A KAM ALGORITHM FOR THE RESONANT NON–LINEAR SCHRÖDINGER EQUATION

M. PROCESI*, AND C. PROCESI**.

Abstract. We prove, by applying a KAM algorithm, existence of large families of stable and unstable quasi periodic solutions for the NLS in any number of independent frequencies. The main tools are the existence of a non-degenerate integrable normal form proved in [18] and [20] and a suitable generalization of the quasi-Töplitz functions introduced in [24]

Contents

*Università di Roma, La Sapienza supported by ERC under FP7, ** Università di Roma, La Sapienza
.

1. Introduction

The present paper is devoted to the construction of stable and unstable quasi–periodic solutions for the completely resonant cubic NLS equation on a torus $\mathbb{T}^d$:

$$iu_t - \Delta u = \mathtt{k}|u|^2 u + \partial_{\bar{u}} G(|u|^2). \tag{1}$$

Here $u := u(t, \varphi)$, $\varphi \in \mathbb{T}^d$, Δ is the Laplace operator, $G(a)$ is a real analytic function whose Taylor series starts from degree 3 and $\mathtt{k} = \pm 1$.

Our results are obtained by exploiting the Hamiltonian structure of equation (1) and applying a KAM algorithm. As is well known such algorithms require strong *non-degeneracy* conditions which are not always valid, even for finite dimensional systems and, when valid, are generally proved by performing on the Hamiltonian a few steps of Birkhoff normal form. This is done in [18] (where the results are for for a larger class of NLS with non–linearity $|u|^{2q}u$) and [20] in which we exhibit and study the normal forms for classes of completely resonant non–linear Schrödinger equations.

Let us give a brief overview of the main difficulties in proving existence and stability of small quasi-periodic solutions for PDEs. One starts with a Hamiltonian PDE which has an elliptic fixed point at $u = 0$, and wishes to prove that some of the solutions of the nonlinear equation stay *close* to the linear solutions for all time. One has to deal with two classes of problems:

the *resonances*, namely the equation linearized at $u = 0$ does not have quasi-periodic solutions so we are dealing with a singular perturbation problem;

the *small divisors*, namely the equation linearized at $u = 0$ is described by an operator whose inverse is unbounded so that in order to find small solutions one needs to use some *Generalized Implicit Function Theorem.*

There are two main approaches to these problems: 1. Use a combination of Lyapunov–Schmidt reduction techniques and a Nash–Moser algorithm to solve the small divisor problem. This is the so–called *Craig–Wayne–Bourgain* approach, see [7] ,[5] and for a recent generalization also [4]. 2. Use a combination of Birkhoff normal form and a KAM algorithm, see for instance [15], [1].

In both cases one usually studies simplified models, namely parameter families of PDEs with the parameters chosen in such a way as to avoid resonances. Even under this simplifying hypotheses the problems related to the small divisors are in general quite complicated. Essentially, in order to perform a quadratic iteration scheme to prove the existence of quasi-periodic solutions, one needs some control on the operator (which we denote by $L(u)$) describing the equation linearized on an approximate solution $u(x,t)$ (and not only at $u=0$). In the Nash–Moser scheme one requires very weak hypotheses, in order to ensure that one may define a left inverse for $L(u)$ with some control on the *loss of regularity*. In a KAM scheme instead one imposes lower bounds on the *eigenvalues* of $L(u)$ and on *their differences*, this allows to prove a stronger result namely the NLS operator linearized at a quasi-periodic solution can be diagonalized by an analytic time dependent change of variables. Note that these last hypotheses imply a very good control on the loss of regularity of L^{-1}. The Nash-Moser approach combined with reducibility arguments (together with some novel ideas from pseudo-differential calculus) was used in [16] in order to prove existence and stability of quasi-periodic solutions for a class of fully-nonlinear perturbations of the KdV equation.

For PDEs in dimension $d>1$, where the eigenvalues of $L(0)$ are clearly multiple, the Nash–Moser algorithm is more readily applicable, see for instance [5]. KAM results for PDEs in dimension $d>1$ are few and relatively recent, see for instance [11], and in particular the paper [8] which studies an NLS with external parameters. Note that not only one needs to impose that the eigenvalues are different but one must give a lower bound on the difference. In the case of the NLS of [8] this requires a subtle analysis and the introduction of the class of Töplitz-Lipschitz functions (see also [24]).

In the case of equation (1), before attempting to study the small divisor problem one must deal with the resonances, since there are no external parameters and the only freedom is in the choice of the *initial data*.

In the case of (1) in dimension one this problem is avoided by just performing a step of Birkhoff normal form then applying a KAM algorithm (see [15] or [1]). This is due to the fact that the NLS equation after one step of Birkhoff normal form is integrable and non-degenerate. Unfortunately this very strong property holds only for the cubic NLS in dimension one, indeed for $d>1$ the non-integrability of the NLS normal form has been exploited (see for instance [6] and [14]) to construct diffusive orbits. In order to overcome this problem Bourgain proposed the idea of choosing the initial data wisely. More precisely one looks for a set $S\subset\mathbb{Z}^d$, the tangential sites, such that the Birkhoff normal form Hamiltonian admits quasi-periodic solutions which excite only the modes $\mathtt{j}\in S$. Then, by choosing S appropriately, one may prove existence of true solutions nearby.

This idea was used in [5] to prove the existence of quasi–periodic solutions with two frequencies for the cubic NLS in dimension two. This strategy was generalized by Wang in [21],[22] to study the NLS on a torus $\mathbb{T}^d$ and prove existence of quasi periodic solutions.

A similar idea was exploited in [13] and [12] to look for "wave packet" periodic solutions (i.e. periodic solutions which at leading order excite an arbitrarily large number of "tangential sites") of the cubic NLS in any dimension both in the case of periodic and Dirichlet boundary conditions. All the previous papers only deal with the existence of quasi-periodic solutions and not the linear stability and reducibility of the normal form. Note that the results by Wang on the NLS imply existence of quasi-periodic solutions for equation (1) and indeed her approach to the resonance problem is parallel in various ways to the one of [18]. Her approach is through the Nash–Moser method and hence does not prove reducibility results as explained before. Note however that [22] covers a larger class of cases i.e. non-cubic NLS equations with explicit dependence on the spatial variable.

In the context of KAM theory and normal form, we mention the result of [9] for the NLS in dimension one with the nonlinearity $|u|^4u$.

A strategy similar to the one used in this paper is proposed by Geng You and Xu in [10], to study the cubic NLS in dimension two. In that paper the authors show that one may give constraints on the tangential sites so that the normal form is non–integrable (i.e. it depends explicitly on the angle variables) but block diagonal with blocks of dimension 2. They apply this result to perform a KAM algorithm and prove existence (but not stability) of quasi–periodic solutions. We also mention the paper [19], which studies the non-local NLS and the beam equation both for periodic and Dirichlet boundary conditions.

The present paper is the last of a series of three papers in which we have developed a strategy aimed at the construction of large families of stable and unstable quasi-periodic solutions for the cubic NLS (1) in any dimension.

In the first paper [18] we study the NLS equation after one step of Birkhoff normal form and give "genericity conditions" on the *tangential sites* $S \subset \mathbb{Z}^d$ in order to make the normal form as simple as possible.

The main results of [18] are formulated in Theorem 2, where we prove that, for $|S| < \infty$ and for generic S, one can choose symplectic coordinates in which the normal form is integrable. On the tangential variables the normal form is non-degenerate and the motion is quasi-periodic with frequency $\omega = \omega(\xi)$, where $\omega(\xi)$ is a diffeomorphism and $\xi \in \mathbb{R}^n$ ($n = |S|$) are free parameters modulating the initial data. Moreover, in the "normal variables," the normal form is a block diagonal quadratic form, with blocks of dimension at most $d{+}1$. All blocks have constant coefficients. These infinitely many blocks are explicitly described by a graph Γ_S (cf. §4.1) which contains all the combinatorial difficulties of the structure. This combinatorial structure will influence the KAM–algorithm presented in this paper.

In [20] we address the delicate question of the *non-degeneracy* of the normal form deduced in [18], we obtain precise positive results for the cubic case $q = 1$.

In this paper we address the question of constructing quasi–periodic solutions and present a general solution. We need to analyze three issues

i) The second Melnikov non–degeneracy condition. This we prove by using the results of [20].
ii) The Töplitz–Lipschitz (cf. [8]) or quasi–Töplitz property of the perturbation. This is done by generalizing the quasi–Töplitz functions of [24] to this context; in particular we need to prove that the changes of variables that we perform to integrate the normal form do not destroy the quasi–Töplitz structure.
iii) The KAM algorithm. This is a variation (with some complications) of a well established path; we follow closely the structure of [1] and of [24].

In all these steps we need to combine the analysis of [24] with the special structure of the graph Γ_S. This is the source of most of the specific problems for the NLS which make this case particularly complex.

The final result will be the construction, for any dimension d, of families of linearly stable (and also elliptic for appropriate initial data) quasi–periodic solutions for the cubic NLS.

Following [18] for any $n \in \mathbb{N}$ we introduce the notion of *generic set of frequencies* $S \subset \mathbb{Z}^d$ with $|S| = n$ (see Definition 3.of [18]). The conclusive result of this analysis is:

Theorem 1. *For any $n \in \mathbb{N}$ and any generic set of frequencies $S = \{\mathtt{j}_1, \dots \mathtt{j}_n\} \subset \mathbb{Z}^d$ the NLS equation* (1) *admits small-amplitude, analytic (both in t and φ), quasi-periodic solutions of the form*

$$u(\varphi, t) = \sum_{j \in S} \sqrt{\xi_j} e^{\mathrm{i}(\omega^\infty(\xi)t + j\cdot\varphi)} + o(\sqrt{\xi}), \quad \omega_j^\infty(\xi) \overset{\xi \to 0}{\approx} |j|^2 \tag{2}$$

for all sufficiently small $\xi \in \mathbb{R}^n$ belonging to a "Cantor-like" set of parameters with asymptotical density 1 *at $\xi = 0$. The term $o(\sqrt{\xi})$ in* (2) *is small in an analytic norm. The equations* (1) *linearized at these quasi-periodic solutions are diagonable by an analytic time dependent change of variables. Finally, in a non empty open set of this Cantor set the solutions are also elliptic and linearly stable.*

We prove this result by verifying that the NLS Hamiltonian can be brought into a normal form which satisfies the properties of an abstract KAM Theorem, Theorem 6.

Most of the properties necessary for Theorem 6 have been verified for the NLS in Theorem 1 of [18] and in [20], here we have to prove the quasi–Töplitz property of the NLS, cf. §11.

It is possible to perform a KAM algorithm for any analytic NLS obtaining a weaker result. In this case one has the second Melnikov condition property i) only in a finite *block form.* This will be discussed elsewhere.

1.0.1. *The plan of the paper.* The paper is divided into four parts. In Part 1 we recall all the properties of the normal form proved in [18] and [20] which will be needed. In Part two we start by recalling the geometric formalism developed in [24] and prove that this formalism is compatible with the structure of the graph Γ_S. Having done this we proceed to define quasi-Töplitz functions in our context and prove their basic properties. Parts 3 and 4 are devoted to the KAM algorithm. In Part 3 we discuss the general properties of the type of algorithm that we shall apply to the NLS while in the final Part 4 we verify that the NLS satisfies all the properties of the class of Hamiltonians studied in Part 3. We can finally conclude that the KAM algorithm, applied to the Hamiltonian of the NLS starting from the normal form described in Part 1, leads to a successful construction of a family of quasi–periodic solutions of the NLS parametrized by a set of positive measures of the parameters ξ_i, actions of the initial excited frequencies. We discuss also which solutions are stable or unstable.

Part 1. The normal form

2. Summary of results from [18]

2.0.2. *The Hamiltonian.* In [18] we have studied the NLS on $\mathbb{T}^d$ as an infinite dimensional Hamiltonian system. After rescaling and passing to Fourier representation[1]

$$u(t, \varphi) := \sum_{k \in \mathbb{Z}^d} u_k(t) e^{\mathrm{i}(k, \varphi)} \tag{3}$$

the Hamiltonian is (having normalized κ):

$$H := \sum_{k \in \mathbb{Z}^d} |k|^2 u_k \bar{u}_k + \sum_{k_i \in \mathbb{Z}^d : \sum_{i=1}^4 (-1)^i k_i = 0} u_{k_1} \bar{u}_{k_2} u_{k_3} \bar{u}_{k_4}. \tag{4}$$

[1]In fact one should work in a slightly more general setting where the torus is the quotient of $\mathbb{R}^d$ by any lattice Λ of finite index in $\mathbb{Z}^d$ and write $u(t, \varphi) := \sum_{k \in \Lambda^*} u_k(t) e^{\mathrm{i}(k, \varphi)}$.

The complex symplectic form is $i\sum_k du_k \wedge d\bar{u}_k$, on the scale of complex Hilbert spaces

$$\bar{\ell}^{(a,p)} := \{u = \{u_k\}_{k\in\mathbb{Z}^d} \mid |u_0|^2 + \sum_{k\in\mathbb{Z}^d} |u_k|^2 e^{2a|k|}|k|^{2p} := ||u||^2_{a,p} \le \infty\}, \tag{5}$$

$$a > 0,\ p > d/2.$$

We systematically apply the fact that we have $d+1$ conserved quantities: the d–vector *momentum* $\mathbb{M}$ and the scalar *mass* $\mathbb{L}$:

$$\mathbb{M} := \sum_{k\in\mathbb{Z}^d} k|u_k|^2\,, \qquad \mathbb{L} := \sum_{k\in\mathbb{Z}^d} |u_k|^2\,,$$

with

$$\{\mathbb{M}, u_h\} = \mathrm{i}hu_h,\ \{\mathbb{M}, \bar{u}_h\} = -\mathrm{i}h\bar{u}_h,\ \{\mathbb{L}, u_h\} = \mathrm{i}u_h,\ \{\mathbb{L}, \bar{u}_h\} = -\mathrm{i}\bar{u}_h. \tag{6}$$

The terms in equation (4) commute with $\mathbb{L}$. The conservation of momentum is expressed by the constraints $\sum_{i=1}^4(-1)^i k_i = 0$.

2.0.3. *Choice of the tangential sites.* If in the Hamiltonian H we remove all quartic terms which do not Poisson commute with the quadratic part, we obtain a simplified Hamiltonian denoted H_{Birk}. This has the property that its Hamiltonian vector field is tangent to infinitely many subspaces obtained by setting some of the coordinates equal to 0 (cf. [18], Proposition 1). On infinitely many of them furthermore the restricted system is completely integrable, thus the next step consists in choosing such a subset S which, for obvious reasons, is called of *tangential sites.* Without loss of generality one may assume that S spans $\mathbb{Z}^d$ over $\mathbb{Z}$ (cf. footnote 1).

With this remark in mind we partition

$$\mathbb{Z}^d = S \cup S^c, \quad S := (\mathtt{j}_1, \dots, \mathtt{j}_n) \tag{7}$$

where the elements of S play the role of *tangential sites* and of S^c the *normal sites.* We divide $u \in \bar{\ell}^{a,p}$ in two components $u = (u_1, u_2)$, where u_1 has indexes in S and u_2 in S^c. The choice of S is subject to several constraints which make it *generic* and which are fully discussed in [18] and finally refined in [20]. Here we shall always assume that these constraints are valid so we just refer to the results of these two papers in all the statements.

We often use the map $\pi : \mathbb{R}^n \to \mathbb{R}^d,\ \pi(a_1, \dots, a_n) := \sum_i a_i \mathtt{j}_i$, notice that π maps $\mathbb{Z}^n$ to $\mathbb{Z}^d$, and set

$$\kappa := \max_{j\in S} |\mathtt{j}|. \tag{8}$$

If we use on $\mathbb{R}^n$ the L^1 norm then κ is also the norm of the map π.

We apply a standard *semi-normal form* change of variables with generating function:

$$F_{Birk} = -\mathrm{i} \sum_{\substack{\alpha,\beta\in(\mathbb{Z}^d)^{\mathbb{N}}:|\alpha|=|\beta|=2\,,|\alpha_2|+|\beta_2|\le 2\\ \sum_k(\alpha_k-\beta_k)k=0\,,\ \sum_k(\alpha_k-\beta_k)|k|^2\neq 0}} \binom{2}{\alpha}\binom{2}{\beta} \frac{u^\alpha \bar{u}^\beta}{\sum_k(\alpha_k-\beta_k)|k|^2}. \tag{9}$$

Here the notation α_2, β_2 refers to the exponents for the variable $u_k, \bar{u}_k$ with $k \in S^c$. We use the operator notation $ad(F)$ for the operator $X \mapsto \{F, X\}$. The change of variables by $\Psi^{(1)} := e^{ad(F_{Birk})}$ is well defined and analytic: $B_{\epsilon_0} \times B_{\epsilon_0} \to B_{2\epsilon_0} \times B_{2\epsilon_0}$, for ϵ_0 small enough, see [18]. By construction $\Psi^{(1)}$ brings (4) to the form $H = H_{Birk} + P^4(u) + P^6(u)$

where $P^4(u)$ is of degree 4 but at least cubic in u_2 while $P^6(u)$ is analytic of degree at least 6 in u, finally

$$H_{Birk} := \sum_{k\in\mathbb{Z}^d} |k|^2 u_k \bar u_k + \sum_{\substack{\alpha,\beta\in(\mathbb{Z}^d)^{\mathbb{N}}:|\alpha|=|\beta|=2\,,\,|\alpha_2|+|\beta_2|\le 2\\ \sum_k(\alpha_k-\beta_k)k=0\,,\,\sum_k(\alpha_k-\beta_k)|k|^2=0}} \binom{2}{\alpha}\binom{2}{\beta} u^\alpha \bar u^\beta. \tag{10}$$

The three constraints in the second summand of the previous formula express the conservation of $\mathbb{L}$, $\mathbb{M}$ and of the *quadratic energy*

$$\mathbb{K} := \sum_{k\in\mathbb{Z}^d} |k|^2 u_k \bar u_k. \tag{11}$$

In order to perform perturbation theory from the system given by the tangential sites it is convenient to switch to polar coordinates. We set

$$u_k := z_k \text{ for } k\in S^c\,,\quad u_{\mathtt{j}_i} := \sqrt{\xi_i+y_i}e^{\mathrm{i}x_i} = \sqrt{\xi_i}(1+\frac{y_i}{2\xi_i}+\ldots)e^{\mathrm{i}x_i} \text{ for } i=1,\ldots n, \tag{12}$$

considering the $\xi_i>0$ as parameters $|y_i|<\xi_i$ while $y,x,w:=(z,\bar z)$ are dynamical variables. We denote by $\ell^{(a,p)} := \ell_S^{(a,p)}$ the subspace of $\bar\ell^{(\mathbf{a},\mathbf{p})}\times\bar\ell^{(\mathbf{a},\mathbf{p})}$ of the sequences $u_i,\bar u_i$ with indices in S^c and denote the coordinates $w=(z,\bar z)$.

Definition 2.1. Let $\mathfrak{K}\subset\mathbb{R}^n_+$ be a compact domain and let $0<c_1<c_2$ be such that

$$c_1^2 = \min_{\mathfrak{K}}(\min_i \xi_i)\,,\quad c_2^2=\max_{\mathfrak{K}}(\max_i \xi_i)$$

It is convenient to choose as $\mathfrak{K}=\mathfrak{H}\times J$ a product in polar coordinates of a compact domain $\mathfrak{H}$ in the unit sphere and some compact set J in the coordinate ρ. We will consider for all $\rho>0$ the *scaled domain* $\rho\mathfrak{K}$ and notice that $\rho\mathfrak{K}=\mathfrak{H}\times\rho J$.

One can refer to such a domain as a *truncated cone.*

We choose $\rho=\varepsilon^2$ and note that, for all $r<c_1\varepsilon$, formula (12) is an analytic and symplectic change of variables Φ_ξ in the domain

$$D_{a,p}(s,r)=D(s,r):=\{x,y,w\,:\,x\in\mathbb{T}^n_s\,,\,|y|\le r^2\,,\,\|w\|_{a,p}\le r\}\subset\mathbb{T}^n_s\times\mathbb{C}^n\times\ell^{(a,p)}. \tag{13}$$

Here $\varepsilon>0$, $s>0$ and $0<r<\varepsilon c_1$ are auxiliary parameters. $\mathbb{T}^n_s$ denotes the compact subset of the complex torus $\mathbb{T}^n_{\mathbb{C}}:=\mathbb{C}^n/2\pi\mathbb{Z}^n$ where $x\in\mathbb{C}^n$, $|\mathrm{Im}(x)|\le s$. Moreover if

$$\sqrt{2n}c_2\kappa^p e^{(s+a\kappa)}\varepsilon<\epsilon_0\,,\quad (\text{recall}\quad \kappa=\max(|\mathtt{j}_i|)\ ) \tag{14}$$

the change of variables sends $D(s,r)\to B_{\epsilon_0}$ so we can apply it to our Hamiltonian.

We thus assume that the parameters ε, r, s satisfy (14). Formula (12) puts in action angle variables $(y;x)=(y_1,\ldots,y_n;x_1,\ldots,x_n)$ the tangential sites, close to the action $\xi=\xi_1,\ldots,\xi_n$, which are parameters for the system.

The symplectic form is now $dy\wedge dx+i\sum_{k\in S^c}dz_k\wedge d\bar z_k$.

We give degree 0 to the angles x, 2 to y and 1 to w. We use the degree only for handling dynamical variables, as follows. We develop in Taylor expansion, in particular since y is small with respect to ξ we develop $\sqrt{\xi_i+y_i}=\sqrt{\xi_i}(1+\frac{y_i}{2\xi_i}+\ldots)$ as a series in $\frac{y_i}{\xi_i}$.

By abuse of notations we still call H the composed Hamiltonian $H\circ\Psi^{(1)}\circ\Phi_\xi$.

Definition 2.2. We define the *normal form* $\mathcal{N}$ which collects all the terms of H_{Birk} of degree ≤ 2 (dropping the constant terms). We then set $P=H-\mathcal{N}$.

Notice that the Hamiltonian H_{Birk} is different from the corresponding one in [18] (in that paper we performed a full normal form transformation), however the resulting normal form $\mathcal{N}$ is the same since it collects only terms of degree less or equal to two in the variables $z = u_2$.

3. Functional setting

Following [17] we study *regular* functions $F : \varepsilon^2\mathfrak{K} \times D_{a,p}(s,r) \to \mathbb{C}$, that is whose Hamiltonian vector field $X_F(\cdot;\xi)$ is M-analytic from $D(s,r) \to \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_S$. In the variables ξ we require Lipschitz regularity. Let us recall the definitions of M-analytic and majorant norm and their properties proved in [3].

Let us consider the space

$$V := \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_S \tag{15}$$

with (s,r)-weighted norm

$$v = (x, y, z, \bar z) \in V\,, \quad \|v\|_V := \|v\|_{s,r} = \|v\|_{V,s,r} = \frac{|x|_\infty}{s} + \frac{|y|_1}{r^2} + \frac{\|z\|_{a,p}}{r} + \frac{\|\bar z\|_{a,p}}{r} \tag{16}$$

where $0 < s < 1$, $0 < r < c_1\varepsilon$ and $|x|_\infty := \max_{h=1,\dots,n} |x_h|$, $|y|_1 := \sum_{h=1}^n |y_h|$.

For a vector field, i.e. a map $X : D(s,r) \to V$, described by the formal Taylor expansion:

$$X = \sum_{\nu,i,\alpha,\beta} X^{(\mathtt{v})}_{\nu,i,\alpha,\beta} e^{\mathrm{i}(\nu,x)} y^i z^\alpha \bar z^\beta \partial_{\mathtt{v}}\,, \quad \mathtt{v} = x, y, z, \bar z$$

we define the *majorant* and its *norm*:

$$MX := \sum_{\nu,i,\alpha,\beta} |X^{(\mathtt{v})}_{\nu,i,\alpha,\beta}| e^{s|\nu|} y^i z^\alpha \bar z^\beta \partial_{\mathtt{v}}\,, \quad \mathtt{v} = x, y, z, \bar z$$

$$||X||_{s,r} \quad := \quad \sup_{(y,z,\bar z)\in D(s,r)} \|MX\|_V\,. \tag{17}$$

The different weights ensure that, if $\|X_F\|_{s,r} < \frac{1}{2}$, then F generates a close–to–identity symplectic change of variables from $D(s/2, r/2) \to D(s,r)$, Proposition 3.3.

Remark 3.1. The notion of M–analytic can be given in general for any map between separable Hilbert spaces with prescribed bases. It means that the *map* given in coordinates by the corresponding majorant functions is in fact analytic. It is then easy to see that composition of M–analytic maps is M–analytic with the corresponding estimate on norms.

In our algorithm we deal with functions which depend in a Lipschitz way on some parameters ξ in a compact set $\mathcal{O} \subseteq \varepsilon^2\mathfrak{K}$ (Formula (2.1)). To handle this dependence we introduce weighted Lipschitz norms for a map $X : \mathcal{O} \times D(s,r) \to V$ setting:

$$\|X\|^{lip}_{s,r,\mathcal{O}} := \sup_{\xi\neq\eta\in\mathcal{O}\,,\ (x,y,w)\in D(s,r)} \frac{\|X(\eta) - X(\xi)\|_{s,r}}{|\eta - \xi|},$$

$$\|X\|_{s,r,\mathcal{O}} = \|X\|_{s,r} := \sup_{\mathcal{O}\times D(s,r)} \|MX\|_V\,, \|X\|^\lambda_{s,r} = \|X\|_{s,r,\mathcal{O}} + \lambda\|X_f\|^{lip}_{s,r,\mathcal{O}} \tag{18}$$

where λ is a parameter proportional to $|\mathcal{O}|$. Correspondingly for a parameter dependent sequence $f = \{f_m(\xi)\}_{m\in I}$, here I is any index set, we define:

$$|f|_\infty := \sup_{\xi\in\mathcal{O}} \sup_{m\in I} |f_m(\xi)|\,, \quad |f|^{lip}_\infty := \sup_{\xi\neq\eta\in\mathcal{O}} \sup_{m\in I} \frac{|f_m(\xi) - f_m(\eta)|}{|\eta - \xi|_\infty}\,, \tag{19}$$

Definition 3.2. We define by $\mathcal{H}_{s,r,\mathcal{O}} = \mathcal{H}_{s,r}$ the space of regular analytic Hamiltonians depending on a parameter $\xi \in \mathcal{O}$ with the norm[2]

$$\|F\|_{s,r}^{\lambda} := \|X_F\|_{s,r}^{\lambda} < \infty. \tag{20}$$

We denote by $\mathbb{I} = \mathbb{Z}^n \times \mathbb{N}^n \times \mathbb{N}^{S^c} \times \mathbb{N}^{S^c}$ the indexing set of the monomials, that is k, i, α, β is associated to $e^{\mathrm{i}(k,x)} y^i z^\alpha \bar{z}^\beta$. For all $I \subset \mathbb{I}$ we define the projection Π_I as the linear operator which acts as the identity on the monomials associated to I and zero otherwise. In particular we define $\Pi_{|k|<K}$ to be the projection relative to the set I of k, i, α, β with $|k| < K$, same for $\Pi_{|k|\geq K}$, similarly we define $\Pi^{(\ell)}$ as the projection on the k, i, α, β with $2i + |\alpha| + |\beta| = \ell$, same for $\Pi^{(\geq \ell)}$ and $\Pi^{(\leq \ell)}$.

The main properties of the majorant norm are contained in the following statements, proved in [2], Lemma 2.10, 2.15, 2.17.

Proposition 3.3. *Let $H, K \in \mathcal{H}_{s,r}$. Then, for all $r/2 \leq r' < r$, $s/2 \leq s' < s$, $\lambda' \leq \lambda$:*

$$\|X_H\|_{s',r'}^{\lambda'} \leq 4\|X_H\|_{s,r}^{\lambda}\,, \tag{21}$$

$$\|X_{\{H,K\}}\|_{s',r'}^{\lambda} = \|\,[X_H, X_K]\,\|_{s',r'}^{\lambda} \leq 2^{2n+3}\delta^{-1}\|X_H\|_{s,r}^{\lambda}\|X_K\|_{s,r}^{\lambda} \tag{22}$$

where δ is defined by:

$$\delta := \min\left\{1 - \frac{s'}{s}, 1 - \frac{r'}{r}\right\}. \tag{23}$$

Let $r/2 \leq r' < r$, $s/2 \leq s' < s$, and $F \in \mathcal{H}_{s,r}$ with

$$\|X_F\|_{s,r}^{\lambda} < \delta/(2^{2n+6}e) \tag{24}$$

with δ defined in (23)*. Then the time* 1*-Hamiltonian flow*

$$\Phi_F^1 := e^{\mathrm{ad}(F)} : D(s', r') \to D(s, r)$$

is well defined, analytic, symplectic, and, $\forall H \in \mathcal{H}_{s,r}$, we have $H \circ \Phi_F^1 \in \mathcal{H}_{s',r'}$ and

$$\|X_{H\circ\Phi_F^1}\|_{s',r'}^{\lambda} \leq 2\|X_H\|_{s,r}^{\lambda}\,, \quad \|X_{H\circ\Phi_F^1} - X_H\|_{s',r'}^{\lambda} \leq 2\|X_F\|_{s,r}^{\lambda}\|X_H\|_{s,r}^{\lambda}\,. \tag{25}$$

For all $I \subset \mathbb{I}$ and, $\forall H \in \mathcal{H}_{s,r}$, we have

$$\|\Pi_I X_H\|_{s,r}^{\lambda} \leq \|X_H\|_{s,r}^{\lambda}\,. \tag{26}$$

In particular we have the smoothing estimates*: $s' < s$,*

$$\|\Pi_{|k|\geq K} X_H\|_{s',r}^{\lambda} \leq \frac{s}{s'}\, e^{-K(s-s')}\|X_H\|_{s,r}^{\lambda}\,, \tag{27}$$

and the degree estimates

$$\|X_{\Pi^{(l\geq d)}H}\|_{s,r'}^{\lambda} \leq (\frac{r'}{r})^{d-2}\|X_H\|_{s,r}^{\lambda}\,. \tag{28}$$

Remark 3.4. For a diagonal quadratic Hamiltonian $F = \sum_m \vartheta_m(\xi) z_m \bar{z}_m$ we have

$$X_F = i(\sum_m \vartheta_m(\xi) z_m \frac{\partial}{\partial z_m} - \sum_m \vartheta_m(\xi)\bar{z}_m \frac{\partial}{\partial \bar{z}_m})$$

$$MX_F = \sum_m |\vartheta_m(\xi)|(z_m \frac{\partial}{\partial z_m} + \bar{z}_m \frac{\partial}{\partial \bar{z}_m}), \quad \|MX_F\|_{s,r}^{\lambda} = |\vartheta|_\infty + \lambda|\vartheta|_\infty^{lip}.$$

[2]in fact Hamiltonians should be considered up to scalar summands and then this is actually a norm

Lemma 3.5. *For $c_1\varepsilon > r > \varepsilon^3$, the perturbation P of Definition 2.2 is in $\mathcal{H}_{s,r}$ and satisfies the bounds*

$$\|X_P\|^{\lambda}_{s,r} \le C(\varepsilon r + \varepsilon^5 r^{-1})\,, \tag{29}$$

where C does not depend on r and depends on ε, λ only through λ/ε^2.

Proof. This is item iv) of Theorem 1 of [18]. The fact that we are using the majorant norm only changes the constant and not the order of magnitude. □

4. The normal form

We will work with many quadratic Hamiltonians in the variables w (thought as a row vector). We represent a quadratic form $\mathcal{F}$ by a matrix F as

$$\mathcal{F}(w) = \frac{1}{2}(w, wJF^t) = -\frac{1}{2}wFJw^t\,, \tag{30}$$

where $J := -\mathrm{i}\{w^t, w\}$ is the standard matrix of the symplectic form which expresses the action by Poisson bracket.

By explicit computation, and under simple generiticity conditions, the normal form $\mathcal{N}$ of Definition 2.2 is as follows:

$$(\omega(\xi), y) + \sum_{k\in S^c} |k|^2|z_k|^2 + \mathcal{Q}(\xi; x, w)\,, \quad \omega_i(\xi) = |\mathrm{j}_i|^2 - 2\xi_i \tag{31}$$

here $\mathcal{Q}(\xi; x, w)$ is a quadratic Hamiltonian in the variables w with coefficients trigonometric polynomials in x given by Formula (30) of [18]:

$$\mathcal{Q}(\xi, w) = 4 \sum^{*}_{\substack{1\le i\ne j\le m\\ h,k\in S^c}} \sqrt{\xi_i\xi_j}e^{\mathrm{i}(x_i-x_j)}z_h\bar z_k + \tag{32}$$

$$+2\sum^{**}_{\substack{1\le i<j\le m\\ h,k\in S^c}} \sqrt{\xi_i\xi_j}e^{-\mathrm{i}(x_i+x_j)}z_h z_k + 2\sum^{**}_{\substack{1\le i<j\le m\\ h,k\in S^c}} \sqrt{\xi_i\xi_j}e^{\mathrm{i}(x_i+x_j)}\bar z_h\bar z_k.$$

Here $\sum^*$ denotes that (h, k, v_i, v_j) satisfy:

$$\{(h, k, v_i, v_j)\,|\, h + v_i = k + v_j,\ |h|^2 + |v_i|^2 = |k|^2 + |v_j|^2\}.$$

and $\sum^{**}$, that (h, v_i, k, v_j) satisfy:

$$\{(h, v_i, k, v_j)\,|\, h + k = v_i + v_j,\ |h|^2 + |k|^2 = |v_i|^2 + |v_j|^2\}.$$

Notice that in the sums $\sum^{**}$ each term appears twice.

This is a very complicated infinite dimensional quadratic Hamiltonian, by applying the results of [18], we decompose this infinite dimensional system into infinitely many decoupled finite dimensional systems corresponding to the connected components of a graph (which is recalled in §4.1). One of the main results of [18] is the construction of an explicit symplectic change of variables which reduces $\mathcal{N}$ to constant coefficients.

Since this construction is needed in the following we recall quickly Theorem 2 of [18] adapted to the case of the cubic NLS. In the cubic case we also apply the more precise results of [20].

Theorem 2. *For all* generic *choices* $S = \{\mathtt{j}_1, \dots, \mathtt{j}_n\} \in \mathbb{Z}^{nd}$ *of the tangential sites, there exists a map*

$$S^c \ni k \to L(k) \in \mathbb{Z}^n\,, \quad |L(k)| \le d+1$$

such that the analytic symplectic change of variables:

$$z_k = e^{-\mathrm{i}(L(k),x)} z'_k,\ \ y = y' + \sum_{k \in S^c} L(k)|z'_k|^2,\ \ x = x'.$$

$$\Psi : (y', x) \times (z', \bar z') \to (y, x) \times (z, \bar z)$$

from $D(s, r/2) \to D(s,r)$ *has the property that* $\mathcal{N}$ *in the new variables has constant coefficients, namely:*

$$\mathcal{N} \circ \Psi = (\omega(\xi), y') + \sum_{k \in S^c} \tilde\Omega_k |z'_k|^2 + \tilde{\mathcal{Q}}(w')\,, \tag{33}$$

where $\omega(\xi)$ *is defined in* (31) *and furthermore:*

i) **Asymptotic of the normal frequencies:** *We have* $\tilde\Omega_k = |k|^2 + \sum_i |\mathtt{j}_i|^2 L^{(i)}(k)$.
ii) **Reducibility***: The matrix* $\tilde Q(\xi)$ *which represents the quadratic form* $\tilde{\mathcal{Q}}(\xi, w')$ *(see formula* (30)*) depends only on the variables* ξ *and all its entries are homogeneous of degree one in these variables. It is block–diagonal with blocks of dimension* $\le d+1$ *and satisfies the following properties:*

All of the blocks except a finite number are self adjoint.

All the (infinitely many) blocks are chosen from a finite list of matrices $\mathcal{M}(\xi)$.

iii) **Smallness:** *If* $\varepsilon^3 < r < c_1\varepsilon$*, the perturbation* $\tilde P := P \circ \Psi$ *is small, more precisely we have the bounds:*

$$\|X_{\tilde P}\|^{\lambda}_{s,r} \le C(\varepsilon r + \varepsilon^5 r^{-1})\,, \tag{34}$$

where C *is independent of* r *and depends on* ε, λ *only through* λ/ε^2.

The smallness condition implies that, if r is of the order of ε^2 and λ/ε^2 is of order one, then $\|X_{\tilde P}\|^{\lambda}_{s,r}$ is of order ε^3. As we shall see this is exactly a type of smallness required in order to insure the success of the KAM algorithm (cf. Theorem 7).

Warning In $\mathbb{Z}^n$ we always use as norm $|l|$ the L^1 norm $\sum_{i=1}^n |l^{(i)}|$. On the other hand in $\mathbb{Z}^d$, and hence in S^c, we use the euclidean L^2 norm.

