adding a mass term my in the left hand side of (1.14), non zero periodic solutions of the form $\mathtt{y}(t,\mathtt{x}) = c(t)$ or $\mathtt{y}(t,\mathtt{x}) = c(\mathtt{x})$ may occur.

1.2 Ideas of proof: the abstract KAM theorem

The proof of Theorem 1.1 is based on the abstract KAM Theorem 4.1 for reversible infinite dimensional systems which proves the existence of elliptic invariant tori and provides a reducible normal form around them. We now explain the main ideas and techniques of proof.

Complex formulation. We extend (1.4) as a first order system with complex valued variables $(\mathtt{y},\mathtt{v}) \in \mathbb{C}^n \times \mathbb{C}^n$. In the unknowns

$$u^+ := \frac{1}{\sqrt{2}}(D\mathtt{y} - \mathrm{i}\mathtt{v}) , \quad u^- := \frac{1}{\sqrt{2}}(D\mathtt{y} + \mathrm{i}\mathtt{v}) , \quad D := \sqrt{-\partial_{\mathtt{xx}} + \mathrm{m}} , \quad \mathrm{i} := \sqrt{-1} ,$$

systems (1.4) becomes the first order system

$$\begin{cases} u_t^+ = \mathrm{i}Du^+ + \mathrm{i}\mathtt{g}(u^+,u^-) \\ u_t^- = \ \ -\mathrm{i}Du^- - \mathrm{i}\mathtt{g}(u^+,u^-) \end{cases} \tag{1.15}$$

where

$$\mathtt{g}(u^+,u^-) = -\frac{1}{\sqrt{2}} g\Big(\mathtt{x}, D^{-1}\Big(\frac{u^+ + u^-}{\sqrt{2}}\Big), D^{-1}\Big(\frac{u^+_{\mathtt{x}} + u^-_{\mathtt{x}}}{\sqrt{2}}\Big), \frac{u^- - u^+}{\mathrm{i}\sqrt{2}}\Big) . \tag{1.16}$$

In (1.15), the dynamical variables (u^+,u^-) are independent. However, since g is real analytic (real on real), the real subspace

$$\mathtt{R} := \{\overline{u^+} = u^-\} \tag{1.17}$$

is invariant under the flow evolution of (1.15), since

$$\overline{\mathtt{g}(u^+,u^-)} = \mathtt{g}(u^+,u^-) , \quad \forall (u^+,u^-) \in \mathtt{R} , \tag{1.18}$$

and the second equation in (1.15) reduces to the complex conjugated of the first one. Clearly, this corresponds to real valued solutions $(\mathtt{y},\mathtt{v})$ of the real system (1.4). We say that system (1.15) is "real-on-real" (see the more general Definition 2.9). For systems satisfying this property it is customary to use also the shorter notation

$$(u^+,u^-) = (u,\bar{u}) .$$

Moreover the subspace of even functions

$$\mathtt{E} := \Big\{u^+(\mathtt{x}) = u^+(-\mathtt{x}) , \ u^-(\mathtt{x}) = u^-(-\mathtt{x})\Big\} \tag{1.19}$$

(see (1.8)) is invariant under the flow evolution of (1.15), by (1.6). System (1.15) is reversible with respect to the involution

$$S(u^+,u^-) = (u^-,u^+) , \tag{1.20}$$

(which is nothing but (1.5) in the variables (u^+,u^-)), noting that (1.3), (1.16) imply

$$\mathtt{g}(u^+,u^-) = \mathtt{g}(u^-,u^+) . \tag{1.21}$$

Dynamical systems formulation. We introduce infinitely many coordinates by Fourier transform

$$u^+ = \sum_{j\in\mathbb{Z}} u_j^+ e^{\mathrm{i}j\mathtt{x}} , \quad u^- = \sum_{j\in\mathbb{Z}} u_j^- e^{-\mathrm{i}j\mathtt{x}} . \tag{1.22}$$

Then (1.15) becomes the infinite dimensional dynamical system

$$\begin{cases} \dot{u}_j^+ = \mathrm{i}\lambda_j u_j^+ + \mathrm{i}\mathtt{g}_j^+(\ldots, u_j^+, u_j^-, \ldots) \\ \dot{u}_j^- = -\mathrm{i}\lambda_j u_j^- - \mathrm{i}\mathtt{g}_j^-(\ldots, u_j^+, u_j^-, \ldots) \end{cases} \qquad \forall j \in \mathbb{Z} , \tag{1.23}$$

where

$$\lambda_j := \sqrt{j^2 + \mathrm{m}} \tag{1.24}$$

are the eigenvalues of D and

$$\mathtt{g}_j^+ = \frac{1}{2\pi} \int_{\mathbb{T}} \mathtt{g}\Big(\sum_{h\in\mathbb{Z}} u_h^+ e^{\mathrm{i}h\mathtt{x}}, \sum_{h\in\mathbb{Z}} u_h^- e^{-\mathrm{i}h\mathtt{x}}\Big) e^{-\mathrm{i}j\mathtt{x}} d\mathtt{x}\,, \qquad \mathtt{g}_j^- := \mathtt{g}_{-j}^+\,. \tag{1.25}$$

By (1.22), the "real" subset $\mathtt{R}$ in (1.17) reads $\overline{u_j^+} = u_j^-$ (this is the motivation for the choice of the signs in (1.22)) and, by (1.18), the second equation in (1.23) is the complex conjugated of the first one, namely

$$\overline{\mathtt{g}_j^+} = \mathtt{g}_j^- \quad \text{when} \quad \overline{u_j^+} = u_j^-\,,\ \forall j\,. \tag{1.26}$$

The subspace $\mathtt{E}$ of even functions in (1.19) reads, under Fourier transform,

$$E := \Big\{ u_j^+ = u_{-j}^+\,,\ u_j^- = u_{-j}^-\,,\ \forall j \in \mathbb{Z} \Big\} \tag{1.27}$$

and note that E is invariant under (1.23) because

$$(\mathtt{g}_{-j}^{\pm})_{|E} = (\mathtt{g}_j^{\pm})_{|E}\,. \tag{1.28}$$

By (1.22) the involution (1.20) reads

$$S : (u_j^+, u_j^-) \to (u_{-j}^-, u_{-j}^+)\,, \quad \forall j \in \mathbb{Z}\,, \tag{1.29}$$

and (1.23) is reversible with respect to S because

$$\mathtt{g}_j^+(\ldots, u_{-i}^-, u_{-i}^+, \ldots) = \mathtt{g}_{-j}^-(\ldots, u_i^+, u_i^-, \ldots)\,. \tag{1.30}$$

Finally, since g is real analytic, the assumptions (1.3) and (1.6) imply the key property

$$\mathtt{g}_j^{\pm}(\ldots, u_i^+, u_i^-, \ldots)\ \textit{has real Taylor coefficients} \tag{1.31}$$

in the variables (u_i^+, u_i^-).

Remark 1.1. *The previous property is compatible with oscillatory phenomena for* (1.23)*, excluding friction phenomena. This is another strong motivation for assuming* (1.3) *and* (1.6)*.*

Abstract KAM theorem. For every choice of symmetric *tangential sites*

$$\mathcal{I} = \mathcal{I}^+ \cup (-\mathcal{I}^+) \quad \text{with} \quad \mathcal{I}^+ \subset \mathbb{N} \setminus \{0\}\,,\ \sharp\mathcal{I} = n\,, \tag{1.32}$$

system (1.23) where $\mathtt{g} = 0$ (i.e. linear) has the invariant tori

$$\Big\{ u_j \bar{u}_j = \xi_{|j|} > 0,\ \text{for } j \in \mathcal{I}\,,\ u_j = \bar{u}_j = 0 \text{ for } j \notin \mathcal{I} \Big\}$$

parametrized by the actions $\xi = (\xi_j)_{j\in\mathcal{I}^+}$. They correspond to the quasi-periodic solutions in (1.11).

We first analyze the nonlinear dynamics of the PDE close to the origin, via a Birkhoff normal form reduction (see section 7). This step depends on the nonlinearity g and on the fact that the mass $\mathrm{m} > 0$. Here we use (1.7) to ensure that the third order Birkhoff normalized system is (partially) integrable and that the "twist condition" holds.

Then we introduce action-angle coordinates on the tangential variables:

$$u_j^+ = \sqrt{\xi_{|j|} + y_j}\, e^{\mathrm{i}x_j}\,,\ u_j^- = \sqrt{\xi_{|j|} + y_j}\, e^{-\mathrm{i}x_j}\,,\ j \in \mathcal{I}\,,\ (u_j^+, u_j^-) = (z_j^+, z_j^-) \equiv (z_j, \bar{z}_j)\,,\ j \notin \mathcal{I}\,, \tag{1.33}$$

where $|y_j| < \xi_{|j|}$. Then, system (1.23) is conjugated to a parameter dependent family of analytic systems of the form

$$\begin{cases} \dot{x} = \omega(\xi) + P^{(x)}(x, y, z, \bar{z}; \xi) \\ \dot{y} = \qquad\quad P^{(y)}(x, y, z, \bar{z}; \xi) \\ \dot{z}_j = \mathrm{i}\Omega_j(\xi) z_j + P^{(z_j)}(x, y, z, \bar{z}; \xi) \\ \dot{\bar{z}}_j = -\mathrm{i}\Omega_j(\xi) \bar{z}_j + P^{(\bar{z}_j)}(x, y, z, \bar{z}; \xi) \,, \quad j \in \mathbb{Z} \setminus \mathcal{I} \,, \end{cases} \tag{1.34}$$

where $(x, y) \in \mathbb{T}^n_s \times \mathbb{C}^n$, $z, \bar{z}$ are infinitely many variables, $\omega(\xi) \in \mathbb{R}^n$, $\Omega(\xi) \in \mathbb{R}^\infty$. The frequencies $\omega_j(\xi)$, $\Omega_j(\xi)$ are close to the unperturbed frequencies λ_j in (1.24) and satisfy $\omega_{-j} = \omega_j$, $\Omega_{-j} = \Omega_j$. System (1.34) is:

1. REVERSIBLE (see Definition 2.7) with respect to the involution

$$S : (x_j, y_j, z_j, \bar{z}_j) \mapsto (-x_{-j}, y_{-j}, \bar{z}_{-j}, z_{-j}) \,, \ \forall j \in \mathbb{Z} \,, \quad S^2 = I \,, \tag{1.35}$$

   which is nothing but (1.29) in the variables (1.33).

2. REAL-COEFFICIENTS, see Definition 2.8. Indeed, by (1.31), (1.33) and (7.30), all the functions

$$P^{(x)} \,, \ \mathrm{i}P^{(y)} \,, \ \mathrm{i}P^{(z_j)} \,, \ \mathrm{i}P^{(\bar{z}_j)}$$

   have *real* Taylor-Fourier coefficients in the variables $(x, y, z, \bar{z})$.

3. REAL-ON-REAL, see Definition 2.9.

4. EVEN. The vector field $P : E \to E$ and so the subspace

$$E := \Big\{ x_j = x_{-j} \,, \ y_j = y_{-j} \,, \ j \in \mathcal{I} \,, \quad z_j = z_{-j} \,, \ \bar{z}_j = \bar{z}_{-j} \,, \ j \in \mathbb{Z} \setminus \mathcal{I} \Big\} \tag{1.36}$$

   is invariant under the flow evolution of (1.34).

In system (1.34) we think $x_j, y_j, z_j^\pm$, as independent variables and then we look for solutions in the invariant subspace E, which means solutions of (1.15) even in $\mathtt{x}$.

Remark 1.2. *It would seem also natural to work directly in the cosine basis* $\{\cos(j\mathtt{x})\}_{j \geq 0}$ *instead of the Fourier representation* (1.22)*, namely to identify* $x_{-j} \equiv x_j$*,* $z^\pm_{-j} \equiv z^\pm_j$*. But, the notion of momentum is not well defined in the space of even functions. For example the vector fields* $z_{-j}\partial_{z_i}$ *and* $z_j \partial_{z_i}$*, that have* DIFFERENT *momentum, would be identified.*

Since the linear frequencies $\omega_{-j} = \omega_j$, $\Omega_{-j} = \Omega_j$, are resonant, along the KAM iteration, the monomial vector fields of the perturbation

$$\begin{aligned} & e^{\mathrm{i}k\cdot x}\partial_{x_j} \,, \quad e^{\mathrm{i}k\cdot x} y^i \partial_{y_j} \,, \ k \in \mathbb{Z}^n_{\mathrm{odd}} \,, \quad |i| = 0, 1, \ j \in \mathcal{I} \,, \\ & e^{\mathrm{i}k\cdot x} z_{\pm j}\partial_{z_j} \,, \quad e^{\mathrm{i}k\cdot x} \bar{z}_{\pm j} \partial_{\bar{z}_j}, \quad \forall k \in \mathbb{Z}^n_{\mathrm{odd}} \,, \ j \in \mathbb{Z} \setminus \mathcal{I} \,, \end{aligned}$$

where $\mathbb{Z}^n_{\mathrm{odd}} := \{k \in \mathbb{Z}^n : k_{-j} = -k_j \,, \forall j \in \mathcal{I}\}$, can not be averaged out. On the other hand, on the invariant subspace E, where we look for the quasi-periodic solutions, the above terms can be replaced by the constant coefficients monomial vector fields, obtained setting $x_{-j} = x_j$, $z^\pm_{-j} = z^\pm_j$. More precisely we proceed as follows: in section 5.1 we replace the nonlinear vector field P with its symmetrized $\mathcal{S}P$ (Definition 5.2) which determines the same dynamics on the invariant subspace E (Corollary 5.1) because

$$P_{|E} = (\mathcal{S}P)_{|E} \,.$$

The vector field $\mathcal{S}P$ is reversible and its weighted and quasi-Töplitz norms are (almost) the same as for P (Proposition 5.2). The homological equations (5.26) for a *symmetric* and *reversible* vector field perturbation (see (5.27)) can be solved, see Proposition 5.1, and the remaining resonant term (5.28) is a diagonal, constant coefficients correction of the normal form (5.1) (also using the *real coefficients* property). This procedure allows the KAM iteration to be carried out (remark 5.1 shows that the symmetrization procedure is required at each KAM step). Note that, after this composite KAM step, the correction to the normal frequencies described in (1.39) comes out from the symmetrized vector field $\mathcal{S}P$ and not from P itself.

As in the Hamiltonian case [4], a major difficulty of the KAM iteration is to fulfill, at each iterative step, the second order Melnikov non-resonance conditions. Actually, following the formulation of the KAM theorem given in [3]-[4] it is sufficient to verify

$$|\omega^\infty(\xi)\cdot k + \Omega_i^\infty(\xi) - \Omega_j^\infty(\xi)| \geq \frac{\gamma}{1+|k|^\tau}, \quad \gamma > 0, \tag{1.37}$$

only for the "final" frequencies $\omega^\infty(\xi)$ and $\Omega^\infty(\xi)$ and not along the inductive iteration.

As in [4] the key idea for verifying the second order Melnikov non-resonance conditions (1.37) for DNLW is to prove the higher order asymptotic decay estimate

$$\Omega_j^\infty(\xi) = j + a(\xi) + \frac{\mathrm{m}}{2j} + O(\frac{\gamma^{2/3}}{j}) \quad \text{for} \quad j \geq O(\gamma^{-1/3}) \tag{1.38}$$

where $a(\xi)$ is a constant independent of j.

This property follows by introducing the notion of *quasi-Töplitz vector field*, see Definition 3.4. The new normal frequencies for a symmetric perturbation $P = \mathcal{S}P$ are $\Omega_j^+ = \Omega_j + \mathrm{i}P^{z_j,z_j}$ where the corrections P^{z_j,z_j} are the diagonal entries of the matrix defined by

$$P^{z,z} z\partial_z := \sum_{i,j} P^{z_i,z_j} z_j \partial_{z_i}, \qquad P^{z_i,z_j} := \int_{\mathbb{T}^n} (\partial_{z_j} P^{(z_i)})(x,0,0,0;\xi)\, dx\,. \tag{1.39}$$

Thanks to the *real-coefficients* property, the corrections $\mathrm{i}P^{z_j,z_j}$ are real. We say that a matrix $P = P^{z,z}$ is quasi-Töplitz if it has the form

$$P = T + R$$

where T is a Töplitz matrix (i.e. constant on the diagonals) and R is a "small" remainder, satisfying in particular $R_{jj} = O(1/j)$. Then (1.38) follows with the constant $a := T_{jj}$ which is independent of j.

The definition of quasi-Töplitz vector field is actually simpler than that of quasi-Töplitz function, used in the Hamiltonian context [4], [31]. In turn, the notion of quasi-Töplitz function is weaker than the Töplitz-Lipschitz property, introduced by Eliasson-Kuksin [16]. The quasi-Töplitz nature of the perturbation is preserved along the KAM iteration (with slightly modified parameters) because the class of quasi-Töplitz vector fields is closed with respect to

1. Lie bracket (Proposition 3.1),
2. Lie series (Proposition 3.2),
3. Solution of the homological equation (Proposition 5.3),

which are the operations along the KAM iterative scheme.

An important difference with respect to [4] is that we do *not* require the conservation of momentum, and so Theorem 4.1 applies to the DNLW equation (1.2) where the nonlinearity g may depend on the space variable $\mathtt{x}$. The properties of quasi-Töpliz functions as introduced in [31], [4], strongly rely on the conservation of momentum. However, we remark that the concept of *momentum* of a vector field is always well defined (see Definition 2.3). Then we overcome the impasse of the non-conservation of momentum introducing the $\mathtt{a}$-weighted majorant norm for vector fields (Definition 2.4) which penalizes the high-momentum monomials (see (2.38)). Hence only the *low-momentum* monomials vector fields are relevant. This fact is crucial, in particular, in order to prove that the class of quasi-Töplitz vector fields is closed with respect to Lie brackets (Proposition 3.1).

Finally, concerning the KAM iteration, we note that we do not follow the same quadratic scheme of [4] for the Hamiltonian case, but a scheme similar to Moser [27] where we eliminate all the linear terms in $(y, z, \bar{z})$, see Definition 2.6. Actually, for a non Hamiltonian system it is more natural to treat the variables $(y, z, \bar{z})$ at the same level: this is realized assigning the same "degree" to these variables, see section 2.4.

1.3 Non existence of quasi-periodic solutions for DNLW

We now consider different nonlinearities for which we can exclude the existence of non-trivial quasi-periodic solutions.

Proposition 1.1. *Let $p \in \mathbb{N}$ be odd. The DNLW equations*

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_{\mathtt{x}}^p + f(\mathtt{y}) \,, \quad \mathtt{x} \in \mathbb{T} \,, \tag{1.40}$$

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \partial_{\mathtt{x}}(\mathtt{y}^p) + f(\mathtt{y}) \,, \quad \mathtt{x} \in \mathbb{T} \,, \tag{1.41}$$

have no smooth quasi-periodic solutions except for trivial periodic solutions of the form $\mathtt{y}(t, \mathtt{x}) = c(t)$. In particular $f \equiv 0$ implies $c(t) \equiv const.$

Proof. We claim that the function $M(\mathtt{y}, \mathtt{v}) := \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}} \, \mathtt{v} \, d\mathtt{x}$ is a Lyapunov function for both (1.40) and (1.41) where $\mathtt{y}_t = \mathtt{v}$. Indeed, along a smooth solution of (1.40) we have

$$\begin{aligned} \frac{d}{dt} M(\mathtt{y}, \mathtt{v}) &= \int_{\mathbb{T}} (\partial_t \mathtt{y}_{\mathtt{x}}) \mathtt{v} + \mathtt{y}_{\mathtt{x}} \, \mathtt{v}_t \, d\mathtt{x} = \int_{\mathbb{T}} (\partial_{\mathtt{x}} \mathtt{v}) \mathtt{v} + \mathtt{y}_{\mathtt{x}} (\mathtt{y}_{\mathtt{xx}} + \mathtt{y}_{\mathtt{x}}^p + f(\mathtt{y})) \, d\mathtt{x} \\ &= \int_{\mathbb{T}} \partial_{\mathtt{x}} \Big(\frac{\mathtt{v}^2 + \mathtt{y}_{\mathtt{x}}^2}{2} + F(\mathtt{y}) \Big) d\mathtt{x} + \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^{p+1} \, d\mathtt{x} = \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^{p+1} \, d\mathtt{x} \,. \end{aligned}$$

As a consequence $M(\mathtt{y}, \mathtt{v})$ is strictly decreasing along the solutions of (1.40) unless $\mathtt{y}_{\mathtt{x}}(t, \mathtt{x}) = 0$, $\forall t$, namely $\mathtt{y}(t, \mathtt{x}) = c$. Indeed, if there is a periodic solution $(\mathtt{y}(t, \mathtt{x}), \mathtt{v}(t, \mathtt{x}))$ of (1.40), with period T, then

$$0 = \int_0^T \Big(\int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^{p+1}(t, \mathtt{x}) \, d\mathtt{x} \Big) dt \quad \Longrightarrow \quad \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^{p+1}(t, \mathtt{x}) \, d\mathtt{x} = 0 \,, \ \forall t \in [0, T] \,,$$

because $p+1$ is even. Hence $\forall t \in [0, T]$, $\mathtt{y}_{\mathtt{x}}(t, \mathtt{x}) = 0$, $\forall \mathtt{x} \in \mathbb{T}$, and we conclude $\mathtt{y}(t, \mathtt{x}) = c(t)$. Similarly we exclude the existence of quasi-periodic solutions, since we would have $T_n \to +\infty$ such that

$$\lim_{n \to +\infty} \int_0^{T_n} \Big(\int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^{p+1}(t, \mathtt{x}) \, d\mathtt{x} \Big) dt = 0 \,.$$

Similarly, along a solution of (1.41) we derive, setting $g(\mathtt{y}) = \mathtt{y}^p$,

$$\frac{d}{dt} M(\mathtt{y}, \mathtt{v}) = \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^2 \, g'(\mathtt{y}) \, d\mathtt{x}$$

and we conclude as above because $g'(\mathtt{y}) = p\mathtt{y}^{p-1} > 0$, $\forall \mathtt{y} \neq 0$. ■

For wave equations, the role of the space variable $\mathtt{x}$ and time variable t is symmetric. A term like $\mathtt{y}_t^p$ for an odd p is a friction term which destroys the existence of quasi-periodic solutions.

Proposition 1.2. *Let $p \in \mathbb{N}$ be odd. The DNLW equation*

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_t^p + f(\mathtt{y}) \,, \quad \mathtt{x} \in \mathbb{T} \,, \tag{1.42}$$

has no smooth quasi-periodic solutions except except for trivial periodic solutions of the form $\mathtt{y}(t, \mathtt{x}) = c(\mathtt{x})$. In particular $f \equiv 0$ implies $c(\mathtt{x}) \equiv const.$

PROOF. The function $H(\mathtt{y},\mathtt{v}) := \int_{\mathbb{T}} \frac{\mathtt{v}^2}{2} + \frac{\mathtt{y}_{\mathtt{x}}^2}{2} - F(\mathtt{y})\, d\mathtt{x}$ (i.e. the Hamiltonian for the semilinear wave equation) is a Lyapunov function for (1.42). Indeed, along a smooth solution of (1.42) we get

$$\frac{d}{dt} H(\mathtt{y},\mathtt{v}) = \int_{\mathbb{T}} \mathtt{v}^{p+1}\, d\mathtt{x}\,.$$

We conclude as above that $\mathtt{y}_t(t,\mathtt{x}) = 0$, $\forall t, \mathtt{x}$, and so $\mathtt{y} = c(\mathtt{x})$. ■

The mass term $m\mathtt{y}$ could be necessary to have existence of quasi-periodic solutions.

Proposition 1.3. *The derivative NLW equation*

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_t^2\,, \quad \mathtt{x} \in \mathbb{T}\,, \tag{1.43}$$

has no smooth solutions defined for all times except the constants.

PROOF. We decompose the solution

$$\mathtt{y}(t,\mathtt{x}) = \mathtt{y}_0(t) + \tilde{\mathtt{y}}(t,\mathtt{x}) \quad \text{where} \quad \mathtt{y}_0 := \int_{\mathbb{T}} \mathtt{y}(t,\mathtt{x}) d\mathtt{x} \quad \text{and} \quad \tilde{\mathtt{y}} := \mathtt{y} - \mathtt{y}_0$$

has zero average in $\mathtt{x}$. Then, projecting (1.43) on the constants, we get

$$\ddot{\mathtt{y}}_0 = \int_{\mathbb{T}} \mathtt{y}_t^2 d\mathtt{x} = \int_{\mathbb{T}} (\dot{\mathtt{y}}_0 + \tilde{\mathtt{y}}_t)^2 d\mathtt{x} = \dot{\mathtt{y}}_0^2 + 2\dot{\mathtt{y}}_0 \int_{\mathbb{T}} \tilde{\mathtt{y}}_t d\mathtt{x} + \int_{\mathbb{T}} \tilde{\mathtt{y}}_t^2 d\mathtt{x} = \dot{\mathtt{y}}_0^2 + \int_{\mathbb{T}} \tilde{\mathtt{y}}_t^2 d\mathtt{x} \geq \dot{\mathtt{y}}_0^2\,. \tag{1.44}$$

Hence $\mathtt{v}_0 := \dot{\mathtt{y}}_0$ satisfies $\dot{\mathtt{v}}_0 \geq \mathtt{v}_0^2$ which blows up unless $\mathtt{v}_0 \equiv 0$. But, in this case, (1.44) implies that $\mathtt{y}_t(t,\mathtt{x}) \equiv 0$, $\forall \mathtt{x}$. Hence $\mathtt{y}(t,\mathtt{x}) = \mathtt{y}(\mathtt{x})$ and (1.43) (and $\mathtt{x} \in \mathbb{T}$) imply that $\mathtt{y}(t,\mathtt{x}) = \text{const}$. ■

For more general even power nonlinearities (both in $\mathtt{y}_{\mathtt{x}}$ and $\mathtt{y}_t$) the mass term $m\mathtt{y}$ (as well as any term depending on $\mathtt{y}$) could be necessary to allow the existence of quasi-periodic solutions:

Proposition 1.4. *Let $p, q \in \mathbb{N}$ be even. Then the derivative NLW equations*

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_{\mathtt{x}}^p\,, \quad \mathtt{x} \in \mathbb{T}\,, \tag{1.45}$$

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_t^p\,, \qquad \mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_{\mathtt{x}}^p + \mathtt{y}_t^q\,, \quad \mathtt{x} \in \mathbb{T}\,,$$

have no smooth periodic/quasi-periodic solutions except the constants.

PROOF. Let us consider for example (1.45). If there exists a periodic solution $(\mathtt{y}(t,\mathtt{x}), \mathtt{v}(t,\mathtt{x}))$ of (1.45), with period T, then

$$\int_0^T \int_{\mathbb{T}} (\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}}) dt d\mathtt{x} = 0 = \int_0^T \int_{\mathbb{T}} \mathtt{y}_{\mathtt{x}}^p(t,\mathtt{x}) d\mathtt{x} dt\,.$$

Hence, $\forall t \in [0,T]$, $\mathtt{y}_{\mathtt{x}}^p(t,\mathtt{x}) = 0$, $\forall \mathtt{x} \in \mathbb{T}$ (because p is even). Hence $\mathtt{y}_{\mathtt{x}}(t,\mathtt{x}) = 0$, $\forall \mathtt{x} \in \mathbb{T}$, that is $\mathtt{y}(t,\mathtt{x}) = c(t)$. Inserting in the equation (1.45) we get $c_{tt}(t) = 0$. Therefore $c(t) = a + bt$ and the only one which is periodic has $b = 0$. For quasi-periodic solutions the argument is the same. ■

We finally remark that all the solutions of the DNLW equation

$$\mathtt{y}_{tt} - \mathtt{y}_{\mathtt{xx}} = \mathtt{y}_t^2 - \mathtt{y}_{\mathtt{x}}^2\,, \quad \mathtt{x} \in \mathbb{T}\,, \tag{1.46}$$

(whose nonlinearity satisfies the "null-condition") are 2π-periodic in time. Indeed we check that

$$\mathtt{y} = -\ln(\alpha(t+\mathtt{x}) + \beta(t-\mathtt{x}))\,, \quad \alpha(t+\mathtt{x}) + \beta(t-\mathtt{x}) > 0 \tag{1.47}$$

are all the solutions of (1.46). The periodicity condition in space $\mathtt{x} \in \mathbb{T}$ implies that $\alpha(\cdot), \beta(\cdot)$ are 2π-periodic. Hence all the solutions (1.47) are 2π periodic.

2 Vector fields formalism

In this section we introduce the main properties of the vector fields that we shall use along the paper (commutators, momentum, norms, reversibility, degree, ...). We shall refer often to section 2 of [4]. The first difference between the present paper and [4] (which applies to Hamiltonian systems) is that we have to work always at the level of vector fields (reversible, see section 2.5) and not of functions (Hamiltonians). An important novelty of this section is the introduction of the $\mathtt{a}$-weighted majorant norm (Definition 2.4) which enables to deal with vector fields without requiring the conservation of momentum as in [4]. Finally, we note that the vector fields that we aim to eliminate along the KAM iteration (see Definition 2.6) are different than in the Hamiltonian case [4].

2.1 Functional setting

For $\mathcal{I} \subset \mathbb{Z}$ (possibly empty) and $a \geq 0, p > 1/2$, we define the Hilbert space

$$\ell^{a,p}_{\mathcal{I}} := \Big\{ z = \{z_j\}_{j \in \mathbb{Z} \setminus \mathcal{I}} \,,\ z_j \in \mathbb{C} \ :\ \|z\|^2_{a,p} := \sum_{j \in \mathbb{Z} \setminus \mathcal{I}} |z_j|^2 e^{2a|j|} \langle j \rangle^{2p} < \infty \Big\} \tag{2.1}$$

that, when $\mathcal{I} = \emptyset$, we denote more simply by $\ell^{a,p}$. Let n be the cardinality of $\mathcal{I}$. We consider

$$V := \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_{\mathcal{I}} \times \ell^{a,p}_{\mathcal{I}} \tag{2.2}$$

(denoted by E in [4]) with (s,r)-weighted norm

$$v = (x, y, z, \bar z) \in V \,, \quad \|v\|_V := \|v\|_{s,r} = \|v\|_{V,s,r} = \frac{|x|_\infty}{s} + \frac{|y|_1}{r^2} + \frac{\|z\|_{a,p}}{r} + \frac{\|\bar z\|_{a,p}}{r} \tag{2.3}$$

where $0 < s, r < 1$, and $|x|_\infty := \max_{h=1,\ldots,n} |x_h|$, $|y|_1 := \sum_{h=1}^n |y_h|$. For all $s' \leq s$, $r' \leq r$,

$$\|v\|_{s',r'} \leq \max\{s/s', (r/r')^2\} \|v\|_{s,r} \,. \tag{2.4}$$

Notice that z and $\bar z$ are *independent* complex variables. We shall also use the notation

$$z^+_j = z_j \,, \quad z^-_j = \bar z_j \,,$$

and we denote the set of variables

$$\mathtt{V} := \Big\{ x_1, \ldots, x_n, y_1, \ldots, y_n, \ldots, z_j, \ldots, \bar z_j, \ldots \Big\} , \quad j \in \mathbb{Z} \setminus \mathcal{I} \,. \tag{2.5}$$

As phase space, we define the toroidal domain

$$D(s,r) := \mathbb{T}^n_s \times D(r) := \mathbb{T}^n_s \times B_{r^2} \times B_r \times B_r \subset V \tag{2.6}$$

where $D(r) := B_{r^2} \times B_r \times B_r$,

$$\mathbb{T}^n_s := \Big\{ x \in \mathbb{C}^n \ :\ \max_{h=1,\ldots,n} |\mathrm{Im}\, x_h| < s \Big\} , \quad B_{r^2} := \Big\{ y \in \mathbb{C}^n \ :\ |y|_1 < r^2 \Big\} \tag{2.7}$$

and $B_r \subset \ell^{a,p}_{\mathcal{I}}$ is the open ball of radius r centered at zero. We think $\mathbb{T}^n$ as the n-dimensional torus $\mathbb{T}^n := 2\pi\mathbb{R}^n/\mathbb{Z}^n$, namely $f : D(s,r) \to \mathbb{C}$ means that f is 2π-periodic in each x_h-variable, $h = 1, \ldots, n$. If $n = 0$ then $D(s,r) \equiv B_r \times B_r \subset \ell^{a,p} \times \ell^{a,p}$.

Remark 2.1. *Let us explain the choice of the "scaling" in* (2.3),(2.6). *We want to prove persistence of elliptic tori for perturbation of the integrable system $\dot x = \omega$, $\dot y = 0$, $\dot z_j = -\mathrm{i}\Omega_j z_j$, $\dot{\bar z}_j = \mathrm{i}\Omega_j \bar z_j$ (recall* (1.34)*). Then it is natural that the variables z and $\bar z$ have the same scaling factor r while the variables y have a smaller scaling factor r'. We choose $r' = r^2$ since in the application to the DNLW y is an action variable, namely it has the same "dimension" of $z\bar z$, recall* (1.33).

We introduce the "real" phase space

$$\mathtt{R}(s,r) := \Big\{ v = (x, y, z^+, z^-) \in D(s,r) \ : \ x \in \mathbb{T}^n \, , \ y \in \mathbb{R}^n \, , \ \overline{z^+} = z^- \Big\} \tag{2.8}$$

where $\overline{z^+}$ is the complex conjugate of z^+.

2.2 Formal vector fields

Along this paper we consider vector fields of the form

$$X(v) = (X^{(x)}(v), X^{(y)}(v), X^{(z)}(v), X^{(\bar z)}(v)) \in V \tag{2.9}$$

where $v \in D(s,r)$ and $X^{(x)}(v), X^{(y)}(v) \in \mathbb{C}^n$, $X^{(z)}(v), X^{(\bar z)}(v) \in \ell^{a,p}_{\mathcal{I}}$. We shall also use the differential geometry notation

$$X(v) = X^{(x)}\partial_x + X^{(y)}\partial_y + X^{(z)}\partial_z + X^{(\bar z)}\partial_{\bar z} = \sum_{\mathtt{v} \in \mathtt{V}} X^{(\mathtt{v})}\partial_{\mathtt{v}} \tag{2.10}$$

(recall (2.5)) where

$$X^{(x)}\partial_x := \sum_{h=1}^{n} X^{(x_h)}\partial_{x_h}$$

and similarly for $X^{(y)}\partial_y, X^{(z)}\partial_z, X^{(\bar z)}\partial_{\bar z}$. Equivalently we write

$$X(v) = \Big(X^{(\mathtt{v})}(v) \Big)_{\mathtt{v} \in \mathtt{V}} . \tag{2.11}$$

Each component is a *formal* scalar power series with infinitely many variables

$$X^{(\mathtt{v})}(v) = \sum_{(k,i,\alpha,\beta) \in \mathbb{I}} X^{(\mathtt{v})}_{k,i,\alpha,\beta} \, e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar z^\beta \tag{2.12}$$

with coefficients $X^{(\mathtt{v})}_{k,i,\alpha,\beta} \in \mathbb{C}$ and multi-indices in

$$\mathbb{I} := \mathbb{Z}^n \times \mathbb{N}^n \times \mathbb{N}^{(\mathbb{Z}\setminus\mathcal{I})} \times \mathbb{N}^{(\mathbb{Z}\setminus\mathcal{I})} \tag{2.13}$$

where

$$\mathbb{N}^{(\mathbb{Z}\setminus\mathcal{I})} := \Big\{ \alpha := (\alpha_j)_{j \in \mathbb{Z}\setminus\mathcal{I}} \in \mathbb{N}^{\mathbb{Z}} \ \text{ with } \ |\alpha| := \sum_{j \in \mathbb{Z}\setminus\mathcal{I}} \alpha_j < +\infty \Big\} . \tag{2.14}$$

In (2.12) we use the standard multi-indices notation $z^\alpha \bar z^\beta := \Pi_{j \in \mathbb{Z}\setminus\mathcal{I}} \, z_j^{\alpha_j} \bar z_j^{\beta_j}$.

We recall that a formal series $\sum_{(k,i,\alpha,\beta)\in\mathbb{I}} c_{k,i,\alpha,\beta}$, $c_{k,i,\alpha,\beta} \in \mathbb{C}$, is *absolutely convergent* if the function $\mathbb{I} \ni (k,i,\alpha,\beta) \mapsto c_{k,i,\alpha,\beta} \in \mathbb{C}$ is in $L^1(\mathbb{I}, \mu)$ where μ is the counting measure of $\mathbb{I}$. Then we set

$$\sum_{(k,i,\alpha,\beta)\in\mathbb{I}} c_{k,i,\alpha,\beta} := \int_{\mathbb{I}} c_{k,i,\alpha,\beta} \, d\mu . \tag{2.15}$$

We consider *monomial* vector fields having all components zero, except one, which is a scalar monomial

$$\mathfrak{m}_{k,i,\alpha,\beta}(v) := e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar z^\beta . \tag{2.16}$$

Definition 2.1. (monomial vector field) *A monomial vector field is* $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}'}(v)$ *where*

$$\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}'}(v) = \mathfrak{m}_{k,i,\alpha,\beta}(v)\partial_{\mathtt{v}'} := e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}'} . \tag{2.17}$$

Each component is

$$\mathfrak{m}^{(\mathtt{v})}_{k,i,\alpha,\beta;\mathtt{v}'}(v) = \begin{cases} e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta & \text{if} \quad \mathtt{v} = \mathtt{v}' \\ 0 & \text{otherwise}. \end{cases}$$

Then a vector field X in (2.10) is decomposed as a formal series of vector field monomials with coefficients in $\mathbb{C}$:

$$X(v) = \sum_{\mathtt{v}\in\mathtt{V}} \sum_{(k,i,\alpha,\beta)\in\mathbb{I}} X^{(\mathtt{v})}_{k,i,\alpha,\beta} \mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}(v) = \sum_{\mathtt{v}\in\mathtt{V}} \sum_{(k,i,\alpha,\beta)\in\mathbb{I}} X^{(\mathtt{v})}_{k,i,\alpha,\beta} e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}} . \tag{2.18}$$

For a subset of indices $I \subset \mathbb{I} \times \mathtt{V}$ we define the projection

$$(\Pi_I X)(v) := \sum_{(k,i,\alpha,\beta,\mathtt{v})\in I} X^{(\mathtt{v})}_{k,i,\alpha,\beta}\, e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}} . \tag{2.19}$$

For any subset of indices $I, I' \subset \mathbb{I} \times \mathtt{V}$ we have

$$\Pi_I \Pi_{I'} = \Pi_{I\cap I'} = \Pi_{I'} \Pi_I . \tag{2.20}$$

Note that a projection on an index set corresponds to the projection on the space generated by the corresponding monomial vector fields.

Definition 2.2. *The formal vector field X is absolutely convergent in V (with norm* (2.3)*) at $v \in V$ if every component $X^{(\mathtt{v})}(v)$, $\mathtt{v} \in \mathtt{V}$, is absolutely convergent in v (see* (2.15)*) and*

$$\left\| \left(X^{(\mathtt{v})}(v)\right)_{\mathtt{v}\in\mathtt{V}} \right\|_V < +\infty .$$

The commutator (or Lie bracket) of two vector fields is

$$[X, Y](v) := dX(v)[Y(v)] - dY(v)[X(v)] \tag{2.21}$$

namely, its $\mathtt{v}$-component is

$$[X, Y]^{(\mathtt{v})} = \sum_{\mathtt{v}'\in\mathtt{V}} \partial_{\mathtt{v}'} X^{(\mathtt{v})} Y^{(\mathtt{v}')} - \partial_{\mathtt{v}'} Y^{(\mathtt{v})} X^{(\mathtt{v}')} . \tag{2.22}$$

Fixed a set of indices

$$\mathcal{I} := \{\mathtt{j}_1, \dots, \mathtt{j}_n\} \subset \mathbb{Z} , \tag{2.23}$$

we set $\mathtt{j} := (\mathtt{j}_1, \dots, \mathtt{j}_n) \in \mathbb{Z}^n$ and we define the *momentum vector field*

$$X_M := \Big(\mathtt{j}, 0, \dots, \mathrm{i}j z_j, \dots, -\mathrm{i}j \bar{z}_j, \dots\Big) , \quad j \in \mathbb{Z} \setminus \mathcal{I} .$$

Definition 2.3. *The* MOMENTUM *of the vector field monomial* $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}$ *is*

$$\pi(k,\alpha,\beta;\mathtt{v}) := \begin{cases} \pi(k,\alpha,\beta) & \text{if} \quad \mathtt{v} \in \{x_1,\dots,x_n,y_1,\dots,y_n\} \\ \pi(k,\alpha,\beta) - \sigma j & \text{if} \quad \mathtt{v} = z_j^\sigma \end{cases} \tag{2.24}$$

where

$$\pi(k,\alpha,\beta) := \sum_{i=1}^n \mathtt{j}_i k_i + \sum_{j\in\mathbb{Z}\setminus\mathcal{I}} (\alpha_j - \beta_j) j \tag{2.25}$$

is the momentum of the scalar monomial $\mathfrak{m}_{k,i,\alpha,\beta}(v)$.

The monomial vector fields $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}$ are eigenvectors, with eigenvalues $\mathrm{i}\pi(k,\alpha,\beta;\mathtt{v})$, of the adjoint action ad_{X_M} of the momentum vector field X_M. This shows the convenience of the exponential basis.

Lemma 2.1. *The commutator*

$$[\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}, X_M] = \mathrm{i}\pi(k,\alpha,\beta;\mathtt{v})\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}\,.$$

PROOF. By (2.22) with $Y = X_M$ we have $\partial_{\mathtt{v}'} Y^{(\mathtt{v})} = \mathrm{i}\sigma j$ if $\mathtt{v} = \mathtt{v}' = z_j^\sigma$ and 0 otherwise. ■

We say that a monomial vector field $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}$ satisfies momentum conservation if and only if $\pi(k,\alpha,\beta;\mathtt{v}) = 0$. Similarly, a vector field X satisfies momentum conservation if and only if it is a linear combination of monomial vector fields with zero momentum, i.e. it commutes with X_M.