4.1. **The geometric graph Γ_S.** It is important to recall that the term $\tilde{\mathcal{Q}}(\xi, w')$ comes from the sum of two contributions, the term $\mathcal{Q}(\xi, x, w)$, in the new variables and the contribution $-2\sum_{k \in S^c} \xi \cdot L(k)|z'_k|^2$ (coming from the y variables). Hence $\tilde{\mathcal{Q}}(w') =$

$$-2\sum_{k \in S^c} \xi \cdot L(k)|z'_k|^2 + 4 \sideset{}{^*}\sum_{\substack{1 \le i \ne j \le m\\ h,k \in S^c}} \sqrt{\xi_i\xi_j}\, z'_h \bar z'_k + 2 \sideset{}{^{**}}\sum_{\substack{1 \le i < j \le m\\ h,k \in S^c}} \sqrt{\xi_i\xi_j}\, z'_h z'_k + 2 \sideset{}{^{**}}\sum_{\substack{1 \le i < j \le m\\ h,k \in S^c}} \sqrt{\xi_i\xi_j}\, \bar z'_h \bar z'_k. \tag{35}$$

In its matrix description the two terms will give the off diagonal and the diagonal terms respectively.

The off diagonal terms are described through a simple geometric construction (which gives a complicated combinatorics). Given two distinct elements $\mathtt{j}_i, \mathtt{j}_j \in S$ construct the sphere $S_{i,j}$ having the two vectors as opposite points of a diameter and the two Hyperplanes, $H_{i,j}$, $H_{j,i}$, passing through $\mathtt{j}_i$ and $\mathtt{j}_j$ respectively, and perpendicular to the line though the two vectors $\mathtt{j}_i, \mathtt{j}_j$.

From this configuration of spheres and pairs of parallel hyperplanes we deduce a *geometric colored graph*, denoted by Γ_S, with vertices the points in S^c and two types of edges, which we call *black* and *red*.

- A black edge connects two points $p \in H_{i,j}$, $q \in H_{j,i}$, such that the line p, q is orthogonal to the two hyperplanes, or in other words $q = p + \mathtt{j}_j - \mathtt{j}_i$.
- A red edge connects two points $p, q \in S_{i,j}$ which are opposite points of a diameter ($p + q = \mathtt{j}_i + \mathtt{j}_j$).

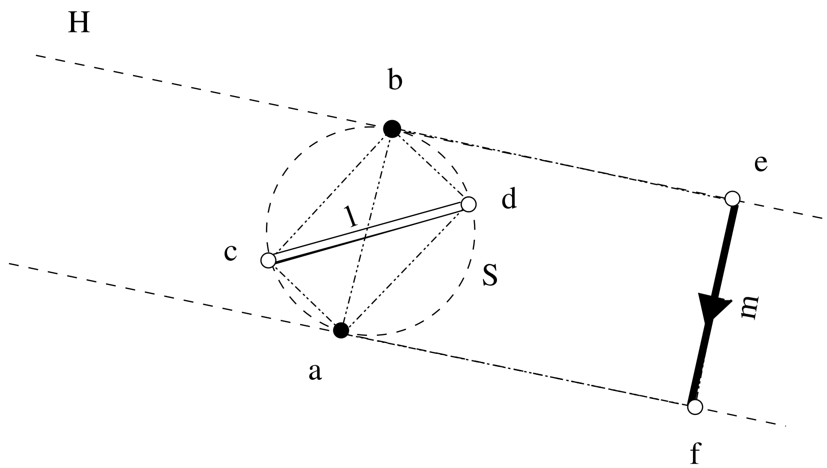


FIGURE 1. the plane $H_{i,j}$ and the sphere $S_{i,j}$. The points $a_1, b_1, \mathtt{j}_j, \mathtt{j}_i$ form the vertices of a rectangle. Same for the points $a_2, \mathtt{j}_j, b_2, \mathtt{j}_i$

The condition for two points p, q to be the vertices of an edge is given by algebraic equations. Visibly $p \in H_{i,j}$ means that $(p - v_i, \mathtt{j}_i - \mathtt{j}_j) = 0$, the corresponding $q = p + \mathtt{j}_j - \mathtt{j}_i$, while $p \in S_{i,j}$ is given by $(p - \mathtt{j}_i, p - \mathtt{j}_j) = 0$ and the corresponding opposite point q is given by $p + q = \mathtt{j}_i + \mathtt{j}_j$.

We thus have two types of constraints describing when two points are joined by an edge, a linear $q - p = \mathtt{j}_j - \mathtt{j}_i$ or $p + q = \mathtt{j}_i + \mathtt{j}_j$ and a quadratic constraint $(p - \mathtt{j}_i, \mathtt{j}_i - \mathtt{j}_j) = 0$ or $(p - \mathtt{j}_i, p - \mathtt{j}_j) = 0$. Given a connected component A of the graph we can choose one vertex $x \in S^c$ and use the linear constraints in order to write all the equations which define A by linear or quadratic equations on x. We keep track of the linear constraints by *marking* the edges by $\mathtt{j}_j - \mathtt{j}_i$ for black edges and $\mathtt{j}_j + \mathtt{j}_i$ for red ones.

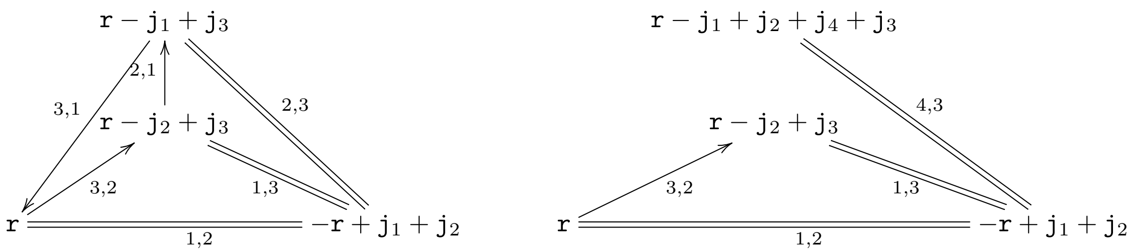
Now each connected component A has a purely combinatorial description which encodes the information on the edges which connect the vertices of A. We obtain an *abstract* graph with two types of edges (black, red) marked with pairs $i, j \in [1, \ldots, n]$.

A connected component of the geometric graph is a solution of a system of equations (associated to the graph) having the vertices as unknowns. It is easily seen that these equations may be all expressed on a single vertex (which we call the root $\mathtt{r}$), more precisely, as seen in [18] we obtain one equation (with unknown $\mathtt{r}$) for each vertex $v \neq \mathtt{r}$. A combinatorial graph of this type is *admissible* if its equations admit a solution $\mathtt{r} \in \mathbb{R}^d$ for generic values of the tangential sites.[3]

In [18] we have seen that such graphs have at most $2d$ vertices hence we have a finite list of combinatorial graphs (which we have described explicitly in terms of a Cayley graph, since we do not need it here we do not recall it). In [20] we have strengthened this estimate, shown that for a generic choice of S the vertices of the geometric graph, corresponding to an admissible combinatorial graph, are *affinely independent* and hence at most $d+1$. This stronger estimate is necessary for the proof of the second Melnikov condition.

We denote by $\mathcal{A}$ the combinatorial graph associated to A, note that $\mathcal{A}$ encodes the information on the equations which the vertices of A must solve so naturally there may be many A wich have the same $\mathcal{A}$.

Example 4.2.



the equations that $\mathtt{r}$ has to satisfy are:

$$\begin{aligned}
(\mathtt{r}, \mathtt{j}_2 - \mathtt{j}_3) &= |\mathtt{j}_2|^2 - (\mathtt{j}_2, \mathtt{j}_3)\\
|\mathtt{r}|^2 - (\mathtt{r}, \mathtt{j}_1 + \mathtt{j}_2) &= -(\mathtt{j}_1, \mathtt{j}_2)\\
(\mathtt{r}, \mathtt{j}_1 - \mathtt{j}_3) &= |\mathtt{j}_1|^2 - (\mathtt{j}_2, \mathtt{j}_3)
\end{aligned}$$

$$\begin{aligned}
(\mathtt{r}, \mathtt{j}_2 - \mathtt{j}_3) &= |\mathtt{j}_2|^2 - (\mathtt{j}_2, \mathtt{j}_3)\\
|\mathtt{r}|^2 + (\mathtt{r}, \mathtt{j}_1 + \mathtt{j}_2) &= -(\mathtt{j}_1, \mathtt{j}_2)\\
(\mathtt{r}, \mathtt{j}_1 - \mathtt{j}_2 - \mathtt{j}_3 - \mathtt{j}_4) &= -|\mathtt{j}_1|^2 + (\mathtt{j}_1, \mathtt{j}_2) + (\mathtt{j}_1, \mathtt{j}_3)\\
&\quad -(\mathtt{j}_2, \mathtt{j}_3) + (\mathtt{j}_1, \mathtt{j}_4) - (\mathtt{j}_2, \mathtt{j}_4) - (\mathtt{j}_3, \mathtt{j}_4)
\end{aligned}$$

In case the graph has no red edges the equations for the vertex x are all linear. This implies that the connected components of Γ_S which correspond to a given combinatorial graph with a chosen vertex are all obtained from a single one by translations by vectors which are orthogonal to the edges of the graph.

By convention we also have chosen a preferred vertex, called the *root*, in each connected component, in such a way that the roots of the translates are the translates of this root.

Formalizing,

- we have a map $\mathtt{r} : S^c \to S^c$ with image the chosen set $S^{c,\mathtt{r}}$ of roots.
- The fibers of this map are the connected components of the graph Γ_S.
- When we walk from the root $\mathtt{r}(k)$ to k (inside the corresponding connected component) we count the parity ± 1 of the number of red edges on the path, this is independent of the path and we denote by $\sigma(k)$ (the *color* of k).
- There are only finitely many elements k with $\sigma(k) = -1$, the finitely many corresponding roots are exactly the roots of the components with red edges.
- In any case the color of the root is always 1 (*black*).

At this point we can explain how to construct the elements $L(k)$ which tell us how to go from the root, of the component A of the graph Γ_S to which k belongs, to k. The equations defining the component A imply that

[3] we are interested only in solutions in S^c but it is more convenient to extend the possible solutions to $\mathbb{R}^d$.

(36)
$$k+\sum_i L_i(k)\mathtt{j}_i=\sigma(k)\mathtt{r}(k)\,,\quad |k|^2+\sum_i L_i(k)|\mathtt{j}_i|^2=\sigma(k)|\mathtt{r}(k)|^2\,,\quad \sigma(k)=1+\sum_i L_i(k).$$

Note that the first of the equations (36) defines the $L(k)$, which *depend only on the combinatorial graph.* The fact that this definition is well posed even if A is not a tree is a consequence of our genericity conditions.

The main fact is that

Proposition 4.3. *The Hamiltonian $\mathcal{Q}(\xi,x,w')$ in the new coordinates z' is the sum $\sum_\ell \mathcal{Q}_\ell(\xi,w')$ over all edges ℓ of the geometric graph of the following elements*

- $\mathcal{Q}_\ell(\xi,w'):=4\sqrt{\xi_i\xi_j}(z'_h\bar z'_k+z'_k\bar z'_h)$ *if h,k are joined by a black edge ℓ marked i,j*
- $\mathcal{Q}_\ell(\xi,w'):=4\sqrt{\xi_i\xi_j}(z'_h z'_k+\bar z'_h\bar z'_k)$ *if h,k are joined by a red edge ℓ marked i,j.*

Form the previous remarks there are only finitely many elements of the second type.

4.3.1. *The matrix blocks of $\tilde{\mathcal{Q}}$ and $ad(\mathcal{N})$.* According to Proposition 4.3, the graph has been constructed in such a way that we can group $\tilde{\mathcal{Q}}=\sum_A\tilde{\mathcal{Q}}_A$ (cf. (35)) where the sum runs over all blocks $A\in\Gamma_S$ and, if $E(A)$ denotes the set of edges in A:

$$\tilde{\mathcal{Q}}_A:=\sum_{k\in A}-2\xi\cdot L(k)|z'_k|^2+\sum_{\ell\in E(A)}\mathcal{Q}_\ell(\xi,w')$$

is a quadratic Hamiltonian in the variables $w'_A=z'_k,\bar z'_k$ with k running over the vertices of A. The matrix of $\tilde{Q}_A$ has a natural block diagonal structure in two conjugated blocks, corresponding to two Lagrangian subspaces in the symplectic space generated by the variables $z'_k,\bar z'_k,\ k\in A$ appearing in it. We can thus divide w'_A into two conjugate components $w'_A=(u',\bar u')$ where $u'_k=(z')_k^{\sigma(k)}$ then $-\frac{\mathrm{i}}{2}\tilde{Q}_A$ has as matrix denoted by $C_A\oplus -C_A$. By convention in the first block the root $\mathtt{r}$ corresponds to $z'_{\mathtt{r}}$.

Given two vertices u'_h,u'_k $h\neq k\in A$ we have that the matrix element $c_{u'_h,u'_k}$ of C_A is non zero if and only if h,k are joined by an edge (marked say (i,j)) and then

$$c_{u'_h,u'_k}=2\sigma(k)\sqrt{\xi_i\xi_j},\qquad c_{u'_k,u'_k}=-\sigma(k)(\xi,L(k)). \tag{37}$$

By definition $L(k)$ depends only on the combinatorial graph $\mathcal{A}$ of which A is a realization, therefore the matrix $C_A=C_{\mathcal{A}}$ depends only on the combinatorial block $\mathcal{A}$.

Remark 4.4. One may choose the root of each combinatorial graph so that any other vertex is connected by a path with at most $[(d+1)/2]$ edges. One deduces the estimates $\sum_i|L_i(k)|\le d+1,$ for all k.

4.4.1. *The space $F^{0,1}$.* In the KAM algorithm we shall need to study in particular the action by Poisson bracket of $\mathcal{N}$ on a special space of functions called $F^{0,1}$, so we recall some of this formalism.

Definition 4.5. We set $F^{0,1}$ to be the space of functions spanned by the basis elements

$$e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x}{z'_k}^{\sigma}=e^{\mathrm{i}\sigma([\sigma(k)\nu+L(k)]\cdot x)}{z_k}^{\sigma}$$

which preserve mass and momentum. [4]

[4] we deviate from the notations of [18] and in $F^{0,1}$ we also impose zero mass

One easily sees that $F^{0,1}$ is a symplectic space under Poisson bracket. The formulas for mass and momentum in the new variables are

$$\{\mathbb{L}, e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma}\} = \mathrm{i}\sigma(k)(\sum_i \nu_i + 1)\, e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma}, \tag{38}$$

$$\{\mathbb{M}, e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma}\} = \mathrm{i}\sigma(k)(\sum_i \nu_i \mathtt{j}_i + \mathtt{r}(k))\, e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma},$$

hence the conservation laws tell us that for an element $e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma} \in F^{0,1}$ the vector $\nu \in \mathbb{Z}^d$ is constrained by the fact that $-\sum_i \nu_i \mathtt{j}_i$ must be in the set of roots in S^c and moreover the mass constraint $\sum_i \nu_i = -1$.

For each connected components A of the graph Γ_S with some root $\mathtt{r}$ any solution ν of $\sum_i \nu_i \mathtt{j}_i + \mathtt{r} = 0$ determines in the space $F^{0,1}$ a block denoted A, ν with basis the elements $e^{\mathrm{i}\sigma(k)\nu\cdot x} {z'_k}^{\sigma}$ with ${z'_k}^{\sigma}$ the corresponding basis of the two Lagrangian blocks corresponding to A.

From the previous formulas we have thus that this space decomposes again into blocks indexed by pairs A, ν with A a connected component of the graph Γ_S and ν any solution of $\sum_i \nu_i \mathtt{j}_i + \mathtt{r} = 0$ where the mass of ν is -1, each such block is a symplectic space decomposed into a pair of $-\frac{\mathrm{i}}{2}\mathcal{N}$ stable Lagrangian subspaces.

Given thus such a pair of a component A and a frequency ν, notice that in fact A is determined by its root which is determined by ν by the conservation law. We have to understand the action of $-\frac{\mathrm{i}}{2}ad(\mathcal{N})$, on the block of $F^{0,1}$ with basis the elements $e^{\mathrm{i}\sum_j \nu_j x_j} {z'_m}^{\sigma(m)}$ with $\mathtt{r}(m) = -\sum_i \nu_i \mathtt{j}_i$, (on its conjugate $\bar{A}$ it is the minus transpose). The action of $\tilde{\mathcal{Q}}$ does not depend on ν and as before it is only through $\tilde{\mathcal{Q}}_A$ and gives the matrix C_A, we need then to understand the elements $(\omega(\xi), y') + \sum_{k\in S^c} \tilde{\Omega}_k |z'_k|^2$. By Formula (36), the term $\sum_{k\in S^c} \tilde{\Omega}_k |z'_k|^2$ contributes on the first block the scalar $|\mathtt{r}(m)|^2$. As of $(\omega(\xi), y')$ it also contributes by a scalar, this time $\sum_i \nu_i |\mathtt{j}_i|^2 - 2\sum_i \nu_i \xi_i$. Summarizing

Proposition 4.6. *The matrix of $-\frac{\mathrm{i}}{2}ad(\mathcal{N})$ on the block A, ν is the sum of the matrix $C_{\mathcal{A}}$ plus the scalar matrix $[\frac{1}{2}(|\mathtt{r}(m)|^2 + \sum_i \nu_i |\mathtt{j}_i|^2) - \sum_i \nu_i \xi_i]\, I_A$.*

4.6.1. *The standard form.* By the rules of Poisson bracket we have on the real space spanned by $z, \bar{z}$ that $\{a, \bar{b}\} = -\{\bar{a}, b\} = \{b, \bar{a}\}$ is imaginary so $\{\bar{a}, \bar{b}\} = -\{a, b\} = \overline{\{a, b\}}$ and

Definition 4.7. $(a, b) := \mathrm{i}\{a, \bar{b}\}$ is a real symmetric form, called the *standard form.*

For the variables we have $(z_h, z_h) = 1,\ (\bar{z}_h, \bar{z}_h) = -1$, so the form is positive definite on the space spanned by the z, negative on the space spanned by the $\bar{z}$ and of course indefinite if we mix the two types of variables. Thus we may say that an element a in the real space spanned by $z, \bar{z}$ is *of type z (resp. $\bar{z}$)* if $(a, a) = 1$ resp. $(a, a) = -1$. Now choose any quadratic real Hamiltonian $\mathcal{H} = \bar{\mathcal{H}}$. We have $\{\mathcal{H}, \{a, \bar{b}\}\} = 0$ by the Jacobi identity, moreover the map $x \mapsto \mathrm{i}\{\mathcal{H}, x\}$ preserves the real subspace spanned by $z, \bar{z}$ hence we have

$$(a, \mathrm{i}\{\mathcal{H}, b\}) = \mathrm{i}\{a, \overline{\mathrm{i}\{\mathcal{H}, b\}}\} = \{a, \{\mathcal{H}, \bar{b}\}\} = -\{\{\mathcal{H}, a\}, \bar{b}\} = \mathrm{i}\{\mathrm{i}\{\mathcal{H}, a\}, \bar{b}\} = (\mathrm{i}\{\mathcal{H}, a\}, b). \tag{39}$$

Formula (39) tells us that the operator $\mathrm{i}\{\mathcal{H}, -\}$ is symmetric with respect to this form.

4.7.1. *The case of $-\frac{\mathrm{i}}{2}\{\tilde{\mathcal{Q}}, -\}$.* We apply the previous analysis to $\mathcal{H} = -\frac{1}{2}\tilde{\mathcal{Q}}$ and its block decomposition. When we have red edges each of the two Lagrangian blocks contains both variables z and $\bar{z}$ and by convention we take as first block the one in which the variable corresponding to the root is of type z. The standard form $(a, b) := \mathrm{i}\{a, \bar{b}\}$ is indefinite.

By assumption the operator $-\frac{\mathrm{i}}{2}\{\tilde{\mathcal{Q}}, -\}$ can be put in normal form by a change of basis preserving the form (a, b). Thus the new basis is formed still by elements which we have called of types z and $\bar{z}$.

From all these considerations one has:

Lemma 4.8. *For all combinatorial blocks $\mathcal{A}$ which do not contain red edges, the matrix $C_{\mathcal{A}}$ is self–adjoint for all $\xi \in \mathbb{R}^n_+$. If $\mathcal{A}$ contains red edges then each vertex k has a sign and corresponds to an element $u'_k = (z')_k^{\sigma(k)}$.*

The diagonal matrix of signs $\sigma_{\mathcal{A}} = diag(\sigma(k))$ is the matrix of the standard form in the basis $u'_k = (z')_k^{\sigma(k)}$ and $C_{\mathcal{A}}$ is self–adjoint with respect to the indefinite form defined by $\sigma_{\mathcal{A}}$.

The orthogonal group of the standard form acts on the entire symplectic block preserving the two Lagrangian subspaces and thus it has a mixed invariant which we still call the standard form

$$\sum_{k \,|\, \mathtt{r}(k)=r} \sigma(k)|z_k|^2 = \sum_{k \,|\, \mathtt{r}(k)=r} \sigma(k)|z'_k|^2. \tag{40}$$

Lemma 4.9 (conservation laws)**.** *In the new variables the conserved quantities are:*

$$\mathbb{L} = \sum_i y'_i + \sum_k \sigma(k)|z'_k|^2\,, \quad \mathbb{M} = \sum_i \mathtt{j}_i y'_i + \sum_k \sigma(k)\mathtt{r}(k)|z'_k|^2\,,$$

$$\mathbb{K} = \sum_i |\mathtt{j}_i|^2 y'_i + \sum_k \sigma(k)|\mathtt{r}(k)|^2|z'_k|^2$$

note that all of these three quadratic Hamiltonians are represented by a scalar *matrix on each component of w'_A.*

Proof. We substitute the new variables and use the identities (36). $\square$

5. Normal form reduction

We now want to simplify $-\frac{\mathrm{i}}{2}\mathcal{N}$ on each pair of stable Lagrangian subspaces described in Proposition 4.6, using the standard Theory of canonical form of symplectic matrices.

The main ingredient we need is :

Theorem 3. *[Proposition 1 of* [20]*] i) For all combinatorial blocks $\mathcal{A}$, $C_{\mathcal{A}}$ has distinct eigenvalues, namely it is regular semisimple for values of the parameters ξ_i outside a real hypersurface (the discriminant).*

ii) For any pair of distinct blocks $(A_1, \nu_1), (A_2, \nu_2)$ the resultant of the characteristic polynomials of the two matrices of the action of $-\frac{\mathrm{i}}{2}ad(\mathcal{N})$ is non–zero, hence outside this hypersurface the eigenvalues of these two blocks are distinct.

The algebraic hypersurface union of all discriminant varieties for all the combinatorial matrices $C_{\mathcal{A}}$ will be denoted by $\mathfrak{A}$ and called *discriminant*. It is given by a homogeneous polynomial equation, thus $(\mathbb{R}_+)^n \setminus \mathfrak{A}$ is a union of finitely many connected open cones $(\mathbb{R}_+)^n_1, \dots, (\mathbb{R}_+)^n_M$ where the number of real, resp. complex eigenvalues of any given combinatorial matrix $C_{\mathcal{A}}$ is constant. On each of these regions $(\mathbb{R}_+)^n_\alpha$ we can thus describe a normal form.

Since our normal form, thought of as operator has possibly also complex eigenvalues let us recall the basic normal form of the simplest Hamiltonians.

Consider $H_\vartheta := a(|z_1|^2 - |z_2|^2) + b(z_1 z_2 + \bar z_1 \bar z_2)$, setting $\vartheta = a + \mathrm{i}b$. On the space with basis $z_1, \bar z_2, \bar z_1, z_2$ (symplectic form J) the operator $-\mathrm{i}ad(H)$ has matrix.

$$M_\vartheta = \begin{vmatrix} a & -b & 0 & 0 \\ b & a & 0 & 0 \\ 0 & 0 & -a & b \\ 0 & 0 & -b & -a \end{vmatrix}, \qquad J = \begin{vmatrix} 0 & 0 & -\mathrm{i} & 0 \\ 0 & 0 & 0 & \mathrm{i} \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \end{vmatrix}$$

with eigenvalues $\pm\vartheta, \pm\bar\vartheta, \quad \vartheta = a + \mathrm{i}b$. One easily sees that a 4×4 real symplectic matrix commuting with this matrix, when the 4 eigenvalues $\pm\vartheta, \pm\bar\vartheta$ are distinct, has the same block form M_β for some complex number $\beta = c + \mathrm{i}d$ and so it is represented by the Hamiltonian $H_\beta = c(|z_1|^2 - |z_2|^2) + d(z_1 z_2 + \bar z_1 \bar z_2)$.

Now we need to decompose the various combinatorial blocks that we previously described. We have already defined the discriminant hypersurface $\mathfrak{A}$. It is now convenient to choose the compact domain $\mathfrak{K}$ of Formula (2.1) to be a union of compact domains

$$\mathfrak{K} = \cup_\alpha \mathfrak{K}_\alpha \tag{41}$$

each contained in the corresponding open connected component $(\mathbb{R}_+)^n_\alpha$ of $(\mathbb{R}_+)^n \setminus \mathfrak{A}$. For each combinatorial matrix $C_{\mathcal{A}}$ the standard form $\sigma_{\mathcal{A}}$ on $(\mathbb{R}_+)^n_\alpha$ has constant signature and we have:

Proposition 5.1 (cf. Williamson [23]). *On each region $(\mathbb{R}_+)^n_\alpha$ the eigenvalues of the $C_{\mathcal{A}}$ are analytic functions of ξ, say $\vartheta_1, \ldots, \vartheta_{dim(\mathcal{A})}$.*

For all $\xi \in (\mathbb{R}_+)^n_\alpha$ there exists a linear symplectic change of coordinates $u' \to U_{\mathcal{A}}(\xi)u' = u''$ such that:

1. *$U_{\mathcal{A}}(\xi)$ is orthogonal with respect to $\sigma_{\mathcal{A}}$*
2. *$U_{\mathcal{A}}(\xi)$ is analytic in ξ.*
3. *$U_{\mathcal{A}}(\xi)$ conjugates $C_{\mathcal{A}}$ into the following normal form:*

For each real eigenvalue ϑ , $C_{\mathcal{A}}$ acts as ϑI on the (one dimensional) eigenspace of ϑ in u_A.

For each pair of conjugate complex eigenvalues $\vartheta_\pm = a \pm \mathrm{i}b$, we have a real two dimensional space such that the two complex eigenvectors lie in its complexification. Then we have a basis of this subspace such that $C_{\mathcal{A}}$ restricted to this subspace is a 2×2 matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

The matrix $\sigma_{\mathcal{A}}$ of the standard form on this basis is $diag(1, -1)$ so one of the variable is a z and the other a $\bar z$.

Remark 5.2. We note that the matrices $C_{\mathcal{A}}$ have entries wich are homogeneous of degree one in ξ. Therefore given any compact domain O which does not intersect $\mathfrak{A}$ the entries of the matrices $U_{\mathcal{A}}(\xi), U_{\mathcal{A}}(\xi)^{-1}, \partial_\xi U_{\mathcal{A}}(\xi)$ can be uniformly bounded in ρO independently of ρ, in particular this applies to each $\mathfrak{K}_\alpha$.

Since the $U_{\mathcal{A}}$ is determined by $\mathcal{A}$ we denote its matrix elements by $[U_{\mathcal{A}}]_{a,b}$ where a, b run over the vertices of $\mathcal{A}$. Namely on a given geometric block A isomorphic to $\mathcal{A}$, $[U_{\mathcal{A}}]_{a,b}$ is the entry relative to the elements z_k^σ associated to a, b respectively.

Given a geometric block A let $\mathcal{A}$ be the corresponding combinatorial block. In each connected component, the non-unique choice of the matrix U, putting in canonical form the matrix $C_{\mathcal{A}}$ determines a symplectic change of variables for all blocks A with combinatorial block $\mathcal{A}$, we do this for all the finitely many $\mathcal{A}$. We may then index the new variables still by S^c and decompose S^c in two sets: an infinite set S^c_r, which indexes the real eigenvalues, namely u_k is an eigenvector of $-\frac{\mathrm{i}}{2}\mathcal{Q}$ of real eigenvalue $\vartheta_k(\xi)$. Then a finite set S_i^c which indexes the complex eigenvalues (note that by the special block structure of the Hamiltonian there are no purely imaginary eigenvalues). By the reality condition each two by two block corresponding to a pair of conjugate eigenvalues is indexed by a pair (h,k) of elements of S_i^c. We write the conjugate eigenvalues as $\vartheta_{h,k},\bar\vartheta_{h,k}$ with $\vartheta_{h,k}=a_{h,k}+\mathrm{i}b_{h,k}$. Note that for $(h,k)\in S_i^c$ we have $\sigma(h)=1$ and $\sigma(k)=-1$. By abuse of notation we still call $x,y,z_k,\bar z_k$ the new variables. We have finally the final *diagonal form of the Hamiltonian*

Theorem 4. *i) For each connected component of $(\mathbb{R}_+)^n_\alpha$ on $\varepsilon^2\mathfrak{K}_\alpha$ we have a symplectic change of variables Ξ (depending analytically on ξ) which puts the Hamiltonian $\mathcal{N}$ in the canonical* diagonal *form $\mathcal{N}=\mathbb{K}+2\mathbb{K}^1$ where*

$$\mathbb{K}^1=-\sum_{i=1}^n\xi_iy_i+\sum_{k\in S_r^c}\sigma(k)\vartheta_k|z_k|^2+\sum_{(h,k)\in S_i^c}a_{h,k}(|z_h|^2-|z_k|^2)+b_{h,k}(z_hz_k+\bar z_h\bar z_k) \tag{42}$$

The elements $\vartheta_k,a_{h,k},b_{h,k}$ are analytic functions of ξ and homogeneous of degree one..

ii) [Elliptic open set] There exists a connected component (hence a non empty open cone) of $\mathbb{R}^n_+$ such that in this region all the eigenvalues are real i.e. S_i^c is empty ([20]*, proposition 1.13).*

Remark 5.3. In order to simplify the notations we shall write the final variables as z_k since no confusion should arise with the initial variables.

We claim that for $\mathbb{L},\mathbb{M},\mathbb{K}$ we have still

$$\mathbb{L}=\sum_i y_i+\sum_k\sigma(k)|z_k|^2\,,\quad \mathbb{M}=\sum_i \mathtt{j}_iy_i+\sum_k\sigma(k)\mathtt{r}(k)|z_k|^2\,, \tag{43}$$
$$\mathbb{K}=\sum_i|\mathtt{j}_i|^2y_i+\sum_k\sigma(k)|\mathtt{r}(k)|^2|z_k|^2.$$

In fact it is enough to compare the contribution of each block of given root $\mathtt{r}$, to these three quantities, it is the sum $\sum_{k\,|\,\mathtt{r}(k)=r}\sigma(k)|z_k|^2$ times $1,\mathtt{r}(k),|\mathtt{r}(k)|^2$ respectively. So it is enough to see that the quadratic expression $\sum_{k\,|\,\mathtt{r}(k)=r}\sigma(k)|z_k|^2$ remains invariant. This expression is in fact the standard form (40) and in our case we can diagonalize the matrix block by an orthogonal transformation with respect to this form, the claim follows.

Corollary 5.4. *i) Let $(\mathbb{R}_+)^n_e$ be the elliptic open region, where all eigenvalues of all combinatorial matrices are real and $\mathfrak{K}_e$ the corresponding domain.*

ii) For all $\xi\in\varepsilon^2\mathfrak{K}_e$ the NLS normal form is

$$(\omega,y)+\sum_k\Omega_k|z_k|^2\,,$$

where

$$\omega_i=|\mathtt{j}_i|^2-2\xi_i\,,\quad \Omega_k=\sigma(k)(|\mathtt{r}(k)|^2+2\vartheta_k).$$

The functions ϑ_k are real valued and analytic.

5.4.1. *Complex coordinates.* Although it is not strictly necessary it is convenient to diagonalize also the blocks with complex eigenvalues, although this implies the introduction of non real symplectic transformations.

We write everything in possibly complex coordinates as $\mathcal{N}=(\omega,y)+\sum_{k\in S^c}\Omega_k|\zeta_k|^2$, where $\zeta_k=z_k$ if $k\in S_r^c$. In order to do this we have to define ζ_k for $k\in S_i^c$. Consider one of the terms , say ${\mathbb{K}_1}^{(h,k)}:=a_{h,k}(|z_h|^2-|z_k|^2)+b_{h,k}(z_hz_k+\bar z_h\bar z_k)$ and set $\vartheta_k:=$

$\overline{a_{h,k} + \mathrm{i} b_{h,k}}$, $\vartheta_h = a_{h,k} - \mathrm{i} b_{h,k}$. Think of $z \mapsto \bar{z}$ as a $\mathbb{C}$ *linear map on polynomials!* so that $\overline{z_h + \mathrm{i}\bar{z}_k} = \bar{z}_h + \mathrm{i} z_k$ and thus setting $\zeta_h := \frac{z_h + \mathrm{i}\bar{z}_k}{\sqrt{2}}$, $\zeta_k := \frac{\bar{z}_h - \mathrm{i} z_k}{\sqrt{2}}$ we have that $\zeta_h, \bar{\zeta}_h$ (resp. ζ_k, $\bar{\zeta}_k$) are eigenvectors with opposite eigenvalues for $ad(\mathbb{K}_1)$:

$$\{\mathbb{K}_1^{(h,k)}, \zeta_h^\sigma\} = \sigma \mathrm{i} \vartheta_h \zeta_h^\sigma, \qquad \{\mathbb{K}_1^{(h,k)}, \zeta_k^\sigma\} = -\sigma \mathrm{i} \vartheta_k \zeta_k^\sigma$$

$$\{\zeta_h^{\sigma_1}, \zeta_k^{\sigma_2}\} = 0, \quad \{\bar{\zeta}_h, \zeta_h\} = \{\bar{\zeta}_k, \zeta_k\} = \mathrm{i}, \ \bar{\alpha}|\zeta_h|^2 + \alpha|\zeta_k|^2 = \mathbb{K}_1^{(h,k)}.$$

Moreover in these coordinates the three quantities $\mathbb{L}, \mathbb{M}, \mathbb{K}$ are still in the form of Formula (43). In fact the same argument that we gave before applies since the complex transformation that we have used is in the orthogonal group of the form.

We can finally claim that we can use the notation z_k also for complex coordinates and write $\mathcal{N} = (\omega, y) + \sum_{k \in S^c} \Omega_k |z_k|^2$ we assume that the Ω_k are all distinct and the complex ones come together with their conjugates according to the previous rules.

Summarizing:

A monomial $\mathfrak{m} = e^{\mathrm{i}(k,x)} y^l z^\alpha \bar{z}^\beta$ has momentum $\mathrm{i}\pi_{\mathtt{r}}(\mathfrak{m})$ with

(44)
$$\pi_{\mathtt{r}}(\mathfrak{m}) := \pi_{\mathtt{r}}(k, \alpha, \beta) = \pi(k) + \sum_{j \in S^c} (\alpha_j - \beta_j) \sigma(j) \mathtt{r}(j) = \pi(k, \alpha, \beta) + \sum_{j \in S^c} (\alpha_j - \beta_j)(\sigma(j)\mathtt{r}(j) - j)$$

and it satisfies momentum conservation if $\pi_{\mathtt{r}}(k, \alpha, \beta) = 0$. Note that given functions f, g which are eingenvectors of momentum we have $\pi_{\mathtt{r}}(\{f, g\}) = \pi_{\mathtt{r}}(f) + \pi_{\mathtt{r}}(g)$. Given $k \in S^c$ (corresponding to the eigenvalue ϑ_k of $C_{\mathcal{A}}$) the monomial $e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma$ is an eigenvector for all our operators with eigenvalues:

$$\{\mathbb{L}, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma\} = \mathrm{i}\sigma\sigma(k)(\sum_i \nu_i + 1)\, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma,$$

$$\{\mathbb{M}, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma\} = \mathrm{i}\sigma\sigma(k)(\sum_i \nu_i \mathtt{j}_i + \mathtt{r}(k)) e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma$$

$$\{\mathbb{K}, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma\} = \mathrm{i}\sigma\sigma(k)(\sum_i \nu_i |\mathtt{j}_i|^2 + |\mathtt{r}(k)|^2)\, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma,$$

$$\{\mathbb{K}^1, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma\} = \mathrm{i}\sigma\sigma(k)(-\sum_i \nu_i \xi_i + \vartheta_k)\, e^{\mathrm{i}\sigma\sigma(k)\nu\cdot x} z_k^\sigma,$$

6. The kernel of $ad(\mathcal{N})$

6.0.2. *Non-degenerate quadratic Hamiltonians.* Consider a quadratic Hamiltonian

$$\mathcal{Q} = (\omega, y) + \sum_{k \in S_r^c} a_k |z_k|^2 + \sum_{(h,k) \in S_i^c} a_{h,k}(|z_h|^2 - |z_k|^2) + b_{h,k}(z_h z_k + \bar{z}_h \bar{z}_k) \tag{45}$$

we want to study the kernel of $ad(\mathcal{Q})$ on the space of Hamiltonians of degree ≤ 2. We set $\vartheta_{h,k} := a_{h,k} + \mathrm{i} b_{h,k}$ so that $\pm\vartheta_{h,k}, \pm\bar{\vartheta}_{h,k}$ are the four eigenvalues of $ad(a_{h,k}(|z_h|^2 - |z_k|^2) + b_{h,k}(z_h z_k + \bar{z}_h \bar{z}_k))$ acting on the space spanned by $z_h, z_k, \bar{z}_h, \bar{z}_k$. It is convenient to write $z^1 = z$, $z^{-1} = \bar{z}$ so if we do not want to specify if a variable is z or $\bar{z}$ we write is as z^σ where σ can be ± 1.