2.3 Weighted majorant norm

In the following we consider a new parameter $\mathtt{a} \geq 0$. Given a formal vector field X as in (2.18) we define its "$\mathtt{a}$-weighted majorant" vector field

$$(M_{\mathtt{a}}X)(v) := \sum_{\mathtt{v}\in\mathtt{V}} \sum_{(k,i,\alpha,\beta)\in\mathbb{I}} e^{\mathtt{a}|\pi(k,\alpha,\beta;\mathtt{v})|} |X^{(\mathtt{v})}_{k,i,\alpha,\beta}| e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}} \tag{2.26}$$

where $\pi(k,\alpha,\beta;\mathtt{v})$ is the momentum of the monomial $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}$ defined in (2.24). When $\mathtt{a} = 0$ we simply write MX instead of M_0X, which coincides with the majorant vector field introduced in [4]-section 2.1.2, see also [2]. The role of the weight $e^{\mathtt{a}|\pi(k,\alpha,\beta;\mathtt{v})|}$ is to "penalize" the high momentum monomials. This is evident from (2.38) which will be exploited in Lemma 3.1 to neglect the high momentum monomial vector fields, slightly decreasing the parameter $\mathtt{a}$.

Definition 2.4. ($\mathtt{a}$-weighted majorant-norm) *The $\mathtt{a}$-weighted majorant norm of a formal vector field X as in* (2.18) *is*

$$\|X\|_{s,r,\mathtt{a}} \;:=\; \sup_{(y,z,\bar{z})\in D(r)} \Big\| \Big(\sum_{k,i,\alpha,\beta} e^{\mathtt{a}|\pi(k,\alpha,\beta;\mathtt{v})|} |X^{(\mathtt{v})}_{k,i,\alpha,\beta}| e^{|k|s} |y^i||z^\alpha||\bar{z}^\beta| \Big)_{\mathtt{v}\in\mathtt{V}} \Big\|_{s,r}. \tag{2.27}$$

When $\mathtt{a} = 0$ the norm $\|\cdot\|_{s,r,0}$ coincides with the "majorant norm" introduced in [4]-Definition 2.6 (where it was simply denoted by $\|\cdot\|_{s,r}$). By (2.27) and (2.26) we get

$$\|X\|_{s,r,\mathtt{a}} = \|M_{\mathtt{a}}X\|_{s,r,0}\,. \tag{2.28}$$

Remark 2.2. *Relation* (2.28) *makes evident that the norm $\|\cdot\|_{s,r,\mathtt{a}}$ satisfies the same properties of the majorant norm $\|\cdot\|_{s,r,0}$. Then the following lemmata on the $\mathtt{a}$-weighted majorant-norm $\|\cdot\|_{s,r,\mathtt{a}}$ follow by the analogous lemmata of [4] for the majorant norm $\|\cdot\|_{s,r,0}$.*

For an absolutely convergent vector field $X : D(s,r) \to V$ we define the sup-norm

$$|X|_{s,r} := \sup_{v\in D(s,r)} \|X(v)\|_{V,s,r}\,. \tag{2.29}$$

Lemma 2.2. *Assume that for some $s, r > 0$, $\mathtt{a} \geq 0$, the weighted majorant-norm*

$$\|X\|_{s,r,\mathtt{a}} < +\infty\,. \tag{2.30}$$

Then the series in (2.18)*, resp.* (2.26)*, absolutely converge to the* analytic *vector field $X(v)$, resp. $M_{\mathtt{a}}X(v)$, for every $v \in D(s,r)$. Moreover*

$$|X|_{s,r},\ |M_{\mathtt{a}}X|_{s,r} \;\leq\; \|X\|_{s,r,\mathtt{a}}\,. \tag{2.31}$$

Then we consider the following Banach subspaces of analytic vector fields

$$\mathcal{V}_{s,r,\mathtt{a}} := \Big\{ X : D(s,r) \to V \ \text{ with norm } \|X\|_{s,r,\mathtt{a}} < +\infty \Big\}$$

and, for $h \in \mathbb{Z}$,

$$\mathcal{V}_{s,r,\mathtt{a}}(h) := \Big\{ X := \sum_{\pi(k,\alpha,\beta;\mathtt{v})=h} X^{(\mathtt{v})}_{k,i,\alpha,\beta}\, e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}} \in \mathcal{V}_{s,r,\mathtt{a}} \Big\}\,.$$

For a vector field $X \in \mathcal{V}_{s,r,\mathtt{a}}(h)$ we define its momentum $\pi(X) = h$.

Lemma 2.3. *Let X, Y have momentum $\pi(X)$, $\pi(Y)$, respectively. Then $\pi([X,Y]) = \pi(X) + \pi(Y)$.*

Finally we consider a family of vector fields

$$X : D(s,r) \times \mathcal{O} \to V \tag{2.32}$$

depending on *parameters* $\xi \in \mathcal{O} \subset \mathbb{R}^n$. For $\lambda \geq 0$, we define the λ-weighted Lipschitz norm

$$\begin{aligned} \|X\|^\lambda_{s,r,\mathtt{a},\mathcal{O}} := \|X\|^\lambda_{s,r,\mathtt{a}} &:= \|X\|_{s,r,\mathtt{a},\mathcal{O}} + \lambda \|X\|^{\mathrm{lip}}_{s,r,\mathtt{a},\mathcal{O}} \\ &:= \sup_{\xi\in\mathcal{O}} \|X(\xi)\|_{s,r,\mathtt{a}} + \lambda \sup_{\xi,\eta\in\mathcal{O},\ \xi\neq\eta} \frac{\|X(\xi)-X(\eta)\|_{s,r,\mathtt{a}}}{|\xi-\eta|} \end{aligned} \tag{2.33}$$

and we set

$$\mathcal{V}^\lambda_{s,r,\mathtt{a}} := \mathcal{V}^\lambda_{s,r,\mathtt{a},\mathcal{O}} := \Big\{ X(\cdot\,;\xi) \in \mathcal{V}_{s,r,\mathtt{a}}\,,\ \forall \xi \in \mathcal{O} \ :\ \|X\|^\lambda_{s,r,\mathtt{a}} < \infty \Big\}\,.$$

Note that, if X is independent of ξ, then $\|X\|^\lambda_{s,r,\mathtt{a}} = \|X\|_{s,r,\mathtt{a}}$, $\forall\lambda$.

The $\|\cdot\|^\lambda_{s,r,\mathtt{a}}$ norm behaves well under projections, see (2.19).

Lemma 2.4. (Projection) *$\forall I \subset \mathbb{I} \times \mathtt{V}$ we have $\|\Pi_I X\|_{s,r,\mathtt{a}} \leq \|X\|_{s,r,\mathtt{a}}$ and $\|\Pi_I X\|^{\mathrm{lip}}_{s,r,\mathtt{a}} \leq \|X\|^{\mathrm{lip}}_{s,r,\mathtt{a}}$.*

Important particular cases are the "ultraviolet" projection

$$(\Pi_{|k|\geq K} X)(v) := \sum_{|k|\geq K, i,\alpha,\beta} X^{(\mathtt{v})}_{k,i,\alpha,\beta}\, e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}}\,, \qquad \Pi_{|k|<K} := \mathrm{Id} - \Pi_{|k|\geq K} \tag{2.34}$$

and the "high momentum" projection

$$(\Pi_{|\pi|\geq K} X)(v) := \sum_{|\pi(k,\alpha,\beta;\mathtt{v})|\geq K} X^{(\mathtt{v})}_{k,i,\alpha,\beta}\, e^{\mathrm{i}k\cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}}\,, \qquad \Pi_{|\pi|<K} := \mathrm{Id} - \Pi_{|\pi|\geq K}\,. \tag{2.35}$$

We also define a further projection Π_{diag} by linearity, setting

$$\Pi_{\mathrm{diag}} \mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}} = \begin{cases} \mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}} & \text{if } k=0,\ i=0,\ \alpha=e_j,\ \beta=0,\ \mathtt{v}=z_j\,,\ j\in\mathbb{Z}\setminus\mathcal{I} \\ \mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}} & \text{if } k=0,\ i=0,\ \alpha=0,\ \beta=e_j,\ \mathtt{v}=\bar{z}_j,\ j\in\mathbb{Z}\setminus\mathcal{I} \\ 0 & \text{otherwise.} \end{cases} \tag{2.36}$$

By (2.27) following smoothing estimates hold:

Lemma 2.5. (Smoothing) *$\forall K \geq 1$ and $\lambda \geq 0$*

$$\|\Pi_{|k|\geq K} X\|^\lambda_{s',r,\mathtt{a}} \leq \frac{s}{s'}\, e^{-K(s-s')} \|X\|^\lambda_{s,r,\mathtt{a}}\,, \qquad \forall\, 0 < s' < s \tag{2.37}$$

$$\|\Pi_{|\pi|\geq K} X\|^\lambda_{s,r,\mathtt{a}'} \leq e^{-K(\mathtt{a}-\mathtt{a}')} \|X\|^\lambda_{s,r,\mathtt{a}}\,, \qquad \forall\, 0 \leq \mathtt{a}' \leq \mathtt{a}\,. \tag{2.38}$$

The space of analytic vector fields with finite $\mathtt{a}$-weighted majorant norm is a Lie algebra:

Proposition 2.1. (Commutator) *Let $X, Y \in \mathcal{V}^\lambda_{s,r,\mathtt{a}}$. Then, for $\lambda \geq 0$, $r/2 \leq r' < r$, $s/2 \leq s' < s$, $[X,Y] \in \mathcal{V}^\lambda_{s',r',\mathtt{a}}$ and*

$$\|[X,Y]\|^\lambda_{s',r',\mathtt{a}} \leq 2^{2n+3}\delta^{-1}\|X\|^\lambda_{s,r,\mathtt{a}}\|Y\|^\lambda_{s,r,\mathtt{a}} \tag{2.39}$$

where

$$\delta := \min\left\{1 - \frac{s'}{s}, 1 - \frac{r'}{r}\right\}. \tag{2.40}$$

Proof. The proof of

$$\|[X,Y]\|_{s',r',\mathtt{a}} \leq 2^{2n+3}\delta^{-1}\|X\|_{s,r,\mathtt{a}}\|Y\|_{s,r,\mathtt{a}} \tag{2.41}$$

follows as in [4], Lemma 2.15, by exploiting Lemma 2.3. Then

$$\begin{aligned}\|[X,Y](\xi) - [X,Y](\eta)\|_{s',r',\mathtt{a}} &= \Big\|[X(\xi) - X(\eta), Y(\xi)] + [X(\eta), Y(\xi) - Y(\eta)]\Big\|_{s',r',\mathtt{a}} \\ &\leq \Big\|[X(\xi) - X(\eta), Y(\xi)]\Big\|_{s',r',\mathtt{a}} + \Big\|[X(\eta), Y(\xi) - Y(\eta)]\Big\|_{s',r',\mathtt{a}} \\ &\overset{(2.41),(2.33)}{\leq} 2^{2n+3}\delta^{-1}|\xi - \eta|\Big(\|X\|^{\rm lip}_{s,r,\mathtt{a}}\|Y\|_{s,r,\mathtt{a}} + \|Y\|^{\rm lip}_{s,r,\mathtt{a}}\|X\|_{s,r,\mathtt{a}}\Big)\end{aligned}$$

and (2.39) follows taking the supremum in the parameters. ■

Given a vector field X, its transformed field under the time 1 flow generated by Y is

$$e^{{\rm ad}_Y}X = \sum_{k\geq 0}\frac{1}{k!}{\rm ad}^k_Y X\,, \qquad {\rm ad}_Y X := [X,Y]\,, \tag{2.42}$$

where ${\rm ad}^k_Y := {\rm ad}^{k-1}_Y{\rm ad}_Y$ and ${\rm ad}^0_Y := {\rm Id}$.

Proposition 2.2. (Flows) *Let $\lambda \geq 0$, $r/2 \leq r' < r$, $s/2 \leq s' < s$, and $Y \in \mathcal{V}^\lambda_{s,r,\mathtt{a}}$ with*

$$\|Y\|^\lambda_{s,r,\mathtt{a}} < \eta := \delta/(2^{2n+5}e) \tag{2.43}$$

and δ defined in (2.40). Then the time 1 flow generated by Y maps $D(s',r') \to D(s,r)$ and, for all $X \in \mathcal{V}^\lambda_{s,r,\mathtt{a}}$, the transformed vector field $e^{{\rm ad}_Y}X \in \mathcal{V}^\lambda_{s',r',\mathtt{a}}$ satisfies

$$\|e^{{\rm ad}_Y}X\|^\lambda_{s',r',\mathtt{a}} \leq \frac{\|X\|^\lambda_{s,r,\mathtt{a}}}{1 - \eta^{-1}\|Y\|^\lambda_{s,r,\mathtt{a}}}\,. \tag{2.44}$$

We conclude this section with two simple lemmata.

Lemma 2.6. *Let $P = \sum_{|k|\leq K,i,\alpha,\beta} P_{k,i,\alpha,\beta;\mathtt{v}}e^{{\rm i}k\cdot x}y^i z^\alpha \bar z^\beta\partial_{\mathtt{v}}$ and $|\Delta_{k,i,\alpha,\beta;\mathtt{v}}| \geq \gamma\langle k\rangle^{-\tau}$, $\forall |k| \leq K, i, \alpha, \beta, \mathtt{v}$. Then*

$$F := \sum_{|k|\leq K,i,\alpha,\beta;\mathtt{v}}\frac{P_{k,i,\alpha,\beta;\mathtt{v}}}{\Delta_{k,i,\alpha,\beta;\mathtt{v}}}e^{{\rm i}k\cdot x}y^i z^\alpha \bar z^\beta\partial_{\mathtt{v}} \quad \text{satisfies} \quad \|F\|_{s,r,\mathtt{a}} \leq \gamma^{-1}K^\tau\|P\|_{s,r,\mathtt{a}}\,.$$

Lemma 2.7. *Let $P = \sum_{j\in\mathbb{Z}\setminus\mathcal{I}} P_j z_j\partial_{z_j}$ with $\|P\|_r < \infty$. Then $|P_j| \leq \|P\|_r$ and $|P_j|^{\rm lip} \leq \|P\|^{\rm lip}_r$. An analogous statement holds for $P = \sum_{j\in\mathbb{Z}\setminus\mathcal{I}} P_j\bar z_j\partial_{\bar z_j}$.*

Proof. By Definition 2.4 we have

$$\|P\|^2_r = 2\sup_{\|z\|_{a,p}<r}\sum_{h\in\mathbb{Z}\setminus\mathcal{I}}|P_h|^2\frac{|z_h|^2}{r^2}e^{2a|h|}\langle h\rangle^{2p} \geq |P_j|^2$$

by evaluating at $z^{(j)}_h := \delta_{jh}e^{-a|j|}\langle j\rangle^p r/\sqrt{2}$. Applying the above estimates to P', where $P' := P(\xi) - P(\eta) = \sum_{j\in\mathbb{Z}\setminus\mathcal{I}}\big(P_j(\xi) - P_j(\eta)\big)z_j\partial_{z_j}$, we get $|P_j(\xi) - P_j(\eta)|/|\xi - \eta| \leq \|P(\xi) - P(\eta)\|_r/|\xi - \eta|$; then the Lipschitz estimate follows. ■

2.4 Degree decomposition

We now define the terms of the vector field that we aim to eliminate along the KAM iteration.

Definition 2.5. *The degree of the monomial vector field* $\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}$ *is*

$$d(\mathfrak{m}_{k,i,\alpha,\beta;\mathtt{v}}) := |i| + |\alpha| + |\beta| - d(\mathtt{v}) \quad \text{where} \quad d(\mathtt{v}) := \begin{cases} 0 & \text{if} \quad \mathtt{v} \in \{x_1, \dots, x_n\} \\ 1 & \text{otherwise.} \end{cases}$$

This notion naturally extends to any vector field by monomial decomposition. We say that a vector field has degree h if it is an absolutely convergent combination of monomial vector fields of degree h. We explicit the degrees of the basic vector field monomials $d(\partial_x) = 0$, $d(\partial_y) = d(\partial_{z_j}) = d(\partial_{\bar{z}_j}) = -1$. The degree d gives to the analytic vector fields the structure of a graded Lie algebra, because, given two vector fields X, Y of degree respectively $d(X)$ and $d(Y)$, then

$$d([X, Y]) = d(X) + d(Y) . \tag{2.45}$$

For $X \in \mathcal{V}_{s,r,\mathtt{a}}$ we define the homogeneous component of degree $l \in \mathbb{N}$,

$$X^{(l)} := \Pi^{(l)} X := \sum_{|i|+|\alpha|+|\beta|-d(\mathtt{v})=l} X^{(\mathtt{v})}_{k,i,\alpha,\beta} \, e^{\mathrm{i} k \cdot x} y^i z^\alpha \bar{z}^\beta \partial_{\mathtt{v}} . \tag{2.46}$$

We also set

$$X^{\leq 0} := X^{(-1)} + X^{(0)} . \tag{2.47}$$

The above projector $\Pi^{(l)}$ has the form Π_I, see (2.19), for a suitable subset $I \subset \mathbb{I} \times \mathtt{V}$.

Definition 2.6. *We denote by* $\mathcal{R}^{\leq 0}$ *the vector fields with degree* ≤ 0*. Using the compact notation*

$$\mathtt{u} := (y, z, \bar{z}) = (y, z^+, z^-) ,$$

a vector field in $\mathcal{R}^{\leq 0}$ *writes*

$$R = R^{\leq 0} = R^{(-1)} + R^{(0)} , \quad R^{(-1)} = R^{\mathtt{u}}(x)\partial_{\mathtt{u}} , \quad R^{(0)} = R^x(x)\partial_x + R^{\mathtt{u},\mathtt{u}}(x)\mathtt{u}\,\partial_{\mathtt{u}} , \tag{2.48}$$

where

$$R^x(x) \in \mathbb{C}^n , \quad R^{\mathtt{u}} \in \mathbb{C}^n \times \ell^{a,p}_{\mathcal{I}} \times \ell^{a,p}_{\mathcal{I}} , \quad R^{\mathtt{u},\mathtt{u}}(x) \in \mathcal{L}(\mathbb{C}^n \times \ell^{a,p}_{\mathcal{I}} \times \ell^{a,p}_{\mathcal{I}}) . \tag{2.49}$$

In more extended notation

$$\begin{aligned} R^{\mathtt{u}}(x)\partial_{\mathtt{u}} &= R^y(x)\partial_y + R^z(x)\partial_z + R^{\bar{z}}(x)\partial_{\bar{z}} \\ R^{\mathtt{u},\mathtt{u}}(x)\mathtt{u}\partial_{\mathtt{u}} &= \big(R^{y,y}(x)y + R^{y,z}(x)z + R^{y,\bar{z}}(x)\bar{z}\big)\partial_y + \big(R^{z,y}(x)y + R^{z,z}(x)z + R^{z,\bar{z}}(x)\bar{z}\big)\partial_z \\ &\quad + \big(R^{\bar{z},y}(x)y + R^{\bar{z},z}(x)z + R^{\bar{z},\bar{z}}(x)\bar{z}\big)\partial_{\bar{z}} . \end{aligned} \tag{2.50}$$

The terms of the vector field that we want to eliminate (or normalize) along the KAM iteration are those in $\mathcal{R}^{\leq 0}$. The graded Lie algebra property (2.45) implies that $\mathcal{R}^{\leq 0}$ is closed by Lie bracket:

Lemma 2.8. *If* $X, Y \in \mathcal{R}^{\leq 0}$ *then* $[X, Y] \in \mathcal{R}^{\leq 0}$.

The above observation is useful for analyzing the new normal form along the KAM step.

Remark 2.3. *In the Hamiltonian KAM theorem [4] we do not eliminate the terms* $R^{z,y}(x)y\partial_z$, $R^{\bar{z},y}(x)y\partial_{\bar{z}}$ *and instead we remove* $R^{y,zz}(x)z^2\partial_y$*. Actually also without eliminating the terms* $R^{z,y}(x)y\partial_z$, $R^{\bar{z},y}(x)y\partial_{\bar{z}}$*, the KAM scheme would be quadratic. In any case it is free to remove such terms. We follow [27].*

2.5 Reversible, real-coefficients, real-on-real, even, vector fields

We now define the class of reversible/anti-reversible vector fields, see [7].

Definition 2.7. (Reversibility) *A vector field* $X = (X^{(x)}, X^{(y)}, X^{(z)}, X^{(\bar z)})$ *(see* (2.9)*) is* REVERSIBLE *with respect to an involution* S *(namely* $S^2 = I$*) if*

$$X \circ S = -S \circ X\,. \tag{2.51}$$

A vector field Y *is* ANTI-REVERSIBLE *if*

$$Y \circ S = S \circ Y\,. \tag{2.52}$$

When the set $\mathcal{I}$ is symmetric as in (1.32) and S is the involution in (1.35), a vector field X is reversible if its coefficients (see (2.12)) satisfy

$$X^{(\mathtt{v})}_{k,i,\alpha,\beta} = \begin{cases} X^{(\hat{\mathtt{v}})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = x_j\,, \quad j \in \mathcal{I}\,, \\ -X^{(\hat{\mathtt{v}})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = y_j\,, \quad j \in \mathcal{I}\,, \\ -X^{(z^{-\sigma}_{-j})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = z^\sigma_j\,, \quad j \in \mathbb{Z}\setminus\mathcal{I} \end{cases} \tag{2.53}$$

where

$$\hat k := (k_{-j})_{j\in\mathcal{I}}\,,\ \hat\imath := (i_{-j})_{j\in\mathcal{I}}\,,\ \hat\beta := (\beta_{-j})_{j\in\mathbb{Z}\setminus\mathcal{I}}\,,\ \hat\alpha := (\alpha_{-j})_{j\in\mathbb{Z}\setminus\mathcal{I}}\,,\ \hat{\mathtt{v}} := (\mathtt{v}_{-j})_{j\in\mathbb{Z}}\,. \tag{2.54}$$

Clearly Y is anti-reversible if

$$Y^{(\mathtt{v})}_{k,i,\alpha,\beta} = \begin{cases} -Y^{(\hat{\mathtt{v}})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = x_j\,, \quad j \in \mathcal{I}\,, \\ Y^{(\hat{\mathtt{v}})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = y_j\,, \quad j \in \mathcal{I}\,, \\ Y^{(z^{-\sigma}_{-j})}_{-\hat k,\hat\imath,\hat\beta,\hat\alpha} & \text{if} \quad \mathtt{v} = z^\sigma_j\,, \quad j \in \mathbb{Z}\setminus\mathcal{I}. \end{cases} \tag{2.55}$$

Definition 2.8. (real-coefficients) *A vector field*

$$X = X^{(x)}\partial_x + X^{(y)}\partial_y + X^{(z^+)}\partial_{z^+} + X^{(z^-)}\partial_{z^-}$$

*is called "*REAL-COEFFICIENTS*" if the Taylor-Fourier coefficients of* $X^{(x)}, \mathrm{i}X^{(y)}, \mathrm{i}X^{(z^+)}, \mathrm{i}X^{(z^-)}$ *are real. The vector field* X *is called "*ANTI-REAL-COEFFICIENTS*" if* $\mathrm{i}X$ *is real-coefficients.*

Definition 2.9. (real-on-real) *A vector field* $X : D(s,r) \to V$ *is* REAL-ON-REAL *if*

$$X^{(x)}(v) = \overline{X^{(x)}(v)}\,,\ X^{(y)}(v) = \overline{X^{(y)}(v)}\,,\ X^{(z^-)}(v) = \overline{X^{(z^+)}(v)}\,, \quad \forall v \in \mathtt{R}(s,r)\,, \tag{2.56}$$

where $\mathtt{R}(s,r)$ *is defined in* (2.8).

On the coefficients in (2.12) the REAL-ON-REAL condition amounts to

$$\overline{X^{(\mathtt{v})}_{k,i,\alpha,\beta}} = \begin{cases} X^{(\mathtt{v})}_{-k,i,\beta,\alpha} & \text{if} \quad \mathtt{v} \in \{x_1,\dots,x_n,y_1,\dots,y_n\} \\ X^{(z_j^{-\sigma})}_{-k,i,\beta,\alpha} & \text{if} \quad \mathtt{v} = z^\sigma_j\,. \end{cases} \tag{2.57}$$

Definition 2.10. (Even) *A vector field* X *is "*EVEN*" if* $X : E \to E$ *(see* (1.36)*).*

On the coefficients in (2.12) the above PARITY condition amounts to

$$X^{(\mathtt{v})}_{k,i,\alpha,\beta} = X^{(\hat{\mathtt{v}})}_{\hat k,\hat\imath,\hat\alpha,\hat\beta} \quad (\text{see } (2.54))\,. \tag{2.58}$$

Lemma 2.9. *Let X, Y be vector fields.*

1. *If X is reversible and Y is anti-reversible then the commutator vector field $[X,Y]$ is reversible as well as the transformed vector field $e^{\mathrm{ad}_Y}X$ (recall (2.42)).*

2. *If X, resp. Y, is real-coefficients, resp. anti-real-coefficients, then $[X,Y]$, $e^{\mathrm{ad}_Y}X$ are real-coefficients,*

3. *If X,Y are real-on-real, then $[X,Y]$, $e^{\mathrm{ad}_Y}X$ are real-on-real,*

4. *If X,Y are even then $[X,Y]$, $e^{\mathrm{ad}_Y}X$ are even.*

PROOF. Let us prove the first case. We have by (2.51), (2.52) that

$$[X,Y](Sv) = DX(Sv)[Y(Sv)] - DY(Sv)[X(Sv)] = DX(Sv)[SY(v)] + DY(Sv)[SX(v)]\,. \tag{2.59}$$

Now, differentiating (2.51), (2.52), we get $DX(Sv)[Sh] = -SDX(v)[h]$, $DY(Sv)[Sh] = SDY(v)[h]$, $\forall v,h$ Hence, inserting in (2.59) we get

$$[X,Y](Sv) = -SDX(v)[Y(v)] + SDY(v)[X(v)] = -S[X,Y](v)$$

namely $[X,Y]$ is reversible. Iterating in (2.42), we get that $e^{\mathrm{ad}_Y}X$ is reversible.
The other cases follow similarly from the Definitions 2.8-2.10. ■

Definition 2.11. *We denote by*

- *$\mathcal{R}_{rev}$ the vector fields which are reversible (Definition 2.7) real-coefficients (Definition 2.8), real-on-real (Definition 2.9) and even (Definition 2.10).*
- *$\mathcal{R}_{a\text{-}rev}$ the vector fields which are anti-reversible, anti-real-coefficients, real-on-real and even.*
- *$\mathcal{R}^{\leq 0}_{rev} := \mathcal{R}_{rev} \cap \mathcal{R}^{\leq 0}$ and $\mathcal{R}^{\leq 0}_{a-rev} := \mathcal{R}_{a-rev} \cap \mathcal{R}^{\leq 0}$.*

Lemma 2.9 immediately implies

Lemma 2.10. *If $X \in \mathcal{R}_{rev}$ and $Y \in \mathcal{R}_{a-rev}$ then $[X,Y]$, $e^{\mathrm{ad}_Y}X \in \mathcal{R}_{rev}$.*

By (2.46), (2.47) and (2.58) we immediately get (the space E was defined in (1.36))

$$X_{|E} \equiv 0 \qquad \Longrightarrow \qquad (X^{\leq 0})_{|E} \equiv 0\,. \tag{2.60}$$

Lemma 2.11. *If $X_{|E} \equiv 0$ and Y is even (definition 2.10), then $\big([X,Y]\big)_{|E} \equiv 0$, $(e^{\mathrm{ad}_Y}X)_{|E} \equiv 0$.*

PROOF. If $v \in E$ we get

$$[X,Y](v) \stackrel{(2.21)}{=} dX(v)[Y(v)] = \frac{d}{ds}_{|s=0} X\big(v + sY(v)\big) = 0\,,$$

since $v + sY(v) \in E$. The statement on the Lie series follows by induction from (2.42). ■

3 Quasi-Töplitz vector fields

Let $N_0 \in \mathbb{N}$, $\theta, \mu \in \mathbb{R}$ be parameters such that

$$1 < \theta, \mu < 6\,, \quad 12N_0^{L-1} + 2\kappa N_0^{b-1} < 1\,, \quad \kappa := \max_{1\leq l\leq n} |\mathtt{j}_l|\,, \tag{3.1}$$

(the $\mathtt{j}_l$ are defined in (2.23)) where

$$0 < b < L < 1\ . \tag{3.2}$$

In the following we will always take

$$N \geq N_0 \,. \tag{3.3}$$

For a scalar function $f : D(s,r) \to \mathbb{C}$ let

$$\|f\|_{s,r,\mathtt{a}} := \sup_{(y,z,\bar z)\in D(r)} \sum_{k,i,\alpha,\beta} e^{\mathtt{a}|\pi(\alpha,\beta,k)|}|f_{k,i,\alpha,\beta}|e^{s|k|}|y^i||z^\alpha||\bar z^\beta| \,. \tag{3.4}$$

Definition 3.1. *A scalar monomial $e^{\mathrm{i}k\cdot x}y^i z^\alpha \bar z^\beta$ is (N,μ)-**low momentum** if*

$$|k| < N^b \,, \quad \alpha+\beta = \gamma \quad \text{with} \quad \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l < \mu N^L \,. \tag{3.5}$$

*An (N,μ)-low momentum scalar monomial is (N,μ,h)-**low** if*

$$|\pi(k,\alpha,\beta) - h| < N^b \,. \tag{3.6}$$

We denote by $\mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu)$, respectively $\mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,h)$, the closure of the space of (N,μ)–low, resp. (N,μ,h)–low, scalar monomials in the norm $\| \ \|_{s,r,\mathtt{a}}$ in (3.4)*.*
The projection on $\mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,h)$ will be denoted by $\Pi^{L,h}_{N,\mu}$. Note that it is a projection (see (2.19)*) on the subset of indexes $I \subset \mathbb{I}$ satisfying* (3.5) *and* (3.6)*.*

Clearly, the momentum (2.25) of a scalar monomial $\mathfrak{m}(k,i,\alpha,\beta)$, which is (N,μ)-low momentum, satisfies

$$|\pi(k,\alpha,\beta)| \overset{(3.1),(3.5)}{\leq} \kappa N^b + \mu N^L \,.$$

Hence a scalar monomial $\mathfrak{m}(k,i,\alpha,\beta)$ may be (N,μ,h)–low only if

$$|h| < |\pi(k,\alpha,\beta)| + N^b < \mu N^L + (\kappa+1)N^b \overset{(3.1),(3.3)}{<} N \,. \tag{3.7}$$

In particular

$$\mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,h) = \emptyset \,, \qquad \forall\, |h| \geq N \,. \tag{3.8}$$

Definition 3.2. *A vector field monomial $\mathfrak{m}(k,i,\alpha,\beta;\mathtt{v})$ is*

- *(N,μ)-**low** if*

$$|\pi(k,\alpha,\beta;\mathtt{v})|, |k| < N^b \,, \ \alpha+\beta = \gamma \ \text{ with } \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l < \mu N^L \,. \tag{3.9}$$

- *(N,θ,μ)–**linear** if*

$$\mathtt{v} = z_m^\sigma,\ |\pi(k,\alpha,\beta;\mathtt{v})|,\ |k| < N^b,\ \alpha+\beta = e_n+\gamma \text{ with } |m|,|n| > \theta N,\ \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l < \mu N^L \,. \tag{3.10}$$

We denote by $\mathcal{V}^L_{s,r,\mathtt{a}}(N,\mu)$, respectively $\mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$, the closure in $\mathcal{V}_{s,r,\mathtt{a}}$ of the vector space generated by the (N,μ)–low, respectively (N,θ,μ)–linear, monomial vector fields. The elements of $\mathcal{V}^L_{s,r,\mathtt{a}}(N,\mu)$, resp. $\mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$, are called (N,μ)-low, resp. (N,θ,μ)–linear, vector fields.

The projections on $\mathcal{V}^L_{s,r,\mathtt{a}}(N,\mu)$, resp. $\mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$, are denoted by $\Pi^L_{N,\mu}$, resp. $\Pi_{N,\theta,\mu}$. Explicitely $\Pi^L_{N,\mu}$ and $\Pi_{N,\theta,\mu}$, are the projections (see (2.19)*) on the subsets of indexes $I \subset \mathbb{I}\times\mathtt{V}$ satisfying* (3.9) *and* (3.10) *respectively.*

By Definition 3.2, (2.24) and (3.6), a vector field $X \in \mathcal{V}^L_{s,r,\mathtt{a}}(N,\mu)$ has components $X^{(x_h)}, X^{(y_h)} \in \mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,0)$ and $X^{(z^\sigma_m)} \in \mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,\sigma m)$. Moreover, by (3.10) and (3.5), a (N,θ,μ)-linear vector field $X \in \mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$ has the form

$$X(v) = \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>\theta N}} X^{\sigma,m}_{\sigma',n}(v) z^{\sigma'}_n \partial_{z^\sigma_m} \quad \text{where} \quad X^{\sigma,m}_{\sigma',n} \in \mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,\sigma m - \sigma' n)\,. \tag{3.11}$$

Remark 3.1. *By Definition 3.1 and* (3.1)*, the coefficients* $X^{\sigma,m}_{\sigma',n}(v)$ *in* (3.11) *do not depend on* $z_j, \bar z_j$ *with* $|j| \geq 6N^L$.

Here and in the following $\mathtt{s}(m) := \text{sign}(m)$.

Lemma 3.1. *Let* $X \in \mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$*. Then the coefficients in* (3.11) *satisfy*

$$X^{\sigma,m}_{\sigma',n} = 0 \qquad \text{if} \quad \sigma \mathtt{s}(m) = -\sigma' \mathtt{s}(n). \tag{3.12}$$

Proof. By (3.8) and $|\sigma m - \sigma' n| \overset{(3.12)}{=} |m| + |n| \overset{(3.10)}{\geq} 2\theta N \overset{(3.1)}{>} N$ we get $\mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,\sigma m - \sigma' n) = \emptyset$. ■

Lemma 3.2. *Let* $\mathfrak{m}_{k,i,\alpha,\beta}$ *be a scalar monomial (see* (2.16)*) such that*

$$\alpha + \beta =: \gamma \qquad \text{with} \sum_{l \in \mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l < 12N^L\,. \tag{3.13}$$

Then

$$\Pi_{N,\theta,\mu}\Big(\mathfrak{m}_{k,i,\alpha,\beta}\, z^{\sigma'}_n\, \partial_{z^\sigma_m}\Big) = \begin{cases} \Big(\Pi^{L,\sigma m - \sigma' n}_{N,\mu}(\mathfrak{m}_{k,i,\alpha,\beta})\Big)\, z^{\sigma'}_n\, \partial_{z^\sigma_m} & \text{if} \quad |m|,|n| > \theta N \\ 0 & \text{otherwise.} \end{cases}$$

Proof. It directly follows by (3.1)-(3.3), (3.6) and (3.10). ■

3.1 Töplitz vector fields

We define the subclass of (N,θ,μ)-linear vector fields which are *Töplitz*.

Definition 3.3. (Töplitz vector field) *A* (N,θ,μ)*-linear vector field* $X \in \mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$ *is* (N,θ,μ)-Töplitz *if the coefficients in* (3.11) *have the form*

$$X^{\sigma,m}_{\sigma',n} = X^\sigma_{\sigma'}\big(\mathtt{s}(m), \sigma m - \sigma' n\big) \quad \text{for some} \quad X^\sigma_{\sigma'}(\varsigma,h) \in \mathcal{A}^L_{s,r,\mathtt{a}}(N,\mu,h) \tag{3.14}$$

and $\varsigma \in \{+,-\}$, $h \in \mathbb{Z}$*. We denote by*

$$\mathcal{T}_{s,r,\mathtt{a}} := \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu) \subset \mathcal{L}_{s,r,\mathtt{a}}(N,\theta,\mu)$$

the space of the (N,θ,μ)*-Töplitz vector fields.*

Lemma 3.3. *Consider* $X, Y \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ *and* $W \in \mathcal{V}^L_{s,r,\mathtt{a}}(N,\mu_1)$ *with* $1 < \mu, \mu_1 < 6$*. For all* $0 < s' < s$, $0 < r' < r$ *and* $\theta' \geq \theta, \mu' \leq \mu$ *one has*

$$\Pi_{N,\theta',\mu'}[X,W] \in \mathcal{T}_{s',r',\mathtt{a}}(N,\theta',\mu')\,. \tag{3.15}$$

If moreover

$$\mu N^L + (\kappa+1)N^b < (\theta' - \theta)N \tag{3.16}$$

then

$$\Pi_{N,\theta',\mu'}[X,Y] \in \mathcal{T}_{s',r',\mathtt{a}}(N,\theta',\mu')\,. \tag{3.17}$$

Proof of (3.15). By definition (recall (3.10)) we have that $X^{(x)}, X^{(y)}$ and $X^{(z_m^\sigma)}$ vanish if $|m| \leq \theta N$. Arguing as in (3.7) we have that $W^{z_j^\sigma} = 0$ if $|j| \geq \mu_1 N^L + (\kappa+1)N^b$. Note that only the components $[X,W]^{(\mathtt{v})}$ with $\mathtt{v} = z_m^\sigma$ and $|m| > \theta N$ contribute to $\Pi_{N,\theta',\mu'}[X,W]$. Noting that $\theta N > \mu_1 N^L + (\kappa+1)N^b$ (by (3.1) and (3.3)) we have

$$[X,W]^{(z_m^\sigma)} = \partial_x X^{(z_m^\sigma)} W^{(x)} + \partial_y X^{(z_m^\sigma)} W^{(y)} + \sum_{\sigma_1, |j| < \mu_1 N^L + \kappa N^b} \partial_{z_j^{\sigma_1}} X^{(z_m^\sigma)} W^{(z_j^{\sigma_1})} . \tag{3.18}$$

By (3.11) and (3.14) we get $X^{(z_m^\sigma)} = \sum_{\sigma', |n| > \theta N} X_{\sigma'}^\sigma(\mathtt{s}(m), \sigma m - \sigma' n) z_n^{\sigma'}$. Let us consider the first term of the right hand side of (3.18). Since $X_{\sigma'}^\sigma(\mathtt{s}(m), \sigma m - \sigma' n)$, $W^{(x_l)} \in \mathcal{A}_{s,r,\mathtt{a}}^L(N,\mu)$ (recall (3.14)), all the monomials in $\partial_x X_{\sigma'}^\sigma(\mathtt{s}(m), \sigma m - \sigma' n) W^{(x)}$ satisfy (3.13). By Lemma 3.2 we have

$$\Pi_{N,\theta',\mu'} \Big(\partial_x X^{(z_m^\sigma)} W^{(x)} \partial_{z_m^\sigma} \Big) = \begin{cases} \sum_{\sigma', |n| > \theta' N} U_{\sigma',n}^{\sigma,m} \, z_n^{\sigma'} \, \partial_{z_m^\sigma}, & \text{if} \quad |m| > \theta' N \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{where} \qquad U_{\sigma',n}^{\sigma,m} := \Pi_{N,\mu'}^{L, \sigma m - \sigma' n} \Big(\partial_x X_{\sigma'}^\sigma(\mathtt{s}(m), \sigma m - \sigma' n) W^{(x)} \Big) .$$

It is immediate to see that $U_{\sigma',n}^{\sigma,m}$ satisfy (3.14). The other terms in (3.18) are analogous. (3.15) follows.