Next the operator $\mathrm{i}\mathcal{Q}$ acts on the real space spanned by the elements $e^{\mathrm{i}\sigma\nu\cdot x} z_k^\sigma$ (we need some convergence conditions given by its norm). If as in our case $\mathcal{Q}$ commutes with the mass and momentum then it acts also on the subspace $F^{0,1}$ where

Definition 6.1. We denote by $F^{0,1}$ the space of functions spanned by the elements $e^{\mathrm{i}\sigma\nu\cdot x} z_k^\sigma$ with zero mass and momentum. I.e. $\sum_i \nu_i = -1, \sum_i \nu_i \mathtt{j}_i + k = 0$.

Definition 6.2. We say that $\mathcal{Q}$ is *non-degenerate*[5] if the coordinates ω_i are linearly independent over $\mathbb{Q}$ and its eigenvalues for the action on $F^{0,1}$ are all non–zero and distinct.

It is then not difficult to analyze the kernel of $\mathrm{i}\,ad(\mathcal{Q})$, i.e. the elements which Poisson commute with $\mathrm{i}\,ad(\mathcal{Q})$, on the space of Hamiltonians of degree ≤ 2 commuting with mass and momentum, we have:

Proposition 6.3. *If $\mathcal{Q}$ is non-degenerate then a (real) Hamiltonian of degree ≤ 2, which commutes with mass, Poisson commutes with $\mathcal{Q}$ if and only if it is of the form:*

$$\mathcal{Q}' = (\omega', y) + \sum_{k \in S_r^c} a'_k |z_k|^2 + \sum_{(h,k) \in S_i^c} a'_{h,k}(|z_h|^2 - |z_k|^2) + b'_{h,k}(z_h z_k + \bar{z}_h \bar{z}_k) \tag{46}$$

with $a'_k, a'_{h,k}, b'_{h,k} \in \mathbb{R}$.

Proof. It is immediate that a Hamiltonian of the form of Formula (46) commutes with $\mathcal{Q}$ we need to show the converse.

Degree zero in w: Monomials of degree ≤ 2 and of degree 0 in w are of the form $y^\ell e^{\mathrm{i}\nu\cdot x}$, $\ell = 0, 1$ and for the eigenvalues we have:

$$\{\mathcal{Q}, y^\ell e^{\mathrm{i}\nu\cdot x}\} = -\mathrm{i}(\omega, \nu)\, y^\ell e^{\mathrm{i}\nu\cdot x} \tag{47}$$

By hypothesis of linear independence over $\mathbb{Q}$ these eigenvalues are 0 if and only if $\nu = 0$, hence the Kernel of $ad(\mathcal{N})$ is x independent and hence of the form $c + (\omega', y)$.

Degree one in w: By definition the eigenvalues of the adjoint action of $\mathcal{Q}$ on $F^{0,1}$ are non-zero.

Degree two in w: We write everything in possibly complex coordinates as $\mathcal{Q} = (\omega, y) + \sum_{k \in S^c} a_k |\zeta_k|^2$ as in the previous paragraph. We assume that the a_k are all distinct and the complex ones come together with their conjugates according to the previous rules. Then we can write any monomial of degree 2

$$M = e^{\mathrm{i}\nu\cdot x} \zeta_h^{\sigma_1} \zeta_k^{\sigma_2}, \quad \{\mathcal{Q},\, M\} = \mathrm{i}[(\omega, \nu) + \sigma_1 a_h + \sigma_2 a_k] M$$

conservation of mass $\eta(\nu) + \sigma_1 + \sigma_2 = 0$ implies that we can write $\nu = \nu_1 + \nu_2$ with $\eta(\nu_1) + \sigma_1 = \eta(\nu_2) + \sigma_2 = 0$. Then $(\omega, \nu_1) + \sigma_1 a_h$ and $(\omega, \nu_2) + \sigma_2 a_k$ are two eigenvalues of elements in $F^{0,1}$ since the eigenvalues are all distinct we must have that the corresponding eigenvectors are one the conjugate of the other and then M is of the form $|\zeta_k|^2$ for some k. Finally if we assume that the Hamiltonian is real the two terms associated to a pair h, k giving complex eigenvalues must have conjugate coefficients so that in real coordinates they give a term of type $a'_{h,k}(|z_h|^2 - |z_k|^2) + b'_{h,k}(z_h z_k + \bar{z}_h \bar{z}_k)$. $\square$

Warning From now on even if we shall use complex coordinates we shall denote them by z_k and not ζ_k.

6.4. **Eigenvalues and eigenvectors.** The fact that the Hamiltonian $\mathcal{N}$ is non-degenerate for generic values of ξ is essentially a consequence of Theorem 3, we state it as:

Proposition 6.5. *The normal form $\mathcal{N}$ is non-degenerate for all ξ outside countably many algebraic hypersurfaces.*

[5] in the usual language we should say *regular semisimple*.

Proof. We have $\omega_i = |\mathtt{j}_i|^2 - 2\xi_i$ which are linearly independent over the rationals outside countably many algebraic hypersurfaces (the dependency relations).

The eigenvalues of the action on $F^{(0,1)}$ are the roots of the characteristic polynomials of the blocks into which this space decomposes. We have seen that these polynomials are all irreducible and distinct. For a block of size > 1 irreducibility implies that the constant term of the characteristic polynomial is a non–zero polynomial in ξ so the eigenvalues are non–zero outside the hypersurfaces given by these determinants. Moreover it is immediate that for the blocks of size 1 the eigenvalues are non–zero linear polynomials.

In order to have that all the eigenvalues be distinct we have to remove the discriminant $\mathfrak{A}$, and countably many *resultants* which are all non–zero polynomials in ξ by the irreducibility and separation Theorem. □

Remark 6.6. In the course of the KAM algorithm of Part 3 we shall see that the non-degeneracy of the Normal form plays a fundamental role. Imposing the non-degeneracy however requires removing a countable number of proper hypersurfaces hence working on complicated sets in parameter-space. In order to avoid this problem, we impose that the compact domains $\mathfrak{K}_\alpha$ of Formula (41), should be disjoint from finitely many resultants, i.e. we fix an integer S_0 and impose $\forall \xi \in \cup_\alpha \mathfrak{K}_\alpha$:

$$-(\xi, k) + \vartheta_i(\xi) \pm \vartheta_j(\xi) \neq 0\,.$$

for all the couples of eigenvalues of (different) combinatorial matrices and forall $k \in \mathbb{Z}^n\,, |k| < S_0$.

We would like to choose coordinates, independent of ξ, which are eigenvectors for $\mathbb{M}$, $\mathbb{L}$, $\mathcal{N}$ so that

Lemma 6.7. *A regular analytic function F Poisson commutes with $\mathcal{N}$, $\mathbb{M}$, $\mathbb{L}$ for generic values of ξ if and only if each monomial appearing in F Poisson commutes with $\mathbb{M}$, $\mathbb{L}$, $\mathbb{K}$ and $\mathbb{K}^1$.*

Proof. An element commutes with $\mathcal{N}$ if and only if it commutes with both homogeneous parts of degree 0,1 that is $\mathbb{K}$ and $\mathbb{K}^1$. □

Part 2. Quasi Töplitz functions

7. Optimal presentations cuts and Good Points

Let us recall the definition of N-optimal presentations, cuts, good points as given in [24]*, we omit most proofs.*

7.1. **N-optimal presentations.** An affine space A of codimension ℓ in $\mathbb{R}^d$ can be defined by a list of ℓ equations $A := \{x \,|\, (v_i, x) = p_i\}$ where the v_i are independent row vectors in $\mathbb{R}^d$. We will write shortly that $A = [v_i; p_i]_\ell$. We will be interested in particular in the case when v_i, p_i have integer coordinates, i.e. are *integer vectors*[6] and the vectors v_i lie in a prescribed ball B_N of radius some constant κN. Recall that, (8), we have set $\kappa := \max_i |\mathtt{j}_i|$. We denote by

$$\langle v_i \rangle_\ell = \mathrm{Span}(v_1, \dots, v_\ell; \mathbb{R}) \cap \mathbb{Z}^d\,, \quad B_N := \{x \in \mathbb{Z}^d \setminus \{0\} \,|\, |x| < \kappa N\},$$

here N is any large number. In particular we implicitly assume that B_N contains a basis of $\mathbb{R}^d$.

[6]such a subspace is usually called *rational*.

For given $s \in \mathbb{N}$, in the set of vectors $\mathbb{Z}^s$ we can define the *sign lexicographical order* $\prec$ *as in* [24] as follows.

Definition 7.2. Given $a = (a_1, \dots, a_s)$ set $(|a|) := (|a_1|, \dots, |a_s|)$ then we set $a \prec b$ if either $(|a|) < (|b|)$ in the lexicographical order (over $\mathbb{N}$) or if $(|a|) = (|b|)$ and $a > b$ in the lexicographical order in $\mathbb{Z}$.

With this definition every non empty set of elements in $\mathbb{Z}^s$ has a unique minimum.

Notice that, by convention, among the finite number of vectors with a given prescribed value of $(|a|)$ we have chosen as minimum the one with non negative coordinates.

In particular consider a fixed but large enough N.

Definition 7.3. We set $\mathcal{H}_N$ to be the set of all affine spaces A which can be presented as $A = [v_i; p_i]_\ell$ for some $0 < \ell \le d$ so that that $v_i \in B_N$, $p_i \in \mathbb{N}$.

We denote the subset of $\mathcal{H}_N$ formed by the subspaces of codimension ℓ by $\mathcal{H}_N^\ell$.

We display as $(p_1, \dots, p_\ell; v_1, \dots, v_\ell)$ a given presentation, so that it is a vector in $\mathbb{Z}^{\ell(d+1)}$. Then we can say that $[v_i; p_i]_\ell \prec [w_i; q_i]_\ell$ if $(p_1, \dots, p_\ell; v_1, \dots, v_\ell) \prec (q_1, \dots, q_\ell; v_1, \dots, v_\ell)$.

Definition 7.4. The N–optimal presentation $[l_i; q_i]_\ell$ of $A \in \mathcal{H}_N^\ell$ is the minimum, in the sign lexicographical order, of the presentations of A which satisfy the previous bounds.

Given an affine subspace $A := \{x \,|(v_i, x) = p_i\,,\ i = 1, \dots, \ell\}$ by the notation $A \stackrel{N}{\to} [v_i; p_i]_\ell$ we mean that the subspace has codimension ℓ and the given presentation is N–optimal.

Remark 7.5. i) Note that each point $m = (m_1, \dots, m_d) \in \mathbb{Z}^d$ has a N–optimal presentation (this presentation is usually not the naive one $[e_i, m_i]_d$ where the e_i form the standard basis of $\mathbb{Z}^d$).

ii) Thus we may use the ordering given by N–optimal presentations of points in order to define a new lexicographic order on $\mathbb{Z}^d$ which we shall denote by $a \prec_N b$ or $a \prec b$ when N is understood.

Remark 7.6. At this point we extend the definition of $\prec$ to the elements of $\mathcal{H}_N$ by using their N–optimal presentation.

Lemma 7.7. *i) If the presentation $A = [v_i; p_i]_\ell$ is N–optimal, we have*

$$0 \le p_1 \le p_2 \le \dots \le p_\ell. \tag{48}$$

ii) For all $j < \ell$ and for all $v \in B_N \cap (\langle v_1, \dots, v_\ell\rangle \setminus \langle v_1, \dots, v_j\rangle)$, one has:

$$|(v, r)| \ge p_{j+1}\,, \quad \forall r \in A. \tag{49}$$

iii) Given $j < \ell$ set $A_j := \{x \,|\, (v_i, x) = p_i,\ i \le j\}$, then the presentation $A_j = [v_i, p_i]_j$ is N–optimal.

iv) Finally $-A$ has a N–optimal presentation $-A = [v'_i, p_i]$ with the same constants p_i and $(|v'_i|) = (|v_i|)$.

Remark 7.8. For fixed N, ℓ, p the number of affine spaces in $\mathcal{H}_N^\ell$ of codimension ℓ and such that $p_\ell \le p$ is bounded by $(2\kappa N + 1)^{\ell d}(p + 1)^\ell$.

7.9. **A decomposition of $\mathbb{Z}^d$.** We shall need several auxiliary parameters in the course of our proof. We start by fixing some numbers

$$\tau_0 > \max(d^2 + n, 12), \quad \tau_1 := (4d)^{d+1}(\tau_0 + 1)\,, \tag{50}$$

$$\mathtt{c} \le \frac{1}{2}\,,\ \mathtt{C} \ge 4\,,\ N_0 = (d+1)!\kappa^{d+1}\mathtt{C}\mathtt{c}^{-1}.$$

In what follows N will always denote some large number, in particular $N > N_0$.

Using the fixed parameters $\mathtt{c}, \mathtt{C}$ and the notion of optimal presentation, for each $N > N_0$ we want to construct a decomposition

$$\mathbb{Z}^d = \cup_{i=0}^d A_i(N) \tag{51}$$

of $\mathbb{Z}^d$ which will be crucial for the estimates of small denominators (cf. §10.7.1) and given by the following

Definition 7.10. i) A subspace $A \xrightarrow{N} [v_i; p_i]_\ell \in \mathcal{H}_N^\ell$ with $1 \le \ell < d$ is called N*–good* if $p_\ell \le \mathtt{c} N^{\frac{\tau_1}{4d}}$. The set of N–good subspaces of codimension $\ell < d$ is denoted by $\mathcal{H}_N^{\ell, g}$.

ii) Given $A \in \mathcal{H}_N^{\ell, g}$ the set:

$$A^g := \left\{ x \in A \cap \mathbb{Z}^d \mid \quad |\mathtt{r}(x)| > \mathtt{C} N^{\tau_1}\,,\ |(v, x)| \ge \mathtt{C} \max(N^{4d\tau_0}, \mathtt{c}^{-4d} p_\ell^{4d}), \forall v \in B_N \setminus \langle v_i \rangle_\ell \right\} \tag{52}$$

will be called the $N-$*good* portion of the subspace A.

Remark 7.11. Notice that every $v \in B_N \setminus \langle v_i \rangle_\ell$ gives a non constant linear function (v, x) on A. Thus the good points of A form a non empty open set complement of a finite union of strips around subspaces of codimension 1 in A. Note moreover that we are interested only in integral points and the integral points in A which are not good are formed, by the finitely many points with $|\mathtt{r}(x)| \le \mathtt{C} N^{\tau_1}$, plus a finite union of affine subspaces of codimension one in A.

We construct a decomposition of $\mathbb{Z}^d$ using the following Proposition which is a variation of [24] Proposition 1).

Proposition 7.12. *Each point $m \xrightarrow{N} [v_i, p_i]$ with $|\mathtt{r}(m)| > \mathtt{C} N^{\tau_1}$ and $p_1 < \mathtt{C} N^{4d\tau_0}$ belongs to the set $[v_i; p_i]_\ell^g$ for some choice $0 < \ell < d$.*

Proof. Consider $A_1 := [v_1; p_1]$, since $\mathtt{C} N^{4d\tau_0} \le \mathtt{c} N^{\frac{\tau_1}{4d}}$ we see that it is N–good, so if $m \in A_1^g$ we are done, otherwise we have that $p_2 < \mathtt{C} \max(N^{4d\tau_0}, \mathtt{c}^{-4d} p_1^{4d}) \le \mathtt{c} N^{\frac{\tau_1}{4d}}$. Consider $A_2 := [v_1, v_2; p_1, p_2]$, we see again that it is N–good, so if $m \in A_2^g$ we are done otherwise we have that $p_3 < \mathtt{C} \max(N^{4d\tau_0}, \mathtt{c}^{-4d} p_2^{4d}) \le \mathtt{c} N^{\frac{\tau_1}{4d}}$ we continue in this way and either we show that $m \in A_i^g$, $i < d$ or we have a sequence of inequalities $p_j < \mathtt{C} \max(N^{4d\tau_0}, \mathtt{c}^{-4d} p_{j-1}^{4d})$. Let $k \ge 1$ be the maximum index such $p_k < \mathtt{C} N^{4d\tau_0}$, if $k < d$ we have for all $p_j, j = k+1, \dots, d$ that $p_j < (\mathtt{c}^{-4d}\mathtt{C})^{j-k} N^{(4d)^{j-k+1}\tau_0}$. We then compute the coordinates of m by Cramer's rule and by just estimating the numerator we have a sum of $d!$ terms each of which can be bounded by $N^{d-1} N^{(4d)^d \tau_0} (\mathtt{c}^{-4d}\mathtt{C})^{d-1}$. This sum is then bounded by $N^{5d(d-1)+(4d)^d\tau_0}$, it follows that $|m| \le \sqrt{d} N^{5d(d-1)+(4d)^d\tau_0}$ and $\mathtt{r}(m) \le \sqrt{d} N^{5d(d-1)+(4d)^d\tau_0} + d\kappa$. Recall $\tau_1 := (4d)^{d+1}(\tau_0 + 1)$ thus we easily see that $|\mathtt{r}(m)| < \mathtt{c} N^{\tau_1}$. □

From this proposition to a point $m \xrightarrow{N} [v_i, p_i]$ with $|\mathtt{r}(m)| > \mathtt{C} N^{\tau_1}$ and $p_1 < \mathtt{C} N^{4d\tau_0}$ we have associated an affine space $A \in \mathcal{H}_N^{\ell, g}$ such that $m \in A^g$ and we set $A_\ell(N)$ to be the set of points of previous type for which A has codimension ℓ.

The remaining points may be distributed in two sets:

$$A_d = A_d(N) \subset \{ m \in \mathbb{Z}^d \ : |\mathtt{r}(m)| \le \mathtt{C} N^{\tau_1} \}$$

and

$$A_0 := \{ m \in \mathbb{Z}^d \ : \ m \xrightarrow{N} [v_i; p_i] \quad \text{with} \quad |\mathtt{r}(m)| > \mathtt{C} N^{\tau_1}, \quad p_1 \ge \mathtt{C} K^{4d\tau_0} \}. \tag{53}$$

In this way we construct a decomposition of $\mathbb{Z}^d = \cup_{\ell=0}^{d} A_\ell(N)$.

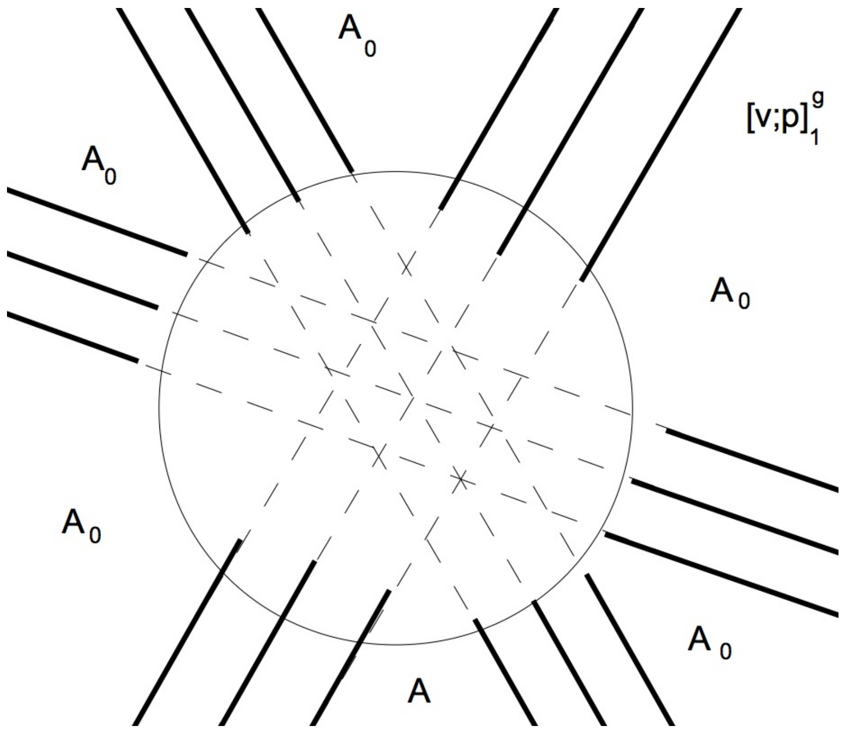


FIGURE 2. A drawing of the standard decomposition in $\mathbb{Z}_1^2$. A_0 is $\mathbb{Z}_1^2$ minus the dashed lines (each dashed line is described by an equation $[v;p]_1$). On each dashed line the set $[v;p]_1^g$ is signed in solid boldface. Note that $[v;p]_1^g$ is $[v;p]_1 \cap \mathbb{Z}_1^2$ minus a finite number of subspaces of codimension two, i.e. points.

7.13. **Cuts.** In the previous paragraph a point m which is a good point has an associated affine space A of some codimension ℓ which, as we have seen in the proof of Proposition 7.12 corresponds to a jump, or as we shall say a *cut*, in the optimal presentation. In fact for technical reasons having to do with the structure of Poisson bracket we need to refine this notion, introducing auxiliary parameters μ, θ which give a better control on the cut and allow some flexibility in the constructions. This is the topic of this paragraph.

We assume that N has been fixed. Given a point m we write $m \xrightarrow{N} [v_i; p_i]$ for its optimal presentation dropping the index ℓ which for a point equals d, we implicitly also mean that $m \in \mathbb{Z}^d$. Set by convention $p_0 = 0$ and $p_{d+1} = \infty$.

We then make a definition involving three more parameters:

Definition 7.14. The parameters N, θ, μ, τ are called *allowable* if

$$\tau_0 \leq \tau \leq \tau_1/(4d), \quad \mathtt{c} < \theta, \mu < \mathtt{C}, \quad N > N_0.$$

We need to analyze certain *cuts*, for the values p_i associated to an optimal presentation of a point. This will be an index ℓ where the values of the p_i jump according to the following:

Definition 7.15. The point $m\xrightarrow{N}[v_i;p_i]$ has a *cut* at $\ell \in \{0,1,\dots,d\}$ with the parameters $\underline{p} = (N,\theta,\mu,\tau)$, if ℓ is such that $p_\ell < \mu N^\tau$, $p_{\ell+1} > \theta N^{4d\tau}$.

The space $A := \{x \,|\, (v_i,x) = p_i,\ i = 1,\dots,\ell\}$ has $[v_i;p_i]_\ell$ as optimal presentation and it is called the *affine space associated to the cut* of m.

Remark 7.16. Note that, by Lemma 7.7, if $m\xrightarrow{N}[v_i;p_i]$ has a (N,θ,μ,τ) cut at ℓ then $|(m,v)| > \theta N^{4d\tau}$ for all $v \in B_N \setminus \langle v_i\rangle_l$.

Consider a subspace $A \in \mathcal{H}_N$ of codimension ℓ such that in its optimal presentation $p_\ell < \mu N^\tau$. The set of points $m \in A$ which have ℓ as a cut with the parameters N,θ,μ,τ have A as associated affine space, if furthermore $|\mathtt{r}(m)| > \theta N^{\tau_1}$ we call them (N,θ,μ,τ)–*good points of* A, we write $m \in A^g_{(N,\theta,\mu,\tau)}$.

By definition the other affine spaces have no good points with respect to these parameters.

Definition 7.17. Given allowable parameters $\underline{p} = (N,\theta,\mu,\tau)$ we denote by $\mathcal{H}^\ell_{\underline{p}}$ the set of affine subspaces $A \in \mathcal{H}_N$ of codimension ℓ such that in their optimal presentation $p_\ell < \mu N^\tau$. The union of all these sets for $0 < \ell < d$ is denoted by $\mathcal{H}_{\underline{p}}$.

Notice that $\theta N^{4d\tau} > \mu N^\tau$ (since by (50) we have $\mathtt{c}N^{4d\tau} > \mathtt{C}N^\tau$), so for any given $m \in S^c$ there is at most **one** choice of ℓ such that m has a ℓ cut with parameters θ,μ,τ.

Remark 7.18. 1) The purpose of defining a cut ℓ is to separate the numbers p_i into *small* and *large*. The parameters N,θ,μ,τ give a quantitative meaning to this statement.

2) The set of good points $A^g_{(N,\theta,\mu,\tau)}$, if non–empty, is the complement in A of a finite number of codimension one subspaces plus finitely many points.

3) Given any rational affine subspace A (i.e. defined by equations over $\mathbb{Z}$) there is an $\bar N$ so that $\forall N \ge \bar N$ we have $A \in \mathcal{H}_N$, its optimal presentation is independent of N, the set $A^g_{(N,\theta,\mu,\tau)}$ is non-empty.

We need an auxiliary parameter depending on an optimal presentation

Definition 7.19. Given $p \le \mathtt{c}N^{\tau_1/(4d)}$, set $\tau(p)$ so that $N^{\tau(p)} = \max(N^{\tau_0}, \mathtt{c}^{-1}p)$.

Notice that $\tau_0 \le \tau(p) \le \tau_1/(4d)$. The connection between the notion of good points A^g, defined in (52), of a given subspace A and the notion just introduced is explained by the following Lemma.

Lemma 7.20. *For all* $\mathtt{c} < \theta,\mu < \mathtt{C}$ *and for all affine subspaces* $[v_i;p_i]_\ell \in \mathcal{H}_N$ *such* $p_\ell \le \mathtt{c}N^{\tau_1/(4d)}$, *we have that every point* $m \in [v_i;p_i]^g_\ell$ *is a* $(N,\theta,\mu,\tau(p_\ell))$*–good point for* $[v_i;p_i]_\ell$. *I.e.* $[v_i;p_i]^g_\ell \subset ([v_i;p_i]_\ell)^g_{(N,\theta,\mu,\tau(p_\ell))}$ *for all* $\mathtt{c} < \theta,\mu < \mathtt{C}$.

Remark 7.21. 1) If $m\xrightarrow{N}[v_i;p_i]$ has a cut at ℓ for the parameters θ',μ',τ then it has also a cut at ℓ for parameters θ,μ,τ with $\theta \le \theta',\ \mu' \le \mu$ provided θ,μ,τ are allowable.

If $\theta \le \theta',\ \mu' \le \mu$ we shall say that the allowable parameters θ,μ are *less restrictive* than θ',μ'.

2) If for a given ℓ,τ we have $p_\ell \le \mathtt{c}N^\tau$, $p_{\ell+1} \ge \mathtt{C}N^{4d\tau}$, then ℓ is a cut with parameters θ,μ,τ for every choice of allowable θ,μ.

Lemma 7.22. *[Neighborhood property] Consider* $m,r \in \mathbb{Z}^d$ *with* $m\xrightarrow{N}[v_i;p_i]$, $r\xrightarrow{N}[w_i;q_i]$. *Suppose that* m *has a cut at* ℓ *for the parameters* N,θ',μ',τ, *and suppose there exist allowable parameters* $\theta < \theta',\mu' < \mu$:

$$|r-m| < \min(\kappa^{-1}(\mu-\mu')N^{\tau-1},\ \kappa^{-1}(\theta'-\theta)N^{4d\tau-1}). \tag{54}$$

then:

(1) The point r has a cut at ℓ for all allowable parameters θ, μ, τ for which (54) *holds.*

(2) $\langle w_1, \dots, w_\ell \rangle = \langle v_1, \dots, v_\ell \rangle$.

(3) $[w_i; q_i]_\ell$ *is the N–optimal presentation of* $[v_i; p_i]_\ell + r - m$.

Corollary 7.23. *Consider* $m \xrightarrow{N} [v_i; p_i]$, $r \xrightarrow{N} [w_i; q_i]$ *such that*

$$|r - m| < \kappa^{-1} c (N^{4d\tau - 1} - \mathtt{c}\mathtt{C}^{-1} N^{\tau - 1}) \tag{55}$$

and both m, r have a cut with parameters $\underline{p} = (N, \theta, \mu, \tau)$ then we can deduce:

i) the vectors m, r have the cut at the same ℓ;

ii) the space B associated to the cut of r is the one parallel to $A = [v_i; p_i]_\ell$ and passing through r namely

$$B = A + r - m. \tag{56}$$

The previous results explain why we wanted to introduce the parameters μ, θ to define cuts, in fact

Remark 7.24. With the above lemma we are stating that if m has a ℓ cut with parameters θ', μ', τ then, for all choices of $\theta < \theta', \mu' < \mu$, for which θ, μ are allowable parameters, we have described a spherical neighborhood B of m such that all points $r \in B$ have a ℓ cut with parameters θ, μ, τ. The radius of B is determined by Formula (54). Note moreover that if r has a cut at ℓ for some parameters then so has $-r$ and with the same parameters. Then lemma 7.22 holds verbatim if in formula (54) we substitute $|m - r|$ with $|m + r|$.

We finally combine 7.20 and 7.23

Lemma 7.25. *For all affine subspaces $[v_i; p_i]_\ell$ with $p_\ell \le \mathtt{c} N^{\tau_1/(4d)}$ the following holds. For all $m \in \mathbb{Z}^d$ with $m \in [v_i; p_i]_\ell^g$, for all $r \in \mathbb{Z}^d$ and for all parameters $\mathtt{c} < \theta, \mu < \mathtt{C}$ such that*

$$|r - m| < \kappa^{-1}(\mu - \mathtt{c}) N^{\tau_0 - 1}, \kappa^{-1}(\mathtt{C} - \theta) N^{4d\tau_0 - 1}, \tag{57}$$

r, m have the same cut ℓ with parameters $\theta, \mu, \tau(p_\ell)$ with parallel corresponding affine spaces.

The definitions which we have given are sufficient to define and analyze the quasi–Töplitz functions, which are introduced in section 8. In the next subsection we collect some definitions which are useful for the measure estimates and which are independent of the auxiliary parameters θ, μ.

7.26. **Graphs and cuts.** Recall that the choice of the vectors $S := \{\mathtt{j}_i\}$ determines a colored marked graph Γ_S with vertices in the set S^c.

This graph has finitely many components containing red edges.[7] The remaining set will be denoted by $\bar{S}^c$ and it is a union of connected components each combinatorially isomorphic to a combinatorial graph out of a finite list $\mathcal{G} := \{\mathcal{A}_1, \dots, \mathcal{A}_N\}$ formed only of black edges. It will be enough to concentrate our analysis only on these *black graphs.*

Definition 7.27. Given $\mathcal{A} \in \mathcal{G}$ we set $\Sigma_{\mathcal{A}}$ to be the union of all connected components of Γ_S isomorphic to $\mathcal{A}$.

[7] A rough estimate of a bound on the norm of these points is $(2d + 3)\kappa$.

The set $\mathcal{G}$ is partially ordered by setting $\mathcal{A}_i \leq \mathcal{A}_j$ if $\mathcal{A}_i$ is isomorphic as marked graph to a subgraph of $\mathcal{A}_j$.

From the theory developed it follows that, if $\mathcal{A} \in \mathcal{G}$ has $d_{\mathcal{A}} + 1$ vertices we have that $\Sigma_{\mathcal{A}}$ is a union of translates of any of its components (of the graph Γ_S). Moreover $\Sigma_{\mathcal{A}}$ is the portion of $\bar{S}^c$ in a union of $d_{\mathcal{A}} + 1$ parallel affine subspaces of codimension $d_{\mathcal{A}}$ minus a union of finitely many affine subspaces of higher codimension whose points lie in $\bigcup_{\mathcal{A}_i > \mathcal{A}} \Sigma_{\mathcal{A}_i}$.

Let us recall how we arrived at this statement. We choose a *root* in each $\mathcal{A}_i$. Using the root a geometric realization of some $\mathcal{A} = \mathcal{A}_i$ is an isomorphic graph, with vertices in $\bar{S}^c$, in which the image $\mathtt{r}$ of the root solves a certain set of $d_{\mathcal{A}}$ independent linear equations (Formula (61) of [18]) and the other vertices are determined by the labels on the graph $\mathcal{A}$. The components A of Γ_S which are isomorphic to $\mathcal{A}$ are exactly those geometric realizations of $\mathcal{A}$ which are not properly contained in a larger component. Recall that we have imposed generic conditions so that the vertices in a component are affinely independent. Therefore, a component A associated to $\mathcal{A}$, spans an affine subspace $\langle A \rangle$ of dimension exactly $d_{\mathcal{A}}$. All other geometric realizations of $\mathcal{A}$ are obtained from a given A by translating the graph A with the integral vectors orthogonal to $\langle A \rangle$. This set is thus a union of $d_{\mathcal{A}}+1$ parallel affine subspaces of codimension $d_{\mathcal{A}}$ passing each through an element m of A, image of a point $a \in \mathcal{A}$. It may well happen that a translate of $\mathtt{r}$ may solve also the linear equations defining a larger graph, the points in the corresponding component lie thus in a stratum $\Sigma_{\mathcal{A}_i}$, $\mathcal{A}_i > \mathcal{A}$.

Example 7.28.

$$\begin{tikzcd} & x - \mathtt{j}_1 + \mathtt{j}_3 \arrow[dl, "3{,}1"'] \\ x \arrow[r, "3{,}2"'] & x - \mathtt{j}_2 + \mathtt{j}_3 \arrow[u, "2{,}1"'] \end{tikzcd}$$

equations

$$(x, \mathtt{j}_1 - \mathtt{j}_3) = |\mathtt{j}_1|^2 - (\mathtt{j}_2, \mathtt{j}_3)$$

$$(x, \mathtt{j}_2 - \mathtt{j}_3) = |\mathtt{j}_2|^2 - (\mathtt{j}_2, \mathtt{j}_3)$$

Notice that this is a subgraph of the graph of example 4.2.

If $a \in \mathcal{A}$ we let $\Sigma_{\mathcal{A},a}$ be the subset of $\Sigma_{\mathcal{A}}$ formed by the corresponding elements so that $\Sigma_{\mathcal{A}}$ is the disjoint union of the strata $\Sigma_{\mathcal{A},a}$ as a runs over the vertices of $\mathcal{A}$.

The set $\Sigma_{\mathcal{A},a}$ spans an affine space $\langle \Sigma_{\mathcal{A},a} \rangle$ of codimension $d_{\mathcal{A}}$ and, in fact, it is the complement in this affine space of the points which belong to graphs which contain strictly $\mathcal{A}$. Thus $\Sigma_{\mathcal{A},a}$ is obtained from the affine space $\langle \Sigma_{\mathcal{A},a} \rangle$ removing a finite union of proper affine subspaces.

Thus for any $\mathcal{A} \in \mathcal{G}$ having chosen a root $\mathtt{r}$ we have the stratum $\Sigma_{\mathcal{A},\mathtt{r}}$ which is the complement in an affine space of a finite union of codimension 1 subspaces. Any point $m \in \bar{S}^c$ lies in a unique stratum $\Sigma_{\mathcal{A},a}$ which is parallel to $\Sigma_{\mathcal{A},\mathtt{r}}$ we thus have for m a corresponding root $\mathtt{r}(m) \in \Sigma_{\mathcal{A},\mathtt{r}}$ which is the intersection of $\Sigma_{\mathcal{A},\mathtt{r}}$ with the connected component of the graph $\mathcal{A}_S$ in which m lies. Thus $m - \mathtt{r}(m)$ depends only upon the stratum $\Sigma_{\mathcal{A},a}$.

Definition 7.29. We denote the vector $m - \mathtt{r}(m)$ as the *type of* m.

Remark 7.30. The possible types run on a finite set $\mathcal{Z}$ of vectors. The type of a vector m, as seen in Formula (36) is a linear combination of the elements $\mathtt{j}_i$ with coefficients the coordinates of $L(m)$. Hence each $u \in \mathcal{Z}$ has $|u| \leq d\kappa$. Note that when $m \in \Sigma_{\mathcal{A},a}$ the element $L(m)$ is fixed (by a), the corresponding type will be denoted by u_a. Notice that $\Sigma_{\mathcal{A},a} = \Sigma_{\mathcal{A},\mathtt{r}} + u_a$.

Remark 7.31. Among the combinatorial graphs we have the graph $\{0\}$ formed by a single vertex. The corresponding *open* stratum $\Sigma_{\{0\},0}$ obviously spans $\mathbb{Z}^d$, it is formed of all the points in S^c which do not belong to any of the proper strata $\Sigma_{\mathcal{A}}$.

We have thus finitely many affine subspaces $\langle \Sigma_{\mathcal{A},a}\rangle$ associated to the pairs $(\mathcal{A}, a)$, these subspaces can be presented using the linear equations associated to the geometric realization by formulas (61) of [18]. We have a finite number of possible systems of equations with coefficients depending linearly or quadratically from the set S. We verify that all the constant coefficients of these equations are $< \mathtt{c}N_0$ (in fact the coefficients can be bound by $(2d\kappa)^2$). Thus by the bounds chosen each of these subspace lies in $\mathcal{H}_N$, $\forall N > N_0$ and thus has an N–optimal presentation $[w_i; q_i]_{d_{\mathcal{A}}}$, furthermore each $q_i < \mathtt{c}N_0$.

Lemma 7.32. *Assume that $m \in \Sigma_{\mathcal{A},a}$ has a cut at ℓ with parameters θ, μ, τ and associated space $[v_i, p_i]_\ell$. Then $[v_i, p_i]_\ell$ is contained in the affine space $\langle \Sigma_{\mathcal{A},a}\rangle = [w_i; q_i]_{d_{\mathcal{A}}}$. In particular $\ell \geq d_{\mathcal{A}}$.*

Proof. Since m is in both spaces it is enough to prove that $\langle w_i\rangle_{d_{\mathcal{A}}} \subset \langle v_i\rangle_\ell$. By contradiction if some $w_j \notin \langle v_i\rangle_\ell$ then we have that $|(w_j, m)| > cN^{4d\tau_0}$. This is incompatible with the estimates on the q_i. □

Theorem 5. *If m, n have the same cut ℓ and the same associated affine space $[v_i, p_i]_\ell$ then they belong to the same stratum $\Sigma_{\mathcal{A},a}$ and hence have the same type, i.e. $m - \mathtt{r}(m) = n - \mathtt{r}(n)$.*

Proof. If both m, n form an isolated component then they belong to the open stratum. Assume that $m \in \Sigma_{\mathcal{A},a}$ with $d_{\mathcal{A}} > 0$. Thus m satisfies the equations $[w_i; q_i]_{d_{\mathcal{A}}}$. By the previous Lemma since $n \in [v_i, p_i]_\ell$ we know that $n \in [w_i; q_i]_{d_{\mathcal{A}}}$. This implies $n \in \Sigma_{\mathcal{A}'}$ where $\mathcal{A}'$ contains $\mathcal{A}$ so that $d_{\mathcal{A}'} \geq d_{\mathcal{A}}$.

Exchanging the roles of m, n we see that $m \in \Sigma_{\mathcal{A}''}$ where $\mathcal{A}''$ contains $\mathcal{A}'$. This implies that $\mathcal{A} = \mathcal{A}'$. Since the equations $[w_i; q_i]_{d_{\mathcal{A}}}$ define the affine space spanned by $\Sigma_{\mathcal{A},a}$ the claim follows. □

8. Functions

8.1. **Töplitz approximation.**

8.1.1. *Piecewise Töplitz functionsn.* Given a parameter $N > N_0$ we will call *low momentum variables* relative to N, denoted by w^L, the z_j^σ such that $|\mathtt{r}(j)| < \mathtt{C}N^3$. Similarly we call *high momentum variables*, denoted by w^H, the z_j^σ such that $|\mathtt{r}(j)| > \mathtt{c}N^{\tau_1}$. Notice that by (50) the low and high variables are separated. Furthermore all variables z_j belonging to blocks with red edges are low by choice of the parameters. The remaining variables will be denoted by w^R. Given any set X of conjugate variables by $\{\cdot, \cdot\}^X$ we mean the Poisson bracket performed only respect to the variables X (keeping the other variables as parameters).

Our definitions will depend on several parameters which in turn depend upon the particular problem treated. We shall denote them by a compact symbol $\underline{p}$.

Recall first that a monomial $e^{\mathrm{i}(k,x)} y^i z^\alpha \bar{z}^\beta$ has momentum

$$\mathrm{i}(\sum_{i=1}^n k_i \mathtt{j}_i + \sum_{j \in S^c} \sigma(j)\mathtt{r}(j)(\alpha_j - \beta_i)).$$

Thus we make a:

Definition 8.2. (Low-momentum) A monomial $e^{\mathrm{i}(k,x)}y^i z^\alpha \bar z^\beta$ is (N,μ)-low momentum if[8]

$$\sum_{j\in S^c} |\mathtt{r}(j)|(\alpha_j+\beta_j) < \mu N^3\,, \quad |k|<N\,. \tag{58}$$

We denote by

$$\mathcal{L}_{s,r}(N,\mu)\subset \mathcal{H}_{s,r}$$

the subspace of functions

$$g=\sum g_{k,i,\alpha,\beta} e^{\mathrm{i}(k,x)}y^i z^\alpha \bar z^\beta \in \mathcal{H}_{s,r} \tag{59}$$

whose monomials are (N,μ)-low momentum. The corresponding projection

$$\Pi^L_{N,\mu} : \mathcal{H}_{s,r}\to \mathcal{L}_{s,r}(N,\mu) \tag{60}$$

is defined as $\Pi^L_{N,\mu} := \Pi_I$ where I is the subset of indexes (k,α,β) satisfying (58) (notice that the exponent i of y plays no role). Finally, given $h\in\mathbb{Z}^d$, we denote by

$$\mathcal{L}_{s,r}(N,\mu,h)\subset \mathcal{L}_{s,r}(N,\mu)$$

the subspace of functions of momentum $-\mathrm{i}\,h$, i.e. whose monomials satisfy (cf. (44)):

$$\pi_{\mathtt{r}}(k,\alpha,\beta)+h=0\,. \tag{61}$$

By (58), any function in $\mathcal{L}_{s,r}(N,\mu)$, $\mathtt{c}<\mu<\mathtt{C}$, only depends on x,y,w^L and therefore

$$g,g'\in\mathcal{L}_{s,r}(N,\mu) \implies gg',\ \{g,g'\},\ \{g,g'\}^{x,y},\ \{g,g'\}^L \ \text{ do not depend on } w^H\,. \tag{62}$$

Moreover if

$$|h|\ge \mu N^3+\kappa N \implies \mathcal{L}_{s,r}(N,\mu,h)=0\,. \tag{63}$$

Definition 8.3. Given N and allowable parameters $\underline{p}=(N,\theta,\mu,\tau)$ we say that a monomial

$$\mathfrak{m}=\mathfrak{m}_{k,l,\alpha,\beta,m,n,\sigma,\sigma'} := e^{\mathrm{i}(k,x)}y^l z^\alpha \bar z^\beta z_m^\sigma z_n^{\sigma'} \tag{64}$$

is $\underline{p}=(N,\theta,\mu,\tau)$–bilinear if:

1) it satisfies momentum conservation, $\pi_{\mathtt{r}}(\mathfrak{m})=0$, (44) and:

$$|k|<N\,,\qquad |\mathtt{r}(n)|,|\mathtt{r}(m)|>\theta N^{\tau_1}\,,\qquad \sum_j |\mathtt{r}(j)|(\alpha_j+\beta_j)<\mu N^3\,. \tag{65}$$

2) There is an $0<\ell<d$ so that both m,n have an ℓ cut with parameters N,θ,μ,τ.

To the high variables m,n of the monomial $\mathfrak{m}$ we associate the two affine subspaces A,B associated to their ℓ–cut. By reordering the variables if necessary, we may assume that $A\prec B$ and associate to the monomial, or equivalently to the pair m,n, only A.

A monomial which is $\underline{p}$ bilinear with associated affine space A is also called $A,\underline{p}$ *restricted*.

We shall often write $m\in\underline{p}-$ cut to mean that there is an $0<\ell<d$ so that m has an ℓ cut with parameters $\underline{p}=(N,\theta,\mu,\tau)$. If we want to stress the affine space A associated to the cut we write $m\in(A,\underline{p})-$ cut.

---

[8]For $k\in\mathbb{Z}^n$, by $|k|$ we always mean the L^1 norm $\sum_i |k_i|$. In $\mathbb{Z}^d$ instead we use the L^2 norm.

Remark 8.4. Note that by momentum conservation and Remark 7.30

$$0=|\pi_{\mathtt{r}}(\mathfrak{m})|=|\sigma\mathtt{r}(m)+\sigma'\mathtt{r}(n)+\sum_j\sigma(j)\mathtt{r}(j)(\alpha_j-\beta_j)+\pi(k)|\geq|\sigma m+\sigma' n|-(2d\kappa+\mu N^3+\kappa N)$$

we deduce

$$|\sigma m+\sigma' n|\leq\mu N^3+3d\kappa N. \tag{66}$$

Thus the monomial $\mathfrak{m}=gz_m^{\sigma}z_n^{\sigma'}$ of Formula (64) has $g\in\mathcal{L}_{s,r}(N,\mu,\sigma\mathtt{r}(m)+\sigma'\mathtt{r}(n))$.

Remark 8.5. Note that under condition 1), in condition 2) it is sufficient to assume that m,n have a cut with the same parameters N,θ,μ,τ. The fact that the cut is at the same ℓ follows from Corollary 7.23 since Formula (66) implies Formula (55).

By Theorem 5, for all $A,\underline{p}$ restricted monomials $\mathfrak{m}$ with given $\sigma m+\sigma' n,\sigma,\sigma'$ we may deduce $\mathtt{r}(m)-m$ from A. By Corollary 7.23 and (66), we deduce

$$-\sigma\sigma' A+\sigma'(\sigma m+\sigma' n)=B. \tag{67}$$

Note that, by hypothesis, $m\in A^g_{\underline{p}},\ n\in B^g_{\underline{p}}$. B in turn fixes $\mathtt{r}(n)-n$ and hence the *type* of the monomial $u(\mathfrak{m})$, defined as

$$u(\mathfrak{m}):=u(A,\sigma m+\sigma' n,\sigma,\sigma')=\sigma(\mathtt{r}(m)-m)+\sigma'(\mathtt{r}(n)-n), \tag{68}$$

depends only on the elements $A,\sigma m+\sigma' n,\sigma,\sigma'$.

Definition 8.6. Set $\underline{p}=(s,r,N,\theta,\mu,\tau)$, in $\mathcal{H}_{s,r}$ we consider the space $\mathcal{B}_{\underline{p}}$ of (N,θ,μ,τ)–bilinear functions, that is whose monomials are all (N,θ,μ,τ)–bilinear. We call

$$\Pi_{\underline{p}}:=\Pi_{(N,\theta,\mu,\tau)}:\mathcal{H}_{s,r}\to\mathcal{B}_{\underline{p}}$$

the projection onto this subspace. A function $f\in\mathcal{B}_{\underline{p}}$ is of the form:

$$f(x,y,z,\bar z)=\sum_{\sigma,\sigma'=\pm}\ \sum_{\substack{|\mathtt{r}(m)|,|\mathtt{r}(n)|>\theta N^{\tau_1},\\ m,n\in\underline{p}-\mathrm{cut}}} f^{\sigma,\sigma'}_{m,n}(x,y,w^L)z_m^{\sigma}z_n^{\sigma'} \tag{69}$$

with $f^{\sigma,\sigma'}_{m,n}(x,y,w^L)\in\mathcal{L}_{s,r}(N,\mu,\sigma\mathtt{r}(m)+\sigma'\mathtt{r}(n))$.

By convention we assume $f^{\sigma,\sigma'}_{m,n}(x,y,w^L)=f^{\sigma',\sigma}_{n,m}(x,y,w^L)$.

Note that, by Definition 8.3, to each element $f^{\sigma,\sigma'}_{m,n}(x,y,w)z_m^{\sigma}z_n^{\sigma'}$ is associated an affine subspace A.

Remark 8.7. Of course, if we take less restrictive parameters θ',μ' with $\theta\leq\theta',\ \mu'\leq\mu$ we have that the set of (N,θ,μ,τ)–bilinear monomials contains the set of (N,θ',μ',τ)–bilinear monomials. In particular we have, for each s,r:

$$\Pi_{(N,\theta',\mu',\tau)}\Pi_{(N,\theta,\mu,\tau)}=\Pi_{(N,\theta,\mu,\tau)}\Pi_{(N,\theta',\mu',\tau)}=\Pi_{(N,\theta',\mu',\tau)}\,.$$

Definition 8.8. Given parameters $\underline{p}$ and an affine space $A\in\mathcal{H}_{\underline{p}}$ (cf. Definition 7.17), the space $\mathcal{T}^{\sigma,\sigma'}_{A,\underline{p}}$ of $A,\underline{p}$-restricted *Töplitz* bilinear functions of signature σ,σ' is formed by the functions $g\in\mathcal{B}_{\underline{p}}$ where the coefficients of Formula 69 satisfy:

$$g=\sum_{m,n}^{(A,\underline{p})}\mathtt{g}(\sigma m+\sigma' n)z_m^{\sigma}z_n^{\sigma'}\,. \tag{70}$$

The apex $(A, \underline{p})$, with $\underline{p} = (N, \theta, \mu, \tau)$, means that the sum is on the $A, \underline{p}$–restricted monomials (cf. 8.3), with bilinear part $z_m^\sigma z_n^{\sigma'}$. Notice that, by definition of $A, \underline{p}$–restricted monomials, the space $\mathcal{T}_{A,\underline{p}}^{\sigma,\sigma'}$ is non–zero only if $A \xrightarrow{N} [v_i; p_i]_\ell$ is of codimension ℓ with $0 < \ell < d$ and $p_\ell < \mu N^\tau$.

For all $h = \sigma m + \sigma' n$ we have:

$$\mathtt{g}(h) \in \mathcal{L}_{s,r}(N, \mu, h + u(A, h, \sigma, \sigma')) \tag{71}$$

where $u(A, h, \sigma, \sigma')$ is the type, see Formula (68).

Remark 8.9. i) Notice that we have a translation invariance property (which justifies the name restricted Töplitz). Indeed given A, σ, σ', h one can choose arbitrarily an element $\mathtt{g}^{\sigma,\sigma'}(A, h)$ satisfying (71), and use formula (70) to define a function in $\mathcal{B}_{\underline{p}}$. One easily sees that indeed such an expression defines a function in $\mathcal{H}_{s,r}$.
ii) Note that condition (71) implies that $\mathtt{g}(sm + \sigma' n)$ in (70) has momentum $\mathrm{i}(\sigma \mathtt{r}(m) + \sigma' \mathtt{r}(n))$ hence g has zero momentum (as required).

Finally we define

Definition 8.10. The space $\mathcal{T}_{\underline{p}}$ of *piecewise Töplitz* bilinear functions

$$g = \sum_{A \in \mathcal{H}_N, \sigma, \sigma' = \pm 1} g^{\sigma,\sigma'}(A), \qquad g^{\sigma,\sigma'}(A) \in \mathcal{T}_{A,\underline{p}}^{\sigma,\sigma'}. \tag{72}$$

Of particular significance are the piecewise Töplitz diagonal functions

$$\mathcal{Q}(z) = \sum_{A \in \mathcal{H}_N} \sum_m^{(A,\underline{p})} \mathcal{Q}(A) z_m \bar{z}_m = \sum_{A \in \mathcal{H}_N} \sum_{m \in (A,\underline{p})-\mathrm{cut}} \mathcal{Q}(A) z_m \bar{z}_m \,, \tag{73}$$

in this formula the elements m run over all vectors which have an ℓ-cut with $0 < \ell < d$ and A is the corresponding affine spaces.

By definition $\mathcal{T}_{\underline{p}} \subset \mathcal{B}_{\underline{p}}$ is a subspace of the (N, θ, μ, τ) bilinear functions.

Lemma 8.11. *Consider $f, g \in \mathcal{T}_{\underline{p}}$ and $q \in \mathcal{L}_{s,r}(N, \mu_1, 0)$, $\mathtt{c} < \mu, \mu_1 < \mathtt{C}$. Given any $\underline{p}' = (s', r', N, \theta', \mu', \tau)$ with $s/2 < s' < s$, $r/2 < r' < r$, $\theta' \geq \theta, \mu' \leq \mu$, one has*

$$\Pi_{N,\theta',\mu',\tau}\{f, q\}^L \,,\ \Pi_{N,\theta',\mu',\tau}\{f, q\}^{x,y} \in \mathcal{T}_{\underline{p}'} \,. \tag{74}$$

If moreover

$$\kappa(\mu N^3 + 3d\kappa N) < (\theta' - \theta) N^{4d\tau - 1}, (\mu - \mu') N^{\tau - 1} \tag{75}$$

then

$$\Pi_{N,\theta',\mu',\tau}\{f, g\} = \Pi_{N,\theta',\mu',\tau}\{f, g\}^H \in \mathcal{T}_{\underline{p}'} \,. \tag{76}$$

Proof. Write $f \in \mathcal{T}_{\underline{p}} \subset \mathcal{B}_{\underline{p}}$ as in (69) where

$$f_{m,n}^{\sigma,\sigma'} = f^{\sigma,\sigma'}(A, \sigma m + \sigma' n) \in \mathcal{L}_{s,r}(N, \mu, \sigma m + \sigma' n + u) = \mathcal{L}_{s,r}(N, \mu, \sigma \mathtt{r}(m) + \sigma' \mathtt{r}(n)) \,, \tag{77}$$

similarly for g (recall that $u = u(A, \sigma m + \sigma' n, \sigma, \sigma')$), with $f_{m,n}^{\sigma,\sigma'} = f_{n,m}^{\sigma',\sigma}$.

Proof of (74). Since the variables z_m^σ, $z_n^{\sigma'}$, $|\mathtt{r}(m)|, |\mathtt{r}(n)| > \theta N^{\tau_1}$, are high momentum,

$$\{f^{\sigma,\sigma'}(A, \sigma m + \sigma' n) z_m^\sigma z_n^{\sigma'} \,,\, q\}^L = \{f^{\sigma,\sigma'}(A, \sigma m + \sigma' n),\, q\}^L \, z_m^\sigma z_n^{\sigma'} \,.$$

The function $\{f^{\sigma,\sigma'}(A, \sigma m + \sigma' n)\,,\ q\}^L$ in in $\mathcal{H}_{s',r'}$ by (22) and does not depend on w^H by (62). Hence the coefficient of $z_m^\sigma z_n^{\sigma'}$ in $\Pi_{N,\theta',\mu'}\{f,q\}^L$ is,

$$\Pi^L_{N,\mu'}\{f^{\sigma,\sigma'}(A, \sigma m + \sigma' n)\,,\, q\}^L \in \mathcal{L}_{s',r'}(N, \mu', \sigma \mathtt{r}(m) + \sigma' \mathtt{r}(n))$$

since $\pi_{\mathtt{r}}(q) = 0$, $\pi_{\mathtt{r}}(f^{\sigma,\sigma'}(A, \sigma m + \sigma' n)) = -\sigma \mathtt{r}(m) - \sigma' \mathtt{r}(n)$.

The proof that $\Pi_{N,\theta',\mu'}\{f,q\}^{x,y} \in \mathcal{T}_{s',r',N,\theta',\mu'}$ is analogous.

Proof of (76). A direct computation gives

$$\{f,g\}^H = \sum_{\substack{|\mathtt{r}(m)|,|\mathtt{r}(n)|>\theta N_1^\tau,\\ m,n\in \underline{p}-\mathrm{cut}}} \sum^{\sigma,\sigma'=\pm} p^{\sigma,\sigma'}_{m,n} z^\sigma_m z^{\sigma'}_n$$

with

$$p^{\sigma,\sigma'}_{m,n} = -2\mathrm{i} \sum_{l\,,\ \sigma_1=\pm} \sigma_1 \left(f^{\sigma,\sigma_1}_{m,l}\, g^{-\sigma_1,\sigma'}_{l,n} + f^{\sigma',\sigma_1}_{n,l}\, g^{-\sigma_1,\sigma}_{l,m} \right), \tag{78}$$

of course l gives a contribution only if suitably restricted by the bilinearity constraint.

By (62) the coefficient $p^{\sigma,\sigma'}_{m,n}$ does not depend on w_H. Therefore

$$\Pi_{N,\theta',\mu'}\{f,g\}^H = \sum_{|\mathtt{r}(m)|,|\mathtt{r}(n)|>\theta' N^{\tau_1},\, \sigma,\sigma'=\pm} q^{\sigma,\sigma'}_{m,n} z^\sigma_m z^{\sigma'}_n \quad \text{with} \quad q^{\sigma,\sigma'}_{m,n} := \Pi^L_{N,\mu'} p^{\sigma,\sigma'}_{m,n}\,. \tag{79}$$

It results $q^{\sigma,\sigma'}_{m,n} \in \mathcal{L}_{s',r'}(N, \mu', \sigma\mathtt{r}(m) + \sigma'\mathtt{r}(n))$ by (79), (78), and momentum conservation.

It remains to prove the (N, θ', μ')-Töplitz property:

$$q^{\sigma,\sigma'}_{m,n} = q^{\sigma,\sigma'}\big(A, \sigma m + \sigma' n\big) \quad \text{for some} \quad q^{\sigma,\sigma'}(A, h) \in \mathcal{L}_{s,r}(N, \mu', h + u))\,, \tag{80}$$

where A is the affine space associated to the pair m, n (cf. 8.3, 2)).

Let us consider in (78)-(79) the term with m, n fixed and $\sigma = +1, \sigma' = -1, \sigma_1 = +1$ (the other cases are analogous)

$$\Pi^L_{N,\mu'} \sum_l f^{+,+}_{m,l}\, g^{-,-}_{l,n}\ , \tag{81}$$

Since by hypothesis $f, g \in \mathcal{T}_{\underline{p}}$ we have that the l that give a contribution have a $\underline{p}$-cut at ℓ and $|\mathtt{r}(l)| > \theta N^{\tau_1}$. By Remark 8.5 the affine space associated to the cut of l is $A' = -A + l + m$; (while the affine space associated to the cut of n is $B = A + n - m$). We may assume without loss of generality that $A \prec B$.

We then divide (81) in 3 parts according to the relative position of A' in the $\prec$ order with respect to $A \prec B$. Note that all these constraints depend only upon $A, m - n, j := l + m$. We treat the case $A \prec A' \prec B$, the other 2 cases are similar.

Since $f, g \in \mathcal{T}_{\underline{p}}$ we have

$$f^{+,+}_{m,l} = f^{+,+}\big(A, m + l\big) \in \mathcal{L}_{s,r}(N, \mu, \mathtt{r}(m) + \mathtt{r}(l)) \tag{82}$$

$$g^{-,-}_{l,n} = g^{-,-}\big(A', -l - n\big) \in \mathcal{L}_{s,r}(N, \mu, -\mathtt{r}(l) - \mathtt{r}(n))\,, \tag{83}$$

for all m, n, l which satisfy the bilinearity constraints with parameters $\underline{p}$.

By construction m, n satisfy the bilinearity constraints with the more restrictive parameters $\underline{p}'$. Set $j := m + l$, by formula (66) we have that the elements j, which come from the elements l which contribute, have $|j| < \mu N^3 + 3d\kappa N$. On the other hand by condition (75) we see that these j satisfy:

$$|j| < \kappa^{-1}(\mu - \mu')N^{\tau - 1},\ \kappa^{-1}(\theta' - \theta)N^{4d\tau - 1}. \tag{84}$$

thus, by Lemma 7.22 the fact that j satisfies (84) implies that m, n, l satisfy the bilinearity constraints with parameters $\underline{p}$.

Then

$$\Pi^L_{N,\mu'} \sum_{A \prec A' \prec B} f^{+,+}_{m,l} g^{--}_{l,n} \overset{(67)}{=} \Pi^L_{N,\mu'} \sum_{\substack{j \in \mathbb{Z}^d : |j| < \mu N^3 + 3d\kappa N \\ A \prec A+j \prec A+n-m}} f^{++}(A,j) g^{--}\big(-A+j, m-n-j\big)$$

depends only on A and $m-n$, i.e. (80). □

8.12. **Quasi–Töplitz functions.** Given $f \in \mathcal{H}_{s,r}$ and $\mathcal{F} \in \mathcal{T}_{\underline{p}}$, we define

$$\bar{f} = \bar{f}(\mathcal{F}) := N^{4d\tau}\big(\Pi_{(N,\theta,\mu,\tau)} f - \mathcal{F}\big), \tag{85}$$

and for $K > N_0$ set

$$\|X_f\|^{K,\theta,\mu}_{s,r} := \sup_{\substack{N,\tau : N \geq K \\ \underline{p}=(s,r,N,\theta,\mu,\tau)}} \big[\inf_{\mathcal{F} \in \mathcal{T}_{\underline{p}}} (\max(\|X_{\mathcal{F}}\|_{s,r}, \|X_{\bar{f}}\|_{s,r}))\big] \tag{86}$$

Remark 8.13. If we take new parameters K', θ', μ' with $K \leq K', \theta \leq \theta', \ \mu' \leq \mu$ we have by Remark 8.7 that

$$\|X_f\|^{K',\theta',\mu'}_{s,r} \leq \|X_f\|^{K,\theta,\mu}_{s,r}.$$

Definition 8.14. We say that $f \in \mathcal{H}_{s,r}$ is *quasi- Töplitz* of parameters (K, θ, μ) if $\|X_f\|^{K,\theta,\mu}_{s,r} < \infty$ and we call this number the *quasi-Töplitz norm* of f.

Remark 8.15. Given $f \in \mathcal{H}_{s,r}$ with finite quasi-Töplitz norm and parameters $\underline{p}$ we say that a function $\mathcal{F} \in \mathcal{T}_{\underline{p}}$, approximates f at order ε if

$$\|X_{\mathcal{F}}\|_{s,r}, N^{4d\tau}\|X_{\Pi_{(N,\theta,\mu,\tau)}f - \mathcal{F}}\|_{s,r} < (1+\varepsilon)\|f\|^{K,\theta,\mu}_{s,r}. \tag{87}$$

Note that by our definitions such approximating functions exist for all allowable parameters $\underline{p}$ and for all positive ε.

Since in our algorithm we deal with functions which depend in a Lipschitz way on some parameters $\xi \in \mathcal{O}$ (a compact set) we take finally a norm which includes also the weighted Lipschitz norm (cf. (18)) with λ a positive number:

Definition 8.16. We set $\vec{p} = (s, r, K, \theta, \mu, \lambda, \mathcal{O})$. We define

$$\|X_f\|^T_{\vec{p}} := \max(\|X_f\|^{K,\theta,\mu}_{s,r}, \|X_f\|^{\lambda}_{s,r}) \tag{88}$$

We denote by $\mathcal{Q}^T_{\vec{p}} \subset \mathcal{H}_{s,r}$ the set of functions with finite norm $\|X_f\|^T_{\vec{p}}$.

Remark 8.17. Notice that our definition includes the Töplitz and anti-Töplitz functions, setting $\mathcal{F} = \Pi_{(N,\theta,\mu,\tau)} f$ and hence $\bar{f} = 0$. In the case of Töplitz functions one trivially has $\|X_f\|^T_{\vec{p}} = \|X_f\|^{\lambda}_{s,r}$.

The definition includes also functions with fast decay in the coefficients so that, taking always as Töplitz approximation $\mathcal{F} = 0$, we still have $\sup_N N^{4d\tau}\|X_{\Pi_{(N,\theta,\mu,\tau)}f}\|^{\lambda}_{s,r} < \infty$.

8.17.1. *Some basic properties.* The following Lemmas are proved in [2].

Lemma 8.18. (Projections 1) *Set* $\vec{p} = (s, r, K, \theta, \mu, \lambda, \mathcal{O})$ *and* $\underline{p} = (s, r, N, \theta, \mu, \tau)$ *with* $N \geq K$. *Consider a subset of monomials* I *such that the projection (see* (26)*) maps*

$$\Pi_I : \mathcal{T}_{\underline{p}} \to \mathcal{T}_{\underline{p}}, \quad \forall N \geq K. \tag{89}$$

Then $\Pi_I : \mathcal{Q}^T_{\vec{p}} \to \mathcal{Q}^T_{\vec{p}}$ *and*

$$\|X_{\Pi_I F}\|^T_{\vec{p}} \leq \|X_F\|^T_{\vec{p}}, \quad \forall F \in \mathcal{Q}^T_{\vec{p}}. \tag{90}$$

Moreover, if $F \in \mathcal{Q}^T_{\vec p}$ satisfies $\Pi_I F = F$, then, $\forall N \geq K$, $\forall \varepsilon > 0$, there exists a decomposition $\Pi_{N,\theta,\mu,\tau} F = \tilde F + N^{-4d\tau}\hat F$ with a Töplitz approximation $\tilde F \in \mathcal{T}_{\underline{p}}$ satisfying $\Pi_I \tilde F = \tilde F$, $\Pi_I \hat F = \hat F$ and $\|X_{\tilde F}\|_{s,r}, \|X_{\hat F}\|_{s,r} < \|X_F\|^T_{\vec p} + \varepsilon$.

Lemma 8.19. (Projections 2) *For all $l \in \mathbb{N}$, $K \in \mathbb{N}$, $N \geq K$, the projections*

$$\Pi^{(l)}, \Pi_{|k|<K}, \Pi_{\rm diag} : \mathcal{T}_{\underline{p}} \to \mathcal{T}_{\underline{p}} \,. \tag{91}$$

here $\Pi^{(l)}$ maps to the space of homogeneous functions of degree l, $\Pi_{\rm diag} := \Pi^{(2)}\Pi_{k=0}$.

If $F \in \mathcal{Q}^T_{\vec p}$ then,

$$\|X_{\Pi^{(l)}F}\|^T_{\vec p}\,,\ \|X_{\Pi_{|k|<K}F}\|^T_{\vec p}\,,\ \|X_{\Pi_{\rm diag}F}\|^T_{\vec p} \leq \|X_F\|^T_{\vec p}\,, \tag{92}$$

$$\|X_{F^{(\leq 2)}}\|^T_{\vec p}\,,\ \|X_F - X_{F^{(\leq 2)}_{|k|<K}}\|^T_{\vec p} \leq \|X_F\|^T_{\vec p}\,. \tag{93}$$

Moreover, $\forall\, 0 < s' < s$, set $\vec p' := (s', r, K, \theta, \mu, \lambda, \mathcal{O})$:

$$\|X_{\Pi_{|k|\geq K}F}\|^T_{\vec p'} \leq e^{-K(s-s')}\frac{s}{s'}\|X_F\|^T_{\vec p} \tag{94}$$

Finally, $\forall\, 0 < r' < r$, set $\vec p' := (s, r', K, \theta, \mu, \lambda, \mathcal{O})$:

$$\|X_{\Pi^{(l\geq D)}F}\|^T_{\vec p'} \leq (\frac{r'}{r})^{D-2}\|X_F\|^T_{\vec p} \tag{95}$$

Lemma 8.20. *Let $Q(z) = \sum_m Q_m z_m \bar z_m$ be a quasi-Töplitz* diagonal quadratic function *in the variables $z, \bar z$ with constant coefficients. For all allowable choices of $\underline{p}$ and $\varepsilon > 0$, there exists a* diagonal quadratic function $\mathcal{Q}(z) \in \mathcal{T}_{\underline{p}}$ *which approximates Q at order ε:*

$$\mathcal{Q}(z) = \sum_{A \in \mathcal{H}_{\underline{p}}} \sum_m^{(A,\underline{p})} \mathcal{Q}(A) z_m \bar z_m\,, \tag{96}$$

so that setting $\bar Q(z)$:

$$N^{-4d\tau}\bar Q(z) = \Pi_{(N,\theta,\mu,\tau)}Q(z) - \mathcal{Q}(z)\,,$$

for all m which have a cut at ℓ with parameters (N,θ,μ,τ) associated to A one has

$$Q_m = \mathcal{Q}(A) + N^{-4d\tau}\bar Q_m, \tag{97}$$

and

$$|Q_m|, |\mathcal{Q}(A)|, |\bar Q_m| \leq (1+\varepsilon)|X_Q|^T_{\vec p}\,. \tag{98}$$

For all m, m' with a cut at ℓ with parameters (N,θ,μ,τ) associated to A we have

$$|Q_m - Q_{m'}| \leq N^{-4d\tau} 2\|X_Q\|^T_{\vec p}. \tag{99}$$

Proof. Since Q is quasi-Töplitz we may approximate it by a function $\mathcal{F} \in \mathcal{T}_{\underline{p}}$; moreover since Q is quadratic and diagonal by the previous discussion we may choose $\mathcal{F}$ of the same form.

Hence we can we can find a quadratic and diagonal function $\mathcal{Q} \in \mathcal{T}_{\underline{p}}$ so that, with $\bar Q = N^{4d\tau}(\Pi_{N,\theta,\mu,\tau}Q - \mathcal{Q})$, we have $\|X_{\mathcal{Q}}\|_r, \|X_{\bar Q}\|_r \leq 2\|X_Q\|^T_r$. To conclude, by Formula (73), we have that a quadratic, diagonal and piecewise Töplitz $\mathcal{Q}$ is of the form (96).

Our last statement is proved by noting that (see (17) for the norm of a vector field)

$$\|X_Q\|^2_r = 2 \sup_{\|z\|_{a,p} \leq r} \sum_{h \in S^c} |Q_h|^2 \frac{|z_h|^2}{r^2} e^{2a|h|}|h|^{2p} \geq |Q_j|^2$$

by evaluating at $z^{(j)}_h := \delta_{jh} e^{-a|j|}|j|^{-p} r/\sqrt{2}$. The same holds for $\mathcal{Q}$ and $\bar Q$. $\square$

Corollary 8.21. *Let Q be as above. Given parameters $\underline{p}$ for every $A \in \mathcal{H}_{\underline{p}}$ choose a point $m_A \in A^g_{\underline{p}}$. The the function $\mathcal{Q}(z) \in \mathcal{T}_{\underline{p}}$ defined as in formula (96) with $\mathcal{Q}(A) = Q_{m_A}$ is a Töplitz approximant of order $\varepsilon = 1$ for Q.*

9. Poisson bracket

9.1. **Poisson bracket estimate.** Analytic quasi-Töplitz functions are closed under Poisson bracket and respect the Cauchy estimates.

More precisely fix allowable $\vec{p} = (s, r, K, \theta, \mu, \lambda, \mathcal{O})$:

Proposition 9.2. *(i) Given $f^{(1)}, f^{(2)} \in \mathcal{Q}^T_{\vec{p}}$, quasi-Töplitz with parameters $\vec{p}$ we have that $\{f^{(1)}, f^{(2)}\} \in Q^T_{\vec{p}'}$, for all allowable parameters $\vec{p}' := (s', r', K', \theta', \mu', \lambda, \mathcal{O})$, with $\vec{p}'$ satisfying :*

$$r/2 < r' < r, \quad s/2 < s' < s, \quad 2\kappa < (\mu - \mu'){K'}^2, \; \kappa \mathtt{C} < (\theta' - \theta){K'}^{4d\tau_0 - 4},$$

$$e^{-(s-s')K'}{K'}^{\tau_1} < 1. \tag{100}$$

We have the bound

$$\|X_{\{f^{(1)}, f^{(2)}\}}\|^T_{\vec{p}'} \le 12\, 2^{2n+3} \delta^{-1} \|X_{f^{(1)}}\|^T_{\vec{p}} \|X_{f^{(2)}}\|^T_{\vec{p}} \tag{101}$$

where $\delta := \min(1 - \frac{s'}{s}, 1 - \frac{r'}{r}) < 1$.

(ii) Given $f^{(1)}, f^{(2)}$ as in item (i), assume that

$$3\, 2^{2n+8}\, e \delta^{-1} \|X_{f^{(1)}}\|^T_{\vec{p}} < 1/2, \tag{102}$$

then the function $f^{(2)} \circ \phi^t_{f^{(1)}} := e^{t\, ad(f^{(1)})}(f^{(2)})$, for $|t| \le 1$, is quasi-Töplitz for the parameters $\vec{p}'$. More precisely $f^{(2)} \circ \phi^t_{f^{(1)}} \in \mathcal{Q}^T_{\vec{p}'}$ for all parameters (K', θ', μ') for which

$$e^{-(s-s')\frac{K'}{(\ln K')^2}}{K'}^{\tau_1} < 1\,, \quad 2\kappa < (\mu - \mu'){K'}^2 \ln(K')^{-2}\,, \; \kappa \mathtt{C} < (\theta' - \theta){K'}^{4d\tau_0 - 4} \ln(K')^{-2}, \tag{103}$$

Finally we have

$$\|X_{f^{(2)} \circ \phi^t_{f^{(1)}}}\|^T_{\vec{p}'} \le 2 \|X_{f^{(2)}}\|^T_{\vec{p}} \tag{104}$$

$$\|X_{f^{(2)} \circ \phi^t_{f^{(1)}}} - X_{f^{(2)}}\|^T_{\vec{p}'} \le 2 \|X_{f^{(1)}}\|^T_{\vec{p}} \|X_{f^{(2)}}\|^T_{\vec{p}} \tag{105}$$

Proof. In order to prove this proposition we need some preliminaries. □

9.3. **A technical Lemma.** We take allowable parameters $\underline{p} := (N, \theta, \mu, \tau)$ and $\underline{p}' := (N, \theta', \mu', \tau)$ such that

$$\kappa \mathtt{C} < (\theta' - \theta) N^{4d\tau - 4}\,, \quad 2\kappa < (\mu - \mu') N^2. \tag{106}$$

Let us set up some notation. In Definition 8.2 we have introduced the notion of a monomial $\mathfrak{m} = e^{i(k,x)} y^l z^\alpha \bar{z}^\beta$ of low momentum with respect to the parameters $\underline{p}$ and denoted by $\Pi^L_{N,\mu}$ the projection on this subspace. Recall that one of the conditions is also that $|k| < N$, that is it has also a *low frequency*.

We shall say instead that $\mathfrak{m}$ is of N*–high frequency* if $|k| \ge N$ and denote Π^U_N the projection on this subspace.

We denote its degree in the high variables to be $d_H(\mathfrak{m})$. We further set

$$m(M) := \sum |\mathtt{r}(j)|(\alpha_j+\beta_j), \quad m_L(M) := \sum_{j \text{ low}} |\mathtt{r}(j)|(\alpha_j+\beta_j).$$

The projection symbol $\Pi_{N,\theta,\mu,\tau}$ is given in definition 8.3.