Proof of (3.17). We have by (2.22)

$$[X,Y] =: Z - Z' , \qquad \text{where} \quad Z := \sum_{\sigma, |m| > \theta N} \Big(\sum_{\sigma_1, |j| > \theta N} \partial_{z_j^{\sigma_1}} X^{(z_m^\sigma)} Y^{(z_j^{\sigma_1})} \Big) \partial_{z_m^\sigma} \tag{3.19}$$

and Z' is analogous exchanging the role of X and Y. We have to prove that $\Pi_{N,\theta',\mu'} Z \in \mathcal{T}_{s',r',\mathtt{a}}(N,\theta',\mu')$. By (3.11) and (3.14) we get

$$Z^{(z_m^\sigma)} = \sum_{\sigma_1, |j| > \theta N} \sum_{\sigma', |n| > \theta N} X_{\sigma_1}^\sigma(\mathtt{s}(m), \sigma m - \sigma_1 j) Y_{\sigma'}^{\sigma_1}(\mathtt{s}(j), \sigma_1 j - \sigma' n) z_n^{\sigma'} .$$

Since both $X_{\sigma_1}^\sigma(\mathtt{s}(m), \sigma m - \sigma_1 j)$ and $Y_{\sigma'}^{\sigma_1}(\mathtt{s}(j), \sigma_1 j - \sigma' n)$ belong to $\mathcal{A}_{s,r,\mathtt{a}}^L(N,\mu)$ (recall (3.14)), all the monomials in their product satisfy (3.13). By Lemma 3.2 we get

$$\Pi_{N,\theta',\mu'} Z = \sum_{\sigma,\sigma', \ |m|,|n| > \theta' N} Z_{\sigma',n}^{\sigma,m} \, z_n^{\sigma'} \, \partial_{z_m^\sigma}$$

where

$$Z_{\sigma',n}^{\sigma,m} := \Pi_{N,\mu'}^{L,\sigma m - \sigma' n} \Big(\sum_{\sigma_1, |j| > \theta N} X_{\sigma_1}^\sigma(\mathtt{s}(m), \sigma m - \sigma_1 j) Y_{\sigma'}^{\sigma_1}(\mathtt{s}(j), \sigma_1 j - \sigma' n) \Big) . \tag{3.20}$$

Note that $X^{\sigma,\sigma_1}(\mathtt{s}(m), \sigma m - \sigma_1 j) \in \mathcal{A}^L(N, \mu, \sigma m - \sigma_1 j)$, formula (3.7) and condition (3.16) imply that if $|m| > \theta' N$ then automatically $|j| > |m| - |\sigma m - \sigma_1 j| > \theta' N - \mu N^L - (\kappa+1)N^b > \theta N$ or $X^{\sigma,\sigma_1}(\mathtt{s}(m), \sigma m - \sigma_1 j) = 0$. Then the summation in (3.20) runs over $j \in \mathbb{Z}$. By (3.12) we have $\mathtt{s}(j) = \sigma\sigma_1 \mathtt{s}(m)$. Therefore

$$Z_{\sigma',n}^{\sigma,m} := \Pi_{N,\mu'}^{L,\sigma m - \sigma' n} \Big(\sum_{\sigma_1, h} X_{\sigma_1}^\sigma(\mathtt{s}(m), h) Y_{\sigma'}^{\sigma_1}(\sigma\sigma_1 \mathtt{s}(m), \sigma m - \sigma' n - h) \Big)$$

satisfying (3.14). ■

3.2 Quasi-Töplitz vector fields

Given a vector field $X \in \mathcal{V}_{s,r,\mathtt{a}}$ and a Töplitz vector $\tilde{X} \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ we define

$$\hat{X} := N(\Pi_{N,\theta,\mu}X - \tilde{X})\,. \tag{3.21}$$

Definition 3.4. (Quasi-Töplitz) *A vector field $X \in \mathcal{V}_{s,r,\mathtt{a}}$ is called (N_0,θ,μ)-quasi-Töplitz if the quasi-Töplitz norm*

$$\|X\|^T_{s,r,\mathtt{a}} := \|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} := \sup_{N\geq N_0}\Big[\inf_{\tilde{X}\in\mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)}\Big(\max\{\|X\|_{s,r,\mathtt{a}}, \|\tilde{X}\|_{s,r,\mathtt{a}}, \|\hat{X}\|_{s,r,\mathtt{a}}\}\Big)\Big] \tag{3.22}$$

is finite. We define

$$\mathcal{Q}^T_{s,r,\mathtt{a}} := \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu) := \Big\{X \in \mathcal{V}_{s,r,\mathtt{a}} \,:\, \|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} < \infty\Big\}\,.$$

In other words, a vector field X is (N_0,θ,μ)-quasi-Töplitz with norm $\|X\|^T_{s,r,\mathtt{a}}$ if, for all $N \geq N_0$, $\forall\varepsilon > 0$, there is $\tilde{X} \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ such that

$$\Pi_{N,\theta,\mu}X = \tilde{X} + N^{-1}\hat{X} \quad \text{and} \quad \|X\|_{s,r,\mathtt{a}}\,,\ \|\tilde{X}\|_{s,r,\mathtt{a}}\,,\ \|\hat{X}\|_{s,r,\mathtt{a}} \leq \|X\|^T_{s,r,\mathtt{a}} + \varepsilon\,. \tag{3.23}$$

We call $\tilde{X} \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ a "*Töplitz approximation*" of X and $\hat{X}$ the "*Töplitz-defect*". Note that, by Definition 3.3 and (3.21) $\Pi_{N,\theta,\mu}\tilde{X} = \tilde{X}$, $\Pi_{N,\theta,\mu}\hat{X} = \hat{X}$.

By the definition in (3.22) we get

$$\|X\|_{s,r,\mathtt{a}} \leq \|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{3.24}$$

If $s' \leq s, r' \leq r, \mathtt{a}' \leq \mathtt{a}, N_0' \geq N_0, \theta' \geq \theta, \mu' \leq \mu$ then $\mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu) \subseteq \mathcal{Q}^T_{s',r',\mathtt{a}'}(N_0',\theta',\mu')$ and

$$\|\cdot\|^T_{s',r',\mathtt{a}',N_0',\theta',\mu'} \leq \max\{s/s', (r/r')^2\}\|\cdot\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{3.25}$$

Lemma 3.4. (Projections 1) *Consider a subset of indices $I \subset \mathbb{I}\times\mathtt{V}$ (see (2.13), (2.5)) such that the projection (see (2.19)) maps*

$$\Pi_I : \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu) \to \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)\,, \quad \forall N \geq N_0\,. \tag{3.26}$$

Then $\Pi_I : \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu) \to \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)$ and

$$\|\Pi_I X\|^T_{s,r,\mathtt{a}} \leq \|X\|^T_{s,r,\mathtt{a}}\,, \quad \forall X \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)\,. \tag{3.27}$$

Moreover, if $X \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)$ satisfies $\Pi_I X = X$, then, $\forall N \geq N_0$, $\forall\varepsilon > 0$, there exists a decomposition $\Pi_{N,\theta,\mu}X = \tilde{X} + N^{-1}\hat{X}$ with a Töplitz approximation $\tilde{X} \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ satisfying $\Pi_I\tilde{X} = \tilde{X}$, $\Pi_I\hat{X} = \hat{X}$ and $\|\tilde{X}\|_{s,r,\mathtt{a}}, \|\hat{X}\|_{s,r,\mathtt{a}} < \|X\|^T_{s,r,\mathtt{a}} + \varepsilon$.

PROOF. By (3.23) and (2.20) (recall that $\Pi_{N,\theta,\mu}$ is a projection on an index subset, see Definition 3.2)

$$\Pi_{N,\theta,\mu}\Pi_I X = \Pi_I\Pi_{N,\theta,\mu}X = \Pi_I\tilde{X} + N^{-1}\Pi_I\hat{X}\,. \tag{3.28}$$

Assumption (3.26) implies that $\Pi_I\tilde{X} \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ and so $\Pi_I\tilde{X}$ is a Töplitz approximation for $\Pi_I X$. Hence (3.27) follows by

$$\|\Pi_I X\|_{s,r,\mathtt{a}}, \|\Pi_I\tilde{X}\|_{s,r,\mathtt{a}}, \|\Pi_I\hat{X}\|_{s,r,\mathtt{a}} \overset{\text{Lemma2.4},(3.23)}{<} \|X\|^T_{s,r,\mathtt{a}} + \varepsilon\,.$$

Now, if $\Pi_I X = X$, then (3.28) shows that $\Pi_I\tilde{X}$ (which satisfies $\Pi_I(\Pi_I\tilde{X}) = \Pi_I\tilde{X}$), is a Töplitz approximation for X. ■

Lemma 3.5. (Projections 2) *For all $l \in \mathbb{N}$, $K \in \mathbb{N}$, $N \geq N_0$, the projections (see* (2.46), (2.34), (2.35), (2.36)*) map*

$$\Pi^{(l)}, \Pi_{|k|<K}, \Pi_{|\pi|<K}, \Pi_{\rm diag} : \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu) \to \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)\,. \tag{3.29}$$

If $X \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)$ then, writing for brevity $\|\cdot\|^T := \|\cdot\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}$,

$$\|\Pi^{(l)}X\|^T\,,\ \|\Pi_{|k|<K}X\|^T\,,\ \|\Pi_{|\pi|<K}X\|^T\,, \|\Pi_{\rm diag}X\|^T \leq \|X\|^T\,, \tag{3.30}$$

$$\|X^{\leq 0}\|^T\,,\ \|X - X^{\leq 0}_{|k|<K}\|^T \leq \|X\|^T\,. \tag{3.31}$$

Moreover, $\forall\, 0 < s' < s$ and $\forall 0 < \mathtt{a}' < \mathtt{a}$:

$$\|\Pi_{|k|\geq K}X\|^T_{s',r,\mathtt{a},N_0,\theta,\mu} \leq e^{-K(s-s')}\frac{s}{s'}\|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} \tag{3.32}$$

$$\|\Pi_{|\pi|\geq K}X\|^T_{s,r,\mathtt{a}',N_0,\theta,\mu} \leq e^{-K(\mathtt{a}-\mathtt{a}')}\|X\|^T_{s,r,\mathtt{a}',N_0,\theta,\mu}\,. \tag{3.33}$$

PROOF. We prove (3.29) for $\Pi_{|\pi|<K}$, the others are analogous. Since $\tilde X \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ then

$$\tilde X(v) = \sum_{\substack{m,n,\sigma,\sigma'\\ |m|,|n|>\theta N}} \tilde X^{\sigma,m}_{\sigma',n}(v) z_n^{\sigma'}\partial_{z_m^\sigma}$$

for some $\tilde X^{\sigma,m}_{\sigma',n}$ satisfying (3.14). Then

$$\big(\Pi_{|\pi|<K}\tilde X\big)(v) = \sum_{\substack{m,n,\sigma,\sigma'\\ |m|,|n|>\theta N}} Y^{\sigma,m}_{\sigma',n}(v) z_n^{\sigma'}\partial_{z_m^\sigma} \quad \text{where} \quad Y^{\sigma,m}_{\sigma',n} := \Pi_{|\pi+\sigma' n-\sigma m|<K}\tilde X^{\sigma,m}_{\sigma',n}$$

(recall Definition 3.1). Therefore $Y^{\sigma,m}_{\sigma',n}$ satisfy (3.14) and $\Pi_{|\pi|<K}\tilde X \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$. The estimates (3.30)-(3.31) follow from (3.29) and Lemma 3.4 (in particular (3.27)). The bounds (3.32)-(3.33) follow by (2.37), (2.38) and similar arguments. ■

Lemma 3.6. *Assume that, $\forall N \geq N_* \geq N_0$, we have the decomposition*

$$Y = Y'_N + Y''_N \quad \text{with} \quad \|Y'_N\|^T_{s,r,\mathtt{a},N,\theta,\mu} \leq K_1\,, \quad N\|\Pi_{N,\theta,\mu}Y''_N\|_{s,r,\mathtt{a}} \leq K_2\,. \tag{3.34}$$

Then $\|Y\|^T_{s,r,\mathtt{a},N_*,\theta,\mu} \leq \max\{\|Y\|_{s,r,\mathtt{a}}, K_1+K_2\}$.

PROOF. By assumption, $\forall N \geq N_*$, we have $\|Y'_N\|^T_{s,r,\mathtt{a},N,\theta,\mu} \leq K_1$. Then, $\forall \varepsilon > 0$, there exist $\tilde Y'_N \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ and $\hat Y'_N$, such that

$$\Pi_{N,\theta,\mu}Y'_N = \tilde Y'_N + N^{-1}\hat Y'_N \quad \text{and} \quad \|\tilde Y'_N\|_{s,r,\mathtt{a}}, \|\hat Y'_N\|_{s,r,\mathtt{a}} \leq K_1+\varepsilon\,. \tag{3.35}$$

Therefore, $\forall N \geq N_*$,

$$\Pi_{N,\theta,\mu}Y = \tilde Y_N + N^{-1}\hat Y_N\,, \quad \tilde Y_N := \tilde Y'_N\,, \quad \hat Y_N := \hat Y'_N + N\Pi_{N,\theta,\mu}Y''_N$$

where $\tilde Y_N \in \mathcal{T}_{s,r,\mathtt{a}}(N,\theta,\mu)$ and

$$\|\tilde Y_N\|_{s,r,\mathtt{a}} = \|\tilde Y'_N\|_{s,r,\mathtt{a}} \stackrel{(3.35)}{\leq} K_1+\varepsilon, \tag{3.36}$$

$$\|\hat Y_N\|_{s,r,\mathtt{a}} \leq \|\hat Y'_N\|_{s,r,\mathtt{a}} + N\|\Pi_{N,\theta,\mu}Y''_N\|_{s,r,\mathtt{a}} \stackrel{(3.35),(3.34)}{\leq} K_1+\varepsilon+K_2\,. \tag{3.37}$$

Then $Y \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_*,\theta,\mu)$ and

$$\begin{aligned}\|Y\|^T_{s,r,\mathtt{a},N_*,\theta,\mu} &\leq \sup_{N\geq N_*}\max\big\{\|Y\|_{s,r,\mathtt{a}}, \|\tilde Y_N\|_{s,r,\mathtt{a}}, \|\hat Y_N\|_{s,r,\mathtt{a}}\big\} \\ &\stackrel{(3.36),(3.37)}{\leq} \max\{\|Y\|_{s,r,\mathtt{a}}, K_1+K_2+\varepsilon\}\,.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary the lemma follows. ■

Proposition 3.1. (Lie bracket) *Assume that* $X^{(1)}, X^{(2)} \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)$ *and* $N_1 \geq N_0$, $\mu_1 \leq \mu$, $\theta_1 \geq \theta$, $s/2 \leq s_1 < s$, $r/2 \leq r_1 < r$, $\mathtt{a}_1 < \mathtt{a}$ *satisfy*

$$(\kappa+1)N_1^{b-L} < \mu - \mu_1, \ \mu_1 N_1^{L-1} + (\kappa+1)N_1^{b-1} < \theta_1 - \theta \,, \tag{3.38}$$

$$2N_1 e^{-N_1^b \min\{\mathtt{a}-\mathtt{a}_1, s-s_1\}/2} < 1 \,, \ b\min\{\mathtt{a}-\mathtt{a}_1, s-s_1\}N_1^b > 2 \,. \tag{3.39}$$

Then $[X^{(1)}, X^{(2)}] \in \mathcal{Q}^T_{s_1,r_1,\mathtt{a}_1}(N_1,\theta_1,\mu_1)$ *and*

$$\|[X^{(1)}, X^{(2)}]\|^T_{s_1,r_1,\mathtt{a}_1,N_1,\theta_1,\mu_1} \leq C(n)\delta^{-1}\|X^{(1)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\|X^{(2)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} \tag{3.40}$$

where $C(n) \geq 1$ *and*

$$\delta := \min\left\{1 - \frac{s_1}{s}, 1 - \frac{r_1}{r}\right\}. \tag{3.41}$$

The proof is based on the following purely algebraic lemma.

Lemma 3.7. (Splitting lemma) *Let* $X^{(1)}, X^{(2)} \in \mathcal{V}_{s,r,\mathtt{a}}$ *and* (3.39) *hold. Then, for all* $N \geq N_1$,

$$\begin{aligned} &\Pi_{N,\theta_1,\mu_1}[X^{(1)}, X^{(2)}] = \\ &\Pi_{N,\theta_1,\mu_1}\Big(\Big[\Pi_{N,\theta,\mu}X^{(1)}, \Pi_{N,\theta,\mu}X^{(2)}\Big] + \Big[\Pi_{N,\theta,\mu}X^{(1)}, \Pi^L_{N,\mu}X^{(2)}\Big] + \Big[\Pi^L_{N,\mu}X^{(1)}, \Pi_{N,\theta,\mu}X^{(2)}\Big] \\ &\quad + \Big[\Pi_{|k|\geq N^b \text{ or } |\pi|\geq N^b}X^{(1)}, X^{(2)}\Big] + \Big[\Pi_{|k|,|\pi|<N^b}X^{(1)}, \Pi_{|k|\geq N^b \text{ or } |\pi|\geq N^b}X^{(2)}\Big]\Big)\,. \end{aligned} \tag{3.42}$$

Proof. We have

$$\begin{aligned} [X^{(1)}, X^{(2)}] &= [\Pi_{|k|,|\pi|<N^b}X^{(1)}, \Pi_{|k|,|\pi|<N^b}X^{(2)}] \\ &+ [\Pi_{|k|\geq N^b \text{ or } |\pi|\geq N^b}X^{(1)}, X^{(2)}] + [\Pi_{|k|,|\pi|<N^b}X^{(1)}, \Pi_{|k|\geq N^b \text{ or } |\pi|\geq N^b}X^{(2)}]\,. \end{aligned} \tag{3.43}$$

The last two terms correspond to the last line in (3.42). We now study the first term in the right hand side of (3.43). It is sufficient to study the case where $X^{(h)}$, $h = 1,2$, are monomial vector fields

$$\mathfrak{m}_h = m_{k^{(h)},i^{(h)},\alpha^{(h)},\beta^{(h)};\mathtt{v}^{(h)}} \ (\text{see}\,(2.17)) \quad \text{with} \quad |k^{(h)}|, |\pi(\mathfrak{m}_h)| < N^b\,,\ h = 1,2\,, \tag{3.44}$$

and analyze under which conditions the projection $\Pi_{N,\theta_1,\mu_1}[\mathfrak{m}_1,\mathfrak{m}_2]$ is not zero.

By the formula of the commutator (2.22) and the definition of the projection Π_{N,θ_1,μ_1} (see Definition 3.2, in particular (3.10)) we have to compute $(D_{\mathtt{v}}\mathfrak{m}_1^{\mathtt{v}'})[\mathfrak{m}_2^{\mathtt{v}}]$ only for $\mathtt{v}' = z_m^\sigma$ with $|m| > \theta_1 N$ and $\mathtt{v} \in \mathtt{V}$, see (2.5).

•) case 1: $\mathtt{v} = x_i$ or $\mathtt{v} = y_i$. By (3.10), in order to have a non trivial projection $\Pi_{N,\theta_1,\mu_1}(D_{\mathtt{v}}\mathfrak{m}_1^{z_m^\sigma})[\mathfrak{m}_2^{\mathtt{v}}]$ it must be

$$\alpha^{(1)} + \beta^{(1)} + \alpha^{(2)} + \beta^{(2)} = e_n + \gamma\,, \quad |n| > \theta_1 N\,, \ \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l < \mu_1 N^L\,. \tag{3.45}$$

We claim that

$$\alpha^{(1)} + \beta^{(1)} = e_n + \gamma^{(1)}\,, \quad \alpha^{(2)} + \beta^{(2)} = \gamma^{(2)}\,, \quad \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\gamma_l^{(h)} < \mu_1 N^L\,,\ h = 1,2\,, \tag{3.46}$$

which implies that $\mathfrak{m}_1$ is (N,θ_1,μ_1)–linear (see (3.10)), hence (N,θ,μ)–linear, and $\mathfrak{m}_2$ is (N,μ_1)–low (see (3.9)), hence (N,μ)–low. Thus $\Pi_{N,\theta,\mu}\mathfrak{m}_1 = \mathfrak{m}_1$ and $\Pi^L_{N,\mu}\mathfrak{m}_2 = \mathfrak{m}_2$ and we obtain the second (and third by commuting indices) term in the right hand side of (3.42). By (3.45), the other possibility instead of (3.46) is

$$\alpha^{(1)} + \beta^{(1)} = \tilde\gamma^{(1)}\,, \quad \alpha^{(2)} + \beta^{(2)} = e_n + \tilde\gamma^{(2)}\,, \quad \sum_{l\in\mathbb{Z}\setminus\mathcal{I}} |l|\tilde\gamma_l^{(h)} < \mu_1 N^L\,,\ h = 1,2\,. \tag{3.47}$$

In such a case, since $|\pi(\mathfrak{m}_2)| < N^b$ we get (recall $\mathfrak{m}_2 = \mathfrak{m}_2^{\mathtt{v}}$ with $\mathtt{v} = x, y$),

$$N^b > |\pi(k^{(2)}, \alpha^{(2)}, \beta^{(2)})| \overset{(2.24),(3.47)}{\geq} |n| - \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \tilde{\gamma}_l^{(2)} - \kappa |k^{(2)}| \overset{(3.45),(3.47),(3.44)}{\geq} \theta_1 N - \mu_1 N^L - \kappa N^b$$

which contradicts (3.1).

•) CASE 2: $\mathtt{v} = z_j^{\sigma_1}$, $j \in \mathbb{Z} \setminus \mathcal{I}$. Only for this case we use (3.39). In order to have a non trivial projection $\Pi_{N,\theta_1,\mu_1}(D_{\mathtt{v}} \mathfrak{m}_1^{z_m^\sigma})[\mathfrak{m}_2^{\mathtt{v}}]$ it must be

$$\alpha^{(1)} + \beta^{(1)} + \alpha^{(2)} + \beta^{(2)} - e_j = e_n + \gamma \,, \quad |n| > \theta_1 N \,, \quad \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \gamma_l < \mu_1 N^L \,. \tag{3.48}$$

We have the two following possible cases:

$$\alpha^{(1)} + \beta^{(1)} = e_j + e_n + \gamma^{(1)} \,, \quad \alpha^{(2)} + \beta^{(2)} = \gamma^{(2)} \,, \quad \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \gamma_l^{(h)} < \mu_1 N^L \,, \ h = 1, 2 \tag{3.49}$$

$$\alpha^{(1)} + \beta^{(1)} = e_j + \tilde{\gamma}^{(1)} \,, \quad \alpha^{(2)} + \beta^{(2)} = e_n + \tilde{\gamma}^{(2)} \,, \quad \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \tilde{\gamma}_l^{(h)} < \mu_1 N^L \,, \ h = 1, 2 \tag{3.50}$$

where $\gamma^{(1)} + \gamma^{(2)} = \tilde{\gamma}^{(1)} + \tilde{\gamma}^{(2)} = \gamma$. Note that, since we differentiate $\mathfrak{m}_1$ with respect to $\mathtt{v} = z_j^{\sigma_1}$ the monomial $\mathfrak{m}_1$ must depend on $z_j^{\sigma_1}$ and so the following case does not arise:

$$\alpha^{(1)} + \beta^{(1)} = \tilde{\gamma}^{(1)} \,, \quad \alpha^{(2)} + \beta^{(2)} = e_j + e_n + \tilde{\gamma}^{(2)} \,, \quad \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \tilde{\gamma}_l^{(h)} < \mu_1 N^L \,, \ h = 1, 2 \,.$$

In the case (3.49), the monomial $\mathfrak{m}_2$ is (N, μ)–low and we claim that $\mathfrak{m}_1$ is (N, θ, μ)–linear. Indeed, since

$$|\pi(\mathfrak{m}_2)| \overset{(2.24)}{=} |\pi(k^{(2)}, \alpha^{(2)}, \beta^{(2)}) - \sigma_1 j| < N^b \tag{3.51}$$

we get $|j| \leq |\pi(k^{(2)}, \alpha^{(2)}, \beta^{(2)})| + N^b$. Hence

$$\begin{aligned} |j| + \sum_l \gamma_l^{(1)} |l| &\leq |\pi(k^{(2)}, \alpha^{(2)}, \beta^{(2)})| + N^b + \sum_l \gamma_l^{(1)} |l| \leq \kappa |k^{(2)}| + \sum_l \gamma_l |l| + N^b \\ &\overset{(3.44),(3.48)}{\leq} (\kappa + 1) N^b + \mu_1 N^L \overset{(3.39)}{\leq} \mu N^L \end{aligned}$$

namely $\mathfrak{m}_1$ is (N, θ, μ)–linear (see (3.10) with $\gamma = e_j + \gamma^{(1)}$). Hence $\Pi_{N,\theta,\mu} \mathfrak{m}_1 = \mathfrak{m}_1$ and $\Pi^L_{N,\mu} \mathfrak{m}_2 = \mathfrak{m}_2$ and we obtain the second term (and third by commuting indices) in the right hand side of (3.42).

In the case (3.50) we claim that both $\mathfrak{m}_1, \mathfrak{m}_2$ are (N, θ, μ)–linear so we obtain the first term in the right hand side of (3.42). Since, by (3.48), $|n| > \theta_1 N > \theta N$ we already know that $\mathfrak{m}_2$ is (N, θ, μ)–linear. Finally, $\mathfrak{m}_1$ is (N, θ, μ)–linear because

$$\begin{aligned} |j| \overset{(3.51)}{>} |\pi(k^{(2)}, \alpha^{(2)}, \beta^{(2)})| - N^b &\overset{(2.24),(3.50)}{\geq} |n| - \sum_{l \in \mathbb{Z} \setminus \mathcal{I}} |l| \tilde{\gamma}_l^{(2)} - \kappa |k^{(2)}| - N^b \\ &\overset{(3.48),(3.50),(3.44)}{>} \theta_1 N - \mu_1 N^L - (\kappa + 1) N^b \overset{(3.39)}{>} \theta N \end{aligned}$$

concluding the proof. ■

PROOF OF PROPOSITION 3.1. Since $X^{(h)} \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$, $h = 1, 2$, for all $N \geq N_1 \geq N_0$ there exist $\tilde{X}^{(h)} \in \mathcal{T}_{s,r,\mathtt{a}}(N, \theta, \mu)$ and $\hat{X}^{(h)}$ such that

$$\Pi_{N,\theta,\mu} X^{(h)} = \tilde{X}^{(h)} + N^{-1} \hat{X}^{(h)} \,, \quad h = 1, 2 \,, \tag{3.52}$$

and

$$\|X^{(h)}\|_{s,r,\mathtt{a}},\ \|\tilde{X}^{(h)}\|_{s,r,\mathtt{a}},\ \|\hat{X}^{(h)}\|_{s,r,\mathtt{a}} < 2\|X^{(h)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{3.53}$$

In order to show that $[X^{(1)}, X^{(2)}] \in \mathcal{Q}^T_{s_1,r_1,\mathtt{a}_1}(N_1,\theta_1,\mu_1)$ and prove (3.40) we have to provide a decomposition

$$\Pi_{N,\theta_1,\mu_1}[X^{(1)}, X^{(2)}] = \tilde{X}^{(1,2)} + N^{-1}\hat{X}^{(1,2)}\,, \quad \forall N \geq N_1\,,$$

with $\tilde{X}^{(1,2)} \in \mathcal{T}_{s_1,r_1,\mathtt{a}_1}(N,\theta_1,\mu_1)$ and

$$\|[X^{(1)}, X^{(2)}]\|_{s_1,r_1,\mathtt{a}_1}, \|\tilde{X}^{(1,2)}\|_{s_1,r_1,\mathtt{a}_1}, \|\hat{X}^{(1,2)}\|_{s_1,r_1,\mathtt{a}_1} < C(n)\delta^{-1}\|X^{(1)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\|X^{(2)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} \tag{3.54}$$

with δ defined in (3.41). The bound (3.54) for the first term follows by (2.41) (as $\mathtt{a}_1 \leq \mathtt{a}$)

$$\|[X^{(1)}, X^{(2)}]\|_{s_1,r_1,\mathtt{a}_1} \leq 2^{2n+3}\delta^{-1}\|X^{(1)}\|_{s,r,\mathtt{a}}\|X^{(2)}\|_{s,r,\mathtt{a}}\,.$$

Considering (3.52) and the splitting (3.42), we define the candidate Töplitz approximation

$$\tilde{X}^{(1,2)} := \Pi_{N,\theta_1,\mu_1}\Big(\Big[\tilde{X}^{(1)}, \tilde{X}^{(2)}\Big] + \Big[\tilde{X}^{(1)}, \Pi^L_{N,\mu}X^{(2)}\Big] + \Big[\Pi^L_{N,\mu}X^{(1)}, \tilde{X}^{(2)}\Big]\Big) \tag{3.55}$$

and the Töplitz-defect

$$\hat{X}^{(1,2)} := N\Big(\Pi_{N,\theta_1,\mu_1}[X^{(1)}, X^{(2)}] - \tilde{X}^{(1,2)}\Big)\,. \tag{3.56}$$

Lemma 3.3 and (3.38) imply that $\tilde{X}^{(1,2)} \in \mathcal{T}_{s_1,r_1,\mathtt{a}}(N,\theta_1,\mu_1) \subset \mathcal{T}_{s_1,r_1,\mathtt{a}_1}(N,\theta_1,\mu_1)$. The estimate (3.54) for $\tilde{X}^{(1,2)}$ follows by (3.55), Lemma 2.4, (2.41), (3.53) and (3.24). Next by (3.42), (3.52), (3.55) the Töplitz defect (3.56) is

$$\begin{aligned}
\hat{X}^{(1,2)} \;&=\; \Pi_{N,\theta_1,\mu_1}\Big(\Big[\tilde{X}^{(1)}, \hat{X}^{(2)}\Big] + \Big[\hat{X}^{(1)}, \tilde{X}^{(2)}\Big] + N^{-1}\Big[\hat{X}^{(1)}, \hat{X}^{(2)}\Big] \\
&\quad + \Big[\hat{X}^{(1)}, \Pi^L_{N,\mu}X^{(2)}\Big] + \Big[\Pi^L_{N,\mu}X^{(1)}, \hat{X}^{(2)}\Big] \\
&\quad + N\Big[\Pi_{|k|\geq N^b \text{ or } |\pi|>N^b}X^{(1)}, X^{(2)}\Big] + N\Big[\Pi_{|k|,|\pi|<N^b}X^{(1)}, \Pi_{|k|\geq N^b \text{ or } |\pi|>N^b}X^{(2)}\Big]\Big)
\end{aligned}$$

and the bound (3.54) follows again by Lemma 2.4, (2.41), (3.24), (3.53), and, for the last two terms, also by (2.37) (2.38) and (3.39). Indeed, let us give the detailed estimate for the term

$$G := N\Big[\Pi_{|k|\geq N^b \text{ or} |\pi|\geq N^b}X^{(1)}, X^{(2)}\Big]\,.$$

We use Proposition 2.1 with $r' \rightsquigarrow r_1$, $r \rightsquigarrow r$, $\mathtt{a} \rightsquigarrow \mathtt{a}_1$, $s' \rightsquigarrow s_1$ and $s \rightsquigarrow s_1 + \sigma/2$, where $\sigma := s - s_1$. Since (recall (2.40))

$$\Big(1 - \frac{s_1}{s_1 + \sigma/2}\Big)^{-1} < 2\Big(1 - \frac{s_1}{s}\Big)^{-1} \leq 2\delta^{-1}$$

with the δ in (3.41), we get, using Lemma 2.5 with $\lambda = 0$,

$$\begin{aligned}
\|G\|_{s_1,r_1,\mathtt{a}_1} \;&\overset{(2.41)}{\leq}\; C(n)\delta^{-1}N\|\Pi_{|k|\geq N^b \text{ or } |\pi|\geq N^b}X^{(1)}\|_{s_1+\sigma/2,r,\mathtt{a}_1}\|X^{(2)}\|_{s,r,\mathtt{a}_1} \\
&\overset{(2.37),(2.38)}{\leq}\; C(n)\delta^{-1}N\frac{s}{s_1}e^{-N^b\min\{\mathtt{a}-\mathtt{a}_1,s-s_1\}/2}\|X^{(1)}\|_{s,r,\mathtt{a}}\|X^{(2)}\|_{s,r,\mathtt{a}} \\
&\overset{(3.39)}{\leq}\; C(n)\delta^{-1}\|X^{(1)}\|_{s,r,\mathtt{a}}\|X^{(2)}\|_{s,r,\mathtt{a}}\,, \qquad \forall\, N \geq N_1\,,
\end{aligned}$$

having used that the function $Ne^{-N^b\min\{\mathtt{a}-\mathtt{a}_1,s-s_1\}/2}$ is decreasing for $N \geq N_1$ by the second assumption in (3.39). The proof of (3.54) -and so of Proposition 3.1- is complete. ■

The quasi-Töplitz character of a vector field is preserved under the flow generated by a quasi-Töplitz vector field. The proof is based on an iteration of Proposition 3.1.

Proposition 3.2. (Lie transform) *Let* $X, Y \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$ *and let* $s/2 \le s' < s$, $r/2 \le r' < r$, $\mathtt{a}' < \mathtt{a}$. *There is* $c(n) > 0$ *such that, if*

$$\|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} \le c(n)\,\delta\,, \tag{3.57}$$

with δ *defined in* (2.40), *then the flow of* X *at time* $t = 1$ *maps* $D(s', r') \to D(s, r)$ *and, for*

$$N_0' \ge \max\{N_0, \bar N\}\,, \quad \bar N := \exp\Big(\max\Big\{\frac{2}{b}, \frac{1}{L-b}, \frac{1}{1-L}, 8\Big\}\Big)\,, \tag{3.58}$$

$\mu' < \mu$, $\theta' > \theta$, *satisfying*

$$(\kappa+1)(N_0')^{b-L}\ln N_0' \le \mu - \mu'\,, \quad (7+\kappa)(N_0')^{L-1}\ln N_0' \le \theta' - \theta\,, \tag{3.59}$$

$$2(N_0')^{-b}\ln^2 N_0' \le b\min\{s - s', \mathtt{a} - \mathtt{a}'\}\,, \tag{3.60}$$

we have $e^{\mathrm{ad}_X} Y \in \mathcal{Q}^T_{s',r',\mathtt{a}'}(N_0', \theta', \mu')$ *with*

$$\|e^{\mathrm{ad}_X} Y\|^T_{s',r',\mathtt{a}',N_0',\theta',\mu'} \le 2\|Y\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{3.61}$$

Moreover, for $h = 0, 1, 2$, *and coefficients* $0 \le b_j \le 1/j!$, $j \in \mathbb{N}$,

$$\Big\|\sum_{j\ge h} b_j \,\mathrm{ad}^j_X(Y)\Big\|^T_{s',r',\mathtt{a}',N_0',\theta',\mu'} \le 2\big(C(n)\delta^{-1}\|X\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\big)^h \|Y\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{3.62}$$

Note that (3.61) is (3.62) with $h = 0$, $b_j := 1/j!$

Proof. Let us prove (3.62). We define

$$Y^{(0)} := Y\,, \quad Y^{(j)} := \mathrm{ad}^j_X(Y) := \mathrm{ad}_X(Y^{(j-1)}) = [Y^{(j-1)}, X]\,, \ j \ge 1\,,$$

and we split, for $h = 0, 1, 2$,

$$Y^{\ge h} := \sum_{j\ge h} b_j Y^{(j)} = \sum_{j=h}^{J-1} b_j Y^{(j)} + \sum_{j \ge J} b_j Y^{(j)} =: Y^{\ge h}_{<J} + Y_{\ge J}\,. \tag{3.63}$$

By Proposition 2.1 (iterated j times) we get

$$\|Y^{(j)}\|_{s',r',\mathtt{a}'} \le \|Y^{(j)}\|_{s',r',\mathtt{a}} \le (2^{2n+3} j\delta^{-1})^j \|X\|^j_{s,r,\mathtt{a}} \|Y\|_{s,r,\mathtt{a}}\,, \quad \forall j \ge 0\,, \tag{3.64}$$

where δ is defined in (2.40). Indeed for $j \ge 0$ and $0 \le i \le j$ set

$$r_i := r - i\,\frac{r - r'}{j}, \qquad s_i := s - i\,\frac{s - s'}{j} \tag{3.65}$$

(note that $s_0 = s, s_j = s', r_0 = r, r_j = r'$) and we inductively have

$$\|Y^{(i+1)}\|_{s_{i+1},r_{i+1},\mathtt{a}} \overset{(2.41)}{\le} 2^{2n+3}\delta_i^{-1}\|X\|_{s_i,r_i,\mathtt{a}}\|Y^{(i)}\|_{s_i,r_i,\mathtt{a}} \tag{3.66}$$

where

$$\delta_i := \min\Big\{1 - \frac{s_{i+1}}{s_i}\,,\, 1 - \frac{r_{i+1}}{r_i}\Big\} \ge \frac{\delta}{j} \tag{3.67}$$

and δ is defined in (2.40). Let

$$\eta := \big(e^3 c(n)\delta\big)^{-1}\|X\|^T \overset{(3.57)}{\le} 1/e^3 \tag{3.68}$$

where, for brevity, $\|\cdot\|^T := \|\cdot\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}$. We have

$$\|Y_{\geq J}\|_{s',r',\mathtt{a}'} \overset{(3.64),(3.24)}{\leq} \sum_{j\geq J} b_j (2^{2n+3} j\delta^{-1}\|X\|^T_{s,r,\mathtt{a}})^j \|Y\|_{s,r,\mathtt{a}} \leq \sum_{j\geq J} \eta^j \|Y\|_{s,r,\mathtt{a}} \tag{3.69}$$

for $j^j b_j \leq j^j/j! < e^j$ and $e^4 2^{2n+3} c(n) \leq 1$. By (3.69) and $\eta < 1/2$ (see (3.68)) we deduce

$$\|Y_{\geq J}\|_{s',r',\mathtt{a}'} \leq 2\eta^J \|Y\|_{s,r,\mathtt{a}}\,. \tag{3.70}$$

In particular, for $J = h = 0,1,2$ (see (3.63)), we get

$$\|Y^{\geq h}\|_{s',r',\mathtt{a}'} \leq 2\eta^h \|Y\|_{s,r,\mathtt{a}}\,. \tag{3.71}$$

For any $N \geq N_0'$ we choose

$$J := J(N) := \ln N \quad \Longrightarrow \quad N\eta^{J-2} \overset{(3.68)}{\leq} e^6/N^2 \tag{3.72}$$

and we rename $Y^{\geq h}$ in (3.63) as

$$Y^{\geq h} = Y_N' + Y_N''\,, \quad Y_N' := Y^{\geq h}_{<J}\,, \quad Y_N'' := Y_{\geq J}\,.$$

Then (3.62) follows by Lemma 3.6 (with $N_* \rightsquigarrow N_0'$, $s \rightsquigarrow s'$, $\mathtt{a} \rightsquigarrow \mathtt{a}'$, $r \rightsquigarrow r'$, $\theta \rightsquigarrow \theta'$, $\mu \rightsquigarrow \mu'$, $Y \rightsquigarrow Y^{\geq h}$) and (3.71) (recall also (3.68)), once we show that

$$\|Y_N'\|^T_{s',r',\mathtt{a}',N,\theta',\mu'} \leq \frac{3}{2}\eta^h \|Y\|^T\,, \quad N\|Y_N''\|_{s',r',\mathtt{a}'} \leq \frac{1}{2}\eta^h\|Y\|^T\,, \quad h = 0,1,2\,. \tag{3.73}$$

The second inequality in (3.73) follows by

$$N\|Y_{\geq J}\|_{s',r',\mathtt{a}'} \overset{(3.70)}{\leq} N2\eta^J \|Y\|_{s,r,\mathtt{a}} \overset{(3.24)}{\leq} \eta^h (N2\eta^{J-h})\|Y\|^T \overset{(3.72)}{\leq} \frac{\eta^h}{2}\|Y\|^T \tag{3.74}$$

for all $N \geq N_0' \geq e^8$ (recall (3.58)). Let us prove the first inequality in (3.73).
Claim: *$\forall j = 0,\ldots,J-1$, we have $Y^{(j)} \in \mathcal{Q}^T_{s',r',\mathtt{a}'}(N,\theta',\mu')$ and*

$$\|Y^{(j)}\|^T_{s',r',\mathtt{a}',N,\theta',\mu'} \leq \big(C(n) j\delta^{-1}\|X\|^T\big)^j \|Y\|^T\,. \tag{3.75}$$

Then the first inequality in (3.73) follows by

$$\Big\|\sum_{j=h}^{J-1} b_j\, Y^{(j)}\Big\|^T_{s',r',\mathtt{a}',N,\theta',\mu'} \overset{(3.75)}{\leq} \sum_{j=h}^{J-1} b_j \big(C(n)j\delta^{-1}\|X\|^T\big)^j \|Y\|^T \overset{(3.68)}{\leq} \sum_{j=h}^{+\infty} \eta^j \|Y\|^T \leq \frac{3}{2}\eta^h \|Y\|^T$$

for $j^j b_j < j^j/j! < e^j$ and the constant $c(n)$ in (3.57) small enough.

Let us prove the claim. Fix $0 \leq j \leq J-1$. We define, $\forall i = 0,\ldots,j$,

$$\mathtt{a}_i := \mathtt{a} - i\,\frac{\mathtt{a}-\mathtt{a}'}{j}\,, \quad \theta_i := \theta + i\frac{\theta'-\theta}{j}\,, \quad \mu_i := \mu - i\,\frac{\mu-\mu'}{j} \tag{3.76}$$

(note that $\mathtt{a}_0 = \mathtt{a}, \mathtt{a}_j = \mathtt{a}'$, $\theta_0 = \theta, \theta_j = \theta'$ and $\mu_0 = \mu, \mu_j = \mu'$) and we prove inductively

$$\|Y^{(i)}\|^T_{s_i,r_i,\mathtt{a}_i,N,\theta_i,\mu_i} \leq \big(C(n)j\delta^{-1}\|X\|^T\big)^i \|Y\|^T\,, \quad \forall i = 0,\ldots,j\,. \tag{3.77}$$

Then (3.77) with $i = j$ gives (3.75) (recall also (3.65)).
Let us prove (3.77). For $i = 0$, formula (3.77) follows because $Y \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0,\theta,\mu)$ and (3.25). Now

we assume that (3.77) holds for $i < j$ and we prove it for $i+1$. We want to apply Proposition 3.1 to the functions f and with $X^{(1)} \rightsquigarrow X$, $X^{(2)} \rightsquigarrow Y^{(i)}$, $N_1 \rightsquigarrow N$, $s \rightsquigarrow s_i, s_1 \rightsquigarrow s_{i+1}$, $\theta \rightsquigarrow \theta_i, \theta_1 \rightsquigarrow \theta_{i+1}$, etc. We have to verify conditions (3.38)-(3.39) that read

$$
\begin{aligned}
&(\kappa+1)N^{b-L} < \mu_i - \mu_{i+1}\,, \quad \mu_{i+1}N^{L-1} + (\kappa+1)N^{b-1} < \theta_{i+1} - \theta_i\,, \\
&2Ne^{-N^b \min\{s_i - s_{i+1}, \mathtt{a}_i - \mathtt{a}_{i+1}\}/2} < 1\,, \quad b\min\{s_i - s_{i+1}, \mathtt{a}_i - \mathtt{a}_{i+1}\}N^b > 2\,.
\end{aligned} \tag{3.78}
$$

Since, by (3.65) and (3.76),

$$
\mu_i - \mu_{i+1} = \frac{\mu - \mu'}{j}\,, \quad \theta_{i+1} - \theta_i = \frac{\theta - \theta'}{j}\,, \quad s_i - s_{i+1} = \frac{s - s'}{j}\,, \quad \mathtt{a}_i - \mathtt{a}_{i+1} = \frac{\mathtt{a} - \mathtt{a}'}{j}
$$

and $j < J = \ln N$ (see (3.72)), $0 < b < L < 1$ (recall (3.2)), $\mu' < \mu \le 6$ (recall (3.1)), the conditions (3.78) are implied by

$$
\begin{aligned}
&(\kappa+1)N^{b-L}\ln N < \mu - \mu'\,, \quad (7+\kappa)N^{L-1}\ln N < \theta' - \theta\,, \\
&2Ne^{-N^b \min\{s - s', \mathtt{a} - \mathtt{a}'\}/(2\ln N)} < 1\,, \quad b\min\{s - s', \mathtt{a} - \mathtt{a}'\}N^b > 2\ln N\,.
\end{aligned} \tag{3.79}
$$

The last two conditions in (3.79) are implied by $b\min\{s-s', \mathtt{a}-\mathtt{a}'\}N^b > 2\ln^2 N$ and since $N \ge N_0' \ge e^{1/1-b}$ (recall (3.58)). Recollecting we have to verify, for all $N \ge N_0'$,

$$
(\kappa+1)N^{b-L}\ln N \le \mu - \mu'\,, \quad (7+\kappa)N^{L-1}\ln N \le \theta' - \theta\,, \quad 2N^{-b}\ln^2 N \le b\min\{s-s', \mathtt{a}-\mathtt{a}'\}\,. \tag{3.80}
$$

Since the function $N \mapsto N^{-\gamma}\ln N$ is decreasing for $N \ge N_0' \ge e^{1/\gamma}$, we have that (3.80) follows by (3.58)-(3.59). Therefore Proposition 3.1 implies that $Y^{(i+1)} \in \mathcal{Q}^T_{s_{i+1}, r_{i+1}, \mathtt{a}_{i+1}}(N, \theta_{i+1}, \mu_{i+1})$ and

$$
\begin{aligned}
\|Y^{(i+1)}\|^T_{s_{i+1}, r_{i+1}, \mathtt{a}_{i+1}, N, \theta_{i+1}, \mu_{i+1}} &\overset{(3.40)}{\le} C(n)\delta_i^{-1}\|X\|^T \|Y^{(i+1)}\|^T_{s_i, r_i, , \mathtt{a}_i, N, \theta_i, \mu_i} \\
&\overset{(3.67)}{\le} C(n) j\delta^{-1}\|X\|^T \|Y^{(i)}\|^T_{s_i, r_i, \mathtt{a}_i, N, \theta_i, \mu_i} \\
&\overset{(3.77)}{\le} (C(n) j\delta^{-1}\|X\|^T)^{i+1}\|Y\|^T
\end{aligned}
$$

proving (3.77) by induction. ■

For a vector field $X \in \mathcal{V}_{s,r,\mathtt{a}}$ depending on parameters $\xi \in \mathcal{O}$, see (2.32), we define the norm

$$
\|X\|^T_{\vec{p}} := \max\Big\{ \sup_{\xi \in \mathcal{O}} \|X(\cdot;\xi)\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,,\ \|X\|^\lambda_{s,r,\mathtt{a},\mathcal{O}} \Big\} \tag{3.81}
$$

where, for brevity,

$$
\vec{p} := (s, r, \mathtt{a}, N_0, \theta, \mu, \lambda, \mathcal{O})\,. \tag{3.82}
$$

We define

$$
\mathcal{Q}^T_{\vec{p}} := \Big\{ X \in \mathcal{V}^\lambda_{s,r,\mathtt{a},\mathcal{O}} \ :\ X(\cdot;\xi) \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)\,,\ \forall\, \xi \in \mathcal{O} \text{ and } \|X\|^T_{\vec{p}} < \infty \Big\}\,. \tag{3.83}
$$

Lemma 3.5 holds true also for the norm $\|\cdot\|^T_{\vec{p}}$. Moreover we have

Corollary 3.1. (Lie bracket) *Assume that $X^{(1)}, X^{(2)} \in \mathcal{Q}^T_{\vec{p}}$ (see* (3.82) *and* (3.83)*) and assume that $\vec{p}_1 := (s_1, r_1, \mathtt{a}_1, N_1, \theta_1, \mu_1, \lambda, \mathcal{O})$ satisfies the hypotheses of Proposition 3.1 (namely* (3.38) *and* (3.39)*). Then $[X^{(1)}, X^{(2)}] \in \mathcal{Q}^T_{\vec{p}_1}$ and*

$$
\|[X^{(1)}, X^{(2)}]\|^T_{\vec{p}_1} \le C(n)\delta^{-1}\|X^{(1)}\|^T_{\vec{p}}\|X^{(2)}\|^T_{\vec{p}} \tag{3.84}
$$

where $C(n) \ge 1$ and δ is defined in (3.41) *.*

Corollary 3.2. (Lie transform) *Let* $X, Y \in \mathcal{Q}^T_{\vec{p}}$ *(see* (3.82) *and* (3.83)*) with*

$$\|X\|^T_{\vec{p}} \le c(n)\,\delta\,, \tag{3.85}$$

with δ *defined in* (2.40) *and* $c(n)$ *in* (3.57)*. Assume that* $\vec{p}\,' := (s', r', \mathtt{a}', N_0', \theta', \mu', \lambda, \mathcal{O})$ *satisfies the hypotheses of Proposition 3.2 (in particular* (3.58), (3.59) *and* (3.60)*). Then* $e^{\mathrm{ad}_X} Y \in \mathcal{Q}^T_{\vec{p}\,'}$ *and*

$$\|e^{\mathrm{ad}_X} Y\|^T_{\vec{p}\,'} \le 2\|Y\|^T_{\vec{p}}\,. \tag{3.86}$$

Moreover, for $h = 0, 1, 2$*, and coefficients* $0 \le b_j \le 1/j!$*,* $j \in \mathbb{N}$*,*

$$\Big\| \sum_{j \ge h} b_j \,\mathrm{ad}^j_X(Y) \Big\|^T_{\vec{p}\,'} \le 2(C\delta^{-1}\|X\|^T_{\vec{p}})^h \|Y\|^T_{\vec{p}}\,. \tag{3.87}$$

4 An abstract KAM theorem

We consider a family of linear integrable vector fields with constant coefficients

$$\mathcal{N}(\xi) := \omega(\xi)\partial_x + \mathrm{i}\Omega(\xi) z \partial_z - \mathrm{i}\Omega(\xi)\bar{z}\partial_{\bar{z}} \tag{4.1}$$

defined on the phase space $\mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p}_{\mathcal{I}} \times \ell^{a,p}_{\mathcal{I}}$, where the tangential sites $\mathcal{I} \subset \mathbb{Z}$ are symmetric as in (1.32), the space $\ell^{a,p}_{\mathcal{I}}$ is defined in (2.1), the tangential frequencies $\omega \in \mathbb{R}^n$ and the normal frequencies $\Omega \in \mathbb{R}^{\mathbb{Z}\setminus\mathcal{I}}$ depend on real parameters

$$\xi \in \mathcal{O} \subset \mathbb{R}^{n/2}\,,$$

(where $n/2 =$ cardinality of $\mathcal{I}^+$, see (1.32)), and satisfy

$$\omega_j(\xi) = \omega_{-j}(\xi)\,,\ \forall j \in \mathcal{I}\,, \quad \Omega_j(\xi) = \Omega_{-j}(\xi)\,,\ \forall j \in \mathbb{Z}\setminus\mathcal{I}\,. \tag{4.2}$$

For each ξ there is an invariant n-torus $\mathcal{T}_0 = \mathbb{T}^n \times \{0\} \times \{0\} \times \{0\}$ with frequency $\omega(\xi)$. In its normal space, the origin $(z, \bar{z}) = 0$ is an elliptic fixed point with proper frequencies $\Omega(\xi)$. The aim is to prove the persistence of a large portion of this family of linearly stable tori under small analytic perturbations

$$\mathcal{P}(x, y, z, \bar{z}; \xi) = \mathcal{P}^{(x)}\partial_x + \mathcal{P}^{(y)}\partial_y + \mathcal{P}^{(z)}\partial_z + \mathcal{P}^{(\bar{z})}\partial_{\bar{z}}\,. \tag{4.3}$$

(**A1**) Parameter dependence. The map $\omega : \mathcal{O} \to \mathbb{R}^n$, $\xi \mapsto \omega(\xi)$, is Lipschitz continuous.

With in mind the application to DNLW we assume

(**A2**) Frequency asymptotics.

$$\Omega_j(\xi) = |j| + a(\xi) + \frac{b(\xi)}{|j|} + O(\frac{1}{j^2}) \quad \text{as} \quad |j| \to +\infty\,. \tag{4.4}$$

Moreover the map $(\Omega_j - |j|)_{j \in \mathbb{Z}\setminus\mathcal{I}} : \mathcal{O} \to \ell_\infty$ is Lipschitz continuous.

By (A1) and (A2), the Lipschitz semi-norms of the frequency maps satisfy, for some $1 \le M_0 < \infty$,

$$|\omega|^{\mathrm{lip}} + |\Omega|^{\mathrm{lip}}_\infty \le M_0 \qquad \text{where} \qquad |\Omega|^{\mathrm{lip}}_\infty := \sup_{\xi \ne \eta \in \mathcal{O}} \frac{|\Omega(\xi) - \Omega(\eta)|_\infty}{|\eta - \xi|} \tag{4.5}$$

and $|z|_\infty := \sup_{j \in \mathbb{Z}\setminus\mathcal{I}} |z_j| < +\infty$. Note that by the Kirszbraun theorem (see e.g. [26]) applied componentwise we can extend ω, Ω on the whole $\mathbb{R}^{n/2}$ preserving the bound (4.5).

(**A3**) Regularity. The perturbation vector field $\mathcal{P}$ in (4.3) maps

$$\mathcal{P} : D(s,r) \times \mathcal{O} \to \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_{\mathcal{I}} \times \ell^{a,p}_{\mathcal{I}}$$

and, for some $s, r > 0$. Moreover $\mathcal{P}$ is
- reversible, see Definition 2.7,
- real-coefficients, see Definition 2.8,
- real-on-real, see Definition 2.9,
- Even, see Definition 2.10.

Finally, in order to obtain the asymptotic expansion for the perturbed frequencies we also assume

(**A4**) Quasi-Töplitz. The perturbation vector field $\mathcal{P}$ is quasi-Töplitz, see Definition 3.4.

Recalling (4.3) and the notations in (2.50), (2.46), we define

$$\mathcal{P}^y(x)\partial_y := \Pi^{(-1)}\mathcal{P}^{(y)}\partial_y\,, \qquad \mathcal{P}_* := \mathcal{P} - \mathcal{P}^y(x)\partial_y \tag{4.6}$$

and we denote $\mathcal{P}^{(-1)}_*, \mathcal{P}^{(0)}_*$ are the terms of degree -1 and 0 respectively of $\mathcal{P}_*$, see (2.46). Let

$$\vec{\omega}(\xi) := (\omega_j(\xi))_{j \in \mathcal{I}^+} \in \mathbb{R}^{n/2}\,, \qquad \text{then} \quad \omega = (\vec{\omega}, \vec{\omega}) \;\; \text{by} \;\; (4.2)\,.$$

Theorem 4.1. **(KAM theorem)** *Fix $s, r, \mathtt{a} > 0$, $1 < \theta, \mu < 6$, $N_0 \geq \bar{N}$ (defined in (3.58)). Let $\gamma \in (0, \gamma_*)$, where $\gamma_* = \gamma_*(n, s, \mathtt{a}) < 1$ is a (small) constant. Let $\lambda := \gamma/M_0$ (see (4.5)) and $\vec{p} := (s, r, \mathtt{a}, N_0, \theta, \mu, \lambda, \mathcal{O})$. Suppose that the vector field $\mathcal{X} = \mathcal{N} + \mathcal{P}$ satisfies* (A1)-(A4). *If*

$$\gamma^{-1}\|\mathcal{P}_*\|^T_{\vec{p}} \leq 1 \quad \text{and} \quad \varepsilon := \max\Big\{\gamma^{-2/3}\|\mathcal{P}^y(x)\partial_y\|^{\lambda}_{s,r,\mathtt{a},\mathcal{O}}, \gamma^{-1}\|\mathcal{P}^{(-1)}_*\|^T_{\vec{p}}, \gamma^{-1}\|\mathcal{P}^{(0)}_*\|^T_{\vec{p}}\Big\} \tag{4.7}$$

is small enough, then

• **(Frequencies)** *There exist Lipschitz functions $\omega^\infty : \mathbb{R}^{n/2} \to \mathbb{R}^n$, $\Omega^\infty : \mathbb{R}^{n/2} \to \ell_\infty$, $a^\infty : \mathbb{R}^{n/2} \to \mathbb{R}$ such that*

$$|\omega^\infty - \omega| + \lambda|\omega^\infty - \omega|^{\rm lip}\,, \quad |\Omega^\infty - \Omega|_\infty + \lambda|\Omega^\infty - \Omega|^{\rm lip}_\infty \;\leq\; C\gamma\varepsilon\,, \;\; |a^\infty| \leq C\gamma\varepsilon\,, \tag{4.8}$$

$$\omega^\infty_j(\xi) = \omega^\infty_{-j}(\xi)\,, \;\; \forall j \in \mathcal{I}\,, \quad \Omega^\infty_j(\xi) = \Omega^\infty_{-j}(\xi)\,, \;\; \forall j \in \mathbb{Z} \setminus \mathcal{I}\,, \tag{4.9}$$

hence $\omega^\infty = (\vec{\omega}^\infty, \vec{\omega}^\infty)$, $\vec{\omega}^\infty := (\omega^\infty_j)_{j \in \mathcal{I}^+} \in \mathbb{R}^{n/2}$, and

$$\sup_{\xi \in \mathbb{R}^{n/2}} |\Omega^\infty_j(\xi) - \Omega_j(\xi) - a^\infty(\xi)| \leq \gamma^{2/3}\varepsilon\frac{C}{|j|}\,, \quad \forall |j| \geq C_\star\gamma^{-1/3}\,. \tag{4.10}$$

• **(KAM normal form)** *for every ξ belonging to*

$$\begin{aligned}
\mathcal{O}_\infty \;\; := \;\; \Big\{ & \xi \in \mathcal{O} \;:\; \forall h \in \mathbb{Z}^{n/2},\, i, j \in \mathbb{Z} \setminus \mathcal{I},\, p \in \mathbb{Z}\,, \\
& |\vec{\omega}^\infty(\xi) \cdot h + \Omega^\infty_j| \geq 2\gamma\langle h\rangle^{-\tau}\,, \quad |\vec{\omega}^\infty(\xi) \cdot h + \Omega^\infty_i(\xi) + \Omega^\infty_j(\xi)| \geq 2\gamma\langle h\rangle^{-\tau}\,, \\
& |\vec{\omega}^\infty(\xi) \cdot h - \Omega^\infty_i(\xi) + \Omega^\infty_j(\xi)| \geq 2\gamma\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0 \text{ or } i \neq \pm j\,, \\
& |\vec{\omega}^\infty(\xi) \cdot h + p| \geq 2\gamma^{2/3}\langle h\rangle^{-\tau}, \;\text{if}\; (h,p) \neq (0,0) \\
& |\vec{\omega}(\xi) \cdot h| \geq 2\gamma^{2/3}\langle h\rangle^{-n/2}, \; \forall\, 0 < |h| < \gamma^{-1/(7n)} \Big\}
\end{aligned} \tag{4.11}$$

there exists an even, analytic, close to the identity diffeomorphism

$$\Phi(\cdot;\xi) : D(s/4, r/4) \ni (x_\infty, y_\infty, z_\infty, \bar{z}_\infty) \mapsto (x, y, z, \bar{z}) \in D(s,r)\,, \tag{4.12}$$

(Lipschitz in ξ) such that the transformed vector field

$$\mathcal{X}_\infty = \mathcal{N}_\infty + \mathcal{P}_\infty := \Phi_\star(\cdot;\xi)\mathcal{X} = (D\Phi(\ ;\xi))^{-1}\mathcal{X}\circ\Phi(\ ;\xi) \quad has \quad \big(\mathcal{P}_\infty^{\leq 0}\big)_{|E} = 0\,, \tag{4.13}$$

see (2.47), (1.36). *Moreover $\mathcal{N}_\infty$ is a constant coefficients linear normal form vector field as* (4.1) *with frequencies $\omega^\infty(\xi)$, $\Omega^\infty(\xi)$, and $\mathcal{P}_\infty$ is reversible, real-coefficients, real-on-real, even. Finally $(\mathcal{X}_\infty)_{|E} = (\mathcal{S}\mathcal{X}_\infty)_{|E}$.*

For all the parameters $\xi \in \mathcal{O}_\infty$, we deduce by (4.13) (and (4.9)) that the torus

$$\{x_\infty \in \mathbb{T}^n\,, y_\infty = 0\,, z_\infty = \bar{z}_\infty = 0\} \cap E \tag{4.14}$$

is an $n/2$-dimensional invariant torus of $\mathcal{X}_\infty : E \to E$ which supports the quasi-periodic solutions $t \mapsto (\omega^\infty(\xi)t + x_*, 0, 0, 0)$ with frequency $\omega^\infty(\xi) = \big(\vec{\omega}^\infty(\xi), \vec{\omega}^\infty(\xi)\big)$ and initial datum $(x_*, 0, 0, 0) \in E$, namely $x_* = (\vec{x}_*, \vec{x}_*)$, $\vec{x}_* \in \mathbb{T}^{n/2}$. Then, by (4.12), we have that

Corollary 4.1. *For all $\xi \in \mathcal{O}_\infty$, the map $\mathbb{T}^{n/2} \ni \vec{x}_\infty \mapsto \Phi\big((\vec{x}_\infty, \vec{x}_\infty), 0, 0, 0; \xi\big) \in E$ defines an $n/2$-dimensional analytic invariant torus of the vector field $\mathcal{X} = \mathcal{N} + \mathcal{P}$. Such torus is linearly stable on E and, in particular, it has zero Lyapunov exponents on E.*

Proof. It remains to prove the stability conclusion. Linearizing the system $\dot{v} = (\mathcal{X}_\infty)_{|E}(v)$, $v \in E$, at the solution $t \mapsto (\omega^\infty(\xi)t, 0, 0, 0) \in E$ we get, since $(\mathcal{X}_\infty)_{|E} = (\mathcal{S}\mathcal{X}_\infty)_{|E}$,

$$\begin{cases} \dot{x} &= \mathcal{S}\mathcal{P}_\infty^{xy}(\omega^\infty(\xi)t)y + \mathcal{S}\mathcal{P}_\infty^{xz}(\omega^\infty(\xi)t)z + \mathcal{S}\mathcal{P}_\infty^{x\bar{z}}(\omega^\infty(\xi)t)\bar{z} \\ \dot{y} &= 0 \\ \dot{z} &= \mathrm{i}\Omega^\infty(\xi)z \\ \dot{\bar{z}} &= -\mathrm{i}\Omega^\infty(\xi)\bar{z}\,. \end{cases} \tag{4.15}$$

System (4.15) is conjugated to the constant coefficient system

$$\begin{cases} \dot{x} &= \langle \mathcal{S}\mathcal{P}_\infty^{xy}\rangle y \\ \dot{y} &= 0 \\ \dot{z} &= \mathrm{i}\Omega^\infty(\xi)z \\ \dot{\bar{z}} &= -\mathrm{i}\Omega^\infty(\xi)\bar{z} \end{cases} \tag{4.16}$$

via a change of variables close to the identity of the form

$$I + \Psi \quad \text{with} \quad \Psi = \begin{pmatrix} 0 & \Psi_1(\omega^\infty(\xi)t) & \Psi_2(\omega^\infty(\xi)t) & \Psi_3(\omega^\infty(\xi)t) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where the terms of Ψ satisfy the homological equations

$$\begin{cases} \omega^\infty(\xi)\cdot\partial_x\Psi_1(x) &= \mathcal{S}\mathcal{P}_\infty^{xy}(x) - \langle\mathcal{S}\mathcal{P}_\infty^{xy}\rangle \\ \omega^\infty(\xi)\cdot\partial_x\Psi_2(x) - \mathrm{i}\Omega^\infty(\xi)\Psi_2(x) &= \mathcal{S}\mathcal{P}_\infty^{xz}(x) \\ \omega^\infty(\xi)\cdot\partial_x\Psi_3(x) + \mathrm{i}\Omega^\infty(\xi)\Psi_3(x) &= \mathcal{S}\mathcal{P}_\infty^{x\bar{z}}(x)\,. \end{cases}$$

The last equations admit solutions (arguing as in the proof of Lemma 5.1) for all $\xi \in \mathcal{O}_\infty$ since we are considering a symmetric, reversible, real-coefficients, real-on-real, even vector field (recall hypothesis (5.27)). As a consequence of (4.16) the torus (4.14) is *linearly stable on E* and, in particular, it has *zero Lyapunov exponents on E*. ■

In the next Theorem 4.2 we verify the Melnikov non-resonance conditions thanks to the asymptotic decay (4.10) of the perturbed frequencies. As in [3]-[4], the Cantor set of "good" parameters $\mathcal{O}_\infty$ in (4.11) are expressed in terms of the final frequencies ω^∞, Ω^∞ (and of the initial tangential frequencies ω), and not inductively. This simplifies the measure estimates.

Theorem 4.2. (Measure estimate) *Let*

$$\mathcal{O} := \mathcal{O}_\rho := \Big\{ \xi := (\xi_j)_{j \in \mathcal{I}^+} \in \mathbb{R}^{n/2} \ : \ 0 < \frac{\rho}{2} \leq |\xi_j| \leq \rho \Big\} \tag{4.17}$$

and assume that

$$\vec{\omega}(\xi) = \bar{\omega} + A\xi\,, \ \bar{\omega} = (\lambda_j)_{j \in \mathcal{I}^+} \in \mathbb{R}^{n/2}\,, \quad \Omega_j(\xi) = \lambda_j + \lambda_j^{-1} \vec{a} \cdot \xi\,, \forall j \notin \mathcal{I}\,, \tag{4.18}$$

for some $A \in \mathrm{Mat}(n/2 \times n/2)$ *and* $\vec{a} \in \mathbb{R}^{n/2}$. *Assume also that*

$$A \ \textit{is invertible, and} \quad (\lambda_i^{-1} \pm \lambda_j^{-1})(A^T)^{-1}\vec{a}\,, \ \ \lambda_j^{-1}(A^T)^{-1}\vec{a} \notin \mathbb{Z}^{n/2} \setminus \{0\}\,, \ \forall i, j \in \mathbb{Z} \setminus \mathcal{I}\,. \tag{4.19}$$

Then the Cantor like set $\mathcal{O}_\infty$ *defined in* (4.11)*, with exponent*

$$\tau > \max\{n+3, 1/b\} \tag{4.20}$$

(b *is fixed in* (3.2)*), satisfies, for* $\rho \in (0, \rho_0(\mathrm{m}))$ *small,*

$$|\mathcal{O} \setminus \mathcal{O}_\infty| \leq C(\tau) \rho^{\frac{n}{2}-1} \gamma^{2/3}\,. \tag{4.21}$$

We assume that $\Omega_j(\xi)$ have the particular structure in (4.18) in view of the application to the DNLW; this implies the more general form (4.4). Moreover, since

$$|(A^T)^{-1}\vec{a}(\lambda_i^{-1} \pm \lambda_j^{-1})|, |(A^T)^{-1}\vec{a}\lambda_j^{-1}| \leq C\Big(\frac{1}{|i|} + \frac{1}{|j|}\Big) \to 0 \quad \text{for } |i|, |j| \to +\infty\,, \tag{4.22}$$

the hypothesis (4.19) amounts to finitely many conditions only. Theorem 4.2 is proved in section 6.4.

5 Homological equations

The integers $k \in \mathbb{Z}^n$ have indexes in $\mathcal{I}$ (see (1.32)), namely $k = (k_h)_{h \in \mathcal{I}}$.

Definition 5.1. (Normal form vector fields) *The normal form vector fields are*

$$\mathcal{N} := \partial_\omega + \mathtt{Nu}\partial_{\mathtt{u}} = \partial_\omega + \mathrm{i}\Omega z \partial_z - \mathrm{i}\Omega \bar{z} \partial_{\bar{z}} = \omega(\xi) \cdot \partial_x + \mathrm{i} \sum_{j \in \mathbb{Z} \setminus \mathcal{I}} \Omega_j(\xi) z_j \partial_{z_j} - \mathrm{i} \sum_{j \in \mathbb{Z} \setminus \mathcal{I}} \Omega_j(\xi) \bar{z}_j \partial_{\bar{z}_j} \tag{5.1}$$

where the frequencies $\omega_j(\xi), \Omega_j(\xi) \in \mathbb{R}$, $\forall\, \xi \in \mathcal{O} \subseteq \mathbb{R}^{n/2}$, *are real and symmetric Lipschitz functions*

$$\omega_{-j} = \omega_j\,, \ \forall j \in \mathcal{I}\,, \qquad \Omega_{-j} = \Omega_j\,, \ \forall j \in \mathbb{Z} \setminus \mathcal{I}\,, \tag{5.2}$$

the matrix $\mathtt{N}$ *is diagonal*

$$\mathtt{N} = \begin{pmatrix} \mathbf{0}_n & 0 & 0 \\ 0 & \mathrm{i}\Omega & 0 \\ 0 & 0 & -\mathrm{i}\Omega \end{pmatrix}\,, \quad \Omega := \mathrm{diag}_{j \in \mathbb{Z} \setminus \mathcal{I}}(\Omega_j)\,, \tag{5.3}$$

and there exists $j_* > 0$ *such that (recall* (4.4)*)*

$$\sup_{\xi \in \mathcal{O}} \Big| \Omega_j(\xi) - \Omega_j(\xi) - a(\xi) \Big| \lessdot \frac{\gamma}{|j|}\,, \quad \forall\, |j| \geq j_*\,, \tag{5.4}$$

(see (4.4)*) for some Lipschitz function* $a : \mathcal{O} \to \mathbb{R}$*, independent of* j*.*

Note that $\mathcal{N} \in \mathcal{R}^{\leq 0}_{rev}$, see Definition 2.11.

The symmetry condition (5.2) implies the resonance relations $\Omega_{-j} - \Omega_j = 0$ and $\omega \cdot k = 0$ for all

$$k \in \mathbb{Z}^n_{\mathrm{odd}} := \Big\{ k \in \mathbb{Z}^n \ : \ k_{-j} = -k_j\,, \ \forall j \in \mathcal{I} \Big\}\,. \tag{5.5}$$

As a consequence, along the KAM iteration there are "resonant" monomial vector fields of the perturbation which can not be averaged out and which are not in the new normal form vector field. However these further resonant monomials are naturally identified with monomials of the normal form, on the symmetric subspace E defined in (1.36), by identifying $x_{-j} = x_j$, $z_{-j} = z_j$, $\bar{z}_{-j} = \bar{z}_j$. The next section makes rigorous this procedure, by defining the "symmetrized" vector field $\mathcal{S}(X)$.

5.1 Symmetrization

For a vector field X, we define its "symmetrized" $\mathcal{S}(X) := \mathcal{S}X$ by linearity on the monomial vector fields:

Definition 5.2. *The symmetrized monomial vector fields are defined by*

$$\mathcal{S}(e^{\mathrm{i}k\cdot x}\partial_{x_j}) := \partial_{x_j}\,, \quad \mathcal{S}(e^{\mathrm{i}k\cdot x}y^i\partial_{y_j}) := y^i\partial_{y_j}\,,\ \forall k\in\mathbb{Z}^n_{\mathrm{odd}}\,,\ |i|=0,1,\ j\in\mathcal{I}\,, \tag{5.6}$$

$$\mathcal{S}(e^{\mathrm{i}k\cdot x}z_{\pm j}\partial_{z_j}) := z_j\partial_{z_j}\,, \quad \mathcal{S}(e^{\mathrm{i}k\cdot x}\bar z_{\pm j}\partial_{\bar z_j}) := \bar z_j\partial_{\bar z_j}\,,\ \forall k\in\mathbb{Z}^n_{\mathrm{odd}}\,,\ j\in\mathbb{Z}\setminus\mathcal{I}\,, \tag{5.7}$$

and $\mathcal{S}$ is the identity on the other monomial vector fields.

By (5.6)-(5.7) we write

$$\mathcal{S}X = X + X' + X'' \tag{5.8}$$

where

$$X' := \sum_{k\in\mathbb{Z}^n_{\mathrm{odd}},j\in\mathbb{Z}\setminus\mathcal{I}} X^{(z_j)}_{k,0,e_j,0}(1-e^{\mathrm{i}k\cdot x})z_j\partial_{z_j} + X^{(\bar z_j)}_{k,0,0,e_j}(1-e^{\mathrm{i}k\cdot x})\bar z_j\partial_{\bar z_j} \tag{5.9}$$

and

$$\begin{aligned} X'' := & \sum_{k\in\mathbb{Z}^n_{\mathrm{odd}},k\neq0,j\in\mathcal{I}} X^{(x_j)}_{k,0,0,0}(1-e^{\mathrm{i}k\cdot x})\partial_{x_j} + \sum_{k\in\mathbb{Z}^n_{\mathrm{odd}},k\neq0,j\in\mathcal{I},|i|=0,1} X^{(y_j)}_{k,i,0,0}(1-e^{\mathrm{i}k\cdot x})y^i\partial_{y_j} \\ & + \sum_{k\in\mathbb{Z}^n_{\mathrm{odd}},j\in\mathbb{Z}\setminus\mathcal{I}} X^{(z_j)}_{k,0,e_{-j},0}(z_j - e^{\mathrm{i}k\cdot x}z_{-j})\partial_{z_j} + X^{(\bar z_j)}_{k,0,0,e_{-j}}(\bar z_j - e^{\mathrm{i}k\cdot x}\bar z_{-j})\partial_{\bar z_j}\,. \end{aligned} \tag{5.10}$$

The "symmetric" subspace E defined in (1.36) is invariant under the flow evolution generated by the vector field X, because $X : E \to E$. Moreover the vector fields X and $\mathcal{S}(X)$ coincide on E:

Proposition 5.1. $X_{|E} = (\mathcal{S}X)_{|E}$.

Proof. By (5.9) and (5.10) since, if $(x, y, z, \bar z)\in E$ and $k\in\mathbb{Z}^n_{\mathrm{odd}}$ (see (5.5)), then $k\cdot x = 0$ and $z_{-j} = z_j$, $\bar z_{-j} = \bar z_j$. ■

Corollary 5.1. *$v(t)\in E$ is a solution of $\dot v = X(v)$ if and only if it is a solution of $\dot v = (\mathcal{S}X)(v)$.*

Hence, we may replace the vector field X with its symmetrized $\mathcal{S}(X)$ without changing the dynamics on the invariant subspace E. The following lemma shows that both the $\mathtt{a}$-weighted and Töplitz norms of the symmetrized vector field $\mathcal{S}(X)$ are controlled by those of X.

Proposition 5.2. *The norms*

$$\|\mathcal{S}X\|_{s,r,\mathtt{a}} \le \|X\|_{s,r,\mathtt{a}} \tag{5.11}$$

$$\|\mathcal{S}X\|^{\mathrm{lip}}_{s,r,\mathtt{a}} \le \|X\|^{\mathrm{lip}}_{s,r,\mathtt{a}} \tag{5.12}$$

$$\|\mathcal{S}X\|^{T}_{s,r,\mathtt{a},N_1,\theta,\mu} \le 9\|X\|^{T}_{s,r,\mathtt{a},N_1,\theta,\mu} \tag{5.13}$$

for $N_1 \ge N_0$ (defined in (3.1)) satisfying

$$N_1 e^{-N_1^b\min\{s,\mathtt{a}\}} \le 1\,, \quad bN_1^b\min\{s,\mathtt{a}\} \ge 1\,. \tag{5.14}$$

Moreover, if X is reversible, or real-coefficients, or real-on-real, or even, the same holds for $\mathcal{S}X$.

PROOF. In order to prove (5.11) we first note that the symmetrized monomial vector fields ∂_{x_h}, $y^i\partial_{x_h}$, $z_j\partial_{z_j}$, $\bar z_j\partial_{\bar z_j}$ in (5.6)-(5.7) have zero momentum and are independent of x. Hence their contribution to the weighted norm (2.27) is smaller or equal than the contribution of the (not yet symmetrized) monomials $e^{\mathrm{i}k\cdot x}\partial_{x_j}$, $e^{\mathrm{i}k\cdot x}y^i\partial_{x_j}$, $e^{\mathrm{i}k\cdot x}z_{\pm j}\partial_{z_j}$, $e^{\mathrm{i}k\cdot x}\bar z_{\pm j}\partial_{\bar z_j}$ of X. This proves (5.11).

PROOF OF (5.13). The estimate (5.13) follows by (5.8) and

$$\|X'\|^T_{s,r,\mathtt{a}} \le 6\|X\|^T_{s,r,\mathtt{a}} \tag{5.15}$$

$$\|X''\|^T_{s,r,\mathtt{a}} \le 2\|X\|^T_{s,r,\mathtt{a}}\,. \tag{5.16}$$

PROOF OF (5.15). We claim that, for $N \ge N_1$, the projection $\Pi_{N,\theta,\mu}X' = \tilde X' + N^{-1}\hat X'$ with

$$\tilde X' \in \mathcal{T}_{s,r,\mathtt{a}}\,, \quad \|\tilde X'\|_{s,r,\mathtt{a}} \le 6\|X\|^T_{s,r,\mathtt{a}}\,, \quad \|\hat X'\|_{s,r,\mathtt{a}} \le 5\|X\|^T_{s,r,\mathtt{a}}\,, \tag{5.17}$$

implying (5.15) (also because $\|X'\|_{s,r,\mathtt{a}} \le 2\|X\|_{s,r,\mathtt{a}}$). In order to prove (5.17) we write the (N,θ,μ)-projection as

$$\Pi_{N,\theta,\mu}X' = U + U^- + U_\perp + U_\perp^- \tag{5.18}$$

where

$$U := \sum_{\substack{k\in\mathcal{K}_N,\\ |j|>\theta N}} X^{(z_j)}_{k,0,e_j,0}(1-e^{\mathrm{i}k\cdot x})z_j\partial_{z_j}\,, \qquad U^- := \sum_{\substack{k\in\mathcal{K}_N,\\ |j|>\theta N}} X^{(\bar z_j)}_{k,0,0,e_j}(1-e^{\mathrm{i}k\cdot x})\bar z_j\partial_{\bar z_j}\,,$$

$$U_\perp := \sum_{|j|>\theta N}\Big(\sum_{k\in\mathbb{Z}^n_{\mathrm{odd}}\setminus\mathcal{K}_N} X^{(z_j)}_{k,0,e_j,0}\Big)z_j\partial_{z_j}\,, \quad U_\perp^- := \sum_{|j|>\theta N}\Big(\sum_{k\in\mathbb{Z}^n_{\mathrm{odd}}\setminus\mathcal{K}_N} X^{(\bar z_j)}_{k,0,0,e_j}\Big)\bar z_j\partial_{\bar z_j}\,,$$

and

$$\mathcal{K}_N := \Big\{k\in\mathbb{Z}^n_{\mathrm{odd}}\,,\ |\pi(k)|,|k|<N^b\Big\}, \quad \pi(k) := \sum_{j\in\mathcal{I}} jk_j\,.$$

Then (5.15) follows by Step 1) and 2) below.

STEP 1) *The projection* $\Pi_{N,\theta,\mu}(U+U^-) = (\tilde U + \tilde U^-) + N^{-1}(\hat U + \hat U^-)$ *with*

$$\tilde U\,, \tilde U^- \in \mathcal{T}_{s,r,\mathtt{a}}\,,\ \|\tilde U\|_{s,r,\mathtt{a}}, \|\tilde U^-\|_{s,r,\mathtt{a}} \le 6\|X\|^T_{s,r,\mathtt{a}}\,, \quad \|\hat U\|_{s,r,\mathtt{a}}, \|\hat U^-\|_{s,r,\mathtt{a}} \le 6\|X\|^T_{s,r,\mathtt{a}}\,. \tag{5.19}$$

Since X is quasi-Töplitz, Lemma 3.5 implies that the projection

$$\Pi_{\mathrm{diag}}\Pi^{(0)}X = \sum_{k\in\mathbb{Z}^n,\,j\in\mathbb{Z}\setminus\mathcal{I}} X^{(z_j)}_{k,0,e_j,0}e^{\mathrm{i}k\cdot x}z_j\partial_{z_j} + \sum_{k\in\mathbb{Z}^n,\,j\in\mathbb{Z}\setminus\mathcal{I}} X^{(\bar z_j)}_{k,0,0,e_j}e^{\mathrm{i}k\cdot x}\bar z_j\partial_{\bar z_j} =: W + W' \tag{5.20}$$

is quasi-Töplitz as well and ($\|\cdot\|^T_{s,r,\mathtt{a}}$ is short for $\|\cdot\|^T_{s,r,\mathtt{a},N_1,\theta,\mu}$)

$$\|W\|^T_{s,r,\mathtt{a}}, \|W'\|^T_{s,r,\mathtt{a}} \le \|\Pi_{\mathrm{diag}}\Pi^{(0)}X\|^T_{s,r,\mathtt{a}} \overset{(3.30)}{\le} \|X\|^T_{s,r,\mathtt{a}}\,.$$

By (3.29) we have $\Pi_{\mathrm{diag}}\Pi^{(0)}\mathcal{T}_{s,r,\mathtt{a}} \subset \mathcal{T}_{s,r,\mathtt{a}}$, hence Lemma 3.4 applied to W implies that for every $N \ge N_1$ there exist (N-dependent)

$$\tilde W = \sum_{\substack{|\pi(k)|,|k|<N^b,\\ |j|>\theta N}} \tilde W_k e^{\mathrm{i}k\cdot x}z_j\partial_{z_j}\,, \qquad \hat W = \sum_{\substack{|\pi(k)|,|k|<N^b,\\ |j|>\theta N}} \hat W_{k,j}e^{\mathrm{i}k\cdot x}z_j\partial_{z_j} \tag{5.21}$$

(note that $\tilde W$ is (N,θ,μ)-linear and Töplitz) with

$$\Pi_{N,\theta,\mu}W = \sum_{\substack{|\pi(k)|,|k|<N^b,\\ |j|>\theta N}} X^{(z_j)}_{k,0,e_j,0}e^{\mathrm{i}k\cdot x}z_j\partial_{z_j} = \tilde W + N^{-1}\hat W \tag{5.22}$$

and

$$\|\tilde W\|_{s,r,\mathtt a}\,,\ \|\hat W\|_{s,r,\mathtt a}\le \frac32\|W\|^T_{s,r,\mathtt a}\le\frac32\|X\|^T_{s,r,\mathtt a}\,.$$

By (5.18),(5.20),(5.21) and (5.22) we have

$$U=\sum_{\substack{k\in\mathcal K_N,\\ |j|>\theta N}}\tilde W_k(1-e^{\mathrm i k\cdot x})z_j\partial_{z_j}+N^{-1}\sum_{\substack{k\in\mathcal K_N,\\ |j|>\theta N}}\hat W_{k,j}(1-e^{\mathrm ik\cdot x})z_j\partial_{z_j}=:\tilde U+N^{-1}\hat U\,.$$

Note that $\tilde U$ is Töplitz. Moreover

$$\|\hat U\|_{s,r,\mathtt a}\overset{(2.27)}{\le}\sup_{\|z\|_{a,p}<r}\Big\|\Big(\sum_{k\in\mathcal K_N}2e^{\mathtt a|\pi(k)|}e^{s|k|}|\hat W_{k,j}||z_j|\Big)_{|j|>\theta N}\Big\|_{s,r}\overset{(5.21)}{\le}2\|\hat W\|_{s,r,\mathtt a}\le 3\|X\|^T_{s,r,\mathtt a}\,.$$

An analogous estimate holds true for $\tilde U$. A similar decomposition holds for U^- in (5.18).

STEP 2) $N\|U_\perp\|_{s,r,\mathtt a}\,,\ N\|U^-_\perp\|_{s,r,\mathtt a}\le\|X\|_{s,r,\mathtt a}\,.$

We have

$$\begin{aligned}\|U_\perp\|_{s,r,\mathtt a}&\overset{(2.27)}{=}\sup_{\|z\|_{a,p}<r}\Big\|\Big(\Big|\sum_{k\in\mathbb Z^n_{\rm odd}\setminus\mathcal K_N}X^{(z_j)}_{k,0,e_j,0}\Big||z_j|\Big)_{|j|>\theta N}\Big\|_{s,r}\\ &\le\sup_{\|z\|_{a,p}<r}\Big\|\Big(e^{-N^b\min\{s,\mathtt a\}}\sum_{|\pi(k)|\ or\ |k|\ge N^b}e^{\mathtt a|\pi(k)|+s|k|}|X^{(z_j)}_{k,0,e_j,0}||z_j|\Big)_{|j|>\theta N}\Big\|_{s,r}\\ &\le e^{-N^b\min\{s,\mathtt a\}}\|X\|_{s,r,\mathtt a}\overset{(5.14)}{\le}N^{-1}\|X\|_{s,r,\mathtt a}\end{aligned}$$

and similarly for $U_\perp$.

PROOF OF (5.16). The estimate (5.16) follows by

$$\|\Pi_{N,\theta,\mu}X''\|_{s,r,\mathtt a}\le 2N^{-1}\|X\|_{s,r,\mathtt a}\,,\quad\forall\,N\ge N_0\,.\tag{5.23}$$

In order to prove (5.23) we note that the momentum of $e^{\mathrm ik\cdot x}z_{-j}\partial_{z_j}$ with $|k|<N^b$, $|j|>\theta N$, $N\ge N_1\ge N_0$, satisfies

$$|\pi(k,e_{-j},0;z_j)|=\Big|\sum_{h\in\mathcal I}hk_h-2j\Big|\ge 2|j|-\kappa|k|\ge 2\theta N-\kappa N^b\overset{(3.1)}{>}N>N^b\tag{5.24}$$

(where $\kappa:=\max_{h\in\mathcal I}|h|$, recall (3.1)). Then by (5.10) and (3.10) the projection $\Pi_{N,\theta,\mu}X''=V+V'$ with

$$V:=\sum_{|j|>\theta N}\big(\sum_{k\in\mathcal K_N}X^{(z_j)}_{k,0,e_{-j},0}\big)z_j\partial_{z_j}\,,\qquad V':=\sum_{|j|>\theta N}\big(\sum_{k\in\mathcal K_N}X^{(\bar z_j)}_{k,0,0,e_{-j}}\big)\bar z_j\partial_{\bar z_j}\,.$$

We have

$$\begin{aligned}\|V\|_{s,r,\mathtt a}&\overset{(2.27)}{=}\sup_{\|z\|_{a,p}<r}\Big\|\Big(\Big|\sum_{k\in\mathcal K_N}X^{(z_j)}_{k,0,e_{-j},0}\Big||z_j|\Big)_{|j|>\theta N}\Big\|_{s,r}\\ &=\sup_{\|z\|_{a,p}<r}\Big\|\Big(\Big|\sum_{k\in\mathcal K_N}X^{(z_j)}_{k,0,e_{-j},0}\Big||z_{-j}|\Big)_{|j|>\theta N}\Big\|_{s,r}\\ &\overset{(5.24)}{\le}\sup_{\|z\|_{a,p}<r}\Big\|\Big(\sum_{k\in\mathcal K_N}e^{-\mathtt aN}e^{\mathtt a|\pi(k,e_{-j},0;z_j)|}|X^{(z_j)}_{k,0,e_{-j},0}||z_{-j}|\Big)_{|j|>\theta N}\Big\|_{s,r}\\ &\le e^{-\mathtt aN}\|X\|_{s,r,\mathtt a}\overset{(5.14)}{\le}N^{-1}\|X\|_{s,r,\mathtt a}\end{aligned}\tag{5.25}$$

where in (5.25) we have used that the domain $\{\|z\|_{a,p} < r\}$ is invariant under the map $z_j \mapsto z_{-j}$. Since a similar estimate holds for V', (5.23) follows.