We use a mixed decomposition $f = \Pi_{N,\theta,\mu,\tau} f + \Pi^L_{N,3\mu'} f + \Pi^U_N f + \Pi_R f$ where $\Pi_R f$ is by definition the projection on those monomials which are neither (N,θ,μ,τ) bilinear nor of $(N,3\mu')$-low momentum nor of N-high frequency.

We hence divide the Poisson bracket in four terms: $\{\cdot,\cdot\} = \{\cdot,\cdot\}^{y,x} + \{\cdot,\cdot\}^L + \{\cdot,\cdot\}^H + \{\cdot,\cdot\}^R$ where the apices identify the variables in which we are performing the derivatives.

We need the following technical lemmas and definitions.

Lemma 9.4. *Consider a monomial $M = e^{\mathrm{i}\langle k,x\rangle} A z^\sigma_m = e^{\mathrm{i}\langle k,x\rangle} y^j z^\alpha \bar z^\beta z^\sigma_m$ which, with respect to some allowable parameters $\underline{p}$, has $|k| < N$, $m(A) < \mu N^3 + \kappa N$. Assume that z_m is a* high variable *i.e.* $|\mathtt{r}(m)| > \mathtt{c}N^{\tau_1}$. *Then M cannot satisfy conservation of momentum.*

Proof. If we have conservation of momentum $\pi(k) + \sum_j \sigma(j)\mathtt{r}(j)(\alpha_j - \beta_j) + \sigma\mathtt{r}(m) = 0$. By the hypothesis and the triangle inequality we have

$$\mathtt{c}N^{\tau_1} < |\mathtt{r}(m)| \le m(A) + \kappa N < \mu N^3 + 2\kappa N < (\mathtt{C}+\kappa)N^3 \tag{107}$$

a contradiction to Formula (50). $\square$

In order to simplify the notations let us set $\Pi' := \Pi_{N,\theta',\mu',\tau}$, $\Pi := \Pi_{N,\theta,\mu,\tau}$. Assume $f^{(1)}, f^{(2)}$ are two functions satisfying conservation of momentum.

Lemma 9.5. *The following splitting formula holds:*

$$\begin{aligned}\Pi'\{f^{(1)}, f^{(2)}\} = \Pi'(&\{\Pi f^{(1)}, \Pi f^{(2)}\}^H \\ &+\{\Pi f^{(1)}, \Pi^L_{N,3\mu'} f^{(2)}\}^{y,x} + \{\Pi f^{(1)}, \Pi^L_{N,3\mu'} f^{(2)}\}^L \\ &\{\Pi^L_{N,3\mu'} f^{(1)}, \Pi f^{(2)}\}^{y,x} + \{\Pi^L_{N,3\mu'} f^{(1)}, \Pi f^{(2)}\}^L + \\ &+\{\Pi^U_N f^{(1)}, f^{(2)}\} + \{f^{(1)}, \Pi^U_N f^{(2)}\} - \{\Pi^U_N f^{(1)}, \Pi^U_N f^{(2)}\})\end{aligned} \tag{108}$$

Proof. We perform a case analysis: we replace each $f^{(i)}$ with a single monomial to show which terms may contribute non trivially to the projection $\Pi'\{f^{(1)}, f^{(2)}\}$.

Consider the expression

$$\Pi'\{e^{i(k^{(1)},x)} y^{l^{(1)}} z^{\alpha^{(1)}} \bar z^{\beta^{(1)}}, e^{i(k^{(2)},x)} y^{l^{(2)}} z^{\alpha^{(2)}} \bar z^{\beta^{(2)}}\}. \tag{109}$$

If one or both of the $|k^{(i)}| > N$ then one or both monomials are of high frequency and we obtain a term in the last line of (108).

Suppose now that $|k^{(1)}|, |k^{(2)}| \le N$ we wish to understand under which conditions on the $\alpha^{(i)}, \beta^{(i)}$ the expression (109) is not zero.

For a monomial $M := e^{i(k,x)} y^a z^\alpha \bar z^\beta$ if $\Pi'(M) \neq 0$ we must have $d_H(M) = 2$ (plus further conditions). For two monomials M_1, M_2 we see that each term of $\{M_1, M_2\}$ has as degree $d_H(\{M_1, M_2\})$ equal to:

i) $d_H(M_1) + d_H(M_2) - 2$ if we have contracted conjugate high variables $z^\sigma_j, z^{-\sigma}_j$.
ii) $d_H(M_1) + d_H(M_2)$ otherwise.

In case i) in order to have $\Pi'\{M_1, M_2\} \neq 0$ we must have $d_H(M_1) + d_H(M_2) = 4$, this happens either if a) $d_H(M_1) = d_H(M_2) = 2$ or b) $d_H(M_1) = 1, d_H(M_2) = 3$ (resp. $d_H(M_1) = 3, d_H(M_2) = 1$). Let us show that, by conservation of momentum, case b) is not possible. Let us denote by $A_i, i = 1,2$ the part of the monomials M_i in the variables z_h which are not high i.e. $M_1 = e^{i(k_1,x)}y^{a_1}A_1 z_j^{\pm}$, $M_2 = e^{i(k_2,x)}y^{a_2}A_2 z_j^{\mp} z_m^{\sigma} z_n^{\sigma'}$. In $\{M_1, M_2\}$ the part of the monomial in the variables z_h which are not high is A_1A_2 so, if $\Pi'\{M_1, M_2\} \neq 0$ we must have $m(A_1A_2) = m(A_1) + m(A_2) < \mu' N^3$, we are thus in the hypotheses of Lemma 9.4 for M_1 a contradiction.

In case a) we claim that both monomials M_1, M_2 are N, μ, θ, τ bilinear, so that this contribution comes from the first line of (108). For this we only need to verify that, if z_j is the variable we contract, then j has a cut at ℓ for the parameters N, θ, μ, τ. Write $M_1 = e^{i(k_1,x)}y^{a_1}A_1 z_j^{\pm} z_m^{\sigma}$, $M_2 = e^{i(k_2,x)}y^{a_2}A_2 z_j^{\mp} z_n^{\sigma'}$. Since $\Pi'\{M_1, M_2\} \neq 0$ we have $m(A_1) + m(A_2) \leq \mu' N^3$. By conservation of momentum $|\mathtt{r}(j) \pm \mathtt{r}(m)| \leq \mu' N^3 + \kappa N$ hence $|j \pm m| \leq \mu' N^3 + \kappa N + 4d\kappa$, using (106) we have

$$|j \pm m| \leq \mu' N^3 + \kappa N + 4d\kappa < \mu' N^3 + 2\kappa N < \mu N^3.$$

We have that m has a ℓ cut for the parameters μ', θ', τ and the hypotheses in Formula (54) of Lemma 7.22 are satisfied, hence j has the cut and we obtain the first term in formula (108).

In case ii) we can have $d_H(M_1) + d_H(M_2) = 2$ either if a) $d_H(M_1) = 2, d_H(M_2) = 0$ (resp. $d_H(M_1) = 0, d_H(M_2) = 2$) or b) $d_H(M_1) = d_H(M_2) = 1$.

We claim that in case a) we obtain the contributions of lines 2,3 of Formula (108).

In fact say that $d_H(M_1) = 2, d_H(M_2) = 0$ and $M_1 = e^{i(k_1,x)}y^{a}A_1 z_m^{\sigma} z_n^{\sigma'}$, $M_2 = e^{i(k_2,x)}y^{b}A_2$ where A_i do not contain high variables. We have that $|k_i| < N$ by hypothesis, the high variables of M_1 have a N, μ', θ', τ cut also by hypothesis and so also a N, μ, θ, τ cut, finally if we contract variables x, y we have $m(A_1), m(A_2) \leq m_L\{M_1, M_2\} < \mu' N^3$. Assume we contract conjugate variables z_h which are not high, let $A_i = B_i z_h^{\pm}$ so that B_1B_2 is the part of $\{M_1, M_2\}$ in the low variables and $m(B_1) + m(B_2) < \mu' N^3$. By conservation of momentum for M_2 we have $|\mathtt{r}(h)| \leq m(B_2) + \kappa N$, hence $m(A_2) \leq 2m(B_2) + \kappa N < 2\mu' N^3 + \kappa N$. In both cases we deduce that we we obtain the contributions of lines 2,3 of Formula (108).

Let us show finally that, by conservation of momentum case b) is not possible. We are now assuming that for instance $M_1 = e^{i(k_1,x)}y^{a_1}B_1 z_h z_m^{\pm}$ with z_m high while z_h not high, hence $|\mathtt{r}(h)| \leq \mathtt{c}N^{\tau_1}$ and $m(B_1) < \mu' N^3$. We have $m(B_1 z_h) \leq \mu' N^3 + \mathtt{c}N^{\tau_1}$. So by conservation of momentum we have.

$$\theta' N^{\tau_1} < |\mathtt{r}(m)| \leq m(B_1) + |\mathtt{r}(h)| + \kappa N < \mu' N^3 + \theta N^{\tau_1} + \kappa N$$

which implies $(\theta' - \theta)N^{\tau_1} < \mu' N^3 + \kappa N$ contradicting (106). $\square$

9.6. **The proof of Proposition 9.2.** We use all the notations and hypotheses of 9.2. We can first use the standard estimates (22) and obtain

$$\|X_{\{f^{(1)}, f^{(2)}\}}\|^{\lambda}_{s',r'} < 2^{2n+3}\delta^{-1}\|X_{f^{(1)}}\|^{\lambda}_{s,r}\|X_{f^{(2)}}\|^{\lambda}_{s,r} \leq 2^{2n+3}\delta^{-1}\|X_{f^{(1)}}\|^{T}_{\vec{p}}\|X_{f^{(2)}}\|^{T}_{\vec{p}}, \tag{110}$$

here δ is defined in (23).

Since the allowable parameters K', θ', μ' satisfy (100) we have that N, θ', μ', τ satisfy (106) for all $N > K', \tau \geq \tau_0$. In order to show that $\{f^{(1)}, f^{(2)}\}$ is quasi–Töplitz (with respect to the chosen parameters), it is enough to provide, for all $N > K'$ and the allowable parameters $\underline{p}' := (N, \theta', \mu', \tau)$ a decomposition

$$\Pi'\{f^{(1)}, f^{(2)}\} = \Pi_{N,\theta',\mu',\tau}\{f^{(1)}, f^{(2)}\} = \mathcal{F}^{(1,2)} + N^{-4d\tau}\bar{f}^{(1,2)}$$

so that $\mathcal{F}^{(1,2)} \in \mathcal{T}_{\underline{p}'}$ and also

$$\|X_{\mathcal{F}^{(1,2)}}\|_{s',r'}, \|X_{\bar{f}^{(1,2)}}\|_{s',r'} < 12\, 2^{2n+3}\delta^{-1}\|X_{f^{(1)}}\|^{T}_{s,r}\|X_{f^{(1)}}\|^{T}_{s,r}. \tag{111}$$

For $\epsilon > 0$ take $\mathcal{F}^{(i)} \in \mathcal{T}_{\underline{p}}$, $\underline{p} = (s, r, N, \theta, \mu, \tau)$ so that setting

$$\bar{f}^{(i)} := N^{4d\tau}\big(\Pi_{(N,\theta,\mu,\tau)} f^{(i)} - \mathcal{F}^{(i)}\big), \tag{112}$$

we have

$$\max(\|X_{\mathcal{F}^{(i)}}\|_{s,r}, \|X_{\bar{f}^{(i)}}\|_{s,r}) < \|X_{f^{(i)}}\|^{T}_{\vec{p}} + \epsilon. \tag{113}$$

We thus define the function

$$\mathcal{F}^{(1,2)} := \Pi_{N,\theta',\mu',\tau}\Big(\{\mathcal{F}^{(1)}, \mathcal{F}^{(2)}\}^{H} + \{\mathcal{F}^{(1)}, \Pi^{L}_{N,3\mu'} f^{(2)}\}^{(y,x)+L} + \{\Pi^{L}_{N,3\mu'} f^{(1)}, \mathcal{F}^{(2)}\}^{(y,x)+L}\Big)$$

where we have denoted $\{\cdot,\cdot\}^{(y,x)+L} = \{\cdot,\cdot\}^{(y,x)} + \{\cdot,\cdot\}^{L}$. We need to show that it satisfies the required conditions.

Lemma 9.7. *(i) One has $\mathcal{F}^{(1,2)} \in \mathcal{T}_{\underline{p}'}$.*

(ii) Setting $\bar{f}^{(1,2)} = N^{4d\tau}(\Pi'\{f^{(1)}, f^{(2)}\} - \mathcal{F}^{(1,2)})$ one has that the bounds (111) *hold.*

Proof. (i) The constraints (74) are satisfied and we just apply Lemma 8.11.

(ii) The estimate (111) for $\mathcal{F}^{(1,2)}$ follows by Cauchy estimates since

$$\|X_{\mathcal{F}^{(1,2)}}\|_{s',r'} \le \|X_{\{\mathcal{F}^{(1)},\mathcal{F}^{(2)}\}}\|_{s',r'} + \|X_{\{\mathcal{F}^{(1)},f^{(2)}\}}\|_{s',r'} + \|X_{\{\mathcal{F}^{(2)},f^{(1)}\}}\|_{s',r'}$$

$$\le 2^{2n+3}\delta^{-1}\big[\|X_{\mathcal{F}^{(1)}}\|_{r,s}\|X_{\mathcal{F}^{(2)}}\|_{r,s} + \|X_{\mathcal{F}^{(1)}}\|_{r,s}\|X_{f^{(2)}}\|_{r,s} + \|X_{f^{(1)}}\|_{r,s}\|X_{\mathcal{F}^{(2)}}\|_{r,s}\big]$$

$$\le 3\, 2^{2n+3}\delta^{-1}\|X_{f^{(1)}}\|^{T}_{\vec{p}}\|X_{f^{(2)}}\|^{T}_{\vec{p}}.$$

We now estimate $\|X_{\bar{f}^{(1,2)}}\|_{s',r'}$. We have $\bar{f}^{(1,2)} =$

$$N^{4d\tau}\Pi'(\{f^{(1)}, f^{(2)}\} - \{\mathcal{F}^{(1)}, \mathcal{F}^{(2)}\}^{H} - \{\mathcal{F}^{(1)}, \Pi^{L}_{N,3\mu'} f^{(2)}\}^{(y,x)+L} - \{\Pi^{L}_{N,3\mu'} f^{(1)}, \mathcal{F}^{(2)}\}^{(y,x)+L})$$

We substitute in formula (108) $\Pi_{\underline{p}} f^{(i)} = \mathcal{F}^{(i)} + N^{-4d\tau}\bar{f}^{(i)}$.

Thus $\bar{f}^{(1,2)} = \Pi'(\Xi)$ with

$$\Xi = \big[\{\mathcal{F}^{(1)} + N^{-4d\tau}\bar{f}^{(1)}, \bar{f}^{(2)}\}^{H} + \{\bar{f}^{(1)}), \mathcal{F}^{(2)} + N^{-4d\tau}\bar{f}^{(2)}\}^{H} + N^{-4d\tau}\{\bar{f}^{(1)}, \bar{f}^{(2)}\}^{H} \tag{114}$$

$$+\{\bar{f}^{(1)}, \Pi^{L}_{N,3\mu'} f^{(2)}\}^{y,x+L} + \{\Pi^{L}_{N,3\mu'} f^{(1)}, \bar{f}^{(2)})\}^{y,x+L}$$

$$+N^{4d\tau}\big[\{\Pi^{U}_{N} f^{(1)}, f^{(2)}\} + \{f^{(1)}, \Pi^{U}_{N} f^{(2)}\} - \{\Pi^{U}_{N} f^{(1)}, \Pi^{U}_{N} f^{(2)}\}\big]$$

In order to estimate the norm $\|X_{\bar{f}^{(1,2)}}\|_{s',r'}$ we estimate the norm $\|X_{\Xi}\|_{s',r'}$.

If we have chosen ϵ sufficiently small in Formula (113) the first two lines of Ξ can be estimated by $9\, 2^{2n+3}\delta^{-1}\|X_{f^{(1)}}\|^{T}_{\vec{p}}\|X_{f^{(2)}}\|^{T}_{\vec{p}}$ and the last by the smoothing estimates (27).

$$\|X_{\Pi^{U}_{N} f}\|_{s',r'} \le 2e^{-N(s-s')}\|X_f\|_{s,r},$$

$$\|X_{\{\Pi^{U}_{N} f^{(1)}, f^{(2)}\} + \{f^{(1)}, \Pi^{U}_{N} f^{(2)}\}}\|_{s',r'} \le 8\, 2^{2n+3} e^{-N(s-s')}\delta^{-1}\|X_{f^{(1)}}\|_{s,r}\|X_{f^{(2)}}\|_{s,r}.$$

Since by (100) $N^{4d\tau}e^{-N(s-s')} < 1$, the estimate (111) follows. □

*Proof. Conclusion of the proof of (**proposition 9.2**)* Proposition 9.2(i) follows from the previous Lemma. The proof of (ii) is identical to that of Proposition 5 (ii) of [24]. □

Part 3. The KAM algorithm

10. An abstract KAM theorem

The starting point for our KAM Theorem is a class of Hamiltonians H, variation of the Hamiltonians considered in [24]:

$$H := \mathcal{N} + P\,, \quad P = P(x, y, z, \bar{z}, \xi). \tag{115}$$

$$\mathcal{N} := (\omega(\xi), y) + \sum_{k \in S_r^c} \Omega_k(\xi)|z_k|^2 + \sum_{(h,k) \in S_i^c} a_{h,k}(|z_h|^2 - |z_k|^2) + b_{h,k}(z_h z_k + \bar{z}_h \bar{z}_k),$$

defined in $D(s, r) \times \mathcal{O}$, where we take $\mathcal{O} \subseteq \varepsilon^2 \mathfrak{K}_\alpha$ a compact domain of diameter of order ε^2 contained in one of the components of Theorem 4 and subject to the restriction given in Remark 6.6. The functions $\omega(\xi), \Omega_n(\xi), a_{h,k}, b_{h,k}$ are well defined for $\xi \in \mathcal{O}$. In our examples the set S_i^c of *complex eigenvalues* or *hyperbolic terms* is finite.

It is well known that, for each $\xi \in \mathcal{O}$, the Hamiltonian equations of motion for the unperturbed $\mathcal{N}$ admit the special solutions $(x, 0, 0, 0) \to (x + \omega(\xi)t, 0, 0, 0)$ that correspond to invariant tori in the phase space.

Our aim is to prove that, under suitable hypotheses, there is a set $\mathcal{O}_\infty \subset \mathcal{O}$ of positive Lebesgue measure, so that, for all $\xi \in \mathcal{O}_\infty$ the Hamiltonians H still admit invariant tori (close to the ones of the unperturbed system) with some frequency $\omega^\infty(\xi)$ (close to $\omega(\xi)$).

Given a value ξ of the parameters we have the torus given by the equations $y = z = 0$. If the Hamiltonian vector field X_H of a Hamiltonian H is tangent to this torus, and if on this torus it coincides with $\sum_{i=1}^n \omega_i^\infty(\xi) \dfrac{\partial}{\partial x_i}$ then the Hamiltonian evolution is quasi–periodic on this torus, which is called a *KAM torus* for H.

This condition depends only on the terms $H^{\leq 2}$ of H of degree ≤ 2. Denote by $H^{(i,j)}$ the part of degree $2i$ in y and j in z, recall that we give degree 0 to the angles x, 2 to y and 1 to w.:

$$H^{\leq 2} = H^0(x) + H^{0,1}(x; w) + H^2(x; y, w), \quad H^2(x; y, w) = H^{1,0}(x; y) + H^{0,2}(x; w) \tag{116}$$

For a value ξ giving a KAM torus for H we have that the term H^0 must be constant (and we usually drop it), the term $H^{(0,1)} = 0$ and finally $H^{(1,0)} = \sum_{i=1}^n \omega_i^\infty(\xi) y_i$ (there is no condition on $H^{0,2}(x; w)$).

Therefore our goal is to find a change of variables (possibly in a smaller domain) so that we have a large set $\mathcal{O}_\infty$ of parameters defining KAM tori for H. The precise statement is contained in Theorem 6.

We start by describing the class of Hamiltonians to which the method applies.

10.1. Compatible Hamiltonians.

We consider a class of Hamiltonians stable under the KAM algorithm.

In the construction there will appear parameters

$$\underline{p} = (r, s, K, \theta, \mu, \mathcal{O}, a, S_0, M, L), \quad (\varepsilon, \gamma, \bar{\Theta})$$

playing different roles, where $\mathcal{O} \subset \varepsilon^2 \mathfrak{K}_\alpha$ is a compact set of positive measure (of order ε^{2n}) while all the others are positive numbers such that $K > N_0$ will play the role of a *frequency cut* and will grow to ∞ in the recursive algorithm, $\gamma < 1$ is an auxiliary parameter which we fix at the end of the algorithm and should be thought of as of order smaller than ε^2 but larger of the order of the perturbation, and [9]:

$$(\mu - \mathtt{c})K^{\tau_0}, (\mathtt{C} - \theta)K^{4d\tau_0} > \kappa K^4,$$

$$\gamma \le 2\varepsilon^2 M < 1/6\,,\ (8M\varepsilon^2)^{-1} > S_0 > 4\sqrt{n}ML\,, \quad a \le M\,, \quad LM < 4. \tag{117}$$

Recall that $\kappa = \max(|\mathtt{j}_i|)$ and see (50) for the definition of $N_0, \mathtt{c}, \mathtt{C}, \tau_0$.

We consider Hamiltonians, defined in $D(s,r) \times \mathcal{O}$, of the form:

$$H := \mathcal{N} + P,\ \mathcal{N} := (\omega(\xi), y) + \sum_{k \in S^c_r} \Omega_k(\xi)|z_k|^2 + \mathcal{C}, \quad P = P(x, y, z, \bar z, \xi), \tag{118}$$

$$\mathcal{C} = \sum_{(h,k)\in S^c_i} a_{h,k}(|z_h|^2 - |z_k|^2) + b_{h,k}(z_h z_k + \bar z_h \bar z_k), \quad (\omega(\xi), y) = \sum_{i=1}^n \omega_i(\xi) y_i.$$

We also may use for the complex eigenvalues the complex coordinates, as explained in §5.4.1. In that case unless there is a risk of confusion we may write the Hamiltonian in full diagonal form $\mathcal{N} := (\omega(\xi), y) + \sum_{k\in S^c} \Omega_k(\xi)|z_k|^2$ with the understanding that finitely many z_k are complex coordinates which come in pairs and that the corresponding $\Omega_k(\xi)$ are complex and in conjugate pairs.

Definition 10.2. We say that a Hamiltonian (118) is *compatible with the parameters* $\underline{p}$ if the following conditions $(A1)$–$(A5)$ are satisfied:

$(A1)$ *Non–degeneracy:* The map $\xi \to \omega(\xi)$ is a lipeomorphism[10] from $\mathcal{O}$ to its image with $|\omega^{-1}|^{lip}_\infty \le L$. Setting $\mathtt{v}_i := |\mathtt{j}_i|^2$, for $i = 1, \dots, n$ we have $|\omega(\xi) - \mathtt{v}|_\infty \le M\varepsilon^2$.

$(A2)$ *Asymptotics of normal frequency:* For all $n \in S^c$ we have a decomposition:

$$\Omega_n(\xi) = \sigma(n)(|\mathtt{r}(n)|^2 + 2\vartheta_n(\xi)) + \tilde\Omega_n(\xi). \tag{119}$$

We assume that the $\vartheta_n(\xi)$ are chosen in a finite list of analytic functions which are homogeneous of degree one in ξ, moreover the $\tilde\Omega := \{\tilde\Omega_n\}_{n\in S^c}$ are Lipschitz functions from $\mathcal{O} \to l_\infty$ with[11]

$$|\omega|^{lip}_\infty + |\Omega|^{lip}_\infty \le M \quad \text{and} \quad 2|\vartheta|^{lip}_\infty \le M\,, \quad 2|\vartheta|_\infty \le M\varepsilon^2\,. \tag{120}$$

$(A3)$ *Regularity and Quasi–Töplitz property:* the functions P, $\vartheta(z) := \sum_j \vartheta_j |z_j|^2$ and $\tilde\Omega(z) := \sum_j \tilde\Omega_j |z_j|^2$ are M–regular, preserve momentum as in (44), are Lipschitz in the parameters ξ and quasi-Töplitz with parameters (K, θ, μ) (cf. Definition 8.14). Moreover for all $N \ge K$, $\tau_0 \le \tau \le \tau_1/4d$ we have $\Pi_{(N,\theta,\mu,\tau)} \sum_j \vartheta_j |z_j|^2 \in \mathcal{T}_{(N,\theta,\mu,\tau)}$.

We need to control the norms of the above functions, we use the free parameter γ, whose purpose is to estimate the measure of the various Cantor sets which will appear, and set $\vec p = (s, r, K, \theta, \mu, \lambda = \gamma M^{-1}, \mathcal{O})$.

We define:

$$\gamma^{-1}\|X_{P^{(i)}}\|^T_{\vec p} := \epsilon^{(i)},\ i = 0, 1, 2\,, \quad \vec\epsilon = (\epsilon^{(0)}, \epsilon^{(1)}, \epsilon^{(2)}), \quad \gamma^{-1}\|X_P\|^T_{\vec p} := \Theta\,. \tag{121}$$

We require that

---

[9] note that the condition $LM < 4$ is added only in order to simplify notations, any ε^2, r, K independent constat would be acceptable.

[10] in our applications all maps will actually be analytic

[11] recall that on $\mathbb{R}^n$ we use the l^∞ norm.

$(A4)$ *Smallness condition*:[12]

$$\Theta < 1\,,\ \gamma^{-1}\|X_{\tilde\Omega}\|^T_{\vec p} < 1\,,\ 2^{2n+15}|\vec\epsilon|K^{4d\tau_1} < 1. \tag{122}$$

Note that the definition of Töplitz norm for diagonal matrices and (122) imply

$$|\tilde\Omega|_\infty \le \|X_{\tilde\Omega}\|^T_{\vec p} < \gamma \le 2M\varepsilon^2. \tag{123}$$

We note moreover that from condition $(A3)$ and by Remarks 8.17,3.4, we have that

$$\|X_{\vartheta(z)}\|^T_{\vec p} = \|X_{\vartheta(z)}\|^\lambda_{s,r} = |\vartheta|_\infty + \lambda|\vartheta|^{lip}_\infty = |\vartheta|_\infty + \gamma M^{-1}|\vartheta|^{lip}_\infty.$$

Then (117) and (120) imply $\|X_{\vartheta(z)}\|^T_{\vec p} < 2M\varepsilon^2$.

$(A5)$ *Non–degeneracy (*Melnikov conditions*):* We denote by $\Delta_{\xi,\varrho}f = \frac{|f(\xi)-f(\varrho\xi)|}{(1-\varrho)|\xi|}$ the variation of f in the radial direction:

$$\inf_{\substack{\varrho\in\mathbb{R},^+\xi\in\mathcal{O}\\ \varrho\neq 1,\ \varrho\xi\in\mathcal{O}}} |\Delta_{\xi,\varrho}(\langle\omega,k\rangle+\Omega\cdot l)| > a\,,\quad \forall k\in\mathbb{Z}^n,\ l\in\mathbb{Z}^{S^c}\,,|l|\le 2\,,\ |k|\le S_0\,,\ (k,l)\neq 0\,, \tag{124}$$

for all (k,l) compatible with momentum conservation and such that, setting $\mathtt{V} := (\mathtt{V}_n)_{n\in S^c}$, with $\mathtt{V}_n := \sigma(n)|\mathtt{r}(n)|^2$, one has (see $(A1)$):

$$\langle\mathtt{v},k\rangle + \mathtt{V}\cdot l = 0. \tag{125}$$

Observe that $\langle\mathtt{v},k\rangle + \mathtt{V}\cdot l \in\mathbb{Z}$.

Remark 10.3. If H is compatible with the parameters $\underline{p}$ it is also compatible with all choices of parameters $\underline{p}'$ where

$$s'\le s\,,\ r'\le r\,,\ K'\ge K\,,\ \mathcal{O}'\subseteq\mathcal{O}\,,$$

provided that (122) still holds.

In working with the cubic NLS we will also have (at the first step of the algorithm):

$(A2*)$ *Homogeneity*: The functions $\omega(\xi)-\mathtt{v}, \Omega_m - \sigma(m)|\mathtt{r}(m)|^2$ are analytic and homogeneous of degree 1.

Remark 10.4. By the homogeneity of $\omega(\xi)$ and ϑ we may easily see that M, L, a can be taken ε independent. These 3 paramters will remain bounded away from $0,\infty$ in the course of the algorithm. If the condition $(A2*)$ were that the functions are homogeneous of degree $q>1$ then we would only have ML, Ma^{-1} as ε independent constants.

[12]by $|\vec\epsilon| := |\epsilon^{(0)}| + |\epsilon^{(1)}| + |\epsilon^{(2)}|$ we mean its L^1 norm note that $|\vec\epsilon| < 3\Theta$.

10.4.1. *Infinite dimensional KAM theorem.* Now we state our infinite dimensional KAM theorem. We use the symbol $\lessdot$ to mean that we have $\le$ *cost* where *cost* depends only on n, d, κ. We use the same symbols $*$ as in the previous paragraph but with a $*_0$.

Theorem 6. *Assume that a Hamiltonian $\mathcal{N}_0 + P_0$ in (118) is compatible with the parameters $p_0 = (r_0, s_0, K_0, \theta_0, \mu_0, \mathcal{O}_0 = \varepsilon^2 \mathfrak{K}_\alpha, a_0, S_0, M_0, L_0)$, i.e. satisfies $(A1 - A5)$, $(A2*)$, and that:*

$$\mathtt{c} \le \frac{\mu_0}{2}\,, \quad \mathtt{C} \ge 2\theta_0\,, \quad M_0 L_0 = 2\,, \quad S_0 = 8\sqrt{n} M_0 L_0 = 16\sqrt{n} \tag{126}$$

and that K_0 is sufficiently large (depending only on the remaining parameters p_0).

There exists $0 < \bar\Theta < 1$ and a positive constant $B := B(K_0)$ such that if furthermore $\Theta_0 < \bar\Theta$ (cf. (121)) and for all γ such that $B\varepsilon^{2(n-1)}\gamma < |\mathcal{O}_0| a_0$, we may construct:

- **(Frequencies)** *Lipschitz functions $\omega_\infty : \mathcal{O}_0 \to \mathbb{R}^n$, $\Omega^{(\infty)} : \mathcal{O}_0 \to \ell_\infty$, satisfying*

$$|\omega_\infty - \omega_0|^\lambda\,,\ \ |\Omega^{(\infty)} - \Omega^{(0)}|^\lambda_\infty \le \gamma |\vec\epsilon_0| \bar\Theta^{-1} \tag{127}$$

where $\vec{\epsilon_0}$ is defined by Formula (121) and $|\omega_\infty|^{\rm lip}$, $|\Omega^{(\infty)}|^{\rm lip}_\infty \le 2M_0$.

- **(Cantor set)** *A Cantor set $\mathcal{O}_\infty$*

$$|\mathcal{O}_0 \setminus \mathcal{O}_\infty| \le B\varepsilon^{2(n-1)} \frac{\gamma}{a_0}, \tag{128}$$

the smallness condition on γ ensures that $\mathcal{O}_\infty$ has positive Lebesgue measure.

- **(KAM normal form)** *A Lipschitz family of analytic symplectic maps*

$$\Phi : D(s/4, r/4) \times \mathcal{O}_\infty \mapsto D(s, r) \tag{129}$$

of the form $\Phi = \mathrm{I} + \Psi$ with $\|\Psi\|_{r/4, s/4} \lessdot |\vec\epsilon_0|$, where $\mathcal{O}_\infty$ is defined in the previous item, such that,

$$H^\infty(\cdot;\xi) := H \circ \Phi(\cdot;\xi) = \omega_\infty(\xi) y_\infty + \Omega^{(\infty)}(\xi) z_\infty \bar z_\infty + P^\infty \quad \textit{has} \quad P^\infty_{\le 2} = 0. \tag{130}$$

Formula (130) tells us that the final Hamiltonian has invariant KAM tori parametrized by $\xi \in \mathcal{O}_\infty$.

Remark 10.5. We will use the freedom given by Remark 10.3 to choose K_0 large enough so that some necessary bounds, which appear in the proof, hold.

It is important to notice that all the conditions which we shall impose on K are independent of r_0 and $\mathcal{O}_0$.

Theorem 6 is proved by an iterative procedure, which occupies the rest of this long section. We produce a sequence of Hamiltonians $H_\nu = \mathcal{N}_\nu + P_\nu$, each of these Hamiltonians will satisfy the properties (A*) of Section 10.1 for suitable parameters p_ν.

In particular for the compact sets $\mathcal{O}_\nu$ and for the domains $D(s_\nu, r_\nu)$ we have the *telescoping* $\mathcal{O}_\nu \subset \mathcal{O}_{\nu-1}$, $D(s_\nu, r_\nu) \subset D(s_{\nu-1}, r_{\nu-1})$.

We set $\mathcal{O}_\infty = \cap_\nu \mathcal{O}_\nu$, $D(s_\infty, r_\infty) = \cap_\nu D(s_\nu, r_\nu)$ and the construction will be such that $\mathcal{O}_\infty$ has positive measure while $s_\infty = s/4, r_\infty = r/4$.

We have a sequence of symplectic transformations $\Phi_\nu : H_{\nu-1} := H_\nu$, where Φ_ν is the value at 1 of the flow generated by the Hamiltonian vector field $X_{F_{\nu-1}}$ associated to the generating function $F_{\nu-1}$ determined as the unique solution of the *Homological equation*, and depending on $P^{\le 2}_{\nu-1}$. The transformation is well defined on the domain $D(s_\nu, r_\nu) \times \mathcal{O}_\nu$, with $D(s_\nu, r_\nu) \subset D(s_{\nu-1}, r_{\nu-1})$. At each step, the perturbation remains bounded while the part $P_\nu^{\leq 2}$ becomes smaller. We also denote $D(s_\nu, r_\nu) = D_\nu$. The Töpliz norm of a function G defined on $D_\nu \times \mathcal{O}_\nu$ and relative to the parameters extracted from the p_ν will be denoted by $\|G\|_\nu$.

The goal is to pass to a limit Hamiltonian $H_\infty = \mathcal{N}_\infty + P_\infty$ with the property that $P_\infty^{\leq 2} = 0$ so that the Hamiltonian vector field on the family of tori parametrized by the parameters $\xi \in \mathcal{O}_\infty$, where the normal coordinates are 0, coincides with the Hamiltonian vector field of $\mathcal{N}_\infty$.

The relevant estimates to be performed are the following.

- We have to estimate the norm of each $F_{\nu-1}$ so to make sure that the value $\Phi^1_{\nu-1}$ at 1 of the flow generated by the Hamiltonian vector field $X_{F_{\nu-1}}$ is well defined on $D_\nu \times \mathcal{O}_\nu$.

  Here the problem is that of *small denominators* since we have to divide by eigenvalues. Here the quasi-Töpliz properties play a major role and the key is Proposition 10.16.
- While we establish the previous item we have to estimate the measure of the set $\mathcal{O}_\nu$, Lemma 10.12.
- We have to perform all the estimates on the new parameters p_ν, this is done in §10.17.
- We have to estimate the norm of the part $P_\nu^{\leq 2}$, for this we have to control simultaneously the three parts in which it naturally decomposes.
- We need to prove that the set $\mathcal{O}_\infty$ has positive measure, Corollary 10.21.
- We need to prove that on the set $\mathcal{O}_\infty$ we have a limit change of coordinates Φ^1_∞ giving rise to the limit Hamiltonian 10.21.

Warning In order for this Theorem to give a non–empty statement we need to have conditions which ensure that the constraints on γ can be satisfied. These constraints amount to a smallness condition on the perturbation, (cf. Theorem 7). In the applications to the NLS this condition is satisfied by suitably restricting the domain of definition of $\mathcal{H}$.

10.6. **KAM step.**

10.6.1. *Formal KAM step.* The input of a KAM step is a Hamiltonian H of the previous form with parameters $\underline{p}$. Of particular relevance is the parameter $K \geq K_0$ which gives a *frequency cut*. The output must produce a new Hamiltonian H_+ of the previous form with parameters $\underline{p}_+$. Thus we need to start from a subset $\mathcal{O}_+ \subset \mathcal{O}$ of the parameters ξ, two new values for the radii of the domain $s_+ \leq s,\ r_+ \leq r$ and a symplectic transformation $\Psi : D(s_+, r_+) \times \mathcal{O}_+ \to D(s, r)$ of type $\Psi = e^{ad(F)}$, so that finally we have a new Hamiltonian $H_+ = H \circ \Psi$ which we expect to be a *simplified version of* H by evaluating the new parameters $\underline{p}_+$. After iterating infinitely many times the KAM step, we hope to arrive at the desired final Hamiltonian which shows the existence of quasi–periodic orbits as in Theorem 6.