We finally prove the last statement of the proposition. The vector field $\mathcal{S}X$ is even because $\mathcal{S}X_{|E} = X_{|E}$ (Proposition 5.1) and X is even. Since X is real-coefficients, Definition 5.2 immediately implies that $\mathcal{S}X$ is real-coefficients. Finally, since X are reversible and real-on-real, (2.53) and (2.57) enable to check that X', X'' in (5.9)-(5.10) are reversible and real-on-real, and so $\mathcal{S}X$ (see (5.8)). ■

Remark 5.1. *The assumptions $X \in \mathcal{R}_{rev}$, $Y \in \mathcal{R}_{a-rev}$, $X = \mathcal{S}X$, $Y = \mathcal{S}Y$ are not sufficient to imply $[X,Y] = \mathcal{S}[X,Y]$, as the example $X = \mathrm{i}(z_{-1}\partial_{z_2} + z_1\partial_{z_{-2}} - \bar z_1\partial_{\bar z_{-2}} - \bar z_{-1}\partial_{\bar z_2})$, $Y = z_2\partial_{z_1} + z_{-2}\partial_{z_{-1}} + \bar z_{-2}\partial_{\bar z_{-1}} + \bar z_2\partial_{\bar z_1}$ shows.*

5.2 Homological equations and quasi-Töplitz property

We consider the homological equation

$$\mathrm{ad}_{\mathcal{N}} F = R - [R] \tag{5.26}$$

where

$$R \in \mathcal{R}^{\leq 0}_{rev} \ (\text{see Definition 2.11}), \quad R = \mathcal{S}R \ (\text{see Definition 5.2}) \tag{5.27}$$

and

$$[R] := \langle R^x \rangle \partial_x + \sum_{j \in \mathbb{Z}\setminus\mathcal{I}} \langle R^{z_j z_j} \rangle z_j \partial_{z_j} + \langle R^{\bar z_j \bar z_j} \rangle \bar z_j \partial_{\bar z_j}\,, \tag{5.28}$$

where $\langle \cdot \rangle$ denotes the average with respect to the angles x. By Lemmata 2.8 and 2.10 and since $\mathcal{N} \in \mathcal{R}^{\leq 0}_{rev}$ (see Definition 5.1), the action

$$\mathrm{ad}_{\mathcal{N}} : \mathcal{R}^{\leq 0}_{a\text{-}rev} \to \mathcal{R}^{\leq 0}_{rev}\,.$$

The commutator

$$\mathrm{ad}_{\mathcal{N}} F = [F, \mathcal{N}] = \begin{cases} \Big(\partial_\omega F^{\mathtt{u}} - \mathtt{N} F^{\mathtt{u}}\Big)\partial_{\mathtt{u}} & \text{if } F = F^{(-1)} \\ \partial_\omega F^x \partial_x + \Big(\partial_\omega F^{\mathtt{u},\mathtt{u}} + [F^{\mathtt{u},\mathtt{u}}, \mathtt{N}]\Big)\mathtt{u}\partial_{\mathtt{u}} & \text{if } F = F^{(0)} \end{cases} \tag{5.29}$$

(recall the notations in (2.48)-(2.50)) where $[F^{\mathtt{u},\mathtt{u}}, \mathtt{N}] = F^{\mathtt{u},\mathtt{u}}\mathtt{N} - \mathtt{N}F^{\mathtt{u},\mathtt{u}}$ is the usual commutator between matrices (and $\mathtt{N}$ is defined in (5.3)). We solve (5.26) when

$$R = R^{(h)}_K := \Pi_{|k|<K}\Pi_{|\pi|<K} R^{(h)}\,, \quad h = 0, -1\,, \quad K \in \mathbb{N} \tag{5.30}$$

(recall the projections (2.34), (2.35) and (2.46)).

Definition 5.3. (Melnikov conditions) *Let $\gamma > 0$. We say that the frequencies $\omega(\xi) = (\vec\omega(\xi), \vec\omega(\xi))$, $\vec\omega \in \mathbb{R}^{n/2}$, $\Omega(\xi)$ satisfy the Melnikov conditions (up to $K > 0$) at $\xi \in \mathbb{R}^{n/2}$, if: $\forall\, h \in \mathbb{Z}^{n/2}$, $|h| < K$, $i, j \in \mathbb{Z}\setminus\mathcal{I}$,*

$$|\vec\omega(\xi)\cdot h| \geq \gamma\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0\,, \tag{5.31}$$

$$|\vec\omega(\xi)\cdot h + \Omega_j| \geq \gamma\langle h\rangle^{-\tau}\,, \tag{5.32}$$

$$|\vec\omega(\xi)\cdot h + \Omega_i(\xi) + \Omega_j(\xi)| \geq \gamma\langle h\rangle^{-\tau}\,, \tag{5.33}$$

$$|\vec\omega(\xi)\cdot h - \Omega_i(\xi) + \Omega_j(\xi)| \geq \gamma\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0 \quad \text{or} \quad i \neq \pm j\,, \tag{5.34}$$

where $\langle h\rangle := \max\{|h|, 1\}$ and

$$\tau > 1/b\,. \tag{5.35}$$

For $k \in \mathbb{Z}^n$ we set $k_\pm := (k_j)_{j\in\mathcal{I}^\pm} \in \mathbb{Z}^{n/2}$, namely $k = (k_+, k_-)$. Then

$$\omega \cdot k = \vec{\omega} \cdot h\,, \quad \text{with} \quad h := k_+ + k_- \in \mathbb{Z}^{n/2} \quad \text{and} \quad k \notin \mathbb{Z}^n_{\rm odd} \overset{(5.5)}{\Longrightarrow} h \neq 0\,. \tag{5.36}$$

Note that $|h| \le |k_+| + |k_-| = |k|$.

Lemma 5.1. (Solution of homological equations) *Let $s, r, \mathtt{a} > 0$, $K > 0$. Let $\mathcal{O} \subset \mathbb{R}^{n/2}$ and assume that the Melnikov conditions* (5.31)-(5.34) *are satisfied $\forall\, \xi \in \mathcal{O}$. Then, $\forall\, \xi \in \mathcal{O}$, the homological equation* (5.26) *with $R = R(\cdot;\xi)$ as in* (5.27),(5.30) *has a unique solution $F = F(\cdot;\xi)$*

$$F \in \mathcal{R}^{\le 0}_{a\text{-}rev}\,, \qquad F = \mathcal{S}F\,, \qquad F = \Pi_{|k|<K}\Pi_{|\pi|<K}F$$

with $\langle F^y\rangle = 0$, $\langle F^{y,y}\rangle = 0$, $\langle F^{z_i^\pm, z_i^\pm}\rangle = 0$. It satisfies

$$\|F\|_{s,r,\mathtt{a},\mathcal{O}} \le \gamma^{-1}K^\tau \|R\|_{s,r,\mathtt{a},\mathcal{O}} \tag{5.37}$$

$$\|F\|^{\rm lip}_{s,r,\mathtt{a},\mathcal{O}} \lessdot \gamma^{-1}K^\tau \|R\|^{\rm lip}_{s,r,\mathtt{a},\mathcal{O}} + \gamma^{-2}K^{2\tau+1}\big(|\omega|^{\rm lip}_{\mathcal{O}} + |\Omega|^{\rm lip}_{\mathcal{O}}\big)\|R\|_{s,r,\mathtt{a},\mathcal{O}}\,. \tag{5.38}$$

Proof. By (5.29) the homological equation (5.26) splits into

$$\partial_\omega F^{\mathtt{u}} - \mathtt{N}F^{\mathtt{u}} = R^{\mathtt{u}}\,, \quad \partial_\omega F^x = R^x - \langle R^x\rangle\,, \quad \partial_\omega F^{\mathtt{u},\mathtt{u}} + [F^{\mathtt{u},\mathtt{u}}, \mathtt{N}] = R^{\mathtt{u},\mathtt{u}} - [R]^{\mathtt{u},\mathtt{u}}\,. \tag{5.39}$$

Since $R = \mathcal{S}R$ (recall (5.27)), by (5.6) we get

$$R^x(x) = \langle R^x\rangle + \sum_{k\notin\mathbb{Z}^n_{\rm odd}} R^x_k e^{\mathrm{i}k\cdot x}\,, \quad \text{similarly for} \quad R^y(x)\,,\ R^{y,y}(x)\,. \tag{5.40}$$

Hence, by (5.31) and (5.36), the second equation in (5.39) has a solution[2]

$$F^x(x) = \sum_{k\notin\mathbb{Z}^n_{\rm odd}} F^x_k e^{\mathrm{i}k\cdot x}\,, \quad F^x_k := \frac{R^x_k}{\mathrm{i}\omega\cdot k}\,. \tag{5.41}$$

By (5.3) the first equation in (5.39) amounts to

$$\partial_\omega F^y = R^y\,, \quad \partial_\omega F^z - \mathrm{i}\Omega F^z = R^z\,, \quad \partial_\omega F^{\bar z} + \mathrm{i}\Omega F^{\bar z} = R^{\bar z}\,. \tag{5.42}$$

Since R is reversible and even

$$\langle R^y\rangle = R^y_0 \overset{(2.53)}{=} -R^{\hat y}_0 \overset{(2.58)}{=} -R^y_0 = -\langle R^y\rangle \tag{5.43}$$

and so the average

$$\langle R^y\rangle = 0\,. \tag{5.44}$$

By (5.31), (5.36), (5.40) and (5.44), the equation $\partial_\omega F^y = R^y$ admits a unique solution with $\langle F^y\rangle = 0$:

$$F^y = \sum_{k\notin\mathbb{Z}^n_{\rm odd}} F^y_k e^{\mathrm{i}k\cdot x}\,, \quad F^y_k := \frac{R^y_k}{\mathrm{i}\omega\cdot k}\,. \tag{5.45}$$

By the non-resonance assumption (5.32) and (5.36), the other two equations in (5.42) admit (unique) solutions. By (5.3), the third equation in (5.39) splits into

$$\partial_\omega F^{y,y} = R^{y,y}\,, \tag{5.46}$$

$$\partial_\omega F^{y,z} + \mathrm{i}F^{y,z}\Omega = R^{y,z} \tag{5.47}$$

[2] Note that $F^x(x)$ is unique because its average $\langle F^x\rangle = F^x_0 = 0$ by (5.53) and (5.54) below.

and the analogous equations for $F^{y,\bar{z}}$, $F^{z,y}$, $F^{\bar{z},y}$;

$$\partial_\omega F^{z,\bar{z}} - \mathrm{i}F^{z,\bar{z}}\Omega - \mathrm{i}\Omega F^{z,\bar{z}} = R^{z,\bar{z}} \tag{5.48}$$

and the analogous equation for $F^{\bar{z},z}$, and the most difficult equation

$$\partial_\omega F^{z,z} + \mathrm{i}F^{z,z}\Omega - \mathrm{i}\Omega F^{z,z} = R^{z,z} - [R]^{z,z} \tag{5.49}$$

plus the analogous for $F^{\bar{z},\bar{z}}$.

Since $R^{y,y}(x)y\partial_y$ is reversible and even we deduce as in (5.43) that $\langle R^{y,y}\rangle = 0$. Then, by (5.31),(5.36) and (5.40), the equation (5.46) admits a unique solution with $\langle F^{y,y}\rangle = 0$. Also the equations (5.47)-(5.48) have a unique solution thanks to (5.32)-(5.33) (and (5.36)), respectively.

We now consider equation (5.49) that amounts to

$$\partial_\omega F^{z_i z_j} + \mathrm{i}(\Omega_j - \Omega_i)F^{z_i z_j} = R^{z_i z_j} - [R]^{z_i z_j}\,, \quad \forall i,j \in \mathbb{Z}\setminus\mathcal{I}\,. \tag{5.50}$$

Developing in Fourier series $F^{z_i z_j}(x) = \sum_{k\in\mathbb{Z}^n} F_k^{z_i z_j} e^{\mathrm{i}k\cdot x}$, equation (5.50) becomes

$$\mathrm{i}(\omega\cdot k + \Omega_j - \Omega_i)F_k^{z_i z_j} = R_k^{z_i z_j} - [R]_k^{z_i z_j}\,. \tag{5.51}$$

If $i \neq \pm j$ then (5.51) is easily solved thanks to the non-resonance assumption (5.34) and (5.36). Since $R = \mathcal{S}R$, by (5.7)

$$\text{if } i = j \implies R_k^{z_i z_i} = 0\,, \forall\, k \in \mathbb{Z}^n_{\mathrm{odd}}\setminus\{0\}; \quad \text{if } i = -j\ ,\ (i\neq 0) \implies R_k^{z_i z_{-i}} = 0\,,\ \forall\, k\in\mathbb{Z}^n_{\mathrm{odd}}\,. \tag{5.52}$$

Then (5.51) is solved by (5.34) and (5.36).

The properties of anti-reversibility, anti-real-coefficients, real-on-real, and parity for the vector field solution F are easily verified. For example, let us consider $F^x(x)\partial_x$ in (5.41). It is anti-reversible because $R^x(x)\partial_x$ is reversible and so

$$F^{\hat{x}}_{-\hat{k}} \overset{(5.41)}{=} \frac{R^{\hat{x}}_{-\hat{k}}}{\mathrm{i}\omega\cdot(-\hat{k})} \overset{(2.53)}{=} \frac{R^x_k}{\mathrm{i}\omega\cdot(-\hat{k})} = -\frac{R^x_k}{\mathrm{i}\omega\cdot k} = -F^x_k\,, \tag{5.53}$$

which is the anti-reversibility property (2.55). Moreover $F^x(x)\partial_x$ is anti-real-coefficients since $R^x_k\in\mathbb{R}$ (by assumption R is real-coefficients). Next, (2.57) enables to check that $F^x(x)\partial_x$ is real-on-real. Finally, $F^x\partial_x$ is even because $R^x\partial_x$ is even and

$$F^{\hat{x}}_{\hat{k}} \overset{(5.41)}{=} \frac{R^{\hat{x}}_{\hat{k}}}{\mathrm{i}\omega\cdot\hat{k}} \overset{(2.58)}{=} \frac{R^x_k}{\mathrm{i}\omega\cdot\hat{k}} = \frac{R^x_k}{\mathrm{i}\omega\cdot k} = F^x_k \tag{5.54}$$

which is the parity property (2.58).

The estimate (5.37) directly follows by bounds on the small divisors in the Melnikov conditions (5.31)-(5.34) (and (5.36)) and the expression of F.

Let us prove the Lipschitz estimate (5.38) for $F^y\partial_y$ where F^y is defined in (5.45) (the other cases are analogous). For $\xi,\eta\in\mathcal{O}$, $\xi\neq\eta$, set $\Delta_{\xi,\eta}f := |\xi-\eta|^{-1}(f(\xi)-f(\eta))$ then

$$\Delta_{\xi,\eta}F^y_k = \frac{\Delta_{\xi,\eta}R^y_k}{\mathrm{i}\omega(\xi)\cdot k} + \frac{R^y_k(\eta)\Delta_{\xi,\eta}\omega\cdot k}{(\omega(\xi)\cdot k)(\omega(\eta)\cdot k)}\,. \tag{5.55}$$

By (5.55), (5.31) and (5.36), we deduce

$$\|F^y\partial_y\|^{\mathrm{lip}}_{s,r,\mathrm{a},\mathcal{O}} \le \gamma^{-1}K^\tau\|R^y\partial_y\|^{\mathrm{lip}}_{s,r,\mathrm{a},\mathcal{O}} + 2\gamma^{-2}K^{2\tau+1}|\omega|^{\mathrm{lip}}_{\mathcal{O}}\|R^y\partial_y\|_{s,r,\mathrm{a},\mathcal{O}}\,.$$

■

We use following lemma about the asymptotic expansion of the small divisors.

Lemma 5.2. *Assume* (5.4)*. Then, for all* $|i|, |j| \geq j_*$,

$$|\Delta_{k,j,i} - \tilde{\Delta}_{k,j,i}| \lessdot \frac{||i| - |j||}{|j||i|} + \gamma\Big(\frac{1}{|i|} + \frac{1}{|j|}\Big) + \frac{1}{|i|^2} + \frac{1}{|j|^2}\,, \tag{5.56}$$

where $\Delta_{k,j,i} := \omega \cdot k + \Omega_j - \Omega_i$ *and* $\tilde{\Delta}_{k,j,i} := \omega \cdot k + |j| - |i|$.

Proof. By (5.4) and (4.4), see also Lemma 5.1 of [4]. ■

The following key proposition proves that the solution of the homological equation is quasi-Töplitz.

Proposition 5.3. (Quasi-Töplitz) *Let the normal form* $\mathcal{N}$ *be as in Definition 5.1 and assume that* $R \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$*. Let* F *be the (unique) solution of the homological equation* (5.26) *found in Lemma 5.1, for all* $\xi \in \mathcal{O}$ *satisfying the Melnikov conditions* (5.31)-(5.34)*. If, in addition,*

$$|\vec{\omega}(\xi) \cdot h + p| \geq \gamma^{2/3} \langle h \rangle^{-\tau}\,, \quad \forall |h| \leq K,\ p \in \mathbb{Z}\,, \quad (h, p) \neq (0, 0)\,, \tag{5.57}$$

then $F = F(\cdot; \xi) \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0^*, \theta, \mu)$ *with*

$$N_0^* := \max\Big\{N_0\,, j_*,\ \hat{c}\gamma^{-1/3} K^{\tau+1}\Big\} \tag{5.58}$$

for a (suitably large) constant $\hat{c} := \hat{c}(\mathrm{m}, \kappa) \geq 1$*. Moreover*

$$\|F(\cdot; \xi)\|^T_{s,r,\mathtt{a},N_0^*,\theta,\mu} \leq 4\hat{c}\gamma^{-1} K^{2\tau} \|R(\cdot; \xi)\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,. \tag{5.59}$$

Proof. We focus on the solution $F = F_K^{(0)}$ of the homological equation with $R = R_K^{(0)}$ (recall (5.30)), and in particular on the most difficult estimate for the solution of (5.49). For brevity we set

$$\mathcal{R} := R^{z,z} - [R]^{z,z} = \sum_{(k,j,i) \in I_1} R_{k,j,i} e^{\mathrm{i}k \cdot x} z_j \partial_{z_i}\,, \qquad R_{k,j,i} := R_k^{z_i z_j}\,,$$

where (recall (5.52) and (5.30))

$$I_1 := \Big\{ |k|\,, |\pi(k) + j - i| < K\,, \quad j \neq \pm i \text{ if } k \in \mathbb{Z}^n_{\rm odd} \Big\}\,. \tag{5.60}$$

Then the solution of (5.49) is (recall (5.51))

$$\mathcal{F} := F^{z,z} := \sum_{(k,j,i) \in I_1} \frac{R_{k,j,i}}{\mathrm{i}\Delta_{k,j,i}} e^{\mathrm{i}k \cdot x} z_j \partial_{z_i}\,, \qquad \Delta_{k,j,i} := \omega(\xi) \cdot k + \Omega_j(\xi) - \Omega_i(\xi)\,. \tag{5.61}$$

In order to prove the estimate (5.59) on the Töplitz norm we compute $\Pi_{N,\theta,\mu}\mathcal{F}$. Then we have to consider only $|i|, |j| > \theta N \geq \theta N_0^* > j_*$ (since $\theta > 1$ by (3.1)) and, since $|\pi(k) + j - i| < K$ (see (5.60)), the bound (5.58) (and (5.35)) implies that $\mathtt{s}(i) = \mathtt{s}(j)$. Then

$$\Pi_{N,\theta,\mu}\mathcal{R} = \sum_{(k,j,i) \in I_2} R_{k,j,i} e^{\mathrm{i}k \cdot x} z_j \partial_{z_i}$$

where

$$I_2 := \Big\{ |k|\,, |\pi(k) + j - i| < K\,, \quad j \neq i \text{ if } k \in \mathbb{Z}^n_{\rm odd}\,, \quad \mathtt{s}(i) = \mathtt{s}(j)\,, \quad |i|, |j| > \theta N \Big\}\,.$$

By assumption $\mathcal{R} \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$. Condition (3.26) holds for the set I_2 and Lemma 3.4 (applied to $\Pi_{N,\theta,\mu}\mathcal{R}$) that, for all $N \geq N_0^* \geq N_0$, there exists a Töplitz approximation

$$\tilde{\mathcal{R}} := \sum_{I_2} \tilde{R}_k(\mathtt{s}(j), j - i) e^{\mathrm{i}k \cdot x} z_j \partial_{z_i} \in \mathcal{T}_{s,r,\mathtt{a}}(N, \theta, \mu)\,, \quad \tilde{R}_k(\mathtt{s}(j), j - i) \in \mathbb{C} \tag{5.62}$$

and

$$\Pi_{N,\theta,\mu}\mathcal{R} = \tilde{\mathcal{R}} + N^{-1}\hat{\mathcal{R}} \qquad \text{with} \qquad \|\mathcal{R}\|_{s,r,\mathtt{a}}, \|\tilde{\mathcal{R}}\|_{s,r,\mathtt{a}}, \|\hat{\mathcal{R}}\|_{s,r,\mathtt{a}} \le 2\|\mathcal{R}\|^T_{s,r,\mathtt{a}}\,. \tag{5.63}$$

Note that $|k|, |\pi(k) + j - i| < K < (N_0^*)^b \le N^b$ by (5.58) and (5.35). By (5.61) we have

$$\Pi_{N,\theta,\mu}\mathcal{F} = -\mathrm{i}\sum_{I_2} \frac{R_{k,j,i}}{\Delta_{k,j,i}} e^{\mathrm{i}k\cdot x} z_j \partial_{z_i}\,. \tag{5.64}$$

We now prove that

$$\tilde{\mathcal{F}} := -\mathrm{i}\sum_{I_2} \frac{\tilde{R}_k(\mathtt{s}(j), j-i)}{\tilde{\Delta}_{k,j,i}} e^{\mathrm{i}k\cdot x} z_j \partial_{z_i}\,, \quad \tilde{\Delta}_{k,j,i} := \omega(\xi)\cdot k + |j| - |i|\,, \tag{5.65}$$

is a Töplitz approximation of $\mathcal{F}$. Indeed, since $\mathtt{s}(j) = \mathtt{s}(i)$:

$$\tilde{\Delta}_{k,j,i} = \omega(\xi)\cdot k + |j| - |i| = \omega(\xi)\cdot k + \mathtt{s}(j)(j-i)\,,$$

and $\tilde{\mathcal{F}}$ in (5.65) is (N,θ,μ)-Töplitz (see (3.14)). Moreover, since $s(j)(j-i) \in \mathbb{Z}$ and $(0,j,j) \notin I_2$, by (5.57) (and (5.36)), we get

$$|\tilde{\Delta}_{k,j,i}| \ge \gamma^{2/3}\langle k\rangle^{-\tau}\,, \quad \forall (k,j,i) \in I_2\,, \tag{5.66}$$

and Lemma 2.6 and (5.65) imply

$$\|\tilde{\mathcal{F}}\|_{s,r,\mathtt{a}} \le \gamma^{-2/3}K^\tau \|\tilde{\mathcal{R}}\|_{s,r,\mathtt{a}}\,. \tag{5.67}$$

The Töplitz defect is

$$\begin{aligned} N^{-1}\hat{\mathcal{F}} \quad &:= \quad \Pi_{N,\theta,\mu}\mathcal{F} - \tilde{\mathcal{F}} \\ &\overset{(5.64),(5.65)}{=} \sum_{I_2}\Big(\frac{R_{k,j,i}}{\Delta_{k,j,i}} - \frac{\tilde{R}_k(\mathtt{s}(j), j-i)}{\tilde{\Delta}_{k,j,i}}\Big) e^{\mathrm{i}k\cdot x} z_j \partial_{z_i} \\ &= \sum_{I_2}\Big[\Big(\frac{R_{k,j,i}}{\Delta_{k,j,i}} - \frac{R_{k,j,i}}{\tilde{\Delta}_{k,j,i}}\Big) + \Big(\frac{R_{k,j,i} - \tilde{R}_k(\mathtt{s}(j), j-i)}{\tilde{\Delta}_{k,j,i}}\Big)\Big] e^{\mathrm{i}k\cdot x} z_j \partial_{z_i} \\ &\overset{(5.63)}{=} \sum_{I_2}\Big[R_{k,j,i}\Big(\frac{\tilde{\Delta}_{k,j,i} - \Delta_{k,j,i}}{\Delta_{k,j,i}\tilde{\Delta}_{k,j,i}}\Big) + N^{-1}\frac{\hat{R}_{k,j,i}}{\tilde{\Delta}_{k,j,i}}\Big] e^{\mathrm{i}k\cdot x} z_j \partial_{z_i}\,. \end{aligned} \tag{5.68}$$

By (5.56), $|j|, |i| \ge \theta N \ge N$, and $|j-i| \le (\kappa+1)K$ we get, taking $\hat{c}$ large enough,

$$|\tilde{\Delta}_{k,j,i} - \Delta_{k,j,i}| \le \frac{C(\kappa+1)K}{N^2} + C\frac{\gamma}{N} + \frac{C}{N^2} \le \frac{\hat{c}}{4N}\left(\frac{K}{N} + \gamma\right) \overset{(5.58)}{\le} \min\left\{\frac{\hat{c}\gamma^{1/3}}{2N}\,, \ \frac{\gamma^{2/3}}{2K^\tau}\right\}\,. \tag{5.69}$$

Hence, for $\langle k\rangle \le K$ we have

$$|\Delta_{k,j,i}| \ge |\tilde{\Delta}_{k,j,i}| - |\tilde{\Delta}_{k,j,i} - \Delta_{k,j,i}| \overset{(5.66),(5.69)}{\ge} \frac{\gamma^{2/3}}{\langle k\rangle^\tau} - \frac{\gamma^{2/3}}{2K^\tau} \ge \frac{\gamma^{2/3}}{2\langle k\rangle^\tau}\,. \tag{5.70}$$

Therefore (5.69), (5.66), (5.70) imply

$$\frac{|\tilde{\Delta}_{k,j,i} - \Delta_{k,j,i}|}{|\Delta_{k,j,i}||\tilde{\Delta}_{k,j,i}|} \le \frac{\hat{c}\gamma^{1/3}}{2N}\frac{2\langle k\rangle^\tau}{\gamma^{2/3}}\frac{\langle k\rangle^\tau}{\gamma^{2/3}} \le \frac{\hat{c}}{N\gamma}K^{2\tau}$$

and (5.68), (5.66), and Lemma 2.6, imply

$$\|\hat{\mathcal{F}}\|_{s,r,\mathtt{a}} \le \hat{c}\gamma^{-1}K^{2\tau}\|\mathcal{R}\|_{s,r,\mathtt{a}} + \gamma^{-2/3}K^\tau\|\hat{\mathcal{R}}\|_{s,r,\mathtt{a}} \overset{(5.63)}{\le} 4\hat{c}\gamma^{-1}K^{2\tau}\|\mathcal{R}\|^T_{s,r,\mathtt{a}}\,. \tag{5.71}$$

In conclusion (5.37), (5.67), (5.71), (5.63) prove (5.59) for $\mathcal{F}$ (see (5.61)).

The case of the solution $F^{z,\bar{z}}$ of the homological equation (5.48) is simpler. Indeed in this case the small divisors are $\omega \cdot k - \Omega_i - \Omega_j$. The $\Pi_{N,\theta,\mu}$-projection selects only the indices with $|i|, |j| > \theta N$ and so

$$|\omega \cdot k - \Omega_i - \Omega_j| \geq \Omega_i + \Omega_j - |\omega||k| \overset{(5.4),(5.58)}{\geq} const N \,. \tag{5.72}$$

In this case we may take the null vector field as Töplitz approximation of $F^{z,\bar{z}}$. Then the Töplitz defect is exactly $F^{z,\bar{z}}$ and $\|F^{z,\bar{z}}\|_{s,r,\mathtt{a}} \leq const \|R^{z,\bar{z}}\|_{s,r,\mathtt{a}}/N$ by (5.72).

The solutions of the other homological equations (5.42), (5.46), (5.47) are similarly estimated. ■

6 Proof of Theorems 4.1 and 4.2

6.1 First step

We perform a preliminary change of variables in order to improve the smallness conditions of the perturbation. In particular we want to average out the term $\mathcal{P}^y(x)\partial_y$ defined in (4.6). We introduce the simmetrized vector fields (see Definition 5.2)

$$R^y(x)\partial_y := \mathcal{S}\mathcal{P}^y(x)\partial_y \,, \qquad R := \mathcal{S}\mathcal{P} \,, \qquad X := \mathcal{S}\mathcal{X} = \mathcal{N} + R \tag{6.1}$$

(since $\mathcal{S}\mathcal{N} = \mathcal{N}$). By assumption (A3) and the last statement of Proposition 5.2, $R \in \mathcal{R}_{rev}$ (see Definition 2.11). Moreover Proposition 5.1 implies that $X_{|E} = \mathcal{X}_{|E}$.

We now study the homological equation

$$-\,\mathrm{ad}_{\mathcal{N}} F + \Pi_{|k|<\gamma^{-1/(7n)}} R^y \partial_y = \langle R^y \rangle \partial_y \overset{(5.44)}{=} 0 \tag{6.2}$$

because R is reversible and even.

Lemma 6.1. *For all ξ in*

$$\mathcal{O}_* := \Big\{ \xi \in \mathcal{O} \ : \ |\vec{\omega}(\xi) \cdot h| \geq \gamma^{2/3} \langle h \rangle^{-n/2} \,, \forall 0 < |h| < \gamma^{-1/(7n)} \Big\} \,, \tag{6.3}$$

the homological equation (6.2) *admits the unique solution*

$$F = F^y(x)\partial_y \,, \quad F := \sum_{k \notin \mathbb{Z}^n_{\rm odd},\, |k|<\gamma^{-1/(7n)}} \frac{R^y_k e^{\mathrm{i}k\cdot x}}{\mathrm{i}\omega(\xi)\cdot k} \partial_y \,, \quad with \ \langle F \rangle = 0 \,. \tag{6.4}$$

It satisfies

$$\|F\|^T_{3s/4,r,\mathtt{a},N_0,\theta,\mu,\lambda,\mathcal{O}_*} = \|F\|^\lambda_{3s/4,r,\mathtt{a},\mathcal{O}_*} \leq C(s)\varepsilon \,. \tag{6.5}$$

Moreover $F \in \mathcal{R}^{\leq 0}_{a-rev}$ and $\mathcal{S}F = F$.

Proof. Note that (6.2) is the first equation of (5.42) with $R^y \rightsquigarrow \Pi_{|k|<\gamma^{-1/(7n)}} R^y(x)$. By Lemma 5.1, it has the unique solution (6.4), see (5.45). The equality in (6.5) follows by (3.81) noting that the quasi-Töplitz norm of F coincides with its majorant norm since $F = F(x)\partial_y$. We now prove the inequality in (6.5). We have

$$\|F\|_{3s/4,r,\mathtt{a}} \overset{(6.4),(6.3)}{\leq} C(s)\gamma^{-2/3} \|\Pi_{|k|<\gamma^{-1/(7n)}} R^y \partial_y\|_{s,r,\mathtt{a}} \overset{(6.1)(5.11)}{\leq} C(s)\gamma^{-2/3} \|\mathcal{P}^y(x)\partial_y\|_{s,r,\mathtt{a}} \overset{(4.7)}{\leq} C(s)\varepsilon \,.$$

Moreover, setting $\Delta_{\xi,\eta} f := |\xi - \eta|^{-1}(f(\xi) - f(\eta))$, we get

$$\Delta_{\xi,\eta} F^y_k = \frac{\Delta_{\xi,\eta} R^y_k}{\mathrm{i}\omega(\xi)\cdot k} + \frac{R^y_k(\eta) \Delta_{\xi,\eta} \omega \cdot k}{(\omega(\xi)\cdot k)(\omega(\eta)\cdot k)} \,,$$

which, in turn, implies

$$\begin{aligned}\|F\|^{\rm lip}_{3s/4,r,\mathtt a} &\stackrel{(6.3),(4.5)}{\le} C(s)\gamma^{-2/3}\|R^y\partial_y\|^{\rm lip}_{s,r,\mathtt a}+C(s)\gamma^{-4/3}M_0\|R^y\partial_y\|_{s,r,\mathtt a}\\ &\stackrel{(6.1)(5.11),(5.12)}{\le} C(s)\gamma^{-2/3}\|\mathcal P^y\partial_y\|^{\rm lip}_{s,r,\mathtt a}+C(s)\gamma^{-4/3}M_0\|\mathcal P^y\partial_y\|_{s,r,\mathtt a}\\ &\stackrel{\lambda:=\gamma/M_0}{\le} C'(s)\gamma^{-2/3}\lambda^{-1}\|\mathcal P^y\partial_y\|^{\lambda}_{s,r,\mathtt a}\stackrel{(4.7)}{\le} C'(s)\lambda^{-1}\varepsilon\end{aligned}$$

The two above estimates imply (6.5). ■

We now apply Corollary 3.2 with $\vec p \rightsquigarrow (3s/4, r, \mathtt a, N_0, \theta, \mu, \lambda, \mathcal O_*)$ and $\vec p\,' \rightsquigarrow \vec p_0$ with

$$\vec p_0 := (s/2, r/2, \mathtt a/2, N_0^{(0)}, 4\theta/3, 3\mu/4, \lambda, \mathcal O_*) \tag{6.6}$$

where $N_0^{(0)} \ge \max\{N_0, \bar N\}$ (recall (3.58)) is chosen large enough so that (recall (3.59)-(3.60))

$$\begin{aligned}&(\kappa+1)(N_0^{(0)})^{b-L}\ln N_0^{(0)} \le \mu/4\,, \qquad (7+\kappa)(N_0^{(0)})^{L-1}\ln N_0^{(0)} \le \theta/3\,,\\ &2(N_0^{(0)})^{-b}\ln^2 N_0^{(0)} \le b\min\{s/4, \mathtt a/2\}\,.\end{aligned} \tag{6.7}$$

Then (3.58)-(3.59)-(3.60) are satisfied and (6.5) imply condition (3.85) for ε sufficiently small. Let $\bar\Phi$ be the time 1-flow of F (so that $e^{{\rm ad}_F} = \bar\Phi_\star$). Since the quasi-Töplitz norm is non-increasing with the parameter N_0 (see (3.25)) we may also take $N_0 \ge \bar N$ large enough so that (5.14) (with $N_0 \rightsquigarrow N_1$) holds. Hence

$$\begin{aligned}\|e^{{\rm ad}_F}(R-R^y\partial_y)\|^T_{\vec p_0} &\stackrel{(3.86)}{\le} 2\|R-R^y\partial_y\|^T_{s,r,\mathtt a,N_0,\theta,\mu,\lambda,\mathcal O_*}\\ &\stackrel{(6.1),(5.13)}{\le} 18\|\mathcal P-\mathcal P^y(x)\partial_y\|^T_{s,r,\mathtt a,N_0,\theta,\mu,\lambda,\mathcal O_*}\stackrel{(4.6),(4.7)}{<} 18\gamma\,.\end{aligned} \tag{6.8}$$

Similarly (3.87) (with $h \rightsquigarrow 1$, $b_j \rightsquigarrow 1/j!$) implies

$$\begin{aligned}\Big\|\Big(e^{{\rm ad}_F}(R-R^y\partial_y)-(R-R^y\partial_y)\Big)^{(h)}\Big\|^T_{\vec p_0} &\lessdot \|\mathcal P_*\|^T_{s,r,\mathtt a,N_0,\theta,\mu,\lambda,\mathcal O_*}\|F\|^T_{3s/4,r,\mathtt a,N_0,\theta,\mu,\lambda,\mathcal O_*}\\ &\stackrel{(6.5),(4.7)}{\le} C(s)\gamma\varepsilon\,, \quad h=-1,0\,.\end{aligned} \tag{6.9}$$

Since the commutator $[F, R^y(x)\partial_y] = [F^y(x)\partial_y, R^y(x)\partial_y] = 0$ we deduce $e^{{\rm ad}_F}(R^y\partial_y) = R^y\partial_y$, and, using also (6.2), we get $e^{{\rm ad}_F}\mathcal N = \mathcal N + {\rm ad}_F\mathcal N$. Hence

$$\begin{aligned}e^{{\rm ad}_F}X &= e^{{\rm ad}_F}\mathcal N + e^{{\rm ad}_F}R^y\partial_y + e^{{\rm ad}_F}(R-R^y\partial_y) = \mathcal N + {\rm ad}_F\mathcal N + R^y\partial_y + e^{{\rm ad}_F}(R-R^y\partial_y)\\ &\stackrel{(6.2)}{=} \mathcal N + \Pi_{|k|\ge\gamma^{-1/(7n)}}R^y\partial_y + e^{{\rm ad}_F}(R-R^y\partial_y) =: \mathcal N_0 + P_0\end{aligned} \tag{6.10}$$

where $\mathcal N_0 := \mathcal N$. Then we consider the symmetrized vector field

$$X_0 := \mathcal S(e^{{\rm ad}_F}X) = \mathcal N_0 + R_0\,, \qquad R_0 := \mathcal S P_0\,. \tag{6.11}$$

Since $R^y(x)\partial_y$ depends on the variable x only we have

$$\|\mathcal S\Pi_{|k|\ge\gamma^{-1/(7n)}}R^y(x)\partial_y\|^T_{\vec p_0} = \|\mathcal S\Pi_{|k|\ge\gamma^{-1/(7n)}}R^y(x)\partial_y\|^\lambda_{s/2,r,\mathtt a,\mathcal O} \lessdot \gamma\varepsilon\,, \tag{6.12}$$

arguing as for (6.5), using (5.13), (2.37), and for $\gamma < \gamma_*$ small (depending on s and n). Recollecting (6.11), (6.10), (6.8), (6.12) and (6.9) we get

Lemma 6.2. *The constants*

$$\bar\varepsilon_0 := \varepsilon_0^{(-1)} + \varepsilon_0^{(0)}\,, \qquad \varepsilon_0^{(h)} := \gamma^{-1}\|R_0^{(h)}\|^T_{\vec p_0}\,,\ h=-1,0\,, \qquad \Theta_0 := \gamma^{-1}\|R_0\|^T_{\vec p_0}\,. \tag{6.13}$$

satisfy

$$\varepsilon_0^{(h)} \le C(s,n)\varepsilon\,,\ h=-1,0\,, \qquad \Theta_0 \le 2^8\,, \tag{6.14}$$

where ε is defined in (4.7).

The vector fields $P_0, R_0 \in \mathcal{R}_{rev}$ because $F \in \mathcal{R}_{a-rev}$ (Lemma 6.1), $R \in \mathcal{R}_{rev}$, and using Lemma 2.9 and Proposition 5.2. Similarly, since $\mathcal{X} \in \mathcal{R}_{rev}$ (by the hypothesis of Theorem 4.1) the vector field

$$\mathcal{X}_0 := e^{\mathrm{ad}_F}\mathcal{X} = \bar{\Phi}_\star \mathcal{X} \in \mathcal{R}_{rev}\,. \tag{6.15}$$

Proposition 5.1 implies that $X_{|E} = (\mathcal{S}\mathcal{X})_{|E} = \mathcal{X}_{|E}$ (see (6.1)) and $X_{0|E} = (e^{\mathrm{ad}_{\mathrm{F}}}X)_{|E}$ (see (6.11)). Moreover, since F is even, Lemma 2.11 (applied with $Y \rightsquigarrow F$) and (6.15) imply

$$\mathcal{X}_{0|E} = X_{0|E}\,. \tag{6.16}$$

6.2 The KAM step

We now describe the iterative scheme which produces a sequence of quasi-Töplitz vector fields X_ν with parameters $\vec{p}_\nu = (s_\nu, r_\nu, \mathtt{a}_\nu, N_0^{(\nu)}, \theta_\nu, \mu_\nu, \lambda, \mathcal{O}_\nu)$, $\lambda = \gamma/M_0$, and such that $X^{\le 0}_{\nu\;|E}$ tends to zero as $\nu \to +\infty$. For compactness of notation we drop the index ν and write "+" for $\nu+1$.

Iterative hypotheses. Suppose $1 < \theta, \mu < 6$, $N_0 \ge \bar{N}$ (defined in (3.58)), $\mathcal{O} \subseteq \mathbb{R}^{n/2}$. Let $X = \mathcal{N} + R$, where $\mathcal{N}$ is a normal form vector field (see Definition 5.1) with Lipschitz frequencies $\omega(\xi), \Omega(\xi)$, $\xi \in \mathbb{R}^{n/2}$ and (5.4) holds with some $a(\xi)$, $\forall\,|j| \ge 6N_0$ (namely $j_* = 6N_0$). Moreover $|\omega|^{\mathrm{lip}}_{\mathbb{R}^{n/2}}, |\Omega|^{\mathrm{lip}}_{\mathbb{R}^{n/2}} \le M \le 2M_0$. The perturbation R satisfies $\|R\|^T_{\vec{p}} < \infty$, $R \in \mathcal{R}_{rev}$, $\mathcal{S}R = R$. We finally fix some K and we assume that $6N_0 \ge \hat{c}\gamma^{-1/3}K^{\tau+1}$ (where $\hat{c}$ is the constant introduced in (5.58)).

We now describe a *KAM step* namely a change of variables generated by the time-1 flow of a vector field F and such that $X_+ := \mathcal{S}e^{\mathrm{ad}_F}X = \mathcal{N}_+ + R_+$ still satisfies the iterative hypotheses, with slightly different parameters, and a much smaller new perturbation R_+, see (6.43).

The new normal form $\mathcal{N}_+$. Set (recall (2.48))

$$R_K^{\le 0} := \Pi_{|k|<K}\Pi_{|\pi|<K}R^{\le 0} = \Pi_{|k|<K}\Pi_{|\pi|<K}R^{(-1)} + \Pi_{|k|<K}\Pi_{|\pi|<K}R^{(0)} =: R_K^{(-1)} + R_K^{(0)}\,. \tag{6.17}$$

Since $R \in \mathcal{R}_{rev}$ then $R_K^{\le 0} \in \mathcal{R}^{\le 0}_{rev}$ and $\mathcal{S}R_K^{\le 0} = R_K^{\le 0}$. The new normal form is defined for $\xi \in \mathcal{O}$ as

$$\mathcal{N}^+ := \mathcal{N} + \hat{\mathcal{N}}\,,$$

$$\hat{\mathcal{N}} \overset{(5.28)}{:=} [R_K^{\le 0}] = \langle R^x\rangle \partial_x + \sum_{j\in\mathbb{Z}\setminus\mathcal{I}} \langle R^{z_j z_j}\rangle z_j\partial_{z_j} + \langle R^{\bar{z}_j\bar{z}_j}\rangle \bar{z}_j\partial_{\bar{z}_j} = \hat{\omega}\cdot\partial_x + \mathrm{i}\sum_{j\in\mathbb{Z}\setminus\mathcal{I}} \hat{\Omega}_j z_j(\partial_{z_j} - \bar{z}_j\partial_{\bar{z}_j}) \tag{6.18}$$

because, since $R_K^{\le 0}$ is real-coefficients (Definition 2.8) and real-on-real (Definition 2.9),

$$\langle R^{z_j z_j}\rangle = \mathrm{i}\hat{\Omega}_j\,,\ \hat{\Omega}_j \in \mathbb{R}\,,\ \ \langle R^{\bar{z}_j\bar{z}_j}\rangle \overset{(2.57)}{=} -\mathrm{i}\hat{\Omega}_j\,,\ \forall\, j \in \mathbb{Z}\setminus\mathcal{I}\,,\ \ \hat{\omega}_j := \langle R^{x_j}\rangle \in \mathbb{R}\,, \forall\, j \in \mathcal{I}\,. \tag{6.19}$$

Moreover, since R is even, $\hat{\omega}\,, \hat{\Omega}$ satisfy (5.2), namely

$$\hat{\omega}_j \overset{(2.58)}{=} \hat{\omega}_{-j}\,, \qquad \hat{\Omega}_j \overset{(2.58)}{=} \hat{\Omega}_{-j}\,. \tag{6.20}$$

Note that $\hat{\mathcal{N}}$ only depends on $R^{(0)}$.

Lemma 6.3.

$$\sup_{\xi\in\mathcal{O}} |\hat{\omega}|, |\hat{\Omega}|_\infty \le 2\|R^{(0)}\|_{s,r,\mathtt{a}}\,, \qquad |\hat{\omega}|^{\mathrm{lip}}_{\mathcal{O}}, |\hat{\Omega}|^{\mathrm{lip}}_{\infty,\mathcal{O}} \le 2\|R^{(0)}\|^{\mathrm{lip}}_{s,r,\mathtt{a}} \tag{6.21}$$

and there exist $\hat{a} : \mathcal{O} \to \mathbb{R}$ *satisfying*

$$\sup_{\xi\in\mathcal{O}} |\hat{a}(\xi)| \le 2\|R^{(0)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu} \tag{6.22}$$

such that

$$\sup_{\xi\in\mathcal{O}} |\hat{\Omega}_j(\xi) - \hat{a}(\xi)| \le \frac{40}{|j|}\|R^{(0)}\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}\,, \quad \forall\,|j| \ge 6(N_0+1)\,. \tag{6.23}$$

Lemma 6.3 is based on the following elementary lemma (see Lemma 5.3 of [4]).

Lemma 6.4. *Suppose that, $\forall N \geq N_0 \geq 9$, $j \geq 6N$,*

$$\hat{\Omega}_j = a_N + b_{N,j} N^{-1} \quad \text{with} \quad a_N, b_{N,j} \in \mathbb{R}\,, \quad |a_N| \leq c\,, \ |b_{N,j}| \leq c\,, \tag{6.24}$$

for some $c > 0$ (independent of j). Then there exists $a \in \mathbb{R}$, satisfying $|a| \leq c$, such that

$$|\hat{\Omega}_j - a| \leq \frac{20c}{j}\,, \quad \forall\, j \geq 6(N_0 + 1)\,. \tag{6.25}$$

PROOF OF LEMMA 6.3. We first prove (6.21). The estimate on $\hat{\omega}$ is trivial. Regarding $\hat{\Omega}$ we note (recall (6.18) and (2.36))

$$\mathrm{i} \sum_j \hat{\Omega}_j (z_j \partial_{z_j} - \bar{z}_j \partial_{\bar{z}_j}) = \Pi_{\rm diag} R^{(0)}_K = \Pi_{\rm diag} R^{(0)}\,. \tag{6.26}$$

Then (6.21) follows by Lemma 2.7. Now we want to apply Lemma 6.4. Note that by (6.20), we can restrict to the case j positive. Since $\Pi_{\rm diag}\Pi^{(0)} R \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$ (by (3.29) and (3.26)) we can apply Lemma 3.4. Then for all $N \geq N_0$ we can decompose

$$\Pi_{N,\theta,\mu}\Pi_{\rm diag} R^{(0)} = \tilde{R}^N + N^{-1}\hat{R}^N \quad \text{with} \quad \tilde{R}^N = \Pi_{\rm diag}\tilde{R}^N \in \mathcal{T}_{s,r,\mathtt{a}}(N, \theta, \mu)\,, \quad \hat{R}^N = \Pi_{\rm diag}\hat{R}^N \tag{6.27}$$

and

$$\|\tilde{R}^N\|_r\,, \, \|\hat{R}^N\|_r \leq 2\|\Pi_{\rm diag} R^{(0)}\|^T \leq 2\|R^{(0)}\|^T\,, \tag{6.28}$$

where $\|\cdot\|^T$ is short for $\|\cdot\|^T_{s,r,\mathtt{a},N_0,\theta,\mu}$. By (6.27) (recall (2.36))

$$\tilde{R}^N = \mathrm{i} \sum_{j \in \mathbb{Z}\setminus\mathcal{I}, \sigma = \pm} \tilde{R}^N_{j,\sigma} z^\sigma_j \partial_{z^\sigma_j}\,, \qquad \hat{R}^N = \mathrm{i} \sum_{j \in \mathbb{Z}\setminus\mathcal{I}, \sigma = \pm} \hat{R}^N_{j,\sigma} z^\sigma_j \partial_{z^\sigma_j}\,. \tag{6.29}$$

By (6.27) $\tilde{R}^N \in \mathcal{T}_{s,r,\mathtt{a}}(N, \theta, \mu)$ and, therefore, $\tilde{R}^N_{j,+}$ do *not* depend on j (see (3.14)), namely

$$\tilde{R}^N_{j,+} \equiv a_N \qquad \text{and we also set } b_{N,j} := \hat{R}^N_{j,+}\,. \tag{6.30}$$

By (6.26) we get

$$\Pi_{N,\theta,\mu}\Pi_{\rm diag} R^{(0)} = \mathrm{i} \sum_{|j| > \theta N} \hat{\Omega}_j (z_j \partial_{z_j} - \bar{z}_j \partial_{\bar{z}_j})\,.$$

Then, for $j \geq 6N > \theta N$, by (6.27), (6.29), (6.30), we get $\hat{\Omega}_j = a_N + N^{-1} b_{N,j}$. Applying Lemma 2.7 to $\tilde{R}^N$ and $\hat{R}^N$ in (6.29), we obtain

$$|a_N| \leq \|\tilde{R}^N\|_r \overset{(6.28)}{\leq} 2\|R^{(0)}\|^T\,, \quad |b_{N,j}| \leq \|\hat{R}^N\|_r \overset{(6.28)}{\leq} 2\|R^{(0)}\|^T\,.$$

Hence the assumptions of Lemma 6.4 are satisfied with $c = 2\|R^{(0)}\|^T$ and (6.22)-(6.23) follows. ■

The new vector field X_+. We decompose

$$X = \mathcal{N} + R = \mathcal{N} + R^{\leq 0}_K + (R - R^{\leq 0}_K)$$

where $R^{\leq 0}_K$ is defined in (6.17). We apply Lemma 5.1 and Proposition 5.3 with $\mathcal{O} \rightsquigarrow \mathcal{O}_+$ where

$$\mathcal{O}_+ := \Big\{ \xi \in \mathcal{O} \mid (5.31) - (5.34) \text{ and } (5.57) \text{ hold} \Big\}\,. \tag{6.31}$$

Let $F = F^{\leq 0}_K = F^{(-1)}_K + F^{(0)}_K \in \mathcal{R}^{\leq 0}_{a-rev}$ be the unique solution of the homological equation

$$\mathrm{ad}_{\mathcal{N}} F = R^{\leq 0}_K - [R^{\leq 0}_K]\,. \tag{6.32}$$

The bounds (5.38), $|\omega|^{\rm lip}, |\Omega|^{\rm lip} \le M \le 2M_0$, and (5.59) (with $R \rightsquigarrow R_K^{(h)}$, $h=-1,0$) imply

$$\|F^{(h)}\|^T_{\vec p_\star} \lessdot \gamma^{-1} K^{2\tau+1} \|R^{(h)}\|^T_{\vec p}\,, \quad h=-1,0\,, \quad \text{where} \quad \vec p_\star := (s,r,\mathtt{a},6N_0,\theta,\mu,\lambda,\mathcal{O}_+)\,. \tag{6.33}$$

Note that in (5.58)-(5.59) $N_0^* = 6N_0$ because, by the iterative hypothesis, $j_* = 6N_0 \ge \hat c \gamma^{-1/3} K^{\tau+1}$.

We introduce the new parameters

$$\vec p_+ := (s_+, r_+, \mathtt{a}_+, N_0^+, \theta_+, \mu_+, \lambda, \mathcal{O}_+)\,, \tag{6.34}$$

where $s/2 \le s_+ < s$, $r/2 \le r_+ < r$, $0 < \mathtt{a}_+ < \mathtt{a}$, $N_0^+ \ge 7N_0$, $\theta_+ > \theta$, $\mu_+ < \mu$, such that

$$(\kappa+1)(N_0^+)^{b-L} \ln N_0^+ \le \mu - \mu_+, \quad (7+\kappa)(N_0^+)^{L-1} \ln N_0^+ \le \theta_+ - \theta\,, \tag{6.35}$$

$$2(N_0^+)^{-b} \ln^2 N_0^+ \le b \min\{s - s_+, \mathtt{a} - \mathtt{a}_+\}\,, \tag{6.36}$$

and note that $N_0^+ \ge \bar N$ defined in (3.58) (by the iterative hypothesis $N_0 \ge \bar N$). If, moreover, the smallness condition

$$\|F\|^T_{\vec p_\star} \le c(n)\, \delta_+\,, \quad \delta_+ := \min\Big\{1 - \frac{s_+}{s}, 1 - \frac{r_+}{r}\Big\} \tag{6.37}$$

holds (see (3.85)), then Corollary 3.2 (with $\vec p \rightsquigarrow \vec p_\star$, $\vec p\,' \rightsquigarrow \vec p_+$, $\delta \rightsquigarrow \delta_+$) implies that the time 1-flow generated by F maps $D(s_+, r_+)$ into $D(s,r)$. The transformed and symmetrized vector field is

$$X^+ := \mathcal{S} e^{{\rm ad}_F} X \overset{(2.42)}{=} \mathcal{S}\Big(X + {\rm ad}_F(X) + \sum_{j\ge 2} \frac{1}{j!} {\rm ad}_F^j(X)\Big) = \mathcal{N}^+ + R^+ \tag{6.38}$$

with the new normal form $\mathcal{N}^+$ defined in (6.18) and, by (6.32), the new perturbation

$$R^+ := \mathcal{S}\Big(R - R_K^{\le 0} + {\rm ad}_F(R^{\le 0}) + {\rm ad}_F(R^{\ge 1}) + \sum_{j\ge 2} \frac{1}{j!} {\rm ad}_F^j(X)\Big) \tag{6.39}$$

where $R^{\ge 1} := \sum_{j\ge 1} R^{(j)}$, see (2.46), so that $R = R^{\le 0} + R^{\ge 1}$.

We set

$$\varepsilon^{(h)} := \gamma^{-1} \|R^{(h)}\|^T_{\vec p}\,,\ h=-1,0\,, \quad \bar\varepsilon := \varepsilon^{(-1)} + \varepsilon^{(0)}\,, \quad \Theta := \gamma^{-1} \|R\|^T_{\vec p} \tag{6.40}$$

and the corresponding quantities $\varepsilon_+^{(h)}, \bar\varepsilon_+, \Theta_+$ for R^+ with parameters $\vec p_+$ defined in (6.34).

Proposition 6.1. (KAM step) *Assume that $\vec p, \vec p_+$ satisfy* (6.35), (6.36), *and that*

$$\delta_+^{-1} K^{2\tau+1} \bar\varepsilon \quad \textit{is small enough}\,, \quad \Theta \le 2^9\,, \tag{6.41}$$

(δ_+ is defined in (6.37)*). Then, by* (6.33)*, the solution $F \in \mathcal{R}^{\le 0}_{rev}$ of* (6.32) *satisfies* (6.37) *and $e^{{\rm ad}_{\rm F}}$ and X^+ in* (6.38) *are well defined. The perturbation $R^+ \in \mathcal{R}_{rev}$ in* (6.39) *satisfies $R = \mathcal{S}R$ and*

$$\begin{aligned} \varepsilon_+^{(-1)} &\lessdot \delta_+^{-2} K^{4\tau+2} \bar\varepsilon^2 + \varepsilon^{(-1)}\, e^{-K\min\{s-s_+,\,\mathtt{a}-\mathtt{a}_+\}} \\ \varepsilon_+^{(0)} &\lessdot \delta_+^{-2} K^{4\tau+2} \big(\varepsilon^{(-1)} + \bar\varepsilon^2\big) + \varepsilon^{(0)}\, e^{-K\min\{s-s_+,\,\mathtt{a}-\mathtt{a}_+\}} \end{aligned} \tag{6.42}$$

$$\Theta_+ \le \Theta(1 + C\delta_+^{-2} K^{4\tau+2} \bar\varepsilon)\,. \tag{6.43}$$

Proof. The proof is split in several lemmata where we analyze each term of R^+ in (6.39).

We first claim that

$$\Big\|{\rm ad}_F(R^{\le 0})\Big\|^T_{\vec p_+} + \Big\|\sum_{j\ge 2} \frac{1}{j!} {\rm ad}_F^j(X)\Big\|^T_{\vec p_+} \lessdot \delta_+^{-2} \gamma K^{2(2\tau+1)} \bar\varepsilon^2\,. \tag{6.44}$$

We have

$$
\begin{aligned}
\sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^j(X) \quad &= \quad \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^j(\mathcal{N}+R) = \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^{j-1}(\mathrm{ad}_F\mathcal{N}) + \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^j(R) \\
&\overset{(6.32)}{=} \quad \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^{j-1}([R_K^{\leq 0}] - R_K^{\leq 0}) + \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^j(R)\,.
\end{aligned}
$$

As we have already noticed, by (6.35), (6.36), (6.37) we can apply Corollary 3.2 (with $\vec p \rightsquigarrow \vec p_\star$, $\vec p\,' \rightsquigarrow \vec p_+$, $\delta \rightsquigarrow \delta_+$, $h \rightsquigarrow 2$) obtaining

$$
\Big\| \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^j(R)\Big\|_{\vec p_+}^T \overset{(3.87)}{\lessdot} \Big(\delta_+^{-1}\|F\|_{\vec p_\star}^T\Big)^2 \|R\|_{\vec p_\star}^T \overset{(6.33),(6.40)}{\lessdot} \delta_+^{-2}K^{2(2\tau+1)}\bar\varepsilon^2\gamma\,\Theta\,. \tag{6.45}
$$

In the same way we get (with $h \rightsquigarrow 1$)

$$
\begin{aligned}
&\Big\| \sum_{j\geq 2}\frac{1}{j!}\mathrm{ad}_F^{j-1}\big([R_K^{\leq 0}] - R_K^{\leq 0}\big)\Big\|_{\vec p_+}^T = \Big\| \sum_{j\geq 1}\frac{1}{(j+1)!}\mathrm{ad}_F^j\big([R_K^{\leq 0}] - R_K^{\leq 0}\big)\Big\|_{\vec p_+}^T \\
&\overset{(3.87)}{\lessdot} \delta_+^{-1}\|F\|_{\vec p_\star}^T\|[R_K^{\leq 0}] - R_K^{\leq 0}\|_{\vec p_\star}^T \leq \delta_+^{-1}\|F\|_{\vec p_\star}^T\|R_K^{\leq 0}\|_{\vec p_\star}^T \overset{(6.33),(6.40)}{\lessdot} \delta_+^{-1}K^{2\tau+1}\gamma\bar\varepsilon^2\,.
\end{aligned} \tag{6.46}
$$

Finally, by Corollary 3.1, applied with $\vec p \rightsquigarrow \vec p_\star$, $\vec p_1 \rightsquigarrow \vec p_+$, $\delta \rightsquigarrow \delta_+$ (note that conditions (3.38)-(3.39) follow by (6.35)-(6.36)), we get

$$
\Big\|\mathrm{ad}_F(R^{\leq 0})\Big\|_{\vec p_+}^T \overset{(3.84)}{\lessdot} \delta_+^{-1}\|F\|_{\vec p_\star}^T\|R^{\leq 0}\|_{\vec p_\star}^T \overset{(6.33),(6.40)}{\lessdot} \delta_+^{-1}K^{2\tau+1}\gamma\,\bar\varepsilon^2\,. \tag{6.47}
$$

The bounds (6.45), (6.46), (6.47), and $\Theta \leq 2^9$ (see (6.41)), prove (6.44).

We now prove (6.43). Again by Corollary 3.1 we get

$$
\Big\|\mathrm{ad}_F(R^{\geq 1})\Big\|_{\vec p_+}^T \lessdot \delta_+^{-1}\|F\|_{\vec p_\star}^T\|R^{\geq 1}\|_{\vec p}^T \overset{(6.33),(3.31),(6.40)}{\lessdot} \delta_+^{-1}K^{2\tau+1}\gamma\,\bar\varepsilon\,\Theta \tag{6.48}
$$

and (6.43) follows by (6.39), (5.13), (3.31), (6.40) (6.48), (6.44) and $\bar\varepsilon \leq 3\Theta$ (which follows by (6.40) and (3.30)).

We now consider $R_+^{(h)}$, $h = 0, -1$. Recalling the degree decomposition $F = F^{(-1)} + F^{(0)}$, formula (2.45) implies that the term $\mathrm{ad}_F R^{\geq 1}$ in (6.39) does not contribute to $R_+^{(-1)}$. On the other hand, its contribution to $R_+^{(0)}$ is $[R^{(1)}, F^{(-1)}]$. Again by (3.84), (6.33), (6.40) and (3.30), we get

$$
\|[R^{(1)}, F^{(-1)}]\|_{\vec p_+}^T \lessdot \delta_+^{-1}\gamma K^{2\tau+1}\varepsilon^{(-1)}\Theta\,. \tag{6.49}
$$

The contribution of $R - R_K^{\leq 0}$ in (6.39) to $R_+^{(h)}$, $h = 0, -1$, is

$$
\Pi_{|k|<K}\Pi_{|\pi|\geq K}R^{(h)} + \Pi_{|k|\geq K}R^{(h)}\,.
$$

By (3.32)-(3.33) (recall $ss_+^{-1} < 2$), (3.30), and (6.40), we get

$$
\Big\|\Pi_{|k|<K}\Pi_{|\pi|\geq K}R^{(h)} + \Pi_{|k|\geq K}R^{(h)}\Big\|_{\vec p_+}^T \leq 3e^{-K\min\{s-s_+,\,\mathtt{a}-\mathtt{a}_+\}}\gamma\varepsilon^{(h)}\,. \tag{6.50}
$$

In conclusion, (6.42) follows by (6.39), (5.13), (6.44), (6.49), (6.50) and $\Theta \leq 2^9$. ■

6.3 KAM iteration

Lemma 6.5. *Suppose that* $\varepsilon_i^{(-1)}, \varepsilon_i^{(0)} \geq 0$, $i = 0, \ldots, \nu$, *satisfy*

$$\begin{aligned} \varepsilon_{i+1}^{(-1)} &\leq C_* \mathtt{K}^i \bar{\varepsilon}_i^2 + C_* \varepsilon_i^{(-1)} e^{-K_* 2^i} \\ \varepsilon_{i+1}^{(0)} &\leq C_* \mathtt{K}^i \big(\varepsilon_i^{(-1)} + \bar{\varepsilon}_i^2\big) + C_* \varepsilon_i^{(0)} e^{-K_* 2^i}, \quad i = 0, \ldots, \nu - 1, \end{aligned} \tag{6.51}$$

where $\bar{\varepsilon}_i := \varepsilon_i^{(-1)} + \varepsilon_i^{(0)}$, *and* $\mathtt{K}, C_* \geq 1$, $K_* > 0$. *Then there exist* $\bar{\varepsilon}_\star < 1$, $C_\star > e^{K_*}$, $\chi \in (1, 2)$, *depending only on* $\mathtt{K}, C_*, K_*$ *(and not on* ν*) such that, if*

$$\bar{\varepsilon}_0 \leq \bar{\varepsilon}_\star \quad \Longrightarrow \quad \bar{\varepsilon}_i \leq C_\star \bar{\varepsilon}_0 \, e^{-K_* \chi^i}, \quad \forall i = 0, \ldots, \nu. \tag{6.52}$$

Proof. Iterating two times (6.51) we get, $\forall\, 0 \leq j \leq \nu - 2$,

$$\bar{\varepsilon}_{j+2} \leq 2C_* \Big(\mathtt{K}^{j+1}(\varepsilon_{j+1}^{(-1)} + \bar{\varepsilon}_{j+1}^2) + \bar{\varepsilon}_{j+1} e^{-K_* 2^{j+1}}\Big) \leq c_1 C_*^{c_1} \mathtt{K}^{c_1 j} \Big(\bar{\varepsilon}_j^2 + \bar{\varepsilon}_j^4 + \bar{\varepsilon}_j e^{-K_* 2^j}\Big) \tag{6.53}$$

for a suitable constant $c_1 > 1$.

We first claim that (6.52) holds with $\chi := 5/4$ for all $i = 2j \leq \nu$. Setting $a_j := \bar{\varepsilon}_{2j}$, we prove that there exist $C_\star$ large and $\bar{\varepsilon}_\star$ small (as in the statement) such that if $a_0 \leq \bar{\varepsilon}_\star$ then

$$(\mathbf{S})_j \qquad a_j \leq c_2^{j+1} a_0 e^{-K_* \tilde{\chi}^{2j}}, \quad \forall\, 0 \leq j \leq \nu/2$$

for a suitable $c_2 = c_2(\mathtt{K}, C_*, K_*) \geq 1$ large enough and $1 < \tilde{\chi} < \sqrt{2}$, e.g. $\tilde{\chi} := 4/3$. We proceed by induction. The statement $(\mathbf{S})_0$ is true for $c_2 \geq e^{K_*}$. Now suppose $(\mathbf{S})_j$ holds true. Note that $a_j \leq 1$ taking $\bar{\varepsilon}_\star \leq \min_{j \geq 0} e^{K_* \tilde{\chi}^{2j}} / c_2^{j+1}$. Then $(\mathbf{S})_{j+1}$ follows by

$$\begin{aligned} a_{j+1} &= \bar{\varepsilon}_{2j+2} \overset{(6.53)}{\leq} c_1 C_*^{c_1} \mathtt{K}^{2c_1 j} \Big(a_j^2 + a_j^4 + a_j e^{-K_* 2^{2j}}\Big) \overset{a_j \leq 1}{\leq} 2c_1 C_*^{c_1} \mathtt{K}^{2c_1 j} \Big(a_j^2 + a_j e^{-K_* 2^{2j}}\Big) \\ &\overset{(\mathbf{S})_j}{\leq} 2c_1 C_*^{c_1} \mathtt{K}^{2c_1 j} \Big((c_2^{j+1} a_0 e^{-K_* \tilde{\chi}^{2j}})^2 + (c_2^{j+1} a_0 e^{-K_* \tilde{\chi}^{2j}}) e^{-K_* 2^{2j}}\Big) \leq c_2^{j+2} a_0 e^{-K_* \tilde{\chi}^{2j+2}} \end{aligned}$$

since $4c_1 C_*^{c_1} \mathtt{K}^{2c_1 j} c_2^{j+1} a_0 e^{-K_* \tilde{\chi}^{2j}} e^{-K_* 2^{2j}} \leq c_2^{j+2} a_0 e^{-K_* \tilde{\chi}^{2j+2}}$ taking c_2 large enough (use $\tilde{\chi} < 2$) and

$$4c_1 C_*^{c_1} \mathtt{K}^{2c_1 j} (c_2^{j+1} a_0 e^{-K_* \tilde{\chi}^{2j}})^2 \leq c_2^{j+2} a_0 e^{-K_* \tilde{\chi}^{2j+2}}$$

since $\tilde{\chi} < \sqrt{2}$ and taking $a_0 \leq \bar{\varepsilon}_\star$ small enough. We have proved inductively $(\mathbf{S})_j$. Then (6.52) for $i = 2j$ follows since $5/4 =: \chi < \tilde{\chi} := 4/3$ and taking $C_\star$ large enough. The case $i = 2j + 1$ follows analogously noting that $\bar{\varepsilon}_1 \leq C \bar{\varepsilon}_0$ (by (6.51)) taking $\bar{\varepsilon}_\star$ small. ■

We now consider as initial parameters of the iteration

$$\vec{p}_0 = (s_0, r_0, \mathtt{a}_0, N_0^{(0)}, \theta_0, \mu_0, \lambda, \mathcal{O}_*) \overset{(6.6)}{:=} (s/2, r/2, \mathtt{a}/2, N_0^{(0)}, 4\theta/3, 3\mu/4, \lambda, \mathcal{O}_*) \tag{6.54}$$

where $s, r, \mathtt{a}, \theta, \mu, \lambda$ are defined in Theorem 4.1, the set $\mathcal{O}_*$ in (6.3), and

$$N_0^{(0)} := \hat{c} \gamma^{-1/3}/6 \ \text{ satisfies } \ N_0^{(0)} \geq \bar{N} \ \text{ and } \ (6.7), \tag{6.55}$$

taking $\gamma \leq \gamma_*(n, s, \mathtt{a})$ small enough (the constants $\hat{c}, \bar{N}$ are defined in (5.58) and (3.58) respectively).

We start a KAM iterative scheme on the vector field $X_0 = \mathcal{N}_0 + R_0 : D_0 \times \mathcal{O}_* \to V$ defined in (6.11) where the normal form $\mathcal{N}_0 = \mathcal{N}$ is defined in (4.1), the frequencies satisfy $|\omega^{(0)}|^{\mathrm{lip}} + |\Omega^{(0)}|_\infty^{\mathrm{lip}} \leq M_0$ on the whole $\mathbb{R}^{n/2}$, see (4.5), and the domain $D_0 := D(s_0, r_0)$.

We define, for $\nu > 0$, the parameters

- $s_{\nu+1} := s_\nu - s_0 2^{-\nu-2} \searrow \dfrac{s_0}{2}, \quad r_{\nu+1} := r_\nu - r_0 2^{-\nu-2} \searrow \dfrac{r_0}{2}, \quad D_\nu := D(s_\nu, r_\nu)$
- $\mathtt{a}_{\nu+1} = \mathtt{a}_\nu - \mathtt{a}_0 2^{-\nu-2} \searrow \dfrac{\mathtt{a}_0}{2}, \quad M_{\nu+1} := M_\nu + M_0 2^{-\nu-2} \nearrow 3\dfrac{M_0}{2},$
- $\theta_{\nu+1} := \theta_\nu + \theta_0 2^{-\nu-2} \nearrow 3\dfrac{\theta_0}{2}, \quad \mu_{\nu+1} := \mu_\nu - \mu_0 2^{-\nu-2} \searrow \dfrac{\mu_0}{2},$
- $N_0^{(\nu)} := N_0^{(0)} 2^{\nu\rho}, \quad K_\nu := 4^\nu, \quad \rho := \max\Big\{2(\tau+1), \dfrac{1}{L-b}, \dfrac{1}{1-L}\Big\}. \qquad (6.56)$

Note that, by (6.55) and (6.56), for all $\nu \geq 0$,

$$6N_0^{(\nu)} \geq \hat{c}\gamma^{-1/3} K_\nu^{\tau+1} \quad \text{and} \quad N_0^{(\nu)} \geq \bar{N} \ \ \text{(defined in (3.58))}. \tag{6.57}$$

Lemma 6.6. (Iterative lemma) *Let $X_0 := \mathcal{N}_0 + R_0$ be as above and set*

$$\mathcal{O}_0 := \Big\{\xi \in \mathcal{O} \ : \ |\vec{\omega}(\xi) \cdot h| \geq 2\gamma^{2/3}\langle h\rangle^{-n/2}, \, \forall 0 < |h| < \gamma^{-1/(7n)}\Big\}. \tag{6.58}$$

If $\bar{\varepsilon}_0, \Theta_0$, defined in (6.13)*, satisfy*

$$\bar{\varepsilon}_0 \ \text{ is small enough} \quad \text{ and } \quad \Theta_0 \leq 2^8, \tag{6.59}$$

then:

(S1)$_\nu$ $\forall 0 \leq i \leq \nu$*, there exist $\omega^{(i)} = (\vec{\omega}^{(i)}, \vec{\omega}^{(i)})$, $\Omega^{(i)}$, $a^{(i)}$ defined for all $\xi \in \mathbb{R}^{n/2}$, satisfying*

$$|\omega^{(i)} - \omega^{(0)}| + \lambda|\omega^{(i)} - \omega^{(0)}|^{\rm lip}, |\Omega^{(i)} - \Omega^{(0)}|_\infty + \lambda|\Omega^{(i)} - \Omega^{(0)}|_\infty^{\rm lip} \leq C(1-2^{-i})\gamma\bar{\varepsilon}_0 \tag{6.60}$$
$$|a^{(i)}| \leq C(1-2^{-i})\gamma\bar{\varepsilon}_0, \qquad |\omega^{(i)}|^{\rm lip}, \ |\Omega^{(i)}|_\infty^{\rm lip} \leq (2-2^{-i})M_0, \tag{6.61}$$
$$|\Omega_j^{(i)} - \Omega_j^{(0)} - a^{(i)}| \leq C\gamma/|j|, \qquad \forall |j| \geq 6(N_0^{(i-1)}+1), \tag{6.62}$$

uniformly on $\mathbb{R}^{n/2}$. For $1 \leq i \leq \nu$, $\mathcal{N}_i := \omega^{(i)}(\xi) \cdot \partial_x + {\rm i}\Omega^{(i)}(\xi)(z\partial_z - \bar{z}\partial_{\bar{z}})$ is a normal form as in Definition 5.1 satisfying (5.4) *for all $|j| \geq 6(N_0^{(i-1)}+1)$. Set*

$$\begin{aligned}
\mathcal{O}_i := \Big\{\xi \in \mathcal{O}_{i-1} \ : \ \text{for} \qquad & h \in \mathbb{Z}^{n/2}, \ |h| < K_{i-1}, \ p \in \mathbb{Z}, \\
|\vec{\omega}^{(i-1)}(\xi) \cdot h| \ &\geq \ (1-2^{-i})2\gamma\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0, \\
|\vec{\omega}^{(i-1)}(\xi) \cdot h + \Omega_j^{(i-1)}(\xi)| \ &\geq \ (1-2^{-i})2\gamma\langle h\rangle^{-\tau}, \\
|\vec{\omega}^{(i-1)}(\xi) \cdot h + \Omega_j^{(i-1)}(\xi) + \Omega_{j'}^{(i-1)}(\xi)| \ &\geq \ (1-2^{-i})2\gamma\langle h\rangle^{-\tau}, \\
|\vec{\omega}^{(i-1)}(\xi) \cdot h - \Omega_j^{(i-1)}(\xi) + \Omega_{j'}^{(i-1)}(\xi)| \ &\geq \ (1-2^{-i})2\gamma\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0 \quad \text{or} \quad j' \neq \pm j, \\
|\vec{\omega}^{(i-1)}(\xi) \cdot h + p| \ &\geq \ (1-2^{-i})2\gamma^{2/3}\langle h\rangle^{-\tau} \quad \text{if} \quad h \neq 0 \quad \text{or} \quad p \neq 0 \Big\}.
\end{aligned} \tag{6.63}$$

(S2)$_\nu$ $\forall 1 \leq i \leq \nu$ *there exists a close-to-the-identity, analytic, even (Definition 2.10) change of variables $\Phi^i(\cdot;\xi) : D_i \to D_{i-1}$, defined (and Lipschitz) for $\xi \in \mathcal{O}_i$, such that*[3]

$$X_i := \mathcal{S}\Phi^i_\star X_{i-1} =: \mathcal{N}_i + R_i : D_i \times \mathcal{O}_i \to D_{i-1}, \quad R_i \in \mathcal{R}_{rev}, \ R_i = \mathcal{S}R_i. \tag{6.64}$$

Set $\vec{p}_i = (s_i, r_i, \mathtt{a}_i, N_0^{(i)}, \theta_i, \mu_i, \lambda_i, \mathcal{O}_i)$ and define

$$\bar{\varepsilon}_i := \varepsilon_i^{(-1)} + \varepsilon_i^{(0)}, \quad \varepsilon_i^{(h)} := \gamma^{-1}\|R_i^{(h)}\|_{\vec{p}_i}^T, \quad h = -1, 0, \quad \Theta_i := \gamma^{-1}\|R_i\|_{\vec{p}_i}^T. \tag{6.65}$$

[3] $\Phi^i_\star$ is the lift to the tangent space (recall (4.13)).

$\forall\, 1 \le i \le \nu$ *we get, uniformly for* $\xi \in \mathbb{R}^{n/2}$,

$$
\begin{aligned}
&|\omega^{(i)}(\xi) - \omega^{(i-1)}(\xi)|\,,\ |\Omega^{(i)}(\xi) - \Omega^{(i-1)}(\xi)|_\infty\,,\ |a^{(i)}(\xi) - a^{(i-1)}(\xi)| \le 2\gamma\bar\varepsilon_{i-1}\,,\\
&|\Omega_j^{(i)}(\xi) - a^{(i)}(\xi) - \Omega_j^{(i-1)}(\xi) + a^{(i-1)}(\xi)| \le 40\gamma\frac{\bar\varepsilon_{i-1}}{|j|}\,, \quad \forall |j| \ge 6(N_0^{(i-1)}+1)
\end{aligned}
\tag{6.66}
$$

and

$$
|\omega^{(i)} - \omega^{(i-1)}|^{\rm lip}_{\mathbb{R}^{n/2}}\,, \quad |\Omega^{(i)} - \Omega^{(i-1)}|^{\rm lip}_{\mathbb{R}^{n/2}} \lessdot 2^i K_{i-2}^{\tau+1} M_0 \bar\varepsilon_{i-1} \tag{6.67}
$$

where $K_{-1} := 1$.

$(\mathbf{S3})_\nu$ $\forall 0 \le i \le \nu-1$, *the* $\varepsilon_i^{(-1)}, \varepsilon_i^{(0)}$ *satisfy (6.51) with* $\mathtt{K} = 4^{4\tau+3}$, $K_* = \min\{s_0, \mathtt{a}_0\}/4$ *and some* $C_* = C_*(n)$.

$(\mathbf{S4})_\nu$ $\forall 0 \le i \le \nu$, *we have* $\bar\varepsilon_i \le C_\star \bar\varepsilon_0 e^{-K_*\chi^i}$ *and* $\Theta_i \le 2\Theta_0$.

Proof. The statement $(\mathbf{S1})_0$ follows by the hypothesis setting $a^{(0)} \equiv 0$ and $N_0^{(-1)} := 0$. $(\mathbf{S2})_0$ and $(\mathbf{S3})_0$ are empty. $(\mathbf{S4})_0$ follows because $C_\star \ge e^{K_*}$ in Lemma 6.5. Note also that, by (6.3) and (4.5), the set $\mathcal{O}_0$ defined in (6.58) satisfies

$$
\mathcal{O}_0 \subseteq \{\xi \in \mathcal{O}_* \ :\ B_\lambda(\xi) \subset \mathcal{O}_*\}\,, \quad \lambda := \gamma/M_0\,. \tag{6.68}
$$

We then proceed by induction.

$(\mathbf{S1})_{\nu+1}$. We start defining the normal form at the step $\nu+1$. For all $\xi \in \mathcal{O}_\nu$ if $\nu \ge 1$ and $\xi \in \mathcal{O}_*$ (see (6.3)) if $\nu = 0$, we set (recall (6.19))

$$
\hat\omega_j^{(\nu)}(\xi) := \langle R_\nu^{x_j}(\cdot;\xi)\rangle\,,\ j \in \mathcal{I}\,, \quad {\rm i}\hat\Omega_j^{(\nu)}(\xi) := \langle R_\nu^{z_j z_j}(\cdot;\xi)\rangle\,, \quad j \in \mathbb{Z}\setminus\mathcal{I}\,. \tag{6.69}
$$

The frequencies $\hat\omega^{(\nu)}, \hat\Omega^{(\nu)}$ satisfy (6.20) (since R_ν is even by $(S2)_\nu$) and by Lemma 6.3 (and (6.65)) there exists $\hat a^{(\nu)}(\xi) \in \mathbb{R}$ such that

$$
|\hat\omega^{(\nu)}(\xi)|\,,\ |\hat\Omega^{(\nu)}(\xi)|_\infty\,,\ |\hat a^{(\nu)}(\xi)| \le 2\gamma\bar\varepsilon_\nu\,,\ \ |\hat\Omega_j^{(\nu)}(\xi) - \hat a^{(\nu)}(\xi)| \le 40\gamma\frac{\bar\varepsilon_\nu}{|j|}\,,\ \ \forall |j| \ge 6(N_0^{(\nu)}+1)\,, \tag{6.70}
$$

uniformly in $\xi \in \mathcal{O}_\nu$ (resp. $\mathcal{O}_*$ if $\nu = 0$). Moreover (6.21), (2.33), (3.81), (6.65) imply

$$
|\hat\omega^{(\nu)}|^{\rm lip}_{\mathcal{O}_\nu}\,,\ |\hat\Omega^{(\nu)}|^{\rm lip}_{\infty,\mathcal{O}_\nu} \le 2M_0\bar\varepsilon_\nu \tag{6.71}
$$

(resp. $\mathcal{O}_*$ if $\nu = 0$). Let

$$
\eta_0 := \lambda = \gamma/M_0\,, \quad \eta_\nu := \gamma/(2^{\nu+3}M_0 K_{\nu-1}^{\tau+1})\,,\ \nu \ge 1\,. \tag{6.72}
$$

Let us define $\mathcal{O}_{\nu+1}$ as in (6.63) (with $i = \nu+1$) and set

$$
\tilde{\mathcal{O}}_{\nu+1} := \bigcup_{\xi \in \mathcal{O}_{\nu+1}} \Big\{\tilde\xi \in \mathbb{R}^{n/2} \ :\ \tilde\xi = \xi + \hat\xi\,,\ |\hat\xi| < \eta_\nu\Big\}\,. \tag{6.73}
$$

We claim that

$$
\tilde{\mathcal{O}}_1 \ \subseteq\ \mathcal{O}_* \quad \text{and} \quad \tilde{\mathcal{O}}_{\nu+1} \ \subseteq\ \mathcal{O}_\nu\,,\ \text{for}\ \ \nu \ge 1\,. \tag{6.74}
$$

The inclusion $\tilde{\mathcal{O}}_1 \subset \mathcal{O}_*$ follows by the definition (6.73) for $\tilde{\mathcal{O}}_1$, the inclusion $\mathcal{O}_1 \subset \mathcal{O}_0$ (see (6.63)), (6.72) and (6.68). Recalling (6.63), the inclusion $\tilde{\mathcal{O}}_{\nu+1} \subseteq \mathcal{O}_\nu$, for $\nu \ge 1$, follows if, for every $\tilde\xi = \xi + \hat\xi$, $\xi \in \mathcal{O}_{\nu+1}$, $|\hat\xi| \le \eta_\nu$, we prove that

$$
|\vec\omega^{(\nu-1)}(\tilde\xi)\cdot h| \ge (1 - 2^{-\nu})2\gamma\langle h\rangle^{-\tau}\,, \quad \forall\, 0 < |h| \le K_{\nu-1}\,, \tag{6.75}
$$

and the analogous estimates for $|\vec{\omega}^{(\nu-1)}(\tilde{\xi}) \cdot h + \Omega_j^{(\nu-1)}(\tilde{\xi})|$, $|\vec{\omega}^{(\nu-1)}(\tilde{\xi}) \cdot h \pm \Omega_j^{(\nu-1)}(\tilde{\xi}) + \Omega_{j'}^{(\nu-1)}(\tilde{\xi})|$, $|\vec{\omega}^{(\nu-1)}(\tilde{\xi}) \cdot h + p|$. By the expression (6.77) (at the previous step) for $\omega^{(\nu)}$, $\Omega^{(\nu)}$, and since $\chi_{\nu-1} \in [0,1]$, we have by definition of $\mathcal{O}_{\nu+1}$

$$\begin{aligned} |\vec{\omega}^{(\nu-1)}(\tilde{\xi}) \cdot h| & \geq |\vec{\omega}^{(\nu)}(\xi) \cdot h| - \big|\big(\vec{\omega}^{(\nu)}(\tilde{\xi}) - \vec{\omega}^{(\nu)}(\xi)\big) \cdot h\big| - \big|\big(\vec{\omega}^{(\nu-1)}(\tilde{\xi}) - \vec{\omega}^{(\nu)}(\tilde{\xi})\big) \cdot h\big| \\ & \overset{(6.61),(6.66)}{\geq} (1 - 2^{-\nu-1}) 2\gamma \langle h \rangle^{-\tau} - 2M_0 \eta_\nu K_{\nu-1} - 2\gamma \bar{\varepsilon}_{\nu-1} K_{\nu-1} \\ & \geq (1 - 2^{-\nu}) 2\gamma \langle h \rangle^{-\tau} \end{aligned}$$

since

$$2M_0 \eta_\nu + 2\gamma \bar{\varepsilon}_{\nu-1} \overset{(6.72),(\mathbf{S4}_\nu)}{\leq} 2^{-\nu-2} \gamma K_{\nu-1}^{-\tau-1} + 2\gamma C_\star \bar{\varepsilon}_0 e^{-K_* \chi^{\nu-1}} \overset{(6.56)}{\leq} 2^{-\nu} \gamma K_{\nu-1}^{-\tau-1}$$

taking $\bar{\varepsilon}_0$ small enough. This proves (6.75) (the other estimates are analogous) and so (6.74).