The function F is obtained by solving the *homological equations*. In order to explain this it will be convenient to write explicitly the terms of $P^{(2)}(x, y, w)$:

$$P^{(1)}(x; w) = \sum_{m \in S^c,\ \sigma = \pm,\ k} P^{(1)}_{k,m,\sigma} e^{\mathrm{i}(k,x)} z_m^\sigma$$

$$P^{(0)}(x) = \sum_k P_k^{(0,0)} e^{\mathrm{i}(k,x)}\,, \quad P^{(1,0)}(x;y) = \sum_k P_k^{(1,0)} \cdot y e^{\mathrm{i}(k,x)},$$

$$P^{(0,2)}(x;w) = \sum_{n,m\in S^c\,,\ \sigma,\sigma'=\pm\,,\ k} P^{(0,2)}_{k,m,\sigma,n,\sigma'} e^{\mathrm{i}(k,x)} z_m^\sigma z_n^{\sigma'}$$

Only those terms which satisfy conservation of mass and momentum may appear. We set

$$[P^{\leq 2}] = (P_0^{(1,0)}, y) + \sum_{m\in S^c_r} P^{(0,2)}_{0,m,+,m,-}|z_m|^2 + \mathcal{P}. \tag{131}$$

Where $\mathcal{P}$ is diagonal only in the complex notation and arises from the term $\mathcal{C}$ of the normal form $\mathcal{N}$ (see (118)).

On the space of quadratic Hamiltonians $ad(\mathcal{N})$ has a basis of eigenvectors described in §6.4. On the space relative to the non-zero eigenvalues $ad(\mathcal{N})$ is formally invertible, hence for those ξ for which the Melnikov resonances do not occur, $[P^{\leq 2}]$ is the projection of $P^{\leq 2}_{\leq K}$ on the kernel of $ad(\mathcal{N})$. We define

$$F := ad(\mathcal{N})^{-1}(P^{\leq 2}_{\leq K} - [P^{\leq 2}]) \quad \implies ad(F)\mathcal{N} = [P^{\leq 2}] - P^{\leq 2}_{\leq K} \tag{132}$$

and since $ad(\mathcal{N})^{-1}$ is diagonal (at least in complex coordinates) this definition can be given degree by degree, thus defining $F^0, F^{(1)}, F^{(2)}$. Notice that even if we use complex coordinates F is always real.

10.6.2. *Estimates.* Formula (132) defines F as a formal expression. We now impose a lower bound on the eigenvalues of $ad(\mathcal{N})$ on the space of functions of degree ≤ 2 which implies that F is analytic. Let us restrict our attention, for instance, to the set $\mathcal{O}'$ of $\xi \in \mathcal{O}$ such that: for all $k \in \mathbb{Z}^n$, $|k| \leq K$ and $l \in \mathbb{Z}^{S^c}$, $|l| \leq 2$ which satisfy momentum conservation, we have

$$|\langle \omega, k\rangle + (l, \Omega)| \geq \gamma K^{-2d\tau_1}. \tag{133}$$

Lemma 10.7. *We have:*

$$\|X_{F^i}\|^\lambda_{s,r,\mathcal{O}'} \leq K^{2d\tau_1}\gamma^{-1}\|X_{P^i}\|^\lambda_{s,r,\mathcal{O}}\,, \quad i = 0,1,2\,. \tag{134}$$

Proof. We first notice that (133) implies that $P^{\leq 2}_{\leq K} - [P^{\leq 2}]$ is a sum of eigenvectors of $ad(\mathcal{N})$ with eigenvalues bounded from below (in absolute value) by $\gamma K^{-2d\tau_1}$, therefore since we are using the Majorant norm we have

$$\|X_{F^i}\|^\lambda_{s,r} \leq K^{2d\tau_1}\gamma^{-1}\|X_{P^{(i)}-[P^{(i)}]}\|^\lambda_{s,r} \leq K^{2d\tau_1}\gamma^{-1}\|X_{P^{(i)}}\|^\lambda_{s,r}\,, \quad i = 0,1,2.$$

□

Then by Proposition 3.3 F defines a symplectic transformation $e^{(ad(F))}$ on a domain $D(s',r') \times \mathcal{O}'$, since by (122) the condition $2^{2n+3}e\delta^{-1}\|X_F\|_{s',r'} < 1$ holds for a suitable choice of s',r' and possibly restricting to a subset $\mathcal{O}'$ of the parameters. More precisely in the next paragraph we will define a set $\mathcal{O}_+ \subset \mathcal{O}$ (see Definition 10.10) and show in Lemma 10.12 that, provided γ is sufficiently small, this set has positive measure. On this set we shall prove in Proposition 10.13 that the inequalities (133) hold. In order to iterate this procedure we need to be sure that F is quasi-Töplitz and estimate its norm, this will be proved in Proposition 10.16. So for the procedure to succeed we need that the perturbation $P^{(i)}$, $i \leq 2$ be rather small.

10.7.1. *KAM step.* For simplicity, below we always use the same symbol *cost* to denote constants independent on the iteration and on the parameters of the Hamiltonian.

We now start from a Hamiltonian in the class of Definition 10.2 and describe a procedure which produces a change of variables under which the Hamiltonian is still in the same class with new parameters which we estimate explicitly.

KAM Step 1. *(1) We Define in 10.10, a compact set $\mathcal{O}_+ \subset \mathcal{O}$ such that*

$$|\mathcal{O} \setminus \mathcal{O}_+| \le \Gamma \frac{\gamma}{a} \varepsilon^{2(n-1)} K^{-\tau_0+n+d/2}, \tag{135}$$

where Γ is a constant depending only on n, d, κ.

(2) We construct, by Formula (132)*, the function F. We prove in Proposition 10.16 that F is quasi-Töplitz with parameters K, θ, μ for all $\xi \in \mathcal{O}_+$.*

For all positive numbers $r_+ < r$ and $s_+ < s$ for which:

$$2^{2n+14}(\min(1 - \frac{r_+}{r}, 1 - \frac{s_+}{s}))^{-1} |\vec{\epsilon}| K^{4d\tau_1} < \frac{1}{2}\,, \tag{136}$$

we show that F generates a 1–parameter group of analytic symplectic transformations $\Phi_F^t : D(s_+, r_+) \to D(s, r)$, well defined for all t, $|t| \le 1$ and for all $\xi \in \mathcal{O}_+$.

(3) Applying Proposition 9.2 ii) with $\underline{p}' \rightsquigarrow \underline{p}^+ = (N_+, \mu_+, \theta_+, s_+, r_+)$ we show that $\Phi_F^1 H := H_+ = \mathcal{N}_+ + P_+$ is quasi-Töplitz for all choices of parameters K_+, θ_+, μ_+ satisfying (103).

Remark 10.8. In order to make sure that $\mathcal{O}_+$ is non–empty we need the corresponding constant $B = \Gamma K^{-\tau_0+n+d/2}$ should satisfy $B\varepsilon^{2(n-1)}\gamma < |\mathcal{O}|a$ which we shall satisfy by imposing a smallness condition on γ since $K \ge K_0$.

The core of the construction is to compute the parameters M_+, L_+, a_+, see (165), and $\Theta_+, \vec{\epsilon}_+$, see (169) relative to the new Hamiltonian. The iterative KAM algorithm is based on the fact that if $\bar{\Theta}$ in Theorem 6 is small enough then H_+ is compatible with the parameters $\vec{p}_+$ and respects the smallness condition $(A3^*)$, so one may iterate the *step*.

10.9. **The set O_+.**

Definition 10.10. $\mathcal{O}_+ = \mathcal{O}_{+,\gamma}$ is defined to be the subset of $\xi \in \mathcal{O}$ where the following *Melnikov non–resonance conditions* are satisfied:

i) For all $k \in \mathbb{Z}^n$, $|k| \le K$ and $h \in \mathbb{Z}$, so that $(h, k) \neq (0, 0)$:

$$|\langle \omega(\xi), k \rangle + h| \ge 2\gamma K^{-\tau_0}\,. \tag{137}$$

ii) For all $k \in \mathbb{Z}^n$, $|k| \le K$, $m \in S^c$, with $\pi(k) \pm \mathtt{r}(m) = 0$:

$$|\langle \omega(\xi), k \rangle \pm \Omega_m| \ge 2\gamma K^{-\tau_0}. \tag{138}$$

iii) For all $|k| \le K, m, n \in S^c$ such that $\min(|\mathtt{r}(m)|, |\mathtt{r}(n)|) \le \mathtt{C} K^{\tau_1}$ and $\pi(k) \pm (\mathtt{r}(m) + \sigma \mathtt{r}(n)) = 0$ with $\sigma = \pm 1$:

$$|\langle \omega(\xi), k \rangle \pm (\Omega_m + \sigma \Omega_n)| \ge 2\gamma K^{-2d\tau_1}. \tag{139}$$

For all $|k| \le K, m, n \in S^c$ $|\mathtt{r}(m)|, |\mathtt{r}(n)| > \mathtt{C} K^{\tau_1}$ and $\pi(k) \pm (\mathtt{r}(m) + \mathtt{r}(n)) = 0$

$$|\langle \omega(\xi), k \rangle \pm (\Omega_m + \Omega_n)| \ge 2\gamma K^{-2d\tau_1} \tag{140}$$

iv) For all affine spaces $A = [v_i, p_i]_\ell$ in $\mathcal{H}_K$ $(1 \le \ell < d)$ with $p_\ell < \mathtt{c} K^{\tau_1/4d}$ we choose a point $m_A \in [v_i; p_i]_\ell^g$ with $|\mathtt{r}(m_A)| > \mathtt{C} K^{\tau_1}$. For all such m_A and for all k such that $|k| \le K$, we require:

$$|\langle \omega(\xi), k \rangle + \Omega_{m_A} - \Omega_{\bar{n}}| \ge 2\gamma \min(K^{-2d\tau_0}, \mathtt{c}^{2d} p_\ell^{-2d}), \tag{141}$$

for all $\bar{n}$ such that $\mathtt{r}(\bar{n}) = \mathtt{r}(m_A) + \pi(k)$.

In order to analyze $\mathcal{O}_+$ we need a first Lemma for the measure estimates. We define

$$\mathcal{R}^\tau_{k,l} := \left\{\xi \in \mathcal{O} | \, |\langle \omega, k\rangle + (l, \Omega)| < \gamma K^{-\tau}\right\}, \quad l \in \mathbb{Z}^{S^c}.$$

Lemma 10.11. *For all $(k,l) \neq (0,0)$ $|k| \le K$ and $|l| \le 2$, which satisfy momentum conservation, one has*

$$|\mathcal{R}^\tau_{k,l}| \lessdot \frac{\gamma}{a}\varepsilon^{2(n-1)}K^{-\tau}. \tag{142}$$

Proof. Let us first state a general fact. Let f be a Lipschitz function on the domain $\mathcal{O} \subset \varepsilon^2\mathfrak{K}$ such that

$$|f(x, \xi_2, \dots, \xi_n) - f(y, \xi_2, \dots, \xi_n)| > a|x-y|$$

for all $x \neq y$ such that $(x, \xi_2, \dots, \xi_n), (y, \xi_2, \dots, \xi_n) \in \mathcal{O}$.

We consider the map $F : \xi \mapsto (f(\xi), \xi_2, \dots, \xi_n)$ which maps $\mathcal{O}$ bijectively to some set B. F is a lipeomorphism and its inverse has Lipschitz constant $< \max(1, a^{-1})$. In B, f is a coordinate and the level surfaces of f are contained in a hypercube of volume $\varepsilon^{2(n-1)}$ therefore the volume of the set where $|f| < c$ can be estimated by $2\varepsilon^{2(n-1)}c$ hence on $\mathcal{O}$ it can be estimated by $2a^{-1}\varepsilon^{2(n-1)}c$. A similar argument is valid if we work in polar coordinates, x, y are radii and the other ξ_i coordinates on the unit sphere.

If $|k| \le S_0$, we have assumed that, for all functions $f_{k,l}(\xi) = \langle\omega, k\rangle + (l, \Omega(\xi))$ which satisfy (125) we have $\inf_{\xi\neq\eta\in\mathcal{O}} |\Delta_{\xi,\varrho} f_{k,l}| > a$. So we may apply the previous argument.

If, on the other hand, $f_{k,l}$ does not satisfy (125) (but $|k| \le S_0$) then we may write

$$f_{k,l}(\xi) = \mathfrak{n} + F_{k,l}(\xi),$$

where $\mathfrak{n}$ is the non–zero integer computed in (125) and

$$|F_{k,l}| \le |\langle\omega(\xi) - \mathtt{v}, k\rangle| + 4\sup_n |\vartheta_n(\xi)| + 2|\tilde{\Omega}|_\infty,$$

by hypothesis $|k| < S_0$ (in particular $S_0 > 1$) while $2|\vartheta|_\infty, |\omega - \mathtt{v}|_\infty \le M\varepsilon^2$. So, by (123) and (117), we have $|f_{k,l}| \ge |\mathfrak{n}| - (S_0 + 4)M\varepsilon^2 > \frac{1}{2}$ and we may deduce that $R^\tau_{k,l}$ is empty. Finally if $|k| \ge S_0$ then we change the variables form ξ to ω and study $G_{k,l}(\omega) := \langle\omega, k\rangle + (\Omega(\xi(\omega)), l)$. Let $\underline{e}_k$ be the versor of k, we may perform an orthogonal change of variables in ω so that $\underline{e}_k$ is the first vector in the standard basis. Then the Lipschitz norm of $\langle\omega, k\rangle$ is the absolute value of the vector k which can be bounded below by $\frac{|k|}{\sqrt{n}} \ge \frac{S_0}{\sqrt{n}} > 4ML$. Then we repeat our argument with respect to ω, indeed

$$|(\Omega(\xi(\omega)), l)|^{lip}_\infty \le |\Omega(\xi)|^{lip}_\infty |\omega^{-1}|^{lip}_\infty \le ML$$

so that

$$|\frac{G_{k,l}(x, \omega_2, \dots, \omega_n) - G_{k,l}(y, \omega_2, \dots, \omega_n)}{x-y}| > 4ML - ML = 3ML\,,$$

for all vectors $(x, \omega_2, \dots, \omega_n) \neq (y, \omega_2, \dots, \omega_n)$ in $\omega(\mathcal{O})$. Thus the volume of the set where $|G_{k,l}(\omega)| < c$ can be estimated by $(ML)^{-1}(M\varepsilon^2)^{(n-1)}c$. The corresponding volume in the space of the parameters ξ is therefore estimated by $L^n(ML)^{-1}(M\varepsilon^2)^{(n-1)}c = (LM)^{n-1}M^{-1}\varepsilon^{2(n-1)}c$. Since by (117) $ML < 4, M > a$ the estimate follows. ☐

Lemma 10.12. *The set $\mathcal{O}_+$ is compact and one has*

$$|\mathcal{O} \setminus \mathcal{O}_+| \le \Gamma\,\frac{\gamma}{a}\varepsilon^{2(n-1)}K^{-\tau_0+n+d+1}. \tag{143}$$

Proof. Since k runs on a finite set the functions $|\langle\omega(\xi),k\rangle|$ are bounded on $\mathcal{O}$, hence the first formula (137) is satisfied for $|h|$ large, for instance $|h|>2|\omega|_\infty K$. So (137) is actually implied by a finite number of inequalities.

Formulas (138) and (139) are a finite number of inequalities by definition. Formula (140) is a priori an infinite list of inequalities, we note however that $|(\omega,k)\pm(\Omega_m+\Omega_n)|>|\omega|_\infty K$ is large when $|\Omega_m+\Omega_n|>2|\omega|_\infty K$.

Next Ω_m has an integral part $\sigma(m)|\mathtt{r}(m)|^2$, but $\sigma(m)=1$ as soon as $|\mathtt{r}(m)|>\mathtt{C}K^{\tau_1}$. This implies that $\Omega_m+\Omega_n$ is large, and hence no condition is imposed, except possibly for finitely many values of m,n. Finally (141) is given only for finitely many elements; in fact in *iv)*, for each $[v_i,p_i]^g_\ell$ and k, we impose only a fixed finite number of condition by choosing a point m_A and a type $u=\mathtt{r}(\bar n)-\bar n\in\mathcal{Z}$. Finally by Remark 7.8 there are a finite number of $[v_i,p_i]^g_\ell$. Thus $\mathcal{O}_+$ is compact.

Let us prove the measure estimates. By definition $\mathcal{O}_+$ is obtained from $\mathcal{O}$ by removing a finite list of strips $\mathcal{R}^{\tau_p}_p$ where p runs in a suitable set of pairs k,l. For a given set of indices I denote by $\mathcal{R}^\tau_I:=\cup_{p\in I}\mathcal{R}^\tau_p$ and, by Lemma 10.11, we estimate $|\mathcal{R}^\tau_I|\le|I|\gamma a^{-1}\varepsilon^{2(n-1)}K^{-\tau}$. The Lemma will thus follow from an estimate on the cardinality of I for the various cases considered.

Recall that the elements $m\in\mathbb{Z}^d$ with $\sigma(m)=-1$ are a finite set of some cardinality depending only on κ,n similarly their norm can be bounded by some $\bar\kappa$ of the order of κ^2. By Hypothesis $(A1)$, $|\omega|_\infty<\kappa^2+1$, and we may assume that K is large so that $\bar\kappa\le\sqrt{2(\kappa^2+1)K}$.

i) Is previously remarked we have to impose (137) with $|h|\le|\omega|_\infty K\le(\kappa^2+1)K$ we have:

$$I_0:=\{(k,h)\,|,|k|\le K\,,\,|h|\le(\kappa^2+1)K\},\quad |I_0|\le(\kappa^2+1))(2K)^{n+1}$$
$$|\mathcal{R}^\tau_{I_0}|\lessdot\varepsilon^{2(n-1)}\gamma a^{-1}K^{-\tau_0+n+1}.$$

ii) In (138), by momentum conservation $l=\pm e_m$ implies that $\pm\mathtt{r}(m)=-\pi(k)$. Hence to impose (138) we have to remove the list indexed by I_1:

$$I_1:=\{(k,l)\,|,|k|\le K\,,\,l=\pm e_m,\ \exists\,u\in\mathcal{Z}:m+u=\mp\pi(k)\}\,,\quad |I_1|\le(2K)^nd.$$
$$|\mathcal{R}^\tau_{I_1}|\lessdot\varepsilon^{2(n-1)}\gamma a^{-1}K^{-\tau_0+n+1}.$$

iii) If $l=\pm(e_m+e_n)$ and $\sigma(m)=\sigma(n)=1$, the index (k,l) can contribute only if we have the condition

$$|\pm\langle\omega,k\rangle+|\mathtt{r}(m)|^2+|\mathtt{r}(n)|^2+2\vartheta_m+2\vartheta_n+\tilde\Omega_m+\tilde\Omega_n|<\frac12.$$

From (120) and (123) this condition implies $|\pm\langle\omega,k\rangle+|\mathtt{r}(m)|^2+|\mathtt{r}(n)|^2|<1$ hence $|\mathtt{r}(m)|^2+|\mathtt{r}(n)|^2<2|\omega|_\infty K$. Setting

$$I_2:=\{(k,l)\,|,|k|\le K\,,\,l=\pm(e_m+e_n)\,,|\mathtt{r}(m)|\le\sqrt{2(\kappa^2+1)K}\,,$$
$$\exists u,v\in\mathcal{Z}:\ m+n+v+u=\mp\pi(k)\},$$
$$|I_2|<(2\sqrt{2(\kappa^2+1)K}+1)^dd^2,\quad |\mathcal{R}^\tau_{I_2}|\lessdot\varepsilon^{2(n-1)}\gamma a^{-1}K^{-\tau_0+n+d/2}.$$

iv) Setting

$$I_3:=\{(k,l)\,|,|k|\le K\,,\,l=e_m-e_n\,,|\mathtt{r}(m)|\le\mathtt{C}K^{\tau_1}\,,\exists u,v\in\mathcal{Z}:\ m-n+u-v=\mp\pi(k)\}$$

One has

$$|I_3|\lessdot K^{d\tau_1+n}\implies|\mathcal{R}^{2d\tau_1}_{I_3}|\lessdot\gamma a^{-1}\varepsilon^{2(n-1)}K^{-d\tau_1+n}.$$

v) To deal with the last case, for all affine subspaces $[v_i;p_i]_\ell \in \mathcal{H}_K$ with $p_\ell \le \mathtt{C}K^{\frac{\tau_1}{4d}}$, for all $|k|\le K$ and for all types $u\in\mathcal{Z}$ we set $\bar n = \mathtt{r}(m)+u+\pi(k)$ and define

$$\mathcal{R}_{k,[v_i;p_i]_\ell,u} := \{\xi \mid |\langle\omega,k\rangle+\Omega_{m_I}-\Omega_{\bar n}| < 2\gamma\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d})\} \tag{144}$$

Following Lemma 10.11, $|\mathcal{R}_{k,[v_i;p_i]_\ell,u}| \lessdot \gamma a^{-1}\varepsilon^{2(n-1)}\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d})$.

We distinguish the two cases. First when $\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d}) = K^{-2d\tau_0}$ we are in

$$I_1^K := \{(k,[v_i;p_i]_\ell,u)\,||k|<K\,,[v_i;p_i]_\ell\in\mathcal{H}_K:\, p_\ell\le \mathtt{c}K^{\tau_0}\,,\, u\in\mathcal{Z}\}$$

when $\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d}) = \mathtt{c}^{2d}p_\ell^{-2d}$ we are in

$$I_2^K := \{(k,[v_i;p_i]_\ell,u)\,||k|<K\,,[v_i;p_i]_\ell\in\mathcal{H}_K:\, \mathtt{c}K^{\tau_0}< p_\ell\le \mathtt{C}K^{\frac{\tau_1}{4d}}\,,\, u\in\mathcal{Z}\}$$

By Remark 7.8 we have $|I_1^K| \lessdot K^{(d+\tau_0)(d-1)+n}$ hence

$$|\mathcal{R}_{I_1^K}| \lessdot \gamma a^{-1}\varepsilon^{2(n-1)}K^{d(d-1)+\tau_0(-d-1)+n} \le \gamma a^{-1}\varepsilon^{2(n-1)}K^{-d\tau_0}$$

as for I_2^K we use again Remark 7.8 to bound with $(2\kappa K)^{d^2}(2p)^{d-1}$ the number of subspaces with given $p=p_\ell$ for all ℓ.

$$|\mathcal{R}_{I_2^K}| \lessdot \gamma a^{-1}\varepsilon^{2(n-1)}\sum_{p_\ell>\mathtt{c}K^{\tau_0}} p_\ell^{-2d-1+d}K^{d^2}K^n \lessdot \gamma a^{-1}\varepsilon^{2(n-1)}K^{-d\tau_0+d^2+n}$$

(from (50)). Summing all these contributions, since we can bound all the factors by $\varepsilon^{2(n-1)}\gamma a^{-1}K^{-\tau_0+n+d+1}$, our Lemma is proved. □

We arrive now to the key estimate which handles *small denominators* and for which we have introduced all the formalism of cuts and quasi–Töplitz functions.

Proposition 10.13. *For all $\xi\in\mathcal{O}_+$, for all $k\in\mathbb{Z}^n$, $|k|\le K$ and $l\in\mathbb{Z}^{S^c}$, $|l|\le 2$ which satisfy momentum conservation, we have*

$$|\langle\omega,k\rangle+(l,\Omega)| \ge \gamma K^{-2d\tau_1}. \tag{145}$$

Proof. By Definition 10.10 the cases i), ii), iii) follow trivially since $2d\tau_1$ is large with respect to τ_0.

We are left with the case $\ell = e_m - e_n$ with $|\mathtt{r}(m)|,|\mathtt{r}(n)| > \mathtt{C}K^{\tau_1}$. To start we have

$$|\langle\omega,k\rangle+\Omega_m-\Omega_n| \ge |\langle\omega,k\rangle+|\mathtt{r}(m)|^2-|\mathtt{r}(n)|^2| - 4|\vartheta|_\infty - 2|\tilde\Omega|_\infty.$$

We bound the two terms $4|\vartheta|_\infty+2|\tilde\Omega|_\infty| \le M\varepsilon^2+4M\varepsilon^2$ by (120). We need to estimate $|\langle\omega,k\rangle+|\mathtt{r}(m)|^2-|\mathtt{r}(n)|^2|$, by momentum conservation $\mathtt{r}(n)=\mathtt{r}(m)+\pi(k)$. First note that, setting $v:=m-\mathtt{r}(m)$ the type of m, we have:

$$\begin{aligned}\langle\omega,k\rangle+|\mathtt{r}(m)|^2-|\mathtt{r}(n)|^2 &= \langle\omega,k\rangle-|\pi(k)|^2-2\langle\pi(k),\mathtt{r}(m)\rangle = \\ &= \langle\omega,k\rangle-|\pi(k)|^2+2\langle\pi(k),v\rangle-2\langle\pi(k),m\rangle.\end{aligned} \tag{146}$$

Note that $\pi(k)\in B_K\cup\{0\}$. Let $m\xrightarrow{K}[v_i;p_i]$, we distinguish two cases: $p_1\ge \mathtt{C}K^{4d\tau_0}$ or $p_1<\mathtt{C}K^{4d\tau_0}$.

Case 1: $p_1\ge\mathtt{C}K^{4d\tau_0}$**.**

If $p_1\ge \mathtt{C}K^{4d\tau_0}$ then m has a cut at $\ell=0$. By Lemma 7.32 we have that m is on the open stratum and $\mathtt{r}(m)=m, v=0$. If $\pi(k)=0$ we have $m=n$ and the denominator is covered by the bound (137) with $h=0$. If $\pi(k)\ne 0$ then by definition $|\langle\pi(k),\mathtt{r}(m)\rangle| = |\langle\pi(k),m\rangle| \ge p_1$. $(A1)$ implies $|\omega|_\infty<2\kappa$ so:

$$|\langle\omega,k\rangle+\Omega_m-\Omega_n| \ge |\langle\omega,k\rangle+|\mathtt{r}(m)|^2-|\mathtt{r}(n)|^2| - 4|\vartheta|_\infty-2|\tilde\Omega|_\infty >$$

$$|\langle\omega,k\rangle-|\pi(k)|^2-2\langle\pi(k),m\rangle|-5M\varepsilon^2>2\mathtt{C}K^{4d\tau_0}-2\kappa K-\kappa^2K^2-5M\varepsilon^2>1.$$

Case 2: $p_1<\mathtt{C}K^{4d\tau_0}$**.** By hypothesis the point m has $|\mathtt{r}(m)|>\mathtt{C}K^{\tau_1}$ and $p_1<\mathtt{C}K^{4d\tau_0}$, thus by Proposition 7.12 m belongs to some A^g where $A\overset{K}{\to}[v_i;p_i]_\ell,\ 1\le\ell<d$. Write $m=\mathtt{r}(m)+v$ and $n=\mathtt{r}(n)+u=\mathtt{r}(m)+\pi(k)+u$ for two types $u,v\in\mathcal{Z}$.

Let us first notice that (133) with $l=e_m-e_n$ is surely satisfied if $|(\pi(k),\mathtt{r}(m))|\ge K^3$ because in that case the absolute value of (146) is greater than $2K^3-\kappa^2K^2-|\omega|K-8dK>K^3$ by assumption (117) ($K>N_0$) and since $(A1)$ implies $|\omega|_\infty<2\kappa$.

If on the other hand $|(\pi(k),\mathtt{r}(m))|<K^3$, then $\pi(k)\in B_K\cup\{0\}$ is in $\langle v_i\rangle_\ell$. In fact otherwise we would have $|(\pi(k),m)|>\mathtt{c}K^{4d\tau_0}$ by definition of A^g, hence $|(\pi(k),\mathtt{r}(m))|>\mathtt{c}K^{4d\tau_0}-2d\kappa^2K>K^3$ by Formula (117) and $(A1)$, a contradiction.

In A^g we have chosen a point m_A, to the points m,m_A we can apply Lemma 7.20, thus they have a cut at ℓ for parameters $(K,\theta,\mu,\tau(p_\ell))$ where θ,μ are only restricted to be allowable. Then they satisfy the hypotheses of Theorem 5 hence we have $m_A=\mathtt{r}(m_A)+v$.

Consider then $n=m+\pi(k)-v+u$ and let $n\overset{K}{\to}[w_i;q_i]$. We have $|m-n|=|\pi(k)-v+u|\le\kappa(K+2d)\le 2\kappa K$. We now can impose that the allowable θ,μ satisfy the constraints given by Formula (117) hence we have the inequality (57) for n in place of r hence, by Lemma 7.25, n has an ℓ cut $[w_i;q_i]_\ell$ with parameters $N,\theta,\mu,\tau(p_\ell)$ and moreover $[w_i;q_i]_\ell=A+\pi(k)-v+u$. The same argument shows that, setting $\bar n:=m_A+\pi(k)-v+u=m_A+n-m$, both n and $\bar n$ have a cut with the same parameters and the same associated subspace $[w_i;q_i]_\ell$.

We thus can apply again Theorem 5 and see that $\mathtt{r}(\bar n)=\mathtt{r}(m_A)+\pi(k)$. By $(A3)$ we know that $\Pi_{K,\theta,\mu,\tau}\sum_a\vartheta_a|z_a|^2$ is Töplitz, for the chosen parameters $K,\theta,\mu,\tau(p_\ell)$, hence we deduce that $\vartheta_m=\vartheta_{m_A}$ and $\vartheta_n=\vartheta_{\bar n}$. We deduce that if $\pi(k)\in\langle v_i\rangle_\ell$:

$$|\mathtt{r}(m)|^2-|\mathtt{r}(n)|^2-|\mathtt{r}(m_A)|^2+|\mathtt{r}(\bar n)|^2=-|\pi(k)|^2-2\langle\pi(k),\mathtt{r}(m)\rangle+|\pi(k)|^2+2\langle\pi(k),\mathtt{r}(m_A)\rangle=0$$

Finally since $\vartheta_m=\vartheta_{m_A}$ and $\vartheta_n=\vartheta_{\bar n}$ we have:

$$|\Omega_m-\Omega_n-\Omega_{m_A}+\Omega_{\bar n}|=|\tilde\Omega_m-\tilde\Omega_n-\tilde\Omega_{m_A}+\tilde\Omega_{\bar n}|\implies$$

$$|\langle\omega,k\rangle+\Omega_m-\Omega_n|\ge|\langle\omega,k\rangle+\Omega_{m_A}-\Omega_{\bar n}|-|\tilde\Omega_m-\tilde\Omega_n-\tilde\Omega_{m_A}+\tilde\Omega_{\bar n}|. \tag{147}$$

By $(A4)$ we also know that $\tilde\Omega(z):=\sum_b\tilde\Omega_b|z_b|^2$ is quasi-Töplitz with parameters K,θ,μ which satisfy (117) hence, we may apply Lemma 8.20 with $Q(z)=\tilde\Omega(z)$. Lemma 7.20 ensures that, for any allowable θ,μ, all $m\in A^g$ satisfy the conditions needed to obtain formula (97) with $N=K,\tau=\tau(p_\ell)$, and also the estimate (99) (with $\tau=\tau(p_\ell)$):

$$|\tilde\Omega_m-\tilde\Omega_{m_A}|<2\|\tilde\Omega\|^T_{\vec p}K^{-4d\tau(p_\ell)}.$$

Similarly we have

$$|\tilde\Omega_n-\tilde\Omega_{\bar n}|<2\|\tilde\Omega\|^T_{\vec p}K^{-4d\tau(p_\ell)}.$$

In conclusion when $\pi(k)\in\langle v_i\rangle_\ell$ we have

$$|\Omega_m-\Omega_n-\Omega_{m_A}+\Omega_{\bar n}|<4\|\tilde\Omega\|^T_{\vec p}K^{-4d\tau(p_\ell)},$$

where by definition $K^{\tau(p)}=\max(K^{\tau_0},\mathtt{c}^{-1}p)$. We now apply the constraint (141) and hence:

$$\begin{aligned}|\langle\omega,k\rangle+\Omega_m-\Omega_n|&\ge|\langle\omega,k\rangle+\Omega_{m_A}-\Omega_{\bar n}|-4\|\tilde\Omega\|^T_{\vec p}K^{-4d\tau(p_\ell)}\ge\\ &2\gamma\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d})-4\|\tilde\Omega\|^T_{\vec p}\min(K^{-4d\tau_0},\mathtt{c}^{4d}p_\ell^{-4d}).\end{aligned} \tag{148}$$

By (122), $\|\tilde\Omega\|^T_{\vec p}<\gamma$, and clearly $4\min(K^{-4d\tau_0},\mathtt{c}^{4d}p_\ell^{-4d})<\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d})$. Hence

$$|\langle\omega,k\rangle+\Omega_m-\Omega_n|\geq\gamma\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d}),$$

so in order to we get the desired inequality we need to show that $\min(K^{-2d\tau_0},\mathtt{c}^{2d}p_\ell^{-2d})\geq K^{-2d\tau_1}$ i.e. that $\mathtt{c}^{2d}p_\ell^{-2d}\geq K^{-2d\tau_1}$. Since $p_\ell\leq \mathtt{C}K^{\tau_1/4d}\implies \mathtt{c}^{2d}p_\ell^{-2d}\geq(\mathtt{c}\mathtt{C}^{-1})^{2d}K^{-\tau_1/2}$ and (cf. (50)) $K\geq N_0>\mathtt{C}\mathtt{c}^{-1}$ implies $(\mathtt{c}\mathtt{C}^{-1})^{2d}K^{-\tau_1/2}\geq K^{-\tau_1/2-2d}\geq K^{-2d\tau_1}$.

□

Remark 10.14. This Proposition essentially says that, by imposing ***only one*** non resonant condition (141), we impose **all** the conditions (133) with $l=e_m-e_n$ such that $m\in[v_i;p_i]^g_j$ and $n=m+\pi(k)$.

10.15. **The Töplitz property for the generating function F.** The function F has been obtained by solving the homological equation for a hamiltonian H compatible with the parameters (K,θ,μ) and given in Formula (132). Recall we are using parameters $\vec p=(s,r,K,\theta,\mu,\lambda=\gamma^{-1}M,\mathcal{O})$. We now prove:

Proposition 10.16. *For $\xi\in\mathcal{O}_+$ the solution of the homological equation F is quasi-Töplitz for parameters (K,θ,μ), moreover one has the bound (cf.* (121)*):*

$$\|X_{F^{(i)}}\|^T_{\vec p'}\leq\gamma^{-1}\mathtt{K}\|X_{P^{(i)}}\|^T_{\vec p}=\mathtt{K}\epsilon^{(i)},\quad \mathtt{K}=5K^{4d\tau_1+1}\,, \tag{149}$$

where $\vec p'=(s,r,K,\theta,\mu,\lambda=\gamma^{-1}M,\mathcal{O}_+)$.

Proof. We have given in Formula (134) a better bound on the norm $\|X_{F^{(i)}}\|_{s,r}$ hence in order to prove our statement we only need to consider the quasi-Töplitz norm $\|X_{F^{(i)}}\|^{(K,\theta,\mu)}_{s,r}$ and the Lipschitz norm.

The quasi–Töplitz property is a condition for $N\geq K$, on the (N,θ,μ,τ)–bilinear part of $F^{(i)}$. Hence if $i=0,1$ the Töplitz norm coincides with the usual majorant norm and (149) follows from the bounds (134).

We are reduced to proving our statement on the quadratic terms:

$$\Pi_{(N,\theta,\mu,\tau)}F^{(2)}=\sum_{\substack{|k|\leq N\,,\\ \min(|\mathtt{r}(n)|,|\mathtt{r}(m)|)>\theta N^{\tau_1},\ m,n\in\underline{p}-cut}}F_{k,m,n}e^{\mathrm{i}(k,x)}z_m\bar z_n+B_{k,m,n}e^{\mathrm{i}(k,x)}z_mz_n+C_{k,m,n}e^{\mathrm{i}(k,x)}\bar z_m\bar z_n$$

with

$$F_{k,m,n}=\frac{P_{k,0,e_m,e_n}}{\langle k,\omega\rangle+\Omega_m-\Omega_n}\,,\ B_{k,m,n}=\frac{P_{k,0,e_m+e_n,0}}{\langle\omega,k\rangle+\Omega_m+\Omega_n}\,,\ C_{k,m,n}=\frac{P_{k,0,0,e_m+e_n}}{\langle\omega,k\rangle-\Omega_m-\Omega_n}. \tag{150}$$

By hypothesis $\min(|\mathtt{r}(m)|,|\mathtt{r}(n)|)>\theta N^{\tau_1}$ so in the case of $B_{k,m,n}$ one has

$$|B_{k,m,n}|=\frac{|P_{k,0,e_m+e_n,0}|}{|\langle k,\omega\rangle+|\mathtt{r}(m)|^2+|\mathtt{r}(n)|^2+2\vartheta_m+2\vartheta_n+\tilde\Omega_m+\tilde\Omega_n|}\leq\mathtt{c}^{-1}|P_{k,0,e_m+e_n,0}|N^{-\tau_1},$$

since

$$|\langle k,\omega\rangle+|\mathtt{r}(m)|^2+|\mathtt{r}(n)|^2+2\vartheta_m+2\vartheta_n+\tilde\Omega_m+\tilde\Omega_n|>2\mathtt{c}N^{\tau_1}-|\omega|N-4|\vartheta|_\infty+2|\tilde\Omega|_\infty>\mathtt{c}N^{\tau_1}.$$

Since $N^{4d\tau-\tau_1}<1$ this means that $\sum_{k,m,n}B_{k,m,n}e^{\mathrm{i}(k,x)}z_mz_n$ is quasi-Töplitz, and we may take the "Töplitz approximation" equal to zero (cf. Remark 8.17). Since $\mathtt{c}K^{2d\tau_1}>1>\gamma$ the final estimate follows by formula (134)

$$\|\sum_{k,m,n}B_{k,m,n}e^{\mathrm{i}(k,x)}z_mz_n\|^{K,\theta,\mu,\tau}_{s,r}\leq\max(K^{2d\tau_1}\gamma^{-1},\mathtt{c}^{-1})\|X_{P^{(i)}}\|^\lambda_{s,r}=K^{2d\tau_1}\gamma^{-1}\|X_{P^{(i)}}\|^\lambda_{s,r}.$$

Same argument for $\sum_{k,m,n} C_{k,m,n} e^{\mathrm{i}(k,x)} \bar{z}_m \bar{z}_n$.

We thus have to study $\sum_{k,m,n} F_{k,m,n} e^{\mathrm{i}(k,x)} z_m \bar{z}_n$. Take $N \geq K$, denote by $\underline{p} := N, \theta, \mu, \tau$, we wish to decompose

$$F_{k,m,n} = \mathcal{F}_k(m-n, [v_i; p_i]_\ell) + N^{-4d\tau} \bar{F}_{k,m,n}, \tag{151}$$

so that $\mathcal{F}_k$ is the k Fourier coefficient of a Töplitz approximation $\mathcal{F} \in \mathcal{T}_{\underline{p}}$.

By momentum conservation we have $\pi(k) + \mathtt{r}(m) - \mathtt{r}(n) = 0$ hence

$$|\mathtt{r}(m)|^2 - |\mathtt{r}(n)|^2 = -|\pi(k)|^2 - 2\langle \pi(k), \mathtt{r}(m)\rangle. \tag{152}$$

For the denominator in the first term of (150) we have

$$\langle k, \omega\rangle + \Omega_m - \Omega_n = \langle k, \omega\rangle + |\mathtt{r}(m)|^2 - |\mathtt{r}(n)|^2 + 2\vartheta_m - 2\vartheta_n + \tilde{\Omega}_m - \tilde{\Omega}_n$$

$$= \langle k, \omega\rangle - |\pi(k)|^2 - 2\langle \pi(k), \mathtt{r}(m)\rangle + 2\vartheta_m - 2\vartheta_n + \tilde{\Omega}_m - \tilde{\Omega}_n. \tag{153}$$

If K is sufficiently large we can estimate $|\langle k, \omega\rangle - |\pi(k)|^2 + 2\vartheta_m - 2\vartheta_n + \tilde{\Omega}_m - \tilde{\Omega}_n| < K^3$. From this we see that if $2|(\pi(k), \mathtt{r}(m))| > \mathtt{c}N^{4d\tau}$ we may again set $\mathcal{F}_k = 0$.

So we are reduced to the case in which $2|(\pi(k), \mathtt{r}(m))| \leq \mathtt{c}N^{4d\tau}$.

By assumption m, n have a cut at ℓ with parameters (N, θ, μ, τ) and

$$|\mathtt{r}(m)|, |\mathtt{r}(n)| \geq \theta N^{\tau_1}, \quad m \xrightarrow{N} [v_i; p_i], \quad n \xrightarrow{N} [w_i; q_i],$$

$$q_\ell, p_\ell \leq \mu N^\tau, \quad q_{\ell+1}, p_{\ell+1} \geq \theta N^{4d\tau}, \quad A := [v_i; p_i]_\ell \prec B := [w_i; q_i]_\ell, \tag{154}$$

hence, by Corollary 7.23, $\langle v_1, \dots, v_\ell\rangle = \langle w_1, \dots, w_\ell\rangle$. We distinguish two cases:

Case 1: $\pi(k) \notin \langle v_i\rangle_\ell$**.** If $\pi(k) \notin \langle v_i\rangle_\ell$ then by the definitions of cut 7.15, and of optimal presentation we have $2|(\pi(k), \mathtt{r}(m))| > \mathtt{c}N^{4d\tau}$ contrary to our hypothesis.

Case 2: $\pi(k) \in \langle v_i\rangle_\ell$**.** We recall that ϑ_m (resp. ϑ_n) are constant on all the m which have the same affine space $A = [v_i; p_i]_\ell$ associated to its ℓ-cut. Moreover, setting $h = n - m$, we know that n has an ℓ-cut with associated affine space $B = A + h = [w_i; q_i]_\ell$. By lemma 8.20 and $\tilde{\Omega}$ has a Töplitz approximation, $\tilde{\mathit{\Omega}}$ see Formula (96). By Corollary 8.21 we can choose a point $m_A \in A^g_{\underline{p}}$ so that $m_A + h \in (A+h)^g_{\underline{p}}$ then we may choose the Töplitz approximant of order one with $\tilde{\mathit{\Omega}}(A) = \tilde{\Omega}_{m_A}$ and $\tilde{\mathit{\Omega}}(A+h) = \tilde{\Omega}_{m_A+h}$:

$$|\tilde{\Omega}_m - \tilde{\mathit{\Omega}}(A)|, |\tilde{\Omega}_n - \tilde{\mathit{\Omega}}(A+h)| < 2\|\tilde{\mathit{\Omega}}\|^T_{\vec{p}} N^{-4d\tau}. \tag{155}$$

Denote by $D_{k,m,n} = \langle k, \omega\rangle + \Omega_m - \Omega_n$ the denominator of the term $F_{k,m,n}$, we define

$$\mathfrak{D}_{k,h,A} := D_{k,m_A,m_A+h}, \implies |\mathfrak{D}_{k,h,A}| \geq \gamma K^{-2d\tau_1}, \ (133). \tag{156}$$

Finally, since $P^{(2)}$ is quasi-Töplitz, we may set

$$\mathcal{F}_k(h, A) = \frac{\mathcal{P}^{(2)}_k(h, A)}{\mathfrak{D}_{k,h,A}}, \quad \bar{F} = N^{4d\tau}(F - \mathcal{F})$$

where $\mathcal{P} = \mathcal{P}^{(2)} \in \mathcal{T}_{\underline{p}}$ is a piecewise–Töplitz approximation of $P_{k,0,e_m,e_n}$ so that for $\bar{P}^{(2)} = N^{4d\tau}(P^{(2)} - \mathcal{P}^{(2)})$ we have the bounds $\|\mathcal{P}^{(2)}\|_{r,s}, \|\bar{P}^{(2)}\|_{r,s} \leq \|P^{(2)}\|^T_{\vec{p}} + \epsilon$ where $\epsilon > 0$ can be taken arbitrarily small (see Lemma 8.20).

We notice that by (155)

$$N^{4d\tau}|\mathfrak{D}_{k,h,A} - D_{k,m,n}| = N^{4d\tau}|\tilde{\Omega}_m - \tilde{\mathit{\Omega}}(A) - \tilde{\Omega}_n + \tilde{\mathit{\Omega}}(A+h)| < 4\|\tilde{\mathit{\Omega}}\|^T_{\vec{p}}. \tag{157}$$

If the denominators are bounded away from zero then (cf. (151)):

$$\bar F_{k,m,n} = N^{4d\tau}(F_{k,n,m} - \mathcal{F}_k(h,A)) = \frac{\bar P^{(2)}_{k,m,n}}{D_{k,m,n}} + \mathcal{P}^{(2)}_k(h,A)\frac{N^{4d\tau}(\mathfrak{D}_{k,h,A} - D_{k,m,n})}{\mathfrak{D}_{k,h,A}D_{k,m,n}},$$

is bounded.

Summing over the indexes k,m,n such that m has a cut with parameters (N,θ,μ), we obtain

$$\|\mathcal{F}\|_{s,r} \le \frac{\|\mathcal{P}^{(2)}\|_{s,r}}{\inf_{k,n,m}\mathfrak{D}_{k,h,A}}, \quad \|\bar F\|_{s,r} \le \frac{\|\bar P^{(2)}\|_{s,r}}{\inf_{k,n,m}|\mathfrak{D}_{k,h,A}|} + \frac{4\|\tilde\Omega\|^T_{s,r}\|\mathcal{P}^{(2)}\|_{s,r}}{\inf_{k,n,m}|D_{k,m,n}\mathfrak{D}_{k,h,A}|}$$

This we may rephrase as

$$\|F^{(2)}\|^T_{\vec p'} \le \|P^{(2)}\|^T_{\vec p} \sup_{\xi\in\mathcal{O}_+}\sup_{k,n,m:|k|<K}\left(\frac{1}{|\mathfrak{D}_{k,m,n}|} + \frac{4\|\tilde\Omega\|^T_{\vec p}}{|\mathfrak{D}_{k,m,n}D_{k,m,n}|} + \lambda\frac{M(|k|+2)}{D^2_{k,m,n}}\right). \tag{158}$$

By the *Smallness condition*, $(A4)$ formula (122), we have $\|\tilde\Omega\|^T_{\vec p} \le \gamma$. The denominators $|\mathfrak{D}_{k,m,n}|, |D_{k,m,n}|$ are $> \gamma K^{-2d\tau_1}$ uniformly in $\mathcal{O}_+$ by Formulas (133) and (156). Recalling that $\lambda = \gamma M^{-1}$, we deduce that

$$\|F^{(2)}\|^T_{\vec p'} \le \|P^{(2)}\|^T_{\vec p}\gamma^{-1}(K^{2d\tau_1} + K^{4d\tau_1+1} + 4K^{4d\tau_1}) \le 5\|P^{(2)}\|^T_{\vec p}\gamma^{-1}K^{4d\tau_1+1}.$$

$\square$

10.17. **The new Hamiltonian H^+.** Recall we have set $\vec p = (s,r,K,\theta,\mu,\lambda,\mathcal{O})$, $\vec p_+ = (s_+,r_+,K,\theta_+,\mu_+,\lambda_+,\mathcal{O}_+)$. By Propositions 9.2 and 10.16, F defines a M-analytic symplectic quasi-Toplitz change of variables from $D(s_+,r_+)$ to $D(s,r)$, where r_+,s_+ are determined by (136).

The change of variables is of the form $\Phi = I + \Psi$ with the bounds (cfr. (105))

$$\|\Psi\|^T_{\vec p_+} \le 2\|X_F\|^T_{\vec p} \tag{159}$$

and for any function $f \in \mathcal{T}_{\underline{p}}$ we have that $e^{ad(F)}f \in \mathcal{T}_{\vec p_+}$ where $\vec p_+$ satisfy (103).

We now analyze $H^+ := e^{ad(F)}(H)$, recall that by definition $ad(F)(\mathcal{N}) = -P^{\le 2}_{\le K} + [P^{\le 2}]$.

$$H^+ := e^{ad(F)}(\mathcal{N}+P) = H + ad(F)H + \sum_{j\ge 2}\frac{ad(F)^j}{j!}(H) = \tag{160}$$

$$= \mathcal{N} + P - P^{\le 2}_{\le K} + [P^{\le 2}] + ad(F)P + \sum_{j\ge 2}\frac{ad(F)^j}{j!}(H)$$

We call $\mathcal{N} + [P^{\le 2}] := \mathcal{N}^+$ and the rest of the Hamiltonian P_+, so that

$$P_+ := (P - P^{\le 2}_{\le K}) + \{F,P\} + \sum_{j\ge 2}\frac{ad(F)^j}{j!}P + \sum_{j\ge 2}\frac{ad(F)^{j-1}}{j!}(-P^{\le 2}_{\le K} + [P^{\le 2}_{\le K}]). \tag{161}$$

By formula (131),

$$\omega^+ := \omega + P^{1,0}_0, \quad \tilde\Omega^+_n := \tilde\Omega_n + P^{0,2}_{0,n,+,n,-},$$

$$\Omega^+_n = \sigma(n)|\mathtt{r}(n)|^2 + 2\vartheta_n + \tilde\Omega^+_n = \Omega_n + P^{0,2}_{0,n,+,n,-}, \quad \mathcal{C}^+ = \mathcal{C} + \mathcal{P}. \tag{162}$$

Recall that $\mathcal{C}$ is the finite complex part of the normal form, same for $\mathcal{P}$ as defined in (131). We need to

- Prove that H^+ satisfies conditions $(A1)-(A5)$.
- Estimate all the new parameters $\vec p_+$.

10.17.1. *Lipschitz estimates* M_+, L_+, λ_+. We have

$$|\omega^+|^{lip} + |\Omega_+|^{lip} = |\omega + P_0^{1,0}|^{lip} + \sup_n |\Omega_n + P^{0,2}_{0,n,+,n,-}|^{lip} \le$$

$$|\omega|^{lip} + \sup_n |\Omega_n|^{lip} + |P_0^{1,0}|^{lip} + \sup_n |P^{0,2}_{0,n,+,n,-}|^{lip} \le M + |P_0^{1,0}|^{lip} + \sup_n |P^{0,2}_{0,n,+,n,-}|^{lip},$$

by definition of M.

In the same way

$$(\omega^+)^{-1} = (\omega + P_0^{1,0})^{-1} = (\omega \circ (Id + P_0^{1,0} \circ \omega^{-1}))^{-1} = (Id + P_0^{1,0} \circ \omega^{-1})^{-1} \circ \omega^{-1},$$

so that ω^+ is invertible as a Lipschitz function provided that $L|P_0^{1,0}|^{lip} < 1$ with the bound

$$|(\omega^+)^{-1}|^{lip} \le \frac{L}{1 - L|P_0^{1,0}|^{lip}}. \tag{163}$$

In order to estimate a_+, defined in Formula (124), we note

$$|\Delta_{\xi,\varrho}(\langle \omega^+(\xi), k\rangle + (\Omega_+(\xi), l))| \tag{164}$$

$$\ge |\Delta_{\xi,\varrho}(\langle \omega(\xi), k\rangle + (\Omega(\xi), l))| - |\Delta_{\xi,\varrho}(\langle P_0^{1,0}(\xi), k\rangle + (P^{0,2}_{0,n,+,n,-}, l))|$$

$$\ge a - S_0 |P_0^{1,0}(\xi)|^{lip} + 2|P^{0,2}_{0,n,+,n,-}|^{lip}$$

We recall that $\|\cdot\|^\lambda = \|\cdot\| + \lambda \|\cdot\|^{lip}$, hence

$$|P_0^{1,0}|^{lip}, |P^{0,2}_{0,n,+,n,-}|^{lip} \le \lambda^{-1} \|P^{(2)}\|^\lambda_{s,r} \le M|\vec\epsilon|$$

since $\lambda = \gamma M^{-1}$. We define

$$M_+ := M(1 + 2|\vec\epsilon|)\,, \quad L_+ := L(1 - ML|\vec\epsilon|)^{-1}, \quad a_+ := a - (S_0 + 2)M|\vec\epsilon| \tag{165}$$

notice that L_+ is well defined since by (A4) $LM|\vec\epsilon| < 1$. By construction

$$|\omega^+|^{lip} + |\Omega_+|^{lip} \le M_+\,, \quad |(\omega^+)^{-1}|^{lip} \le L_+\,,$$

finally we have:

$$|\omega^+ - \mathtt{v}| \le |\omega - \mathtt{v}| + |P_0^{1,0}| \le M\varepsilon^2 + \gamma|\vec\epsilon| \le M_+\varepsilon^2,$$

since $\gamma < 2M\varepsilon^2$. We finally set $\lambda_+ := \gamma(M_+)^{-1}$ we have $\lambda_+ < \lambda$ since $M_+ > M$.

10.17.2. *Töplitz estimates* $\vec\epsilon_+, \Theta_+$. We wish to show P_+ is quasi–Töplitz and bound

$$\epsilon_+^{(h)} := \gamma^{-1} \|X_{P_+^h}\|^T_{\vec p_+}\,, \quad \text{for } h = 0, 1, 2\,; \quad \Theta_+ := \gamma^{-1} \|X_{P_+}\|^T_{\vec p_+}.$$

We have that (cf. (161)) $P_+ = (P - P^{\le 2}_{\le K}) + A + B$ where

$$A := \sum_{j \ge 2} \frac{ad(F)^{j-1}}{j!}(-P^{\le 2}_{\le K} + [P^{\le 2}_{\le K}]), \quad B = \{F, P\} + \sum_{j \ge 2} \frac{ad(F)^j}{j!} P,$$

We argue as in Proposition 5 of [24] or in Proposition 9.2. By Formula (149) $\|X_F\|^T_{\vec p} \le 5K^{4d\tau_1}|\vec\epsilon|$, the hypothesis (136) implies that we have the conditions of (102), that is $2^{2n+14}\delta^{-1}\|X_F\|^T_{\vec p} < 1/2$, where $\delta = \min(1 - \frac{s_+}{s}, 1 - \frac{r_+}{r})$. Therefore we have that

$$\|X_A\|^T_{\vec p_+} \le 2\delta^{-1} \|X_F\|^T_{\vec p} \|X_{P^{\le 2}}\|^T_{\vec p}, \quad \|X_B\|^T_{\vec p_+} \le 2\delta^{-1} \|X_F\|^T_{\vec p} \|X_P\|^T_{\vec p} \tag{166}$$

Using (149) we rewrite (166) as

$$\gamma^{-1}\|X_A\|^T_{\vec p_+} \lessdot \delta^{-1} \mathtt{K} |\epsilon|^2\,, \quad \gamma^{-1}\|X_B\|^T_{\vec p_+} \lessdot \delta^{-1} \mathtt{K} |\epsilon| \Theta\,, \tag{167}$$

We obtain:

$$\|X_{P_+}\|^T_{\vec{p}_+} \le \|X_P\|^T_{\vec{p}} + 4\delta^{-1}\|X_F\|^T_{\vec{p}}\|X_P\|^T_{\vec{p}} \quad i.e. \quad \Theta_+ - \Theta \lessdot \delta^{-1}\mathtt{K}|\vec{\epsilon}|\Theta\,. \tag{168}$$

Let us now compute the terms of order ≤ 2 in P^+, we have:

$$P_+^{(h)} = P_{>K}^{(h)} + (A + \{F, P^{>2}\} + \{F, P^{\le 2}\} + \sum_{j\ge 2}\frac{ad(F)^j}{j!}P)^{(h)}.$$

Again from (149), denoting $E = \sum_{j\ge 2}\frac{ad(F)^j}{j!}P$, we have the bounds:

$$\|X_{\{F,P^{\le 2}\}}\|^T_{\vec{p}_+} \le 2\delta^{-1}\|X_F\|^T_{\vec{p}}\ \|X_{P^{\le 2}}\|^T_{\vec{p}} \lessdot \delta^{-1}\mathtt{K}\gamma|\vec{\epsilon}|^2,$$

$$\|X_E\|^T_{s_+,r_+} \lessdot \delta^{-2}(\|X_F\|^T_{\vec{p}})^2\|X_P\|^T_{\vec{p}} \lessdot \delta^{-2}\mathtt{K}^2|\vec{\epsilon}|^2\Theta$$

The contributions from $\{F, P^{>2}\}$ are

$$\Pi_0\{F, P^{>2}\} = 0\,, \quad \Pi_1\{F, P^{>2}\} = \{F^{(0)}, P^{(3)}\}\,,$$

$$\Pi_2\{F, P^{>2}\} = \{F^0, P^{(4)}\} + \{F^{(1)}, P^{(3)}\}\,,$$

so applying the Cauchy estimates we have, setting $\mathfrak{z} := \text{const.}(\delta^{-1}\mathtt{K})^2$:

$$\begin{aligned}
\epsilon_+^{(0)} &\le \mathfrak{z}|\vec{\epsilon}|^2(1+\Theta) + 2\epsilon^{(0)}e^{-(s-s_+)K}\\
\epsilon_+^{(1)} &\le \mathfrak{z}\big(\Theta\,\epsilon^{(0)} + |\vec{\epsilon}|^2(1+\Theta)\big) + 2\epsilon^{(1)}e^{-(s-s_+)K}\\
\epsilon_+^{(2)} &\le \mathfrak{z}\big(\Theta(\epsilon^{(0)}+\epsilon^{(1)}) + |\vec{\epsilon}|^2(1+\Theta)\big) + 2\epsilon^{(2)}e^{-(s-s_+)K}.
\end{aligned}$$

Note that the terms $2\epsilon^{(h)}e^{-(s-s_+)K}$ come from $P_{>K}^{(h)}$ via the smoothing estimates 94.

We write in matrix form, denoting by $\vec{\epsilon}$ the three dimensional column vector of coordinates $(\epsilon^{(0)}, \epsilon^{(1)}, \epsilon^{(2)})$ and $\underline{1} := (1,1,1)$ we have

$$\begin{array}{rcl}
\vec{\epsilon}_+ & \le & \mathfrak{z}\big(\Theta\mathtt{L}\vec{\epsilon} + |\vec{\epsilon}|^2(1+\Theta)\underline{1}\big) + 2e^{-(s-s_+)K}\vec{\epsilon}\\
\Theta_+ & \le & \Theta + \mathfrak{z}\Theta|\vec{\epsilon}|
\end{array} \tag{169}$$

where the matrix $\mathtt{L}$ is

$$\mathtt{L} = \begin{pmatrix}0&0&0\\1&0&0\\1&1&0\end{pmatrix} \implies \mathtt{L}^2 = \begin{pmatrix}0&0&0\\0&0&0\\2&0&0\end{pmatrix}, \quad \mathtt{L}^3 = 0 \tag{170}$$

and the vector inequality means that the inequality is true for all three coordinates.

10.18. **Iteration.**

10.18.1. *A useful inequality.* Now we need to be able to handle in a recursive way the inequalities obtained so far, we start with a formal inequality which is a variation of Lemma 5.8 of [2].

We *fix* χ such that

$$1 < \chi < 2^{\frac{1}{3}}\,. \tag{171}$$

and $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ three positive numbers satisfying:

$$\mathfrak{a}, \mathfrak{b}, \mathfrak{c} > 1\,, \quad 12\mathfrak{b}^2 \le \min_{i\in\mathbb{N}}\frac{e^{\mathfrak{a}2^{i-2}+\chi^{i-2}-\chi^{i+1}}}{\mathfrak{c}^{2i-1}}\,, \tag{172}$$

and set

$$\mathfrak{i} := \min_{i\in\mathbb{N}}\Big(\frac{e^{\chi^i}}{2\mathfrak{b}(2\mathfrak{c})^i}, \frac{e^{(2-\chi^3)\chi^{i-2}}}{32\mathfrak{b}^3\mathfrak{c}^{3i-3}}\Big) > 0. \tag{173}$$

Lemma 10.19. *For $j \in \mathbb{N}$ consider a sequence $(\vec{\epsilon}_j, \Theta_j)$ with $\vec{\epsilon}_j := (\epsilon_j^{(0)}, \epsilon_j^{(1)}, \epsilon_j^{(2)})$ a vector and Θ_j a number, all with positive components. Set $|\vec{\epsilon}_j| := \epsilon_j^{(0)} + \epsilon_j^{(1)} + \epsilon_j^{(2)}$.*

Suppose that for ($\mathtt{L}$ as in (170)) we have:

$$\begin{cases} \vec{\epsilon}_{j+1} & \leq & \mathfrak{b}\mathfrak{c}^j(\Theta_j \mathtt{L}\vec{\epsilon}_j + |\vec{\epsilon}_j|^2(1+\Theta_j)\underline{1}) + e^{-\mathfrak{a}2^j}\vec{\epsilon}_j \\ \Theta_{j+1} & \leq & \Theta_j + \mathfrak{b}\mathfrak{c}^j\Theta_j|\vec{\epsilon}_j|. \end{cases} \tag{174}$$

There exist $\mathtt{C}_0 := \mathtt{C}_0(\mathfrak{c}, \chi, \mathfrak{a}, \mathfrak{b}) > 1$ such that for all $\bar{\Theta} > 0$ satisfying

$$2\bar{\Theta}\mathtt{C}_0 < \min(1/3, \mathfrak{i}) \tag{175}$$

we have that

$$1/3|\vec{\epsilon}_0|, \Theta_0 < \bar{\Theta} \implies |\vec{\epsilon}_j| \leq \mathtt{C}_0|\vec{\epsilon}_0|e^{-\chi^j}\,, \quad \Theta_j \leq \Theta_0(1 + \mathtt{C}_0\sum_{0<l\leq j} 2^{-l})\,, \quad \forall\, j \geq 0. \tag{176}$$

Let i_0 be the value of i for which the minimum $\mathfrak{i}$ in (173) is achieved. One easily sees that, since $|\vec{\epsilon}_0|, \Theta_0 \leq 3\bar{\Theta} < 1$ we can find a value $\mathtt{C}_0$ (depending only on $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \chi$) for which both relations (176) hold for all $i \leq i_0$ and $|\vec{\epsilon}_0| < 1$.

We now work by induction and suppose that both relations hold up to some $i \geq i_0$. Then $\Theta_i \leq \Theta_0(1 + \mathtt{C}_0\sum_{1\leq l\leq i} 2^{-l}) < 2\Theta_0\mathtt{C}_0$ and, assuming $2\Theta_0\mathtt{C}_0 \leq 2\bar{\Theta}\mathtt{C}_0 \leq \mathfrak{i}$, we have $\mathfrak{b}\mathfrak{c}^i 2^{i+1}e^{-\chi^i} \leq \mathfrak{i}^{-1} \leq (2\Theta_0\mathtt{C}_0)^{-1}$ estimate :

$$\Theta_{i+1} \leq \Theta_0(1 + \mathtt{C}_0\sum_{l\leq i} 2^{-l} + \mathfrak{b}\mathfrak{c}^i 2\Theta_0 A_0^2 e^{-\chi^i}) \leq \Theta_0(1 + \mathtt{C}_0\sum_{1\leq l\leq i+1} 2^{-l}).$$

Notice that the constraint (175) and the inequality (176) imply $\Theta_j < 1$ for all j. We now substitute $\Theta_j < 1$ for all $j \leq i$ in the first relation and get

$$\vec{\epsilon}_{j+1} \leq \mathfrak{b}\mathfrak{c}^j(\mathtt{L}\vec{\epsilon}_j + 2|\vec{\epsilon}_j|^2\underline{1}) + e^{-\mathfrak{a}2^j}\vec{\epsilon}_j$$

we obtain a bound for $\vec{\epsilon}_{i+1}$ in terms of $\vec{\epsilon}_i$, $\vec{\epsilon}_{i-1}$ and $\vec{\epsilon}_{i-2}$.

We now assume by induction that the bounds in (176) are satisfied for all $j \leq i$ and then:

$$\vec{\epsilon}_{i+1} \leq \mathfrak{b}^2\mathfrak{c}^{2i-1}(\mathtt{L}^2\vec{\epsilon}_{i-1} + 2|\vec{\epsilon}_{i-1}|^2\mathtt{L}\underline{1}) + \mathfrak{b}\mathfrak{c}^i e^{-\mathfrak{a}2^{i-1}}\mathtt{L}\vec{\epsilon}_{i-1} + 2\mathfrak{b}\mathfrak{c}^i|\vec{\epsilon}_i|^2\underline{1} + e^{-\mathfrak{a}2^i}\vec{\epsilon}_i \leq$$

$$\vec{\epsilon}_{i+1} \leq \mathfrak{b}^2\mathfrak{c}^{2i-1}\mathtt{L}^2\vec{\epsilon}_{i-1} + 2\mathfrak{b}^2\mathfrak{c}^{2i-1}|\vec{\epsilon}_{i-1}|^2\mathtt{L}\underline{1} + \mathfrak{b}\mathfrak{c}^i e^{-\mathfrak{a}2^{i-1}}\mathtt{L}\vec{\epsilon}_{i-1} + 2\mathfrak{b}\mathfrak{c}^i|\vec{\epsilon}_i|^2\underline{1} + e^{-\mathfrak{a}2^i}\vec{\epsilon}_i \leq$$

$$2\mathfrak{b}^3\mathfrak{c}^{3i-3}|\vec{\epsilon}_{i-2}|^2\mathtt{L}^2\underline{1} + \mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-\mathfrak{a}2^{i-2}}\mathtt{L}^2\vec{\epsilon}_{i-2} + 2\mathfrak{b}^2\mathfrak{c}^{2i-1}|\vec{\epsilon}_{i-1}|^2\mathtt{L}\underline{1}$$

$$+\mathfrak{b}\mathfrak{c}^i e^{-\mathfrak{a}2^{i-1}}\mathtt{L}\vec{\epsilon}_{i-1} + 2\mathfrak{b}\mathfrak{c}^i|\vec{\epsilon}_i|^2\underline{1} + e^{-\mathfrak{a}2^i}\vec{\epsilon}_i.$$

This in turn implies, since $|\mathtt{L}| = 3, |\mathtt{L}^2| = 2$, that

$$|\vec{\epsilon}_{i+1}| \leq 4\mathfrak{b}^3\mathfrak{c}^{3i-3}|\vec{\epsilon}_{i-2}|^2 + 6\mathfrak{b}^2\mathfrak{c}^{2i-1}|\vec{\epsilon}_{i-1}|^2 + 6\mathfrak{b}\mathfrak{c}^i|\vec{\epsilon}_i|^2$$

$$+2\mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-\mathfrak{a}2^{i-2}}|\vec{\epsilon}_{i-2}| + 3\mathfrak{b}\mathfrak{c}^i e^{-\mathfrak{a}2^{i-1}}|\vec{\epsilon}_{i-1}| + e^{-\mathfrak{a}2^i}|\vec{\epsilon}_i| \leq$$

$$|\vec{\epsilon}_{i+1}| \leq |\vec{\epsilon}_0|^2{\mathtt{C}_0}^2(4\mathfrak{b}^3\mathfrak{c}^{3i-3}e^{-2\chi^{i-2}} + 6\mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-2\chi^{i-1}} + 6\mathfrak{b}\mathfrak{c}^i e^{-2\chi^i}) +$$

$$|\vec{\epsilon}_0|\mathtt{C}_0(2\mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-\mathfrak{a}2^{i-2}-\chi^{i-2}} + 3\mathfrak{b}\mathfrak{c}^i e^{-\mathfrak{a}2^{i-1}-\chi^{i-1}} + e^{-\mathfrak{a}2^i-\chi^i}) \leq$$

$$|\vec{\epsilon}_0|\mathtt{C}_0[16|\vec{\epsilon}_0|\mathtt{C}_0\mathfrak{b}^3\mathfrak{c}^{3i-3}e^{-2\chi^{i-2}} + 6\mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-\mathfrak{a}2^{i-2}-\chi^{i-2}}] \leq |\vec{\epsilon}_0|\mathtt{C}_0 e^{-\chi^{i+1}}$$

This is achieved provided that,

$$[16\bar{\Theta}\mathtt{C}_0\mathfrak{b}^3\mathfrak{c}^{3i-3}e^{-2\chi^{i-2}} + 6\mathfrak{b}^2\mathfrak{c}^{2i-1}e^{-\mathfrak{a}2^{i-2}-\chi^{i-2}}] \leq e^{-\chi^{i+1}}$$

and this in turn is valid if we assume the constraint (172).

We will apply this to

$$\mathfrak{c} := 4^{1+4d\tau_1}\,, \quad \mathfrak{b} := cost\, K^{8d\tau_1}\,, \quad \mathfrak{a} := \frac{Ks_0}{32}\,. \tag{177}$$

We note that, with this choice of parameters, condition (172) amounts to a largeness condition on K.

10.19.1. *Parameters in the iteration.* Let $H_0 = \mathcal{N}_0 + P_0 : D_0 \times \mathcal{O}_0 \to \mathbb{C}$ be as in Theorem 6. Define

$$\epsilon_0^{(h)} := \frac{\|X_{P_0^{(h)}}\|_{\vec p_0}^T}{\gamma} \le \Theta_0 := \frac{\|X_{P_0}\|_{\vec p_0}^T}{\gamma} = \frac{\|X_{P_0}\|_0}{\gamma}. \tag{178}$$

We have the estimate $e^{-1} \le \max_\nu e^{-\chi^\nu} 2^{\nu(1+4d\tau_1)} < \infty$ since, for any $p > 0$, we have that $\lim_{\nu\to\infty} 2^{p\,\nu} e^{-\chi^\nu} = 0$. We define:

$$C_\star = 2^{2n+10} e \mathtt{C}_0 M_0 a_0^{-1} \kappa K_0^{4d\tau_1} \max_\nu e^{-\chi^\nu} 2^{\nu(1+4d\tau_1)}\,, \tag{179}$$

here $\mathtt{C}_0$ is the constant of Lemma 10.19 with the choice of parameters (177), note that it depends only on $K_0, \kappa, d, n, \tau_1, \tau_0, s_0$ and $M_0 a_0^{-1}$. Recall that, as we have stated in Remark 10.4 $M_0 a_0^{-1} > 1$ is an ε independent constant of the problem.

We now fix $\bar\Theta := \bar\Theta(K_0, \kappa, d, n, \tau_1, \tau_0, s_0, a_0)$ in order to ensure the *smallness conditions*:

$$C_\star \bar\Theta < (12\, e)^{-1}, \quad \prod_\nu (1 + C_\star \bar\Theta e^{-\chi^{\nu-1}}) < \sqrt{2}. \tag{180}$$

together with the condition (175) of Lemma 10.19.

We now need to estimate all the parameters in the iteration. For the parameters which increase we exhibit bounds from above and for the ones which decrease bounds from below. Thus for $\nu \in \mathbb{N}$ we define

- $\delta_\nu := 2^{-\nu-3}\,, \quad r_{\nu+1} := (1-\delta_\nu) r_\nu\,, \quad s_{\nu+1} := (1-\delta_\nu) s_\nu\,, \quad D_\nu := D(s_\nu, r_\nu)\,,$
- $M_\nu := M_{\nu-1}(1 + C_\star |\vec\epsilon_0| e^{-\chi^{\nu-1}}) \le \sqrt{2} M_0$, $\lambda_\nu := \frac{\gamma}{M_\nu}$,
- $L_\nu := L_{\nu-1}(1 - C_\star |\vec\epsilon_0| e^{-\chi^{\nu-1}})^{-1} \le \sqrt{2} L_0$
- $K_\nu := 4^\nu K_0\,,\,, \quad \theta_\nu = \theta_0(1 + \sum_{j\le\nu} 2^{-j})\,, \quad \mu_\nu = \mu_0(1 - \sum_{j\le\nu} 2^{-j})\,,$
- $a_\nu := a_0(1 - C_\star |\vec\epsilon_0| \sum_{j\le\nu} 2^{-j}),$
- $\mathfrak{z}_\nu = \text{const.}\delta_\nu^{-2} \mathtt{K}_\nu^2 = \mathfrak{b}\mathfrak{c}^\nu.$

We have made our definitions so that $\min(1 - \frac{r_{\nu+1}}{r_\nu}, 1 - \frac{s_\nu}{s_{\nu+1}}) = \delta_\nu$. Note that $r_{\nu+1} \searrow r_0 \prod_{\nu=0}^\infty (1-\delta_\nu) > \frac{r_0}{2}\,, \quad s_{\nu+1} \searrow s_0 \prod_{\nu=0}^\infty (1-\delta_\nu) > \frac{s_0}{2}$, $\theta_\nu \nearrow 3\theta_0/2 < \mathtt{C} = 2\theta_0\,, \quad \mu_\nu \searrow 2\mu_0/3 > \mathtt{c} = \mu_0/2$.

For compactness of notation we will denote

$$\|\cdot\|_j := \|\cdot\|_{\vec p_j}^T\,, \quad \underline{p}_j := (r_j, s_j, \mathcal{O}_j, K_j, S_0 = 16\sqrt{n}, \theta_j, \mu_j, a_j, M_j, L_j, \mathtt{c}, \mathtt{C}), \tag{181}$$

where $\mathcal{O}_j$ is defined in the course of the proof of Lemma 10.20 .