We define a smooth cut-off function $\chi_\nu : \mathbb{R}^{n/2} \to [0,1]$ which takes value 1 on $\mathcal{O}_{\nu+1}$ and value 0 outside $\tilde{\mathcal{O}}_{\nu+1}$, in particular χ_ν vanishes outside $\mathcal{O}_\nu$ (resp. $\mathcal{O}_*$ if $\nu = 0$) by (6.74). Recalling (6.72) we can construct χ_ν, $\nu \geq 0$, in such a way that

$$|\chi_\nu|^{\mathrm{lip}}_{\mathbb{R}^{n/2}} \lessdot \gamma^{-1} M_0 2^\nu K_{\nu-1}^{\tau+1} \tag{6.76}$$

(recall $K_{-1} := 1$). We extend $\hat{\omega}^{(\nu)}, \hat{\Omega}^{(\nu)}, \hat{a}^{(\nu)}$ to zero outside $\mathcal{O}_\nu$ for $\nu \geq 1$ and, for $\nu = 0$ outside $\mathcal{O}_*$. Then we define on the *whole* $\mathbb{R}^{n/2}$

$$\omega^{(\nu+1)} := \omega^{(\nu)} + \chi_\nu \hat{\omega}^{(\nu)}, \quad \Omega^{(\nu+1)} := \Omega^{(\nu)} + \chi_\nu \hat{\Omega}^{(\nu)}, \quad a^{(\nu+1)} := a^{(\nu)} + \chi_\nu \hat{a}^{(\nu)}. \tag{6.77}$$

The estimates (6.66) with $i = \nu + 1$ directly follow by (6.70). By (6.77), (6.74), (6.76), (6.71), (6.70), we get

$$|\omega^{(\nu+1)} - \omega^{(\nu)}|^{\mathrm{lip}}_{\mathbb{R}^{n/2}} \leq |\chi_\nu|^{\mathrm{lip}}_{\mathbb{R}^{n/2}} |\hat{\omega}^{(\nu)}|_{\mathbb{R}^{n/2}} + |\chi_\nu|_{\mathbb{R}^{n/2}} |\hat{\omega}^{(\nu)}|^{\mathrm{lip}}_{\mathcal{O}_\nu} \lessdot 2^\nu K_{\nu-1}^{\tau+1} M_0 \bar{\varepsilon}_\nu$$

proving the first estimate in (6.67) with $i = \nu + 1$. The other estimate is analogous.

Estimates (6.60),(6.61),(6.62) with $i = \nu + 1$ follow by $(\mathbf{S1})_\nu$ and by (6.66)-(6.67) with $i = \nu + 1$. This completes the proof of $(\mathbf{S1})_{\nu+1}$.

$(\mathbf{S2})_{\nu+1}$. We apply the KAM step Proposition 6.1 with $\vec{p} \rightsquigarrow \vec{p}_\nu$, $\vec{p}_+ \rightsquigarrow \vec{p}_{\nu+1}$, $X \rightsquigarrow X_\nu$, $X_+ \rightsquigarrow X_{\nu+1}$. The parameters defined in (6.56) (and $\tau > 1/b$) satisfy the conditions (6.35)-(6.36), for all $\nu \in \mathbb{N}$, taking γ small enough (recall (6.55)). Again by (6.56)

$$\delta_{\nu+1} \overset{(6.37)}{:=} \min\Big\{1 - \frac{s_{\nu+1}}{s_\nu}, 1 - \frac{r_{\nu+1}}{r_\nu}\Big\} \quad \text{so that} \quad 2^{-\nu-2} \leq \delta_{\nu+1} \leq 2^{-\nu-1}, \tag{6.78}$$

and the condition (6.41) is satisfied by $(S4)_\nu$, taking $\bar{\varepsilon}_0$ small enough and since $\Theta_0 \leq 1$, see (6.59). Hence Proposition 6.1 applies. Let $\Phi^{\nu+1}$ be the time-1-flow generated by the solution $F_{\nu+1} \in \mathcal{R}^{\leq 0}_{a-rev}$ of the homological equation (6.32). Then $\Phi^{\nu+1}_\star = e^{\mathrm{ad}_{F_{\nu+1}}}$. The flow $\Phi^{\nu+1}$ is even because $F_{\nu+1}$ is even. We have that (6.64) holds with $i = \nu+1$. The estimates (6.66)-(6.67) for $i = \nu+1$ have already been proved. We define $\bar{\varepsilon}_{\nu+1}, \varepsilon^{(-1)}_{\nu+1}, \varepsilon^{(0)}_{\nu+1}$ as in (6.65).

$(\mathbf{S3})_{\nu+1}$ follows by (6.42), (6.56) and (6.78).

$(\mathbf{S4})_{\nu+1}$. By $(S3)_\nu$ we can apply Lemma 6.5 and (6.52) implies $\bar{\varepsilon}_{\nu+1} \leq C_\star \bar{\varepsilon}_0 e^{-K_* \chi^{\nu+1}}$. Moreover

$$\Theta_{\nu+1} \overset{(6.43)}{\leq} \Theta_0 \Pi_{i=0}^\nu \Big(1 + C \delta_{i+1}^{-2} K_i^{4\tau+2} \bar{\varepsilon}_i\Big) \overset{(6.78),(S4)_\nu}{\leq} 2\Theta_0$$

for $\bar{\varepsilon}_0$ small enough. ■

Proof of the KAM Theorem 4.1 completed. Assumption (6.59) holds by (6.14) taking ε in (4.7) small enough. Then the iterative Lemma 6.6 applies. We define

$$\omega^\infty := \lim_{\nu \to \infty} \omega^{(\nu)}, \quad \Omega^\infty := \lim_{\nu \to \infty} \Omega^{(\nu)}, \quad a^\infty := \lim_{\nu \to \infty} a^{(\nu)}.$$

It could happen that $\mathcal{O}_{\nu_0} = \emptyset$ for some ν_0. In such a case $\mathcal{O}_\infty = \emptyset$ and the iterative process stops after finitely many steps. However, we can always set $\omega_\nu := \omega_{\nu_0}$, $\Omega_\nu := \Omega_{\nu_0}$, $a_\pm^{(\nu)} := a^{(\nu_0)}$, $\forall \nu \geq \nu_0$, and ω^∞, Ω^∞, a^∞ are always well defined and the bounds (4.8)-(4.10) hold. Indeed (4.8) follows by (6.60) recalling that $\mathcal{N}_0 = \mathcal{N}$. Then (4.9) follows by (5.2) because the normal forms $\mathcal{N}_\nu$ satisfy Definition 5.1. We now prove (4.10). For all $\forall \nu \geq 0$, $j \geq 6(N_0^{(\nu)} + 1)$, we have (recall that $a^{(0)} = 0$)

$$
\begin{aligned}
|\Omega_j^\infty - \Omega_j^{(0)} - a^\infty| &\leq \sum_{0 \leq i \leq \nu} |\Omega_j^{(i+1)} - a^{(i+1)} - \Omega_j^{(i)} + a^{(i)}| + \sum_{i > \nu} |\Omega_j^{(i+1)} - \Omega_j^{(i)}| + |a^{(i+1)} - a^{(i)}| \\
&\overset{(6.66)}{\leq} 40\gamma \sum_{0 \leq i \leq \nu} \frac{\bar{\varepsilon}_i}{j} + 4\gamma \sum_{i > \nu} \bar{\varepsilon}_i \overset{(\mathbf{S4})_\nu}{\lessdot} \frac{\bar{\varepsilon}_0 \gamma}{j} + \gamma \sum_{i > \nu} \bar{\varepsilon}_i \, .
\end{aligned}
$$

Therefore, $\forall \nu \geq 0$, $6(N_0^{(\nu)} + 1) \leq j < 6(N_0^{(\nu+1)} + 1)$,

$$
|\Omega_j^\infty - \Omega_j^{(0)} - a^\infty| \lessdot \frac{\bar{\varepsilon}_0 \gamma}{j} + \gamma \frac{N_0^{(\nu+1)}}{j} \sum_{i > \nu} \bar{\varepsilon}_i \overset{(6.56),(6.55)}{\lessdot} \frac{\bar{\varepsilon}_0 \gamma}{j} + \frac{\gamma}{j} \gamma^{-1/3} 2^{\rho(\nu+1)} \sum_{i > \nu} \bar{\varepsilon}_i
$$

and (4.10) follows by $(S4)_\nu$ and since $\Omega_j^{(0)} = \Omega_j$.

The transformation Φ in (4.12) is defined by

$$
\Phi := \lim_{\nu \to \infty} \bar{\Phi} \circ \Phi^0 \circ \Phi^1 \circ \cdots \circ \Phi^\nu \tag{6.79}
$$

where $\bar{\Phi}$ is defined in section 6.1 as the time 1-flow of F in (6.4). The map Φ is even because Φ^i, $i \geq 0$, and $\bar{\Phi}$ are even. We now verify that Φ is defined for parameters $\xi \in \mathcal{O}_\infty$ defined in (4.11).

Lemma 6.7. *$\mathcal{O}_\infty \subset \cap_{i \geq 0} \mathcal{O}_i$ (defined in* (6.63) *and* (6.58)*).*

Proof. We have $\mathcal{O}_\infty \subseteq \mathcal{O}_0$ by (4.11) and (6.58). For $i \geq 0$, if $\xi \in \mathcal{O}_\infty$ then, for all $|h| \leq K_i$,

$$
\begin{aligned}
|\vec{\omega}^{(i)}(\xi) \cdot h| &\geq |\vec{\omega}^\infty(\xi) \cdot h| - |h| \sum_{n \geq i} |\vec{\omega}^{(n+1)}(\xi) - \vec{\omega}^{(n)}(\xi)| \\
&\overset{(4.11),(6.66)}{\geq} 2\gamma \langle h \rangle^{-\tau} - K_i 2\gamma \sum_{n \geq i} \bar{\varepsilon}_n \geq (1 - 2^{-i}) 2\gamma \langle h \rangle^{-\tau}
\end{aligned}
$$

by the definition of K_i in (6.56), $(S4)_\nu$ and $\bar{\varepsilon}_0$ (that is ε) small enough. The other inequalities in (6.63) are verified analogously. ■

It remains to prove (4.13). By the definition of Φ in (6.79), the final vector field

$$
\mathcal{X}_\infty = \lim_{i \to \infty} \mathcal{X}_i \quad \text{where} \quad \mathcal{X}_i := \Phi_\star^i \mathcal{X}_{i-1} \, , \ i \geq 1 \, , \ \mathcal{X}_0 \text{ defined in (6.15)} \, . \tag{6.80}
$$

The vector field $\mathcal{X}_\infty \in \mathcal{R}_{rev}$ because $\mathcal{X}_0 \in \mathcal{R}_{rev}$ (see (6.15)) and each $\mathcal{X}_i \in \mathcal{R}_{rev}$ because $\Phi_\star^i = e^{\mathrm{ad}_{F_i}}$ with $F_i \in \mathcal{R}_{a-rev}$ (then use Lemma 2.10). Let

$$
X_\infty := \lim_{i \to \infty} X_i = \mathcal{N}_\infty + R_\infty \quad \text{where} \quad \mathcal{N}_\infty := \lim_{i \to \infty} \mathcal{N}_i \, , \quad R_\infty := \lim_{i \to \infty} R_i \, . \tag{6.81}
$$

By (6.64), $(S_4)_\nu$ and (6.65) we get

$$
R_\infty = \mathcal{S} R_\infty \, , \qquad R_\infty^{\leq 0} = 0 \, . \tag{6.82}
$$

Lemma 6.8. $(\mathcal{X}_\infty)_{|E} = (X_\infty)_{|E}$.

PROOF. The lemma follows by proving $(\mathcal{X}_i)_{|E} = (X_i)_{|E}$, $\forall\, i \geq 0$. The inductive basis for $i = 0$ is (6.16). Let us assume that $(\mathcal{X}_{i-1})_{|E} = (X_{i-1})_{|E}$. Then

$$(\mathcal{X}_i)_{|E} - (X_i)_{|E} \overset{(6.64),(6.80)}{=} (\Phi^i_\star \mathcal{X}_{i-1})_{|E} - (\mathcal{S}\Phi^i_\star X_{i-1})_{|E} = \Big(\Phi^i_\star(\mathcal{X}_{i-1} - X_{i-1})\Big)_{|E} \equiv 0$$

by Proposition 5.1 and Lemma 2.11 (used with $X \rightsquigarrow \mathcal{X}_{i-1} - X_{i-1}$, $Y \rightsquigarrow F_i$ defined in Lemma 6.6 so that $e^{\mathrm{ad}_{F_i}} = \Phi^i_\star$). ■

We have already chosen $\mathcal{N}_\infty$ in (6.81), then $\mathcal{P}_\infty$ in (4.13) is $\mathcal{P}_\infty = \mathcal{X}_\infty - \mathcal{N}_\infty$. It is now simple to show that $\big(\mathcal{P}^{\leq 0}_\infty\big)_{|E} = 0$. Indeed

$$\big(\mathcal{P}^{\leq 0}_\infty\big)_{|E} = \big((\mathcal{X}_\infty - \mathcal{N}_\infty)^{\leq 0}\big)_{|E} \overset{(6.81)}{=} \big((\mathcal{X}_\infty - X_\infty + R_\infty)^{\leq 0}\big)_{|E} \overset{(6.82)}{=} \big((\mathcal{X}_\infty - X_\infty)^{\leq 0}\big)_{|E} \overset{(2.60)}{\equiv} 0\,.$$

by Lemma 6.8. Finally $\mathcal{P}_\infty \in \mathcal{R}_{rev}$ because $\mathcal{N}_\infty \in \mathcal{R}_{rev}$ (obvious) and $\mathcal{X}_\infty \in \mathcal{R}_{rev}$. This concludes the proof of Theorem 4.1.

6.4 Measure estimates: proof of Theorem 4.2

Recalling the definitions of $\mathcal{O}$ and $\mathcal{O}_\infty$, given in (4.17) and (4.11) respectively, we have to estimate the measure of

$$\mathcal{O} \setminus \mathcal{O}_\infty = \mathcal{C}^1 \cup \mathcal{C}^{2+} \cup \mathcal{C}^{2-} \cup \tilde{\mathcal{C}} \cup (\mathcal{O} \setminus \mathcal{O}_0) \tag{6.83}$$

where

$$\mathcal{O}_0 := \Big\{\xi \in \mathcal{O} \ :\ |\vec{\omega}(\xi) \cdot h| \geq 2\gamma^{2/3}\langle h\rangle^{-n/2}\,,\ \ \forall 0 < |h| < \gamma^{-1/7n}\Big\}$$

$$\mathcal{C}^1 := \bigcup_{h\in\mathbb{Z}^{n/2}, j\in\mathbb{Z}\setminus\mathcal{I}} \mathcal{R}^{1,\tau}_{h,j}(\gamma)\,, \qquad \mathcal{R}^{1,\tau}_{h,j}(\gamma) := \Big\{\xi \in \mathcal{O} \,:\, |\vec{\omega}^\infty(\xi)\cdot h + \Omega^\infty_j| < 2\gamma\langle h\rangle^{-\tau}\Big\}$$

$$\mathcal{C}^{2+} := \bigcup_{h\in\mathbb{Z}^{n/2}, i,j\in\mathbb{Z}\setminus\mathcal{I}} \mathcal{R}^{2+,\tau}_{h,i,j}(\gamma)\,, \quad \mathcal{R}^{2+,\tau}_{h,i,j}(\gamma) := \Big\{\xi \in \mathcal{O} \,:\, |\vec{\omega}^\infty(\xi)\cdot h + \Omega^\infty_i(\xi) + \Omega^\infty_j(\xi)| < 2\gamma\langle h\rangle^{-\tau}\Big\}$$

$$\mathcal{C}^{2-} := \bigcup_{\substack{h\in\mathbb{Z}^{n/2}, i,j\in\mathbb{Z}\setminus\mathcal{I},\\ \text{if}\, i=\pm j,\, h\neq 0}} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\,, \quad \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma) := \Big\{\xi \in \mathcal{O} \,:\, |\vec{\omega}^\infty(\xi)\cdot h + \Omega^\infty_i(\xi) - \Omega^\infty_j(\xi)| < 2\gamma\langle h\rangle^{-\tau}\Big\}$$

$$\tilde{\mathcal{C}} := \bigcup_{\substack{h\in\mathbb{Z}^{n/2}, p\in\mathbb{Z}\\ (h,p)\neq(0,0)}} \tilde{\mathcal{R}}^\tau_{hp}(\gamma^{2/3})\,, \quad \tilde{\mathcal{R}}^\tau_{hp}(\gamma^{2/3}) := \Big\{\xi \in \mathcal{O} \,:\, |\vec{\omega}^\infty(\xi)\cdot h + p| < 2\gamma^{2/3}\langle h\rangle^{-\tau}\Big\}\,.$$

We first consider the most difficult estimate

$$|\mathcal{C}^{2-}| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}\,. \tag{6.84}$$

Lemma 6.9. *If $\mathcal{R}^{2-,\tau}_{h,i,j}(\gamma) \neq \emptyset$ then*

$$||i| - |j|| \leq C\,\langle h\rangle\,. \tag{6.85}$$

PROOF. By the definition of $\mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)$, (4.8) and (4.18). ■

As a consequence, we have to estimate

$$\mathcal{C}^{2-} = \bigcup_{(h,i,j)\in\mathtt{I}} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)$$

where

$$\mathtt{I} := \Big\{(h,i,j) \in \mathbb{Z}^{n/2} \times (\mathbb{Z}\setminus\mathcal{I})^2 \,:\, (6.85) \text{ holds and, if } i = \pm j\,, \text{ then } h \neq 0\Big\}.$$

By (4.18), (4.19), (4.8), for ε small enough, $\vec{\omega}^\infty : \mathcal{O} \to \vec{\omega}^\infty(\mathcal{O})$ is invertible,

$$\xi = (\vec{\omega}^\infty)^{-1}(\zeta) = A^{-1}(\zeta - \bar{\omega}) + r_\varepsilon(\zeta)\,, \quad |(\vec{\omega}^\infty)^{-1}|^{\mathrm{lip}} \leq 2\|A^{-1}\| \tag{6.86}$$

where $|r_\varepsilon| = O(\varepsilon\gamma)$, $|r_\varepsilon|^{\mathrm{lip}} = O(\varepsilon)$.

Lemma 6.10. *For* $(h,i,j) \in \mathtt{I}$, $0 < \eta < \eta_0(\mathrm{m})$ *small enough, we have*

$$|\mathcal{R}^{2-,\tau}_{h,i,j}(\eta)| \lessdot \eta \rho^{\frac{n}{2}-1} \langle h \rangle^{-\tau} \, . \tag{6.87}$$

Proof. By (4.8), (4.18), we have

$$\vec{\omega}^\infty(\xi) \cdot h + \Omega^\infty_i(\xi) - \Omega^\infty_j(\xi) = \vec{\omega}^\infty(\xi) \cdot h + \lambda_i - \lambda_j + \vec{a} \cdot \xi (\lambda_i^{-1} - \lambda_j^{-1}) + r_{i,j}(\xi) \tag{6.88}$$

where

$$|r_{i,j}(\xi)| = O(\varepsilon\gamma) \, , \ |r_{i,j}|^{\mathrm{lip}} = O(\varepsilon) \, . \tag{6.89}$$

Case 1: $h = 0$, $i \neq \pm j$. By (6.88), (6.89), we have, $\forall \xi \in \mathcal{O}$ (see (4.17)), $\forall i \neq \pm j$,

$$|\vec{\omega}^\infty(\xi) \cdot h + \Omega^\infty_i(\xi) - \Omega^\infty_j(\xi)| \geq |\lambda_i - \lambda_j| - C|\xi| - C\varepsilon\gamma \geq \bar{c} > 2\eta \langle h \rangle^{-\tau} \, ,$$

hence $\mathcal{R}^{2-,\tau}_{h,i,j}(\eta) = \emptyset$.

Case 2: $h \neq 0$. Introducing the final frequencies $\zeta := \vec{\omega}^\infty(\xi) \in \mathbb{R}^{n/2}$ as parameters in (6.88) we get

$$\begin{aligned} f_{h,i,j}(\zeta) \quad &:= \quad \vec{\omega}^\infty(\xi) \cdot h + \Omega^\infty_i(\xi) - \Omega^\infty_j(\xi) \\ &\overset{(6.86)}{=} \quad \Big(h + (A^T)^{-1} \vec{a} (\lambda_i^{-1} - \lambda_j^{-1}) \Big) \cdot \zeta - A^{-1} \bar{\omega}_+ \cdot \vec{a} (\lambda_i^{-1} - \lambda_j^{-1}) + \lambda_i - \lambda_j + \tilde{r}_{i,j}(\zeta) \end{aligned}$$

where $\tilde{r}_{i,j}$ satisfies (6.89) as well. By (4.19) and (4.22) the vector

$$\tilde{a} := h + (A^T)^{-1} \vec{a} (\lambda_i^{-1} - \lambda_j^{-1}) \quad \text{satisfies} \quad |\tilde{a}| \geq c = c(A, \vec{a}) > 0 \, , \ \forall h \neq 0, \ \forall i,j \in \mathbb{Z} \setminus \mathcal{I} \, . \tag{6.90}$$

In the direction $\zeta = s\tilde{a}|\tilde{a}|^{-1} + w$, $w \cdot \tilde{a} = 0$, the function $\tilde{f}_{k,i,j}(s) := f_{k,i,j}(s\tilde{a}|\tilde{a}|^{-1} + w)$ satisfies

$$\tilde{f}_{k,i,j}(s_2) - \tilde{f}_{k,i,j}(s_1) \overset{(6.89)}{\geq} (s_2 - s_1)(|\tilde{a}| - C\varepsilon) \geq (s_2 - s_1)|\tilde{a}|/2 \, .$$

By Fubini theorem we get $|\{\zeta \in \vec{\omega}^\infty(\mathcal{O}) \, : \, |f_{h,i,j}(\zeta)| \leq 2\eta \langle h \rangle^{-\tau}\}| \lessdot \eta \rho^{\frac{n}{2}-1} \langle h \rangle^{-\tau}$ which implies (6.87) thanks to (6.86). ■

We split

$$\mathtt{I} = \mathtt{I}_> \cup \mathtt{I}_< \quad \text{where} \quad \mathtt{I}_> := \Big\{ (h,i,j) \in \mathtt{I} \, : \, \min\{|i|,|j|\} > C_\sharp \gamma^{-1/3} \langle h \rangle^{\tau_0} \Big\} \tag{6.91}$$

where $C_\sharp > C_\star$ in (4.10) for $\tau_0 := 2 + (n/2)$. We set $\mathtt{I}_< := \mathtt{I} \setminus \mathtt{I}_>$.

Lemma 6.11. *For all* $(h,i,j) \in \mathtt{I}_>$ *we have*

$$\mathcal{R}^{2-,\tau_0}_{h,i,j}(\gamma^{2/3}) \subset \mathcal{R}^{2-,\tau_0}_{h,i_0,j_0}(2\gamma^{2/3}) \tag{6.92}$$

where $i_0, j_0 \in \mathbb{Z} \setminus \mathcal{I}$ *satisfy*

$$|i_0| - |j_0| = |i| - |j| \qquad \text{and} \qquad \min\{|j_0|, |i_0|\} \geq C_\sharp \gamma^{-1/3} \langle h \rangle^{\tau_0} \, . \tag{6.93}$$

Proof. Since $|j| \geq \gamma^{-1/3} C_\star$, by (4.10) and (4.18) we have the frequency asymptotic

$$\Omega^\infty_j(\xi) = |j| + \frac{\mathrm{m}}{2|j|} + \vec{a} \cdot \xi \lambda_j^{-1} + \hat{a}^\infty(\xi) + O\left(\frac{1}{|j|^3}\right) + O\left(\varepsilon \frac{\gamma^{2/3}}{|j|}\right) \, . \tag{6.94}$$

By (6.93) and (6.85) we have $||i|-|j|| = ||i_0|-|j_0|| \le C\langle h\rangle$. If $\xi \in \mathcal{O}\setminus \mathcal{R}^{2-,\tau_0}_{h,i_0,j_0}(2\gamma^{2/3})$, since $|i|,|j|,|i_0|,|j_0| \ge \mu_0 := C_\sharp \gamma^{-1/3}\langle h\rangle^{\tau_0}$ (recall (6.91) and (6.93)), we have

$$\begin{aligned}
|\omega^\infty_+(\xi)\cdot h + \Omega^\infty_i(\xi) - \Omega^\infty_j(\xi)| \quad & \geq \quad |\omega^\infty_+(\xi)\cdot k + \Omega^\infty_{i_0}(\xi) - \Omega^\infty_{j_0}(\xi)| \\
& \quad\; -|\Omega^\infty_i(\xi) - \Omega^\infty_{i_0}(\xi) - \Omega^\infty_j(\xi) + \Omega^\infty_{j_0}(\xi)| \\
& \overset{(6.94)}{\geq} \frac{4\gamma^{2/3}}{\langle h\rangle^{\tau_0}} - ||i|-|i_0|-|j|+|j_0|| \\
& \quad\; -|\vec a\cdot\xi(\lambda_j^{-1}-\lambda_i^{-1})| + |\vec a\cdot\xi(\lambda_{j_0}^{-1}-\lambda_{i_0}^{-1})| \\
& \quad\; -C\varepsilon\frac{\gamma^{2/3}}{\mu_0} - \frac{C}{\mu_0^3} - \frac{\mathrm{m}}{2}\frac{||i|-|j||}{|i|\,|j|} - \frac{\mathrm{m}}{2}\frac{||i_0|-|j_0||}{|i_0|\,|j_0|} \\
& \overset{(6.93),(6.85)}{\geq} \frac{4\gamma^{2/3}}{\langle h\rangle^{\tau_0}} - C\varepsilon\frac{\gamma^{2/3}}{\mu_0} - C\frac{\langle h\rangle}{\mu_0^2} \overset{(6.93)}{\geq} \frac{2\gamma^{2/3}}{\langle h\rangle^{\tau_0}}
\end{aligned}$$

taking $C_\sharp$ in (6.93) large enough, and having used $|\lambda_j^{-1}-\lambda_i^{-1}| \le ||i|-|j|||i|^{-1}|j|^{-1}$. Therefore $\xi\in\mathcal{O}\setminus\mathcal{R}^{\tau_0}_{h,i,j}(\gamma^{2/3})$ proving (6.92). ■

As a corollary we deduce:

Lemma 6.12. $\Big|\bigcup_{(h,i,j)\in \mathtt{I}_>} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\Big| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}$.

Proof. Since $0<\gamma\le 1$ and $\tau\ge\tau_0$ (see (4.20)), we have $\mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\subset \mathcal{R}^{2-,\tau_0}_{h,i,j}(\gamma^{2/3})$. Then Lemma 6.11 and (6.87) imply that, for each $p\in\mathbb{Z}$,

$$\Big|\bigcup_{(h,i,j)\in \mathtt{I}_>,\,|i|-|j|=p} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\Big| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}\langle h\rangle^{-\tau_0}.$$

Therefore

$$\Big|\bigcup_{(h,i,j)\in \mathtt{I}_>} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\Big| \lessdot \sum_{h,|p|\le C\langle h\rangle} \gamma^{2/3}\rho^{\frac{n}{2}-1}\langle h\rangle^{-\tau_0} \lessdot \sum_{h\in\mathbb{Z}^{n/2}} \gamma^{2/3}\rho^{\frac{n}{2}-1}\langle h\rangle^{-\tau_0+1}$$

(since $\tau_0 > 1+(n/2)$) proving the lemma. ■

Lemma 6.13. $\Big|\bigcup_{(h,i,j)\in \mathtt{I}_<} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\Big| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}$.

Proof. For all $(h,i,j)\in\mathtt{I}_<$ we have (see (6.85))

$$\min\{|i|,|j|\} < C\gamma^{-1/3}\langle h\rangle^{\tau_0}\,,\ ||i|-|j||\le C\langle h\rangle \implies \max\{|i|,|j|\} < C'\gamma^{-1/3}\langle h\rangle^{\tau_0}.$$

Therefore, using also Lemma 6.10 and (6.85)

$$\Big|\bigcup_{(h,i,j)\in \mathtt{I}_<} \mathcal{R}^{2-,\tau}_{h,i,j}(\gamma)\Big| \ \lessdot\ \sum_h \sum_{|i|\le C'\gamma^{-1/3}\langle h\rangle^{\tau_0}} \sum_{||i|-|j||\le C\langle h\rangle} \frac{\gamma\rho^{\frac{n}{2}-1}}{\langle h\rangle^\tau} \lessdot \sum_h \frac{\gamma^{2/3}\rho^{\frac{n}{2}-1}}{\langle h\rangle^{\tau-\tau_0-1}}$$

which, by (4.20) and $\tau_0 := 2+(n/2)$, gives the lemma. ■

Lemmata 6.12, 6.13 imply (6.84). Let us consider the other cases. By (4.19), arguing as in Lemma 6.10, we get that for $0<\gamma<\gamma_0(\mathrm{m})$ small, the measure

$$|\mathcal{R}^{1,\tau}_{h,j}(\gamma)|,\ |\mathcal{R}^{2+,\tau}_{h,i,j}(\gamma)| \lessdot \gamma\rho^{\frac{n}{2}-1}\langle h\rangle^{-\tau}, \quad |\tilde{\mathcal{R}}^\tau_{h,p}(\gamma^{2/3})| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}\langle h\rangle^{-\tau},$$

and, by standard arguments,

$$\Big|\mathcal{C}^1\cup\mathcal{C}^{2+}\Big| \lessdot \gamma\rho^{\frac{n}{2}-1}, \quad |\tilde{\mathcal{C}}| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}, \quad |\mathcal{O}\setminus\mathcal{O}_0| \lessdot \gamma^{2/3}\rho^{\frac{n}{2}-1}. \tag{6.95}$$

Finally (6.83), (6.84), (6.95) imply (4.21).

7 Proof of Theorem 1.1

By hypothesis the nonlinearity $g(\mathtt{x},\mathtt{y},\mathtt{y_x},\mathtt{v}) = \mathtt{y}\mathtt{y}_\mathtt{x}^2 +$ h.o.t. has the convergent Taylor expansion

$$g(\mathtt{x},\mathtt{y},\mathtt{y}_x,\mathtt{v}) = \mathtt{y}\mathtt{y}_\mathtt{x}^2 + \sum_{k,h,l\in\mathbb{N}\,:\,k+h+l>3} g^{(k,h,l)}(\mathtt{x})\mathtt{y}^k\mathtt{y}_\mathtt{x}^h\mathtt{v}^l$$

where

$$\|g^{(k,h,l)}\|_{a_0,p} < C^{k+h+l} \quad \text{for some} \quad a_0>0\,,\ p>1/2\,. \tag{7.1}$$

Note that we are identifying the functions $g^{(k,h,l)}(\mathtt{x})$ with their Fourier series $\{g^{(k,h,l)}_{j_0}\}_{j_0\in\mathbb{Z}} \in \ell^{a_0,p}$ (recall (2.1)). Let

$$a := a_0/2\,, \qquad \{u_j^\pm\}_{j\in\mathbb{Z}} \in \ell^{a,p}\,. \tag{7.2}$$

The function $\mathtt{g}$ in (1.16) is

$$\begin{aligned}\mathtt{g}(u^+,u^-) &= -\frac{1}{4}\Big(D^{-1}(u^+ + u^-)\Big)\Big(D^{-1}(u^+_\mathtt{x} + u^-_\mathtt{x})\Big)^2 + O_5\\ &= \frac{1}{4}\sum_{\substack{\sigma_1,\sigma_2,\sigma_3=\pm\,,\\ j_1,j_2,j_3\in\mathbb{Z}}} \sigma_2\sigma_3\frac{j_2j_3}{\lambda_{j_1}\lambda_{j_2}\lambda_{j_3}}u^{\sigma_1}_{j_1}u^{\sigma_2}_{j_2}u^{\sigma_3}_{j_3}e^{\mathrm{i}(\sigma_1j_1+\sigma_2j_2+\sigma_3j_3)\mathtt{x}} + O_5\,. \end{aligned}\tag{7.3}$$

and $\mathtt{g}_j^+$ in (1.25) is

$$\mathtt{g}_j^+ := \mathtt{g}_j = -\sum_{d\geq3}\sum_{\substack{\sigma_1,\ldots,\sigma_d=\pm1\,,\\ j_0,j_1,\ldots j_d\in\mathbb{Z}\,,\\ j_0+\sum_{i=1}^d\sigma_ij_i=j}} (\sqrt{2})^{-d-1}\mathtt{g}_{\vec\sigma,\vec\jmath,j_0}u^{\vec\sigma}_{\vec\jmath} \tag{7.4}$$

$$= \frac{1}{4}\sum_{\sigma_1j_1+\sigma_2j_2+\sigma_3j_3=j}\frac{j_2j_3}{\lambda_{j_1}\lambda_{j_2}\lambda_{j_3}}\sigma_2\sigma_3\,u^{\sigma_1}_{j_1}u^{\sigma_2}_{j_2}u^{\sigma_3}_{j_3} - \sum_{d>3}\sum_{\substack{\sigma_1,\ldots,\sigma_d=\pm1\,,\\ j_0,j_1,\ldots j_d\in\mathbb{Z}\,,\\ j_0+\sum_{i=1}^d\sigma_ij_i=j}}(\sqrt{2})^{-d-1}\mathtt{g}_{\vec\sigma,\vec\jmath,j_0}u^{\vec\sigma}_{\vec\jmath}$$

$$=: \mathtt{g}_j^{(=3)} + \mathtt{g}_j^{(\geq5)}\,, \tag{7.5}$$

where $\vec\jmath = (j_1,\ldots,j_d)$, $\vec\sigma = (\sigma_1,\ldots,\sigma_d)$ and $u^{\vec\sigma}_{\vec\jmath} = \prod_{i=1}^d u^{\sigma_i}_{j_i}$. The coefficients $g_{\vec\sigma,\vec\jmath,j_0}$ are explicitly

$$g_{\vec\sigma,\vec\jmath,j_0} = \sum_{k,h,l\in\mathbb{N}\,:\,h+k+l=d}(-1)^l\mathrm{i}^{h+l}\sigma_{k+1}\cdots\sigma_{k+h+l}\frac{j_{k+1}\cdots j_{k+h}}{\lambda_{j_1}\cdots\lambda_{j_{k+h}}}g^{(k,h,l)}_{j_0}\,.$$

We consider (1.23) as the equations of motion of the vector field $\mathcal{N}_0 + G$ where

$$\mathcal{N}_0 := \sum_{\sigma=\pm,\,j\in\mathbb{Z}}\sigma\mathrm{i}\lambda_ju_j^\sigma\partial_{u_j^\sigma} \tag{7.6}$$

is the linear *normal form* vector field, and (recall (1.25))

$$\begin{aligned} G &= \sum_{\sigma=\pm,j\in\mathbb{Z}}G^{(u_j^\sigma)}\partial_{u_j^\sigma}\,, \quad G^{(u_j^\sigma)} := \mathrm{i}\sigma\mathtt{g}_{\sigma j}\,, \qquad G = G^{(=3)} + G^{(\geq5)}\,, \\ G^{(=3)} &= \sum_{\sigma=\pm,j\in\mathbb{Z}}\mathrm{i}\sigma\mathtt{g}^{(=3)}_{\sigma j}\partial_{u_j^\sigma}\,, \qquad G^{(\geq5)} = \sum_{\sigma=\pm,j\in\mathbb{Z}}\mathrm{i}\sigma\mathtt{g}^{(\geq5)}_{\sigma j}\partial_{u_j^\sigma}\end{aligned}\tag{7.7}$$

is a nonlinear perturbation. Note that

$$G^{(u^+_{-j})} = -G^{(u_j^-)} \tag{7.8}$$

and that $G^{(=3)}$ *has zero momentum* by (7.4) and (7.7). Moreover, by (1.26), (1.28), (1.30), (1.31) we have

Lemma 7.1. *G is reversible (w.r.t. the involution S in (1.29)), real-coefficients, real-on-real, even, namely $G \in \mathcal{R}_{rev}$ (recall Definition 2.11 in absence of x, y-variables).*

Lemma 7.2. *For $a_0 - a \geq \mathtt{a} := a_0/2$ (recall (7.2)) and $R > 0$ small enough*

$$\|G\|_{R,\mathtt{a}}\,, \quad \|G^{(=3)}\|_{R,\mathtt{a}} \lessdot R^2\,, \qquad\qquad \|G^{(\geq 5)}\|_{R,\mathtt{a}} \lessdot R^4\,. \tag{7.9}$$

Moreover for N_0 satisfying (3.1) we have that $G, G^{(=3)}, G^{(\geq 5)}$ belong to $\mathcal{Q}^T_{R,\mathtt{a}}(N_0, 3/2, 4)$ with

$$\|G\|^T_{R,\mathtt{a},N_0,3/2,4}\,, \quad \|G^{(=3)}\|^T_{R,\mathtt{a},N_0,3/2,4} \lessdot R^2 \qquad\qquad \|G^{(\geq 5)}\|^T_{R,\mathtt{a},N_0,3/2,4} \lessdot R^4\,. \tag{7.10}$$

Proof. We first note that for $d, h, k, l \in \mathbb{N}$

$$\Big\|\Big(\sum_{\substack{\sigma_1,\ldots,\sigma_d=\pm 1\,,\\ j_0,j_1,\ldots j_d \in \mathbb{Z}\,,\\ j_0+\sum_{i=1}^d \sigma_i j_i = j}} e^{\mathtt{a}|j_0|} |g^{(k,h,l)}_{j_0}| |u^{\vec\sigma}_{\vec j}| \Big)_{j\in\mathbb{Z}} \Big\|_{a,p} \lessdot \|g^{(k,h,l)}\|_{a_0,p} (\|u\|_{a,p} + \|\bar u\|_{a,p})^d\,. \tag{7.11}$$

Indeed

$$\sum_{\substack{\sigma_1,\ldots,\sigma_d=\pm 1\,,\\ j_0,j_1,\ldots j_d \in \mathbb{Z}\,,\\ j_0+\sum_{i=1}^d \sigma_i j_i = j}} e^{\mathtt{a}|j_0|} |g^{(k,h,l)}_{j_0}| |u^{\vec\sigma}_{\vec j}| \leq \big(\tilde g^{(k,h,l)} * \tilde u * \tilde u * \cdots * \tilde u\big)_j\,, \qquad \forall\, j \in \mathbb{Z}\,,$$

where $\tilde g^{(k,h,l)} := (e^{\mathtt{a}|j_0|} g^{(k,h,l)}_{j_0})_{j_0\in\mathbb{Z}}$, $\tilde u := (\tilde u_n)_{n\in\mathbb{Z}}$, $\tilde u_n := |u_n| + |\bar u_n|$, $*$ denotes the convolution of sequences and

$$\|\tilde g^{(k,h,l)} * \tilde u * \tilde u * \cdots * \tilde u\|_{a,p} \lessdot \|\tilde g^{(k,h,l)}\|_{a,p} \|\tilde u\|^d_{a,p} \lessdot \|g^{(k,h,l)}\|_{a_0,p} (\|u\|_{a,p} + \|\bar u\|_{a,p})^d$$

by the Hilbert algebra property of $\ell^{a,p}$ and since $\mathtt{a} \leq a_0 - a$.