Lemma 10.20. (Iterative Lemma) *Let $\mathtt{C}_0, C_\star, \bar\Theta$ be fixed as in formulas* (179) *and* (180). *Let Γ be as in Formula* (135) *and $B = 4\Gamma K_0^{-\tau_0+n+d/2}$. If for the Hamiltonian H_0 we can choose γ so that if for Θ_0 defined in* (178) *we have:*

$$\Theta_0 \le \bar\Theta\,, B\gamma\varepsilon^{2n-2}a_0^{-1} < |\mathcal{O}_0| \tag{182}$$

are satisfied, then we can construct recursively sets $\mathcal{O}_j \subset \mathcal{O}$ and a Hamiltonian $H_j = \mathcal{N}_j + P_j : D_j \times \mathcal{O}_j \to \mathbb{C}$, $\mathcal{N}_j := (\omega^{(j)}(\xi), y) + \sum_{k\in S^c}\Omega^{(j)}_k(\xi)|z_k|^2$ with $\Omega^{(j)}_n(\xi) = \sigma(n)(|\mathtt{r}(n)|^2 + 2\vartheta_n(\xi)) + \tilde\Omega^{(j)}_n(\xi)$. So that, if we define

$$\epsilon_j^{(h)} := \frac{\|X_{P_j^{(h)}}\|_j}{\gamma}\,, \vec\epsilon_j := (\epsilon_j^{(0)}, \epsilon_j^{(1)}, \epsilon_j^{(2)}), \quad \Theta_j = \frac{\|X_{P_j}\|_j}{\gamma}, \tag{183}$$

the following properties are satisfied for all j:

$(\mathbf{S1})_\mathbf{j}$ *For $j > 0$, $\mathcal{O}_j \subset \mathcal{O}_{j-1}$ is defined by* (137)-(141) *with $\omega \rightsquigarrow \omega^{(j-1)}$ and $\Omega_n \rightsquigarrow \Omega_n^{(j-1)}$. We have that $H_j = H_{j-1}\circ\Phi_j$ where $\Phi_j : D_j \times \mathcal{O}_j \to D_{j-1}$ is a Lipschitz family of real analytic symplectic maps of the form $\Phi_j = I + \Psi_j$ with $\|\Psi_j\|_{D_j}^{\lambda_j} < C_\star\bar\Theta 2^{-j}$.*

$(\mathbf{S2})_\mathbf{j}$ *The Hamiltonian H_j is compatible with the parameters $\underline{p}_j$ (Definition 10.2). The parameters $r_+ = r_{j+1}$, $s_+ = s_{j+1}$ satisfy the hypotheses* (136) *of the KAM step and the set $\mathcal{O}_{j+1} \subset \mathcal{O}_j$ satisfies Formula* (135)*, namely, using* (50)*:*

$$|\mathcal{O}_j \setminus \mathcal{O}_{j+1}| \le \Gamma\gamma a_j^{-1}\varepsilon^{2(n-1)}K_j^{-\tau_0+n+d/2} \implies |\mathcal{O}_0 \setminus \mathcal{O}_{j+1}| \le B\gamma a_0^{-1}\varepsilon^{2(n-1)}.$$

$(\mathbf{S3})_\mathbf{j}$ *There exist Lipschitz extensions $\hat\omega^{(j)}$, $\hat\Omega^{(j)}$ of $\omega^{(j)}$, $\tilde\Omega^{(j)}$ defined on $\mathcal{O}_0$ and, for $j \ge 1$:*

$$|\hat\omega^{(j)} - \hat\omega^{(j-1)}| + \lambda_j|\hat\omega^{(j)} - \hat\omega^{(j-1)}|^{\mathrm{lip}} \le \gamma\epsilon_j^{(2)}\,,\ \|\hat\Omega^{(j)} - \hat\Omega^{(j-1)}\|_\infty + \lambda_j\|\hat\Omega^{(j)} - \hat\Omega^{(j-1)}\|_\infty^{\mathrm{lip}} \le \gamma\epsilon_j^{(2)} \tag{184}$$

$$|\hat\omega^{(j)}|^{\mathrm{lip}} + \|\hat\Omega^{(j)}\|_\infty^{\mathrm{lip}} \le M_j\,, \quad |\hat\omega^{(j)} - \mathtt{v}| < M_j\varepsilon^2. \tag{185}$$

$(\mathbf{S4})_\mathbf{j}$ $\vec\epsilon_j, \Theta_j$ *satisfy (174) and* (176) *hence $|\vec\epsilon_j| \le \mathtt{C}_0|\vec\epsilon_0|e^{-\chi^j}$.*

$(\mathbf{S5})_\mathbf{j}$ *For $j > 0$ the sequence of composed maps $\tilde\Phi_j := \Phi_1\circ\Phi_2\circ\cdots\circ\Phi_j = I + \tilde\Psi_j$ satisfies $\|\tilde\Psi_{\nu+1} - \tilde\Psi_\nu\|_{D_j}^\lambda \le C_\star|\vec\epsilon_0|2^{-\nu}$, $\|\tilde\Psi_j\|_{D_j}^\lambda \le 2C_\star|\vec\epsilon_0|$.*

Proof. We proceed by induction the conditions $(\mathbf{Si})_\mathbf{0}$ are satisfied by the hypotheses of Theorem 6 except that, for $(\mathbf{S2})_\mathbf{0}$, once we have chosen K satisfying the constraints of the previous Lemmas, we have to impose a further smallness condition on $\bar\Theta$ deduced by formula (136).

Then, by induction, we prove the statements $(\mathbf{Si})_{\nu+\mathbf{1}}$, $i = 1,\dots,5$, by assuming the validity of $(\mathbf{Si})_\mathbf{j}$ for $j \le \nu$.

$(\mathbf{S1})_{\nu+\mathbf{1}}$. We apply the KAM step with $H = H_\nu$. By $(\mathbf{S2})_\nu$, we have that H_ν is compatible with the parameters p_ν. In order to implement the KAM step, and deduce $(\mathbf{S1})_{\nu+\mathbf{1}}$ we need to verify the constraints of Formulas (100), (103) and (136) are satisfied for $r_+, s_+, K_+, \theta_+, \mu_+ = r_{\nu+1}, s_{\nu+1}, K_{\nu+1}, \theta_{\nu+1}, \mu_{\nu+1}, K = K_\nu, \vec{\epsilon} = \vec{\epsilon}_\nu$. Substituting in (100), (103) we easily see that this amounts to a lower bound on K, depending only on τ_0, τ_1 and the remaining parameters in p_0. This we have imposed at the beginning of the algorithm, as we explained in Remark 10.5.

As for (136), we have by induction the inequality on $|\vec{\epsilon}_\nu|$ and we have to verify $2\kappa\delta_\nu^{-1} e \mathtt{C}_0 \bar{\Theta} e^{-\chi^\nu} {K_\nu}^{4d\tau_1} < \frac{1}{2}$, which is contained in the constraints (180).

Then following the KAM step we construct the set $\mathcal{O}_+$ which coincides by definition with $\mathcal{O}_{\nu+1}$. On $\mathcal{O}_{\nu+1}$, we define the generating function $F = F_{\nu+1}$. Then we construct the real analytic symplectic map $\Phi_{\nu+1} : D_{\nu+1} \times \mathcal{O}_{\nu+1} \to D_\nu$, Lipschitz in $\mathcal{O}_{\nu+1}$, generated by F. We have:

$$H_{\nu+1} := H_+ = H^\nu \circ \Phi_{\nu+1} =: \mathcal{N}_{\nu+1} + P_{\nu+1}\,, \quad \mathcal{N}_{\nu+1} := \mathcal{N}_\nu + [P_\nu]\,.$$

and $P_{\nu+1} := P_+$ defined in (160).

$(\mathbf{S2})_{\nu+\mathbf{1}}$. By construction $H_{\nu+1}$ is of the form given by Formula (10.2). We want to apply the results of §10.17 in order to prove that, $\forall \xi \in \mathcal{O}_{\nu+1}$, the Hamiltonian $H_{\nu+1}$, is compatible with the parameters $p_{\nu+1}$. For this it is enough to show that the constraints found on $\underline{p}_+$ in that section are in this case valid for the parameters $p_{\nu+1}$. First we need to verify (103) which is a largeness condition on K_0:

(186)
$$e^{-s_\nu \frac{2^\nu K_0}{(\ln 4^\nu K_0)^2}} 4^\nu {K_0}^{\tau_1} < 1\,,\ \kappa < \mu_0 2^{3\nu} {K_0}^2 \ln(4^\nu K_0)^{-2}\,,\ \kappa \mathtt{C} < \theta_0 2^{3\nu+1} {K_0}^{4d\tau_0-4} \ln(4^\nu K_0)^{-2}.$$

From Formula (180) we can bound uniformly $M_\nu \le \sqrt{2} M_0, L_\nu \le \sqrt{2} L_0$ so that we have for all ν that $S_0 = 8\sqrt{n} M_0 L_0 > 4\sqrt{n} M_\nu L_\nu$. By exploiting (165) and $|\vec{\epsilon}_\nu| \le \mathtt{C}_0 |\vec{\epsilon}_0| e^{-\chi^\nu}$, since $2\mathtt{C}_0, 2M_0 L_0 \mathtt{C}_0 \le C_\star$, we verify that $M_+ \le M_{\nu+1}$ and $L_+ \le L_{\nu+1}$:

$$M_+ := M_\nu (1 + 2|\vec{\epsilon}_\nu|) \le M_\nu (1 + 2\mathtt{C}_0 |\vec{\epsilon}_0| e^{-\chi^\nu}) \le M_{\nu+1}$$
$$L_+ := L_\nu (1 - M_\nu L_\nu \mathtt{C}_0 |\vec{\epsilon}_0| e^{-\chi^\nu})^{-1} \le L_{\nu+1}.$$

Finally in order to prove that $\omega_{\nu+1}$ is a lipeomorphism we argue as for (163) since $L_{\nu+1} |P_0^{1,0}|^{lip} \le \sqrt{2} L_0 \mathtt{C}_0 |\vec{\epsilon}_0| e^{-\chi^{\mathrm{j}+1} u} < 1$.

The estimate on $\mathcal{O}_{\nu+1}$ follows from Lemma (10.12). We finally show that a_+, defined in Formula (165) is $\ge a_{\nu+1}$, by noting that

$$a_\nu - a_+ \le 4 S_0 M_\nu |\vec{\epsilon}_\nu| \overset{(179)}{\le} 4\sqrt{2} S_0 M_0 \mathtt{C}_0 |\vec{\epsilon}_0| e^{-\chi^\nu} \le a_0 C_\star |\vec{\epsilon}_0| 2^{-\nu-1} = a_\nu - a_{\nu+1}.$$

$(\mathbf{S3})_{\nu+\mathbf{1}}$. The frequency maps $\omega^{(\nu+1)}$, $\Omega^{(\nu+1)}$ are defined on $\mathcal{O}_{\nu+1}$ and, as we have discussed in the previous item, have Lipschitz seminorm bounded by $M_{\nu+1}$. Then we may apply Formula (10.17) to deduce the bound (184) (recall that $\epsilon_\nu^{(2)} = \gamma^{-1} \|P_\nu^{(2)}\|_\nu$), for $\omega^{(\nu+1)}$ and $\tilde{\Omega}^{(\nu+1)}$. By the Kirszbraun theorem (see e.g. [15]), used component–wise, $\omega^{(\nu+1)}$ and $\tilde{\Omega}^{(\nu+1)}$ can be extended to maps– which we denote by $\hat{\omega}^{(\nu+1)}$, $\hat{\Omega}^{(\nu+1)}$– defined on the whole $\mathcal{O}_0$ and preserving the same sup-norm and Lipschitz seminorms,(185) follows. Moreover this extension may be performed so that $\hat{\omega}^{(\nu+1)} = \hat{\omega}^{(\nu)} + \hat{P}^{(1,0)}$ where $\hat{P}^{(1,0)}$ is an extension of $P^{(1,0)}$ which preserves the λ-norm (same for $\Omega^{(\nu+1)}$); this verifies (184).

$(\mathbf{S4})_{\nu+\mathbf{1}}$ $\vec{\epsilon}_{\nu+1} = \vec{\epsilon}_+$ satisfies (169), with $\zeta = \zeta_\nu$ and $(s - s_+)K/2 = \sigma_\nu K_\nu$. Recalling the definition of $\mathfrak{b}, \mathfrak{a}, \mathfrak{c}$ we have that $\vec{\epsilon}_{\nu+1}, \Theta_{\nu+1}$ satisfy the inequality of (174). We are in the Hypotheses of Lemma 10.19, so that also the bounds (176) holds.

$(\mathbf{S4})_{\nu+1}$ follows by (182), $(\mathbf{S3})_\nu$ and Lemma 10.19.

$(\mathbf{S5})_{\nu+1}$. Let us denote by $\mathcal{H}^{\lambda,\mathcal{O}}_{s,r}$ the normed space of functions in $\mathcal{H}_{s,r}$ which depend in a Lipschitz way from parameters in $\mathcal{O}$ with finite $\|.\|^\lambda$ norm. The estimate of the norm of the map $\Psi_{\nu+1} : \mathcal{H}^{\lambda,\mathcal{O}}_{s,r} \to \mathcal{H}^{\lambda_\nu,\mathcal{O}_\nu}_{s_\nu,r_\nu}$ follows, using (149), $\|X_{F_\nu}\|_\nu \le \mathtt{K}_\nu|\vec\epsilon_\nu|$ with $\mathtt{K}_\nu = 5(4^\nu K)^{4d\tau_1}$ and hence from (176) (179) and

$$\|X_{F_\nu}\|_\nu \le 5(4^\nu K)^{4d\tau_1}\mathtt{C}_0|\vec\epsilon_0|e^{-\chi^\nu} \le C_\star|\vec\epsilon_0|2^{-\nu-2n-7}.$$

From (180) and (102)

$$12\,2^{2n+6}e\delta_\nu^{-1}\|X_{F_\nu}\|_\nu \le \frac{1}{2} \implies \quad \|\Psi_{\nu+1}\| \le 2\|X_{F_\nu}\|^{\lambda_\nu}_{s_\nu,r_\nu}.$$

Now we can estimate

$$1+\tilde\Psi_\nu = \prod_{i=1}^{\nu}(1+\Psi_i), \quad \|1+\tilde\Psi_\nu\| \le \prod_{i=1}^{\nu}(1+\|\Psi_i\|) \le \prod_{i=1}^{\nu}(1+2^{-i}) \le 2$$

$$\tilde\Psi_{\nu+1} = \tilde\Psi_\nu + (1+\tilde\Psi_\nu)\Psi_\nu \implies \|\tilde\Psi_{\nu+1}-\tilde\Psi_\nu\| \le C_\star|\vec\epsilon_0|2^{-\nu}. \tag{187}$$

Notice that Ψ_ν also maps $\mathcal{Q}^T_{\underline{p}_{\nu-1}}$ to $\mathcal{Q}^T_{\underline{p}_\nu}$ and we have similar estimates using the Töplitz norms. □

Corollary 10.21. *For all $\xi \in \mathcal{O}_\infty := \cap_{\nu\ge0}\mathcal{O}_\nu$ the sequence $\tilde\Phi_\nu = I + \tilde\Psi_\nu$ converges uniformly on $D(s_0/2, r_0/2)$ to an analytic symplectic map $\Phi = I+\Psi$ such that the essential part of the perturbation $P^\infty_{\le2}(\cdot,\xi) = 0$. Moreover we have*

$$|\mathcal{O}_0 \setminus \mathcal{O}_\infty| \le cost\,\gamma\varepsilon^{2n-2}.$$

Proof. The fact that the $\tilde\Psi_\nu$ give a Cauchy sequence follows from Formula (187), therefore the sequence $\tilde\Phi_\nu$ converges as a sequence of Poisson bracket preserving homomorphisms from $\mathcal{H}_{s,r}$ to $\mathcal{H}_{s_\infty,r_\infty}$ to a Poisson bracket preserving homomorphism $\tilde\Phi_\infty$. The fact that this is induced by a coordinate change follows from the fact that we can construct the *local inverse*, $\lim\tilde\Theta_\nu$ where $\tilde\Theta_\nu = \Theta_\nu\circ\Theta_{\nu-1}\circ\cdots\circ\Theta_1$ and Θ_ν is the flux at time -1.

Finally $P^\infty_{\le2}(\cdot,\xi)=0$, $\forall\xi\in\mathcal{O}_\infty$, follows by (183) and $(S4)_\nu$. □

With this Corollary we finish the proof of Theorem 6.

We can finally conclude that

Theorem 7. *If for the Hamiltonian H_0 we have on the domains $\varepsilon^2\mathfrak{K}_\alpha\times D(s,r)$ a uniform estimate for the perturbation as in (29), i.e. $\|X_{P_0}\|^T_{\vec p_0}, \|X_{\tilde\Omega^0}\|^T_{\vec p} \le C\varepsilon^\beta$ with $\beta>2$, then for ε sufficiently small, the conditions on γ of the the iterative Lemma can be satisfied.*

Proof. The conditions we have imposed on γ are: $\gamma<1, \gamma\le 2\varepsilon^2M$, $\|X_{\tilde\Omega^0}\|^T_{\vec p} \le \gamma$ $B\varepsilon^{2(n-1)}\gamma < |\mathcal{O}_0|a_0$, $\Theta_0 := \frac{\|X_{P_0}\|_0}{\gamma} < \bar\Theta$. We have taken $\mathcal{O}_0 = \varepsilon^2\mathfrak{K}_\alpha$ (for some component of the complement of the discriminant) hence $|\mathcal{O}_0|$ can be estimated by $C_1\varepsilon^{2n}$, hence we impose on γ:

$$(\bar\Theta)^{-1}C\varepsilon^{\beta-2} < \gamma < \min(2M, B^{-1}a_0C_1) \tag{188}$$

as soon as $\varepsilon^{\beta-2} < \bar\Theta C^{-1}\min(2M, B^{-1}a_0C_1)$. □

Part 4. The NLS

In this final part we prove that the NLS is a compatible Hamiltonian (in suitable coordinates) according to Definition 10.2 and therefore we can apply to it the KAM algorithm and arrive at the conclusions of Theorem 6. Most of our work will be in showing the Töplitz property of the NLS.

11. The Töplitz property of the NLS

In fact A_d is Töplitz so $\|X_{A_d}\|^T_{(R,K,\theta,\mu)} = \|X_{A_d}\|_R$ follows from Remark 8.17.

11.1. Semi normal form. We now analyze the Birkhoff normal form change of variables defined in (9) with the purpose of proving that it maintains the quasi-Töplitz property. Note that the initial variables $u, \bar u \in \ell^{a,p}_{S=\emptyset} = \bar\ell^{a,p} \times \bar\ell^{a,p}$. So all the definitions of quasi-Töplitz functions of Part. 2 hold with $S=\emptyset$ and hence $n=0$ (i.e. there are no action-angle variables x, y). Moreover at this step we assume that $\mathtt{r}(m) = m$ for all m.

We fix the parameters $N_0, \mathtt{c}, \mathtt{C}$ as in (50). We start by noting that for all $d>0$

$$A_d := \sum_{k_i\in\mathbb{Z}^n:\ \sum(-1)^i k_i = 0} u_{k_1}\bar u_{k_2} u_{k_3}\bar u_{k_4}\dots u_{k_{2d-1}}\bar u_{k_{2d}} = \sum_{\substack{\alpha,\beta\in(\mathbb{Z}^n)^{\mathbb{N}}:\\ |\alpha|=|\beta|=d}} \binom{d}{\alpha}\binom{d}{\beta} u^\alpha \bar u^\beta, \tag{189}$$

is quasi Töplitz for all allowable (K,θ,μ).

In fact A_d is Töplitz so

$$\|X_{A_d}\|^T_{(R,K,\theta,\mu)} = \|X_{A_d}\|_R \le C(d) R^{d-2} \tag{190}$$

follows from Remark 8.17 and usual dimensional arguments, see (28). Notice that now the parameters $\vec p = (R,K,\theta,\mu)$ do not involve $\lambda, s, \mathcal{O}$.

Proposition 11.2. *For all choices of parameters $0<R<\epsilon_0$ and for allowable (K,θ,μ), the generating function F_{Birk} defined in* (9) *is quasi–Töplitz with $\|X_{F_{Birk}}\|^T_{(R/2,K,\theta,\mu)} \le \|X_{A_2}\|^T_{(R,K,\theta,\mu)} \le$ const R^2. Then for all (K',θ',μ') which respect* (103) *with (K,θ,μ) we have that $H\circ\Psi^{(1)} = H_{Birk} + P^{(4)} + P^{(6)}$ with $\|X_{P^{(i)}}\|^T_{(R/4,K,\theta,\mu)} \le$ const R^{i-2}, $i=4,6$ and $\|X_{H_{Birk}-\mathbb{K}}\|^T_{(R/4,K,\theta,\mu)} \le$ const R^2 (recall that $\mathbb{K}$ is defined in* (11)*)*

Proof. We need to compute the projection of F_{Birk} on the space of N,θ,μ,τ–bilinear functions, namely following formula (69) we compute $F^{\pm,\pm}_{m,n}$ for all m,n such that $m \xrightarrow{N} [v_i;p_i]$ and m,n have the cut ℓ with parameters θ,μ,τ.

By symmetry and reality we may consider just $+,+$ and $+,-$. We need to exhibit for them Töplitz approximations $\mathcal{F}^{+,+}, \mathcal{F}^{+,-}$.

In the case $+,+$ we write $\alpha = \alpha_0 + e_m + e_n, \beta = \beta_0$, in the case $+,-$ we write $\alpha = \alpha_0 + e_m, \beta = \beta_0 + e_n$, by definition α_0, β_0 are the exponents of the low variables, and in our case, since $|\alpha_2| + |\beta|_2 = 2$, the support of α_0, β_0 is in S and we have

$$F^{+,\pm}_{m,n} = -\mathrm{i} \sum^{*}_{\alpha^0,\beta^0\in\mathbb{N}^S} c_{\alpha,\beta} \frac{u^\alpha \bar u^\beta}{\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}} - \beta^0_{\mathtt{j}})|\mathtt{j}|^2 + |m|^2 \pm |n|^2}. \tag{191}$$

here $c_{\alpha,\beta} := \binom{2}{\alpha}\binom{2}{\beta}$ for simplicity of notation, while the symbol $\sum^*$ summarizes the conditions of (9), namely:

$$\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}} - \beta^0_{\mathtt{j}})|\mathtt{j}|^2 + |m|^2 \pm |n|^2 \neq 0\,, \quad \sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}} - \beta^0_{\mathtt{j}})\mathtt{j} + m \pm n = 0\,, \quad |\alpha_0| + |\beta_0| = 2 \tag{192}$$

In the case $+,+$ we claim that the denominator is *big* so that we can choose $\mathcal{F}^{+,+}_{m,n}=0$. Indeed we have $|m|^2+|n|^2>2\mathtt{c}N^{\tau_1}$ while $|\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}}-\beta^0_{\mathtt{j}})|\mathtt{j}|^2|<2\kappa^2$ where $\kappa:=\sup_{\mathtt{j}\in S}|\mathtt{j}|$. Since N is large all these denominators are bounded below by $\mathtt{c}N^{\tau_1}$. So for $\bar F^{+,+}_{m,n}:=N^{4d\tau}F^{+,+}_{m,n}$ we bound $\|X_{\bar F}\|_R\le N^{4d\tau-\tau_1}\mathtt{c}^{-1}\|X_{A_2}\|_R\le\|X_{A_2}\|_R$.

In the case $+,-$ we notice that $n-m=\pi(\alpha^0,\beta^0):=\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}}-\beta^0_{\mathtt{j}})\mathtt{j}\in B_N\cup\{0\}$. If $m=n$ the denomiators in (191) are m,n independent, we can take $\mathcal{F}^{+,+}_{m,n}=F^{+,+}_{m,n}$. When $m\neq n$ we write

$$|m|^2-|n|^2=(m-n,m+n)=(m-n,2m+n-m)=-2(\pi(\alpha^0,\beta^0),m)-|\pi(\alpha^0,\beta^0)|^2. \tag{193}$$

We have to distinguish two types of terms in the sum, that is the ones in which $\pi(\alpha^0,\beta^0):=\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}}-\beta^0_{\mathtt{j}})\mathtt{j}\notin\langle v_i\rangle_\ell$ and the other terms.

$$\begin{aligned}F^{+-}_{m,n}=&-\mathrm{i}\sum^{*}_{\substack{\alpha^0,\beta^0\in\mathbb{N}^S\\ \pi(\alpha^0,\beta^0)\notin\langle v_i\rangle_\ell}}c_{\alpha,\beta}\frac{u^\alpha\bar u^\beta}{\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}}-\beta^0_{\mathtt{j}})|\mathtt{j}|^2-2(\pi(\alpha^0,\beta^0),m)-|\pi(\alpha^0,\beta^0)|^2}\\ &-\mathrm{i}\sum^{*}_{\substack{\alpha^0,\beta^0\in\mathbb{N}^S\\ \pi(\alpha^0,\beta^0)\in\langle v_i\rangle_\ell}}c_{\alpha,\beta}\frac{u^\alpha\bar u^\beta}{\sum_{\mathtt{j}\in S}(\alpha^0_{\mathtt{j}}-\beta^0_{\mathtt{j}})|\mathtt{j}|^2-2(\pi(\alpha^0,\beta^0),m)-|\pi(\alpha^0,\beta^0)|^2}.\end{aligned} \tag{194}$$

In the first terms, since m has a cut at ℓ, we have, by Remark 7.16, $|(v,m)|>\theta N^{4d\tau}$ for all $v\notin\langle v_i\rangle_\ell$ hence the denominator is big and we proceed as in the case of $F^{+,+}$.

In the second terms, the right hand side of formula (193) depends only upon $m-n=\pi(\alpha^0,\beta^0)$ and on the cut $[v_i;p_i]_\ell$. This implies that the constraints in the sum (192) and the denominators in Formula (194) depend only on $m-n$ and on the cut $[v_i;p_i]_\ell$ so the second summand of formula (194) is in $\mathcal{T}_{N,\theta,\mu}$. The bounds follow by recalling that the denominators are non-zero integers. Then the bounds on the transformed Hamiltonian follow from Proposition 9.2 and by the degree considerations (95). ☐

We fix

$$\theta=\mathtt{C}(1-\frac{1}{16}),\quad\theta'=\mathtt{C}(1-\frac{1}{8}),\quad\mu=\mathtt{c}(1+\frac{1}{16}),\quad\mu'=\mathtt{c}(1+\frac{1}{8}) \tag{195}$$

so that (103) holds for all $K=K'>N_0$.

11.3. **Action angle variables.** The results we need are mostly contained in [2], although there are some small notational differences and the results in that paper are stated for $\mathbb{Z}$ instead of $\mathbb{Z}^d$, but the proofs follow verbatim in our case.

We introduce action-angle variables on the tangential sites $S:=\{\mathtt{j}_1,\dots,\mathtt{j}_n\}$ via the analytic and symplectic map

$$\Phi_\xi(x,y,z,\bar z):=(u,\bar u) \tag{196}$$

defined by

$$u_{\mathtt{j}_l}:=\sqrt{\xi_l+y_l}\,e^{\mathrm{i}x_l},\ \bar u_{\mathtt{j}_l}:=\sqrt{\xi_l+y_l}\,e^{-\mathrm{i}x_l},\ l=1,\dots,n\,,\quad u_j:=z_j\,,\ \bar u_j:=\bar z_j\,,\ j\in\mathbb{Z}^d\setminus S\,. \tag{197}$$

Let us consider for $\varepsilon^2>0$ the set $\varepsilon^2\mathfrak{K}_\alpha$ as in Theorem 4

Lemma 11.4. (Domains) *Let $s, r, \varepsilon, R > 0$ satisfy*

$$2c_1 r < \varepsilon\,, \qquad R = C_*\varepsilon \quad \text{with} \quad C_* := 4c_2\sqrt{n}\kappa^p e^{(s+a\kappa)}\,. \tag{198}$$

Then, for all $\xi \in \varepsilon^2\mathfrak{K}_\alpha \cup 2\varepsilon^2\mathfrak{K}_\alpha$, the map

$$\Phi_\xi(\,\cdot\,;\xi) : D(s,2r) \to \mathcal{D}(R/2) := B_{R/2} \times B_{R/2} \subset \bar\ell^{a,p} \times \bar\ell^{a,p} \tag{199}$$

is well defined and analytic (recall that $D(s,2r)$ is defined in (13) *and $\kappa := \sup_{\mathtt{j}\in S}|\mathtt{j}|$).*

For the proof see [2] Lemma 7.5.

Given a function $F : \mathcal{D}(R/2) \to \mathbb{C}$, the previous Lemma shows that the composite map $F\circ\Phi_\xi : D(s,2r) \to \mathbb{C}$ is well defined and regular. The main result of this section is Proposition 11.5: if F is quasi-Töplitz in the variables $(u,\bar u)$ then the composite $F\circ\Phi_\xi$ is quasi-Töplitz in the variables $(x,y,z,\bar z)$ (see Definition 8.14).

We write

$$F = \sum_{\alpha,\beta} F_{\alpha,\beta}\mathfrak{m}_{\alpha,\beta}\,, \quad \mathfrak{m}_{\alpha,\beta} := (u^{(1)})^{\alpha^{(1)}}(\bar u^{(1)})^{\beta^{(1)}}(u^{(2)})^{\alpha^{(2)}}(\bar u^{(2)})^{\beta^{(2)}}\,, \tag{200}$$

where

$$u = (u^{(1)}, u^{(2)})\,, \quad u^{(1)} := \{u_j\}_{j\in S}\,,\ u^{(2)} := \{u_j\}_{j\in\mathbb{Z}^d\setminus S}\,, \quad \text{similarly for } \bar u\,,$$

$$(\alpha^{(1)},\beta^{(1)}) := \{\alpha_j,\beta_j\}_{j\in S}\,, \quad (\alpha^{(2)},\beta^{(2)}) := \{\alpha_j,\beta_j\}_{j\in S^c}\,. \tag{201}$$

We define

$$\mathcal{H}^D_R := \Big\{F \in \mathcal{H}_R \ :\ F = \sum_{\alpha^{(2)}+\beta^{(2)}\ge D} F_{\alpha,\beta}u^\alpha\bar u^\beta\Big\}\,. \tag{202}$$

Proposition 11.5. (Quasi–Töplitz) *Let $\vec p = (r,s,K,\theta,\mu,\lambda,\varepsilon^2\mathfrak{K}_\alpha)$, with K,θ,μ,μ' be admissible parameters and*

$$(\mu'-\mu)K^3 > K\,, \qquad K^{\tau_1}2^{-\frac{K}{2\kappa}+1} < 1\,. \tag{203}$$

If $F \in \mathcal{Q}^T_{R/2,K,\theta,\mu'} \cap \mathcal{H}^D_{R/2}$, then $f := F\circ\Phi_\xi \in \mathcal{Q}^T_{\vec p}$ and

$$\|X_f\|^T_{\vec p} \lessdot (8r/R)^{D-2}\frac{\lambda}{\varepsilon^2}\|X_F\|^T_{R/2,K,\theta,\mu'}\,. \tag{204}$$

For the proof we will need several Lemmas. First let us deduce the main consequence of Proposition 11.5.

Corollary 11.6. *For all $\varepsilon > 0$, $c_1\varepsilon/2 > r > \varepsilon^3$ and $s > 0$ satisfying* (198)*, the perturbation P of Definition 2.2 is in $\mathcal{Q}^T_{\vec p}$ for the parameters $\vec p = (s, r, \theta = \mathtt{C}(1-\frac14), \mu = \mathtt{c}(1+\frac14), \lambda, \varepsilon^2\mathfrak{K}_\alpha)$. Moreover P satisfies the bounds*

$$\|X_P\|^T_{\vec p} \le C\frac{\lambda}{\varepsilon^2}(\varepsilon r + \varepsilon^5 r^{-1})\,, \tag{205}$$

where C does not depend on ε, r.

Proof. We choose $\mu' = \mathtt{c}(1+\frac18)$, the perturbation P has contributions from three terms: 1) The term $P^{(4)}\circ\Phi_\xi$, 2) The term $P^{(6)}\circ\Phi_\xi$ and finally 3) the terms of degree > 2 in $(H_{Birk} - \mathbb{K})\circ\Phi_\xi$. By proposition 11.2 with the choice of parameters (195) all the terms above are quasi-Töplitz with the parameter μ' and the Bounds of Proposition 11.2 hold.

In item 1) we note that, by definition, $P^{(4)} \in \mathcal{H}^3_R$ so, by Propositions (11.5) and (11.2), we have $\frac{\varepsilon^2}{\lambda}\|X_{P^{(4)}}\|^T_{\vec p} \le (r/R)R^2 \le C\varepsilon\, r$. In item 2) we recall that by momentum conservation the first term of $P^{(6)}$ of degree $D = 0$ ($P^{(6)} \in \mathcal{H}^0_R\setminus\mathcal{H}^1_R$) is actually of degree at least 8. Then we divide $P^{(6)} = R+Q$ where $Q \in \mathcal{H}^1_R$ and R is of degree at least 8 in $u, \bar{u}$. By Propositions (11.5) and (11.2), we have $\|X_R\|^T_{\vec{p}} \le (r/R)^{-2} R^6 \le C\varepsilon^8\, r^{-2} \le C\varepsilon^5 r^{-1}$. In the same way $\|X_Q\|^T_{\vec{p}} \le (r/R)^{-1} R^4 \le C\varepsilon^5\, r^{-1}$.

Finally in 3) we collect the terms of degree 3 and 4 in formula (10), we get the estimates $\frac{\varepsilon^2}{\lambda}|X_{\Pi^{\ge 3}(H_{Birk}-\mathbb{K})}|^T_{\vec{p}'} \le C\varepsilon r$. □

The rest of this section is devoted to the proof of Proposition 11.5. Introducing the action-angle variables (197) in (200), and using the Taylor expansion

(206) $$(1+t)^g = \sum_{h\ge 0} \binom{g}{h} t^h\,, \quad \binom{g}{0} := 1\,, \quad \binom{g}{h} := \frac{g(g-1)\dots(g-h+1)}{h!}\,,\ h \ge 1\,,$$

we get

(207) $$f := F \circ \Phi_\xi = \sum_{k,i,\alpha^{(2)},\beta^{(2)}} f_{k,i,\alpha^{(2)},\beta^{(2)}} e^{\mathrm{i}(k,x)} y^i z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}}$$

with Taylor–Fourier coefficients

(208) $$f_{k,i,\alpha^{(2)},\beta^{(2)}} := \sum_{\alpha^{(1)}-\beta^{(1)}=k} F_{\alpha,\beta} \prod_{l=1}^n \xi_l^{\frac{\alpha_l^{(1)}+\beta_l^{(1)}}{2} - i_l} \binom{\frac{\alpha_l^{(1)}+\beta_l^{(1)}}{2}}{i_l}\,.$$

Lemma 11.7. **(M-regularity)** *If $F \in \mathcal{H}^D_{R/2}$ then $f := F \circ \Phi_\xi \in \mathcal{H}_{s,2r}$ and*

(209) $$\|X_f\|_{s,2r,\varepsilon^2\mathfrak{K}_\alpha \cup 2\varepsilon^2\mathfrak{K}_\alpha} \lessdot (8r/R)^{D-2} \|X_F\|_{R/2}\,, \quad \|X_f\|^{lip}_{s,2r,\varepsilon^2\mathfrak{K}_\alpha} \lessdot \varepsilon^{-2}(8r/R)^{D-2}\|X_F\|_{R/2}\,.$$

Moreover if F preserves momentum then so does $F \circ \Phi_\xi$.

Proof. See [2] Lemma 7.7. □

Definition 11.8. For a monomial $\mathfrak{m}_{\alpha,\beta} := (u^{(1)})^{\alpha^{(1)}} (\bar{u}^{(1)})^{\beta^{(1)}} (u^{(2)})^{\alpha^{(2)}} (\bar{u}^{(2)})^{\beta^{(2)}}$ (as in (200)) we set

(210) $$\mathfrak{p}(\mathfrak{m}_{\alpha,\beta}) := \sum_{l=1}^n \langle \mathtt{j}_l \rangle (\alpha^{(1)}_{\mathtt{j}_l} + \beta^{(1)}_{\mathtt{j}_l})\,, \quad \langle j \rangle := \max\{1, |j|\}\,.$$

For any F as in (200), $K \in \mathbb{N}$, we define the projection

(211) $$\Pi_{\mathfrak{p}\ge K} F := \sum_{\mathfrak{p}(\mathfrak{m}_{\alpha,\beta}) \ge K} F_{\alpha,\beta}\mathfrak{m}_{\alpha,\beta}\,, \quad \Pi_{\mathfrak{p}<K} := I - \Pi_{\mathfrak{p}\ge K}\,.$$

Lemma 11.9. *Let $F \in \mathcal{H}_{R/2}$. Then*

(212) $$\|X_{(\Pi_{\mathfrak{p}\ge K}F)\circ\Phi_\xi}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha} \le 2^{-\frac{K}{2\kappa}+1} \|X_{F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha}\,.$$

Proof. See [2] Lemma 7.8. □

Let $K, \theta, \mu, \mu', \tau$ be as in Proposition 11.5. For $N \ge K$ and $F \in \mathcal{H}_{R/2}$ we set

(213) $$f^* := \Pi_{N,\theta,\mu,\tau}\Big((F - \Pi_{N,\theta,\mu',\tau}F) \circ \Phi_\xi\Big)\,.$$

Note that $\Pi_{N,\theta,\mu',\tau}$ is the projection on the bilinear functions in the variables $u, \bar{u}$, while $\Pi_{N,\theta,\mu,\tau}$ in the variables $x, y, z, \bar{z}$.

The next Lemma corresponds to Lemma 7.9 of [2].

Lemma 11.10. *We have*

$$\|X_{f^*}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha} \le 2^{-\frac{N}{2\kappa}+1}\|X_{F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha}\,. \tag{214}$$

Proof. We first claim that if $F = \mathfrak{m}_{\alpha,\beta}$ is a monomial as in (200) with $\mathfrak{p}(\mathfrak{m}_{\alpha,\beta}) < N$ then $f^* = 0$.

CASE 1: $\mathfrak{m}_{\alpha,\beta}$ is (N,θ,μ',τ)–bilinear, see Definition 8.3. Then $\Pi_{N,\theta,\mu',\tau}\mathfrak{m}_{\alpha,\beta} = \mathfrak{m}_{\alpha,\beta}$ and $f^* = 0$, see (213).

CASE 2: $\mathfrak{m}_{\alpha,\beta}$ is *not* (N,θ,μ',τ)–bilinear. Then we have $\Pi_{N,\theta,\mu',\tau}\mathfrak{m}_{\alpha,\beta} = 0$ and hence