Now we can rewrite the sum in (7.4) as $\mathtt{g}_j = \sum_{|\alpha|+|\beta|\geq 3} (\mathtt{g}_j)_{\alpha,\beta} u^\alpha \bar u^\beta$ where $(\mathtt{g}_j)_{\alpha,\beta}$ can be explicitly computed from (7.4) but has a complicated combinatorics. In order to compute the norm $\|G\|_{R,\mathtt{a}}$ we note that

$$\frac{1}{|\lambda_l|} \lessdot 1\,, \quad \frac{|l|}{|\lambda_l|} \leq 1 \tag{7.12}$$

and the momentum of

$$u^{\vec\sigma}_{\vec j} = u^\alpha \bar u^\beta \quad \Longrightarrow \quad \pi(\alpha, \beta; u^\sigma_j) = \sum_{1\leq i\leq d} \sigma_i j_i - \sigma j\,. \tag{7.13}$$

For all $\mathtt{a} \geq 0$ and $R > 0$ we have (recall (7.7))

$$\begin{aligned}
\|G\|_{R,\mathtt{a}} &\overset{(2.27)}{=} \sup_{\|u\|_{a,p}, \|\bar u\|_{a,p} < R} \Big\|\Big(\sum_{|\alpha|+|\beta|\geq 3} e^{\mathtt{a}|\pi(\alpha,\beta;u^\sigma_j)|} |(g_{\sigma j})_{\alpha,\beta}| |u^\alpha| |\bar u^\beta| \Big)_{\sigma=\pm, j\in\mathbb{Z}} \Big\|_R \\
&\overset{(2.3)}{=} \sup_{\|u\|_{a,p}, \|\bar u\|_{a,p} < R} \frac{1}{R} \sum_{\sigma=\pm} \Big\|\Big(\sum_{|\alpha|+|\beta|\geq 3} e^{\mathtt{a}|\pi(\alpha,\beta;u^\sigma_j)|} |(g_{\sigma j})_{\alpha,\beta}| |u^\alpha| |\bar u^\beta| \Big)_{j\in\mathbb{Z}} \Big\|_{a,p} \\
&\overset{(7.13)}{\lessdot} \frac{1}{R} \sup_{\|u\|_{a,p}, \|\bar u\|_{a,p} < R} \Big\|\Big(\sum_{d\geq 3} (\sqrt2)^{-d-1} \sum_{\substack{\sigma_1,\ldots,\sigma_d=\pm 1\,,\\ j_0,j_1,\ldots j_d \in \mathbb{Z}\,,\\ j_0+\sum_{i=1}^d \sigma_i j_i = j}} \sum_{\substack{k,h,l\in\mathbb{N}:\\ h+k+l=d}} e^{\mathtt{a}|j_0|} |g^{(k,h,l)}_{j_0}| |u^{\vec\sigma}_{\vec j}| \Big)_{j\in\mathbb{Z}} \Big\|_{a,p} \\
&\overset{(7.11)}{\lessdot} \frac{1}{R} \sup_{\|u\|_{a,p}, \|\bar u\|_{a,p} < R} \sum_{d\geq 3} \sum_{\substack{k,h,l\\ h+k+l=d}} (\sqrt2)^{-d-1} \|g^{(k,h,l)}\|_{a_0,p} (\|u\|_{a,p} + \|\bar u\|_{a,p})^d \overset{(7.1)}{\lessdot} R^d
\end{aligned}$$

proving (7.9) for $R > 0$ small enough.

Let us now prove the estimate (7.10) for the quasi-Töplitz norm of G (the estimate for $G^{(=3)}$ and $G^{(\geq 5)}$ are analogous). For $N \geq N_0$, by (7.4) and (7.7) we deduce that the linear projection

$$\Pi_{N,3/2,4} G = \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>(3/2)N}} G^{\sigma,m}_{\sigma',n} \, u_n^{\sigma'} \partial_{u_m^\sigma} = \tilde{G} + N^{-1} \hat{G}$$

where (recall (3.10), (3.11), (3.12))

$$G^{\sigma,m}_{\sigma',n} := -\mathrm{i}\sigma \sum_{d \geq 2} \sum_{\substack{\sigma_1,\ldots,\sigma_d = \pm 1, \\ \sum_{i=1}^d |j_i| < 4N^L, |j_0| < N^b \\ j_0 + \sum_{i=1}^d \sigma_i j_i = \sigma m - \sigma' n}} (\sqrt{2})^{-d-2} \mathtt{g}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} u^{\vec{\sigma}}_{\vec{\jmath}}$$

$$\mathtt{g}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} = \sum_{k,h,l \in \mathbb{N} \,:\, h+k+l = d+1} (\mathrm{i})^{h+l} (-1)^l g^{(k,h,l)}_{j_0} \Big(k \frac{1}{\lambda_n} \frac{j_k \ldots j_{k+h-1} \sigma_k \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h-1}}} + \tag{7.14}$$

$$+ h\sigma' \frac{n}{\lambda_n} \frac{j_{k+1} \ldots j_{k+h-1} \sigma_{k+1} \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h-1}}} + \tag{7.15}$$

$$+ l\sigma' \frac{j_{k+1} \ldots j_{k+h} \sigma_{k+1} \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h}}} \Big), \tag{7.16}$$

the Töplitz approximation is

$$\tilde{G} := \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>(3/2)N}} \tilde{G}^{\sigma,m}_{\sigma',n} \, u_n^{\sigma'} \partial_{u_m^\sigma}, \qquad \tilde{G}^{\sigma,m}_{\sigma',n} := -\mathrm{i}\sigma \sum_{d \geq 2} \sum_{\substack{\sigma_1,\ldots,\sigma_d = \pm 1, \\ \sum_{i=1}^d |j_i| < 4N^L, |j_0| < N^b \\ j_0 + \sum_{i=1}^d \sigma_i j_i = \sigma m - \sigma' n}} (\sqrt{2})^{-d-2} \tilde{\mathtt{g}}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} u^{\vec{\sigma}}_{\vec{\jmath}}$$

$$\tilde{\mathtt{g}}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} = \sum_{k,h,l \in \mathbb{N} \,:\, h+k+l = d+1} (\mathrm{i})^{h+l} (-1)^l g^{(k,h,l)}_{j_0} \Big(h\sigma' \mathtt{s}(n) \frac{j_{k+1} \ldots j_{k+h-1} \sigma_{k+1} \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h-1}}}$$
$$+ \; l\sigma' \frac{j_{k+1} \ldots j_{k+h} \sigma_{k+1} \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h}}} \Big)$$

(the term in (7.14) is replaced by 0, in (7.15) the factor n/λ_n replaced by the sign $\mathtt{s}(n)$, and (7.16) is left unchanged) and the corresponding Töplitz defect is

$$\hat{G} := \frac{N}{\lambda_n} \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>(3/2)N}} \hat{G}^{\sigma,m}_{\sigma',n} \, u_n^{\sigma'} \partial_{u_m^\sigma}, \quad \hat{G}^{\sigma,m}_{\sigma',n} := -\mathrm{i}\sigma \sum_{d \geq 2} \sum_{\substack{\sigma_1,\ldots,\sigma_d = \pm 1, \\ \sum_{i=1}^d |j_i| < 4N^L, |j_0| < N^b \\ j_0 + \sum_{i=1}^d \sigma_i j_i = \sigma m - \sigma' n}} (\sqrt{2})^{-d-2} \hat{\mathtt{g}}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} u^{\vec{\sigma}}_{\vec{\jmath}}$$

$$\hat{\mathtt{g}}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0} = \sum_{k,h,l \in \mathbb{N} \,:\, h+k+l = d+1} \mathrm{i}^{h+l} (-1)^l g^{(k,h,l)}_{j_0} \Big(k \frac{j_k \ldots j_{k+h-1} \sigma_k \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h-1}}} +$$
$$+ \; h\sigma' \mathtt{s}(n)(|n| - \lambda_n) \frac{j_{k+1} \ldots j_{k+h-1} \sigma_{k+1} \ldots \sigma_{k+h+l-1}}{\lambda_{j_1} \ldots \lambda_{j_{k+h-1}}} \Big).$$

Using that $0 \leq \lambda_n - |n| \leq c(\mathtt{m}) = \sqrt{\mathtt{m}}$ for all $n \in \mathbb{Z}$ and that $\lambda_n \geq |n| > (3/2)N$ we have that the Taylor coefficients of $\tilde{G}, \hat{G}$ are uniformly bounded. Then, arguing as in the proof of (7.9), for $R > 0$ small enough, we deduce that

$$\|\tilde{G}\|_{R,\mathtt{a}}, \; \|\hat{G}\|_{R,\mathtt{a}} \lessdot R^2.$$

Note that $\tilde{\mathtt{g}}^{\sigma',n}_{\vec{\sigma},\vec{\jmath},j_0}$ depends on n only through $\mathtt{s}(n)$. Since by (3.12) $\mathtt{s}(n) = \sigma\sigma' \mathtt{s}(m)$ we have that $\tilde{G} \in \mathcal{T}_{R,\mathtt{a}}(N, 3/2, 4)$ (recall Definition 3.3). By Definition 3.4 we get (7.10). ■

Proposition 7.1. (Birkhoff normal form) *For any $\mathcal{I}$ as in* (1.32)*, and* $\mathrm{m} > 0$*, there exists* $R_0 > 0$ *and a real analytic change of variables*

$$\Gamma : B_{R/2} \times B_{R/2} \subset \ell^{a,p} \times \ell^{a,p} \ \to \ B_R \times B_R \subset \ell^{a,p} \times \ell^{a,p} , \quad 0 < R < R_0 ,$$

that takes the vector field $\mathcal{N}_0 + G$ *into*

$$\Big(D\Gamma^{-1}[\mathcal{N}_0 + G] \Big) \circ \Gamma = \mathcal{N}_0 + G_1 + G_2 + G_3 \tag{7.17}$$

where G_1, G_2, G_3 *satisfy* (7.8)*,*

$$\sigma 4\mathrm{i}\, G_1^{(u_j^\sigma)} := \begin{cases} -\dfrac{j^2}{\lambda_j^3} u_j^+ u_j^- u_j^\sigma + 2 \displaystyle\sum_{i \in \mathcal{I}} \dfrac{i^2}{\lambda_i^2 \lambda_j} u_i^+ u_i^- u_j^\sigma & \text{if} \quad j \in \mathcal{I} , \\ 2 \displaystyle\sum_{i \in \mathcal{I}} \dfrac{i^2}{\lambda_i^2 \lambda_j} u_i^+ u_i^- u_j^\sigma & \text{if} \quad j \notin \mathcal{I} , \end{cases} \tag{7.18}$$

$$\sigma 4\mathrm{i} G_2^{(u_j^\sigma)} := \begin{cases} 0 & \text{if} \quad j \in \mathcal{I} , \\ - \displaystyle\sum_{\substack{\sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 = \sigma j \\ j_1, j_2, j_3 \notin \mathcal{I}}} \dfrac{j_2 j_3}{\lambda_{j_1} \lambda_{j_2} \lambda_{j_3}} \sigma_2 \sigma_3\, u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{j_3}^{\sigma_3} & \text{if} \quad j \notin \mathcal{I} , \end{cases} \tag{7.19}$$

and $\forall\, \mathtt{a} \geq 0$

$$\|G_1\|_{R,\mathtt{a}} = \|G_1\|_{R,0} , \ \|G_2\|_{R,\mathtt{a}} = \|G_2\|_{R,0} \lessdot R^2 . \tag{7.20}$$

Moreover for N_0' *large enough we have*

$$\|G_1\|^T_{R,\mathtt{a},N_0,3/2,4} \lessdot R^2 , \qquad \|G_2\|^T_{R,\mathtt{a},N_0,3/2,4} \lessdot R^2 , \qquad \|G_3\|^T_{R/2,\mathtt{a}/2,N_0',7/4,3} \lessdot R^4 . \tag{7.21}$$

Finally $\mathcal{N}_0 + G_1 + G_2 + G_3 \in \mathcal{R}_{rev}$ *(recall Definition 2.11 in absence of* x, y*-variables).*

Remark 7.1. *The estimate for* G_3 *in* (7.21) *follows by assumption* (1.7)*. If* $g = \mathtt{y}\mathtt{y}_{\mathtt{x}}^2 + O_4$ *then* $\|G_3\|^T_{R/2,\mathtt{a}/2,N_0',7/4,3} \lessdot R^3$ *which is not enough for a direct application of Theorem 4.1 (see subsection 7.2). The term of order four should be removed by a further step of Birkhoff normal form. For simplicity, we did not pursue this point.*

Proof. The estimates (7.20) and (7.21) for G_1, G_2 follows by (3.27) and the analogous estimates (7.9) and (7.10) for G, since G_1, G_2 are projections (recall (2.19)) of G, satisfying (3.26).

In order to obtain the estimates for G_3 we need the following result proved in [4] (Lemma 7.2 and formula (7.21), see also [29]).

Lemma 7.3. *There exists an absolute constant* $c_* > 0$*, such that, for every* $\mathrm{m} \in (0, \infty)$ *and* $j_i \in \mathbb{Z}$*,* $\sigma_i = \pm$*,* $i = 1, 2, 3, 4$ *satisfying* $\sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0$ *but not satisfying*

$$j_1 = j_2 , \ j_3 = j_4 , \ \sigma_1 = -\sigma_2 , \ \sigma_3 = -\sigma_4 \ \text{(or permutations of the indexes)} , \tag{7.22}$$

we have

$$|\sigma_1 \lambda_{j_1} + \sigma_2 \lambda_{j_2} + \sigma_3 \lambda_{j_3} + \sigma_4 \lambda_{j_4}| \geq \frac{c_* \mathrm{m}}{(n_0^2 + \mathrm{m})^{3/2}} > 0 \quad \textit{where} \quad n_0 := \min\{\langle j_1 \rangle, \langle j_2 \rangle, \langle j_3 \rangle, \langle j_4 \rangle\} . \tag{7.23}$$

Let us define

$$F := \sum_{j \in \mathbb{Z}, \sigma = \pm} F^{(u_j^\sigma)} \partial_{u_j^\sigma} \qquad \text{with}$$

$$F^{(u_j^\sigma)} := \sum_{\substack{\sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 = \sigma j \\ \sigma_1 \lambda_{j_1} + \sigma_2 \lambda_{j_2} + \sigma_3 \lambda_{j_3} - \sigma \lambda_j \neq 0 \\ (j_1, j_2, j_3, j) \notin (\mathcal{I}^c)^4}} -\frac{1}{4} \frac{\sigma}{\sigma_1 \lambda_{j_1} + \sigma_2 \lambda_{j_2} + \sigma_3 \lambda_{j_3} - \sigma \lambda_j} \ \frac{j_2 j_3}{\lambda_{j_1} \lambda_{j_2} \lambda_{j_3}} \sigma_2 \sigma_3 \ u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{j_3}^{\sigma_3} . \tag{7.24}$$

By Lemma 7.3 and arguing as in Lemma 7.2 we get that

$$\forall\, \mathtt{a} \geq 0\,,\ R > 0\,, \qquad \|F\|_{R,\mathtt{a}} = \|F\|_{R,0} \lessdot R^2\,. \tag{7.25}$$

Moreover we claim that

$$\|F\|^T_{R,\mathtt{a},N_0,3/2,4} \lessdot R^2\,. \tag{7.26}$$

For $N \geq N_0$, by (7.24) we have that $\Pi_{N,3/2,4}F = \tilde{F} + N^{-1}\hat{F}$, where (recall (3.10), (3.11) and (3.12)), denoting for brevity $d_1 := \sigma_1\lambda_{j_1} + \sigma_2\lambda_{j_2} + \sigma'\lambda_n - \sigma\lambda_m$ and $d_2 := \sigma_1\lambda_{j_1} + \sigma_2\lambda_{j_2} + \sigma'|n| - \sigma|m|$

$$\begin{aligned}
\tilde{F} \ &:= \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>(3/2)N}} \tilde{F}^{\sigma,m}_{\sigma',n}\, z_n^{\sigma'}\partial_{z_m^\sigma}\,, \qquad \hat{F} := N \sum_{\substack{m,n,\sigma,\sigma' \\ |m|,|n|>(3/2)N}} \hat{F}^{\sigma,m}_{\sigma',n}\, z_n^{\sigma'}\partial_{z_m^\sigma}\,, \\
\tilde{F}^{\sigma,m}_{\sigma',n} \ &:= -\delta_{\sigma,\sigma'}\mathtt{s}(m) \sum_{\substack{\sigma_1 j_1+\sigma_2 j_2=\sigma m-\sigma' n\,,\ d_1\neq 0, \\ |j_1|+|j_2|<4N^L\,,\ j_1 \text{ or } j_2\in\mathcal{I}}} \frac{1}{2d_2}\frac{\sigma_2 j_2}{\lambda_{j_1}\lambda_{j_2}}\, u_{j_1}^{\sigma_1}u_{j_2}^{\sigma_2}\,, \\
\hat{F}^{\sigma,m}_{\sigma',n} \ &:= -\frac{\sigma}{4\lambda_n} \sum_{\substack{\sigma_1 j_1+\sigma_2 j_2=\sigma m-\sigma' n\,,\ d_1\neq 0, \\ |j_1|+|j_2|<4N^L\,,\ j_1 \text{ or } j_2\in\mathcal{I}}} \Big(\sigma_1 j_1 + 2\sigma\mathtt{s}(m)\Big(|n| - \delta_{\sigma,\sigma'}\frac{d_1}{d_2}\lambda_n\Big)\Big)\frac{1}{d_1}\frac{\sigma_2 j_2}{\lambda_{j_1}\lambda_{j_2}}\, u_{j_1}^{\sigma_1}u_{j_2}^{\sigma_2}
\end{aligned}$$

(here $\delta_{\sigma,\sigma'} = 1$ if $\sigma = \sigma'$ and 0 otherwise).

Let us consider first the case $\sigma = \sigma'$. We have

$$|d_1 - d_2| = \big|\lambda_n - \lambda_m - |n| + |m|\big| \lessdot \Big(\frac{1}{|n|} + \frac{1}{|m|}\Big) \lessdot \frac{1}{|n|}\,,$$

noting that $1/2 \leq |n|/|m| \leq 2$ by $\sigma_1 j_1 + \sigma_2 j_2 = \sigma m - \sigma' n$ and $|j_1| + |j_2| < 4N^L$, for $N \geq N_0$ large enough. Then, since by (7.23), $1 \lessdot |d_1|$, for $|n| \geq (3/2)N$ and N_0 large enough, $1 \lessdot |d_1| - |d_2 - d_1| \leq |d_2|$. In particular $|d_2| \geq const. > 0$ and $\tilde{F}, \hat{F}$ are well defined. Moreover

$$\Big|\frac{d_1}{d_2} - 1\Big| = \Big|\frac{1}{d_2}(d_1 - d_2)\Big| \lessdot \frac{1}{|n|} \qquad \text{and} \qquad \big|\lambda_n - |n|\big| \lessdot \frac{1}{|n|}$$

and, therefore, $\big||n| - d_1 d_2^{-1}\lambda_n\big| \lessdot 1$.

In the case $\sigma = -\sigma'$, since $|j_1| + |j_2| < 4N^L$ and $\lambda_m \geq |m| \geq N$, we get $|d_1| \geq |n|$.

Recollecting we have that, both in the case $\sigma = \sigma'$ and $\sigma = -\sigma'$, the Taylor coefficients of $\tilde{F}, \hat{F}$ are uniformly bounded and, arguing as in the proof of Lemma 7.2, we get

$$\|\tilde{F}\|_{R,\mathtt{a}}\,,\ \|\hat{F}\|_{R,\mathtt{a}} \lessdot R^2\,.$$

We note that $\tilde{F} \in \mathcal{T}_{R,\mathtt{a}}(N, 3/2, 4)$; indeed $\sigma = \sigma'$ and by (3.12) $\mathtt{s}(m) = \mathtt{s}(n)$, so that $d_2 := \sigma_1\lambda_{j_1} + \sigma_2\lambda_{j_2} + \mathtt{s}(m)(\sigma' n - \sigma m)$. Then by Definition 3.4 we get (7.26).

With $\mathcal{N}_0$ defined in (7.6) we have

$$\Big[\mathcal{N}_0, u_{j_1}^{\sigma_1}u_{j_2}^{\sigma_2}u_{j_3}^{\sigma_3}\,\partial_{u_j^\sigma}\Big] = \mathrm{i}(\sigma_1\lambda_{j_1} + \sigma_2\lambda_{j_2} + \sigma_3\lambda_{j_3} - \sigma\lambda_j)u_{j_1}^{\sigma_1}u_{j_2}^{\sigma_2}u_{j_3}^{\sigma_3}\,\partial_{u_j^\sigma}\,.$$

Then F in (7.24) solves the homological equation

$$[\mathcal{N}_0, F] + G^{(=3)} = \mathrm{ad}_F(\mathcal{N}_0) + G^{(=3)} = G_1 + G_2 \tag{7.27}$$

since $F^{(u_j^\sigma)} = \mathrm{i}(\sigma_1\lambda_{j_1} + \sigma_2\lambda_{j_2} + \sigma_3\lambda_{j_3} - \sigma\lambda_j)^{-1}(G^{(=3)})^{(u_j^\sigma)}$.

Then we define Γ as the time-1 flow generated by the vector field F. Then

$$\Big(D\Gamma^{-1}[\mathcal{N}_0 + G]\Big)\circ\Gamma = e^{\mathrm{ad}F}(\mathcal{N}_0 + G)\,.$$

By (7.27) we have that

$$G_3 = \sum_{j\geq 1} \frac{1}{j!}\mathrm{ad}_F^j G^{(=3)} + \frac{1}{(j+1)!}\mathrm{ad}_F^j (G_1 + G_2 - G^{(=3)}) + e^{\mathrm{ad}F} G^{(\geq 5)}$$

and the estimate for G_3 in (7.21) follows by Corollary 3.2 taking $R < R_0$ small enough and N_0' large enough.

We claim that $F \in \mathcal{R}_{a-rev}$. Indeed F is real-on-real (recall Definition 2.9) since $\overline{F^{(u_j^\sigma)}_{|\mathtt{R}}} = F^{(u_j^{-\sigma})}_{|\mathtt{R}}$ by (7.24) ($\mathtt{R}$ defined in (1.17)). F is anti-real-coefficients (recall Definition 2.8) since the Taylor coefficients in (7.24) are real. F is anti-reversible (recall Definition 2.7) with respect to the involution S in (1.29) since by (7.24) we have $F^{(u_j^\sigma)} \circ S = F^{(u_{-j}^{-\sigma})}$. Finally F is even (recall Definition 2.10) since, again by (7.24) $F^{(u_j^\sigma)}_{|E} = F^{(u^\sigma_{-j})}_{|E}$ (with E defined in (1.27)).

Then $\mathcal{N}_0 + G_1 + G_2 + G_3 = e^{\mathrm{ad}F}(\mathcal{N}_0 + G) \in \mathcal{R}_{rev}$ by Lemma 2.10. ■

7.1 Action-angle variables

Let us denote by

$$(u^+, u^-) = \Phi(x, y, z, \bar{z}; \xi) = \Phi(x, y, z^+, z^-; \xi) \tag{7.28}$$

the change of variable introduced in (1.33). For $\rho > 0$, let (recall (1.32))

$$\mathcal{O}_\rho := \Big\{\xi \in \mathbb{R}^{n/2} \; : \; \frac{\rho}{2} \leq \xi_j \leq \rho\,, \; j \in \mathcal{I}^+\Big\}\,. \tag{7.29}$$

A vector field $X = (X^{(u^+)}, X^{(u^-)})$ is transformed by the change of variable Φ in

$$\begin{aligned}
&Y := \Phi_\star X = \Big(D\Phi^{-1}[X]\Big) \circ \Phi\,, \qquad \text{with} \\
&Y^{(x_j)} = -\frac{\mathrm{i}}{2}\Big(\frac{1}{u_j^+}X^{(u_j^+)} - \frac{1}{u_j^-}X^{(u_j^-)}\Big) \circ \Phi\,, \quad Y^{(y_j)} = \Big(u_j^- X^{(u_j^+)} + u_j^+ X^{(u_j^-)}\Big) \circ \Phi\,, \quad j \in \mathcal{I}\,, \\
&Y^{(z_j^\sigma)} = X^{(u_j^\sigma)} \circ \Phi\,, \quad \sigma = \pm\,, \quad j \in \mathbb{Z} \setminus \mathcal{I}\,.
\end{aligned} \tag{7.30}$$

Lemma 7.4. (Lemma 7.6 of [4]) *Let us fix*

$$a = a_0/2\,, \;\; p > 1/2\,, \;\; \textit{and take} \;\; 0 < 16r^2 < \rho\,, \quad \rho = C_* R^2 \quad \text{with} \quad C_*^{-1} := 48 n \kappa^{2p} e^{2(s+a\kappa)}\,. \tag{7.31}$$

Then, for all $\xi \in \mathcal{O}_\rho \cup \mathcal{O}_{2\rho}$, the map

$$\Phi(\,\cdot\,; \xi) : D(s, 2r) \to B_{R/2} \times B_{R/2} \subset \ell^{a,p} \times \ell^{a,p} \tag{7.32}$$

is well defined and analytic ($D(s, 2r)$ is defined in (2.6) and κ in (3.1)).

Given a vector field $X : B_{R/2} \times B_{R/2} \to \ell^{a,p} \times \ell^{a,p}$, the previous Lemma and (7.30) show that the transformed vector field $Y := \Phi_\star X : D(s, 2r) \to \ell^{a,p} \times \ell^{a,p}$. It results that, if X is quasi-Töplitz in the variables $(u, \bar{u})$ then Y is quasi-Töplitz in the variables $(x, y, z, \bar{z})$ (see Definition 3.4). We define

$$\mathcal{V}^d_{R,\mathtt{a}} := \Big\{X \in \mathcal{V}_{R,\mathtt{a}} \; : \; X^{(u_j^\sigma)} = \sum_{|\alpha^{(2)} + \beta^{(2)}| \geq d} X^{(u_j^\sigma)}_{\alpha,\beta} u^\alpha \bar{u}^\beta\Big\}\,. \tag{7.33}$$

Proposition 7.2. (Quasi–Töplitz) *Let N_0, θ, μ, μ' satisfy* (3.1) *and*

$$(\mu' - \mu) N_0^L > N_0^b\,, \qquad N_0 2^{-\frac{N_0^b}{2\kappa} + 1} < 1\,. \tag{7.34}$$

If $X \in \mathcal{Q}^T_{R/2,\mathtt{a}}(N_0, \theta, \mu') \cap \mathcal{V}^d_{R/2,\mathtt{a}}$ with $d = 0, 1$, then $Y := \Phi_\star X \in \mathcal{Q}^T_{s,r,\mathtt{a}}(N_0, \theta, \mu)$ and

$$\|Y\|^T_{s,r,\mathtt{a},N_0,\theta,\mu,\mathcal{O}_\rho} \lessdot (8r/R)^{d-2} \|X\|^T_{R/2,\mathtt{a},N_0,\theta,\mu'}\,. \tag{7.35}$$

The proof of Proposition 7.2 follows closely the analogous Proposition 7.2 in [4] (replacing the Hamiltonians with the vector fields) and we omit it.

We also give the following similar lemma (see also Lemma 7.11 in [4]).

Lemma 7.5. *Let* $X \in \mathcal{V}_{R/2,\mathtt{a}}$, $Y := \Phi_\star X$ *and* $Y_0(x,y) := Y(x,y,0,0) - Y^{(y)}(x,0,0,0)\partial_y$. *Then, assuming* (7.31),

$$\|Y_0\|_{s,2r,\mathtt{a},\mathcal{O}_\rho \cup \mathcal{O}_{2\rho}} \lessdot (R/r)\|X\|_{R/2,\mathtt{a}} \,. \tag{7.36}$$

7.2 Conclusion of Proof of Theorem 1.1

Recalling (7.30) the vector field $\mathcal{N}_0 + G_1 + G_2 + G_3$ in (7.17) is transformed by the change of variable (1.33) into

$$\Phi_\star(\mathcal{N}_0 + G_1 + G_2 + G_3) = \mathcal{N} + \mathcal{P} = \mathcal{N} + \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 \,, \tag{7.37}$$

where the normal form $\mathcal{N}$ is as in (4.1) with frequencies

$$\omega_j(\xi) := \lambda_j + \frac{1}{4}\frac{j^2}{\lambda_j^3}\xi_{|j|} - \frac{1}{2}\sum_{i\in\mathcal{I}}\frac{i^2}{\lambda_i^2\lambda_j}\xi_{|i|} \quad \text{for } j \in \mathcal{I} \,, \qquad \Omega_j(\xi) := \lambda_j - \frac{1}{2}\sum_{i\in\mathcal{I}}\frac{i^2}{\lambda_i^2\lambda_j}\xi_{|i|} \quad \text{for } j \notin \mathcal{I} \tag{7.38}$$

($\lambda_j = \sqrt{j^2 + \mathrm{m}}$ defined in (1.24)) and the three terms of the perturbation are:

$$\begin{aligned}
\mathcal{P}_1^{(x_j)} &:= \frac{1}{4}\frac{j^2}{\lambda_j^3}y_j - \frac{1}{2}\sum_{i\in\mathcal{I}}\frac{i^2}{\lambda_i^2\lambda_j}y_j \,, \qquad \mathcal{P}_1^{(y_j)} = 0 \,, \quad \text{for } j \in \mathcal{I} \\
\mathcal{P}_1^{(z_j^\sigma)} &:= -\frac{\sigma \mathrm{i}}{2}\sum_{i\in\mathcal{I}}\frac{i^2}{\lambda_i^2\lambda_j}y_j z_j^\sigma \,, \qquad \sigma = \pm \,, \ j \notin \mathcal{I} \,; \\
\mathcal{P}_2 &:= \Phi_\star G_2 \qquad (\text{note that } \mathcal{P}_2^{(x)} = \mathcal{P}_2^{(y)} = 0 \,, \ \mathcal{P}_2^{(z_j^\pm)} = G_2^{(u_j^\pm)} \,, \ j \notin \mathcal{I}) \,; \\
\mathcal{P}_3 &:= \Phi_\star G_3 \,.
\end{aligned} \tag{7.39}$$

As in (4.6) we decompose the perturbation

$$\mathcal{P} = \mathcal{P}^y(x;\xi)\partial_y + \mathcal{P}_* \,, \qquad \mathcal{P}^y(x;\xi)\partial_y := \Pi^{(-1)}\mathcal{P}^{(y)}\partial_y = \Pi^{(-1)}\mathcal{P}_3^{(y)}\partial_y = \mathcal{P}_3^{(y)}(x,0,0,0;\xi)\partial_y \,. \tag{7.40}$$

Lemma 7.6. *Let* $s, r > 0$ *as in* (7.31) *and* N *large enough (w.r.t.* $\mathrm{m}, \mathcal{I}, L, b$*). Then*

$$\|\mathcal{P}^y\partial_y\|^\lambda_{s,r,\mathtt{a}/2,\mathcal{O}} \lessdot (1 + \lambda/\rho)R^6 r^{-2} \,, \qquad \|\mathcal{P}_*\|^T_{\vec{p}} \lessdot (1 + \lambda/\rho)(r^2 + R^5 r^{-1}) \,, \tag{7.41}$$

where

$$\mathcal{O} = \mathcal{O}(\rho) := \left\{ \xi \in \mathbb{R}^n \ : \ \frac{2}{3}\rho \le \xi_l \le \frac{3}{4}\rho \,, \ \ l = 1, \ldots, n \right\} \subset \mathcal{O}_\rho \tag{7.42}$$

(where $\mathcal{O}_\rho$ *was defined in* (7.29)*)and* $\vec{p} := (s, r, \mathtt{a}/2, N, 2, 2, \lambda, \mathcal{O})$.

Proof. By the definition (7.40) we have

$$\begin{aligned}
\|\mathcal{P}^y\partial_y\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} &= \|\Pi^{(-1)}\mathcal{P}_3^{(y)}\partial_y\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} \overset{\text{Lemma 2.4}}{\le} \|\mathcal{P}_3^{(y)}\partial_y\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} \\
&\overset{(7.35),(7.39)}{\lessdot} \left(\frac{r}{R}\right)^{-2}\|G_3\|^T_{R/2,\mathtt{a}/2,N,7/4,3} \overset{(7.21)}{\lessdot} \frac{R^6}{r^2}
\end{aligned} \tag{7.43}$$

(applying (7.35) with $d \rightsquigarrow 0$, $N_0 \rightsquigarrow N$, $\theta \rightsquigarrow 7/4$, $\mu \rightsquigarrow 2$, $\mu' \rightsquigarrow 3$) and taking N large enough so that (7.34) holds and $N \ge N_0'$ defined in Proposition 7.1.

By (7.37), (7.39) and (7.40) we write

$$\mathcal{P}_* = \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_4 + \mathcal{P}_5 \qquad \text{where} \tag{7.44}$$

$$\mathcal{P}_4 := \mathcal{P}_3(x,y,z,\bar z;\xi) - \mathcal{P}_3(x,y,0,0;\xi)\,, \quad \mathcal{P}_5 := \mathcal{P}_3(x,y,0,0;\xi) - \mathcal{P}_3^{(y)}(x,0,0,0;\xi)\partial_y\,.$$

We claim that

$$\|\mathcal{P}_1\|^T_{s,r,\mathtt{a}/2,N,2,2,\mathcal{O}_\rho}, \quad \|\mathcal{P}_2\|^T_{s,r,\mathtt{a}/2,N,2,2,\mathcal{O}_\rho} \quad \lessdot \quad r^2\,. \tag{7.45}$$

Indeed the estimate on $\mathcal{P}_1$ follows since $\mathcal{P}_1$ is Töplitz and $\|\mathcal{P}_1\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} \lessdot r^2$ by (7.39). On the other hand the estimate on $\mathcal{P}_2$ follows by (7.39) and (7.21) with $N \geq N_0$ large enough to fulfill (3.1).

By (7.39) and (7.35) (with $d \rightsquigarrow 1$, $N_0 \rightsquigarrow N$, $\mu \rightsquigarrow 2$, $\mu' \rightsquigarrow 3$), for N large enough, we get

$$\|\mathcal{P}_4\|^T_{s,r,\mathtt{a}/2,N,2,2,\mathcal{O}_\rho} \lessdot \Big(\frac{r}{R}\Big)^{-1} \|G_3\|^T_{R/2,\mathtt{a}/2,N_0',7/4,3} \overset{(7.21)}{\lessdot} \Big(\frac{r}{R}\Big)^{-1} R^4 = \frac{R^5}{r}\,. \tag{7.46}$$

Since $\mathcal{P}_5$ does not depend on the variables $z^\pm$ we get

$$\|\mathcal{P}_5\|^T_{s,r,\mathtt{a}/2,N,2,2,\mathcal{O}_\rho} = \|\mathcal{P}_5\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} \overset{(7.36)}{\lessdot} \Big(\frac{r}{R}\Big)^{-1} \|G_3\|_{R/2,\mathtt{a}/2} \overset{(7.21)}{\lessdot} \Big(\frac{r}{R}\Big)^{-1} R^4 = \frac{R^5}{r}\,. \tag{7.47}$$

In conclusion, by (7.44), (7.45), (7.46), (7.47) we get

$$\|\mathcal{P}_*\|^T_{s,r,\mathtt{a}/2,N,2,2,\mathcal{O}_\rho} \lessdot r^2 + R^5 r^{-1}\,.$$

In order to prove the estimates (7.41) we have to prove Lipschitz estimates (see (2.33), (3.81)). We first note that the vector fields $\mathcal{P}^y\partial_y$ and $\mathcal{P}_*$ are *analytic* in the parameters $\xi \in \mathcal{O}_\rho$. Then we apply Cauchy estimates in the subdomain $\mathcal{O} = \mathcal{O}(\rho) \subset \mathcal{O}_\rho$ (see (7.42)), noting that $\rho \lessdot \mathrm{dist}(\mathcal{O}, \partial\mathcal{O}_\rho)$. Then

$$\|\mathcal{P}_*\|^{\mathrm{lip}}_{s,r,\mathtt{a}/2,\mathcal{O}} \lessdot \rho^{-1}\|\mathcal{P}_*\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho} \qquad \text{and} \qquad \|\mathcal{P}^y\partial_y\|^{\mathrm{lip}}_{s,r,\mathtt{a}/2,\mathcal{O}} \lessdot \rho^{-1}\|\mathcal{P}^y\partial_y\|_{s,r,\mathtt{a}/2,\mathcal{O}_\rho}\,.$$

and (7.41) are proved. ■

We now verify that the assumptions of Theorems 4.1-4.2 are fulfilled by $\mathcal{N} + \mathcal{P}$ in (7.37) with parameters $\xi \in \mathcal{O}(\rho)$ defined in (7.42). Note that the sets $\mathcal{O} = [\rho/2, \rho]^n$ defined in Theorem 4.2 and $\mathcal{O}(\rho)$ defined in (7.42) are diffeomorphic through $\xi_i \mapsto (7\rho + 2\xi_i)/12$. Next $\vec\omega$ and Ω_j, defined in (7.38) satisfy (4.18) with

$$A := (A_{ji})_{i,j\in\mathcal{I}^+}\,, \quad A_{ji} = \frac14(\delta_{ij} - 2)\frac{i^2}{\lambda_i^2\lambda_j}\,, \quad \vec a := \Big(-\frac{i^2}{2\lambda_i^2}\Big)_{i\in\mathcal{I}^+}\,.$$

Then hypotheses (A1)-(A2) follow. Moreover (A3)-(A4) and the quantitative bound (4.7) follow by (7.41), choosing

$$s = 1,\ r = R^{1+\frac34}\,,\ \rho = C_* R^2 \text{ as in (7.31)},\ \ N \text{ as in Lemma 7.6},\ \theta = 2,\ \mu = 2,\ \gamma = R^{3+\frac15} \tag{7.48}$$

and taking *R small enough*. Hence Theorem 4.1 applies.

Let us verify that also the assumptions of Theorem 4.2 are fulfilled.

Denoting by $\mathbf{1}_{n/2}$ the $(n/2)\times(n/2)$ matrix with all entries equal to 1, we have

$$A = \frac14 D_1(\mathrm{Id}_{n/2} - 2\mathbf{1}_{n/2})D_2 \qquad \text{where} \quad D_1 := \mathrm{diag}_{j\in\mathcal{I}^+}\lambda_j^{-1}\,, \quad D_2 := \mathrm{diag}_{i\in\mathcal{I}^+} i^2\lambda_i^{-2}\,.$$

Since $\mathbf{1}_{n/2}^2 = (n/2)\mathbf{1}_{n/2}$ the matrix A is invertible with

$$A^{-1} = 4D_2^{-1}\Big(\mathrm{Id}_{n/2} - \frac{2}{n-1}\mathbf{1}_{n/2}\Big)D_1^{-1} \quad \text{and} \quad (A^T)^{-1}\vec a = \frac{2}{n-1}\bar\omega$$

where $\bar\omega_j := \lambda_j := \sqrt{j^2 + \mathrm{m}}$, $j \in \mathcal{I}^+$. Therefore, for every choices of $\mathcal{I}^+$ the conditions in (4.19) are fulfilled, excluding at most finitely many values of m (recall (4.22)).

We deduce that the Cantor set of parameters $\mathcal{O}_\infty \subset \mathcal{O}$ in (4.11) has asymptotically full density because

$$\frac{|\mathcal{O}\setminus\mathcal{O}_\infty|}{|\mathcal{O}|} \overset{(4.21)}{\lessdot} \rho^{-1}\gamma^{2/3} \overset{(7.48)}{\lessdot} R^{-2}R^{\frac23(3+\frac15)} = R^{\frac{2}{15}} \to 0\,.$$

The proof of Theorem 1.1 is now completed.

References

[1] Arnold V.I., *Reversible systems*, Nonlinear and turbulent processes in physics, Vol. 3 (Kiev, 1983), 1161-1174, Harwood Academic Publ., Chur, 1984.

[2] Bambusi D., Grébert B., *Birkhoff normal form for partial differential equations with tame modulus* Duke Math. J. 135, no. 3, 507-567, 2006.

[3] Berti M., Biasco L., *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys, 305, 3, 741-796, 2011.

[4] Berti M., Biasco L., Procesi M., *KAM theory for the Hamiltonian derivative wave equation*, to appear on Annales scientifique del' ENS.

[5] Berti M., Bolle P. *Quasi-periodic solutions with Sobolev regularity of NLS on* $\mathbb{T}^d$ *with a multiplicative potential*, Eur. Jour. Math, vol. 15, 2013.

[6] Berti M., Bolle P. *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity, 25, 2579-2613, 2012.

[7] Biasco L., Di Gregorio L. *A Birkhoff-Lewis type theorem for the nonlinear wave equation*, Arch. Ration. Mech. Anal. 196, no. 1, 303-362, 2010.

[8] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, no. 11, 1994.

[9] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of* 2*D linear Schrödinger equations*, Annals of Math. 148, 363-439, 1998.

[10] Bourgain J., *Periodic solutions of nonlinear wave equations*, Harmonic analysis and partial differential equations, 69–97, Chicago Lectures in Math., Univ. Chicago Press, 1999.

[11] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.

[12] Chierchia L., You J., *KAM tori for 1D nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys. 211, 497-525, 2000.

[13] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, 9, Société Mathématique de France, Paris, 2000.

[14] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498, 1993.

[15] Delort J.M., *Quasi-linear perturbations of Hamiltonian Klein-Gordon equations on spheres*, preprint 2012.

[16] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., 172, 371-435, 2010.

[17] Geng J., You J., *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Comm. Math. Phys. 262, 343-372, 2006.

[18] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226, 6, 5361-5402, 2011.

[19] Grebert B., Thomann L., *KAM for the quantum harmonic oscillator*, Comm. Math. Phys. 307, 2, 383-427, 2011.

[20] Liu J., Yuan X., *A KAM Theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations*, Comm. Math. Phys, 307 (3), 629-673, 2011.

[21] Kappeler T., Pöschel J., *KAM and KdV*, Springer, 2003.

[22] Klainermann S., Majda A., *Formation of singularities for wave equations including the nonlinear vibrating string*, Comm. Pure Appl. Math., 33, 241-263, 1980.

[23] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21, no. 3, 22–37, 95, 1987.

[24] Kuksin S., *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math-Math Phys., 10, 3, 1-64, 1998.

[25] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications, 19. Oxford University Press, 2000.

[26] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math, 2, 143, no. 1, 149-179, 1996.

[27] Moser J., *Convergent series expansions for quasi-periodic motions*, Math.Ann. 169, 136-176, 1967.

[28] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Sc. Norm. Pisa, 23, 119-148, 1996.

[29] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv., 71, no. 2, 269-296, 1996.

[30] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, preprint 2012.

[31] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, preprint 2011.

[32] Rabinowitz P., *Periodic solutions of nonlinear hyperbolic partial differential equations*, Comm. Pure Appl. Math., vol. xx, 105-205, 1967.

[33] Rabinowitz, P. *Periodic solutions of nonlinear hyperbolic partial differential equations II*, Comm. Pure Appl. Math. 22, 15-39, 1968.

[34] Rabinowitz, P. *A curious singular perturbation problem*, Differential equations, Proc. Conf., Birmingham/Ala. 1983, North-Holland Math. Stud. 92, 455-464 (1984).

[35] Sevryuk M.B., *Reversible Systems*, Lecture Notes in Math, 1211, Springer-Verlag, New York/Berlin, 1986.

[36] Wang W. M., *Supercritical nonlinear Schrödinger equations I: quasi-periodic solutions*, preprint 2010.

[37] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528, 1990.

[38] Zhang J., Gao M., Yuan X. *KAM tori for reversible partial differential equations*, Nonlinearity 24, 1189-1228, 2011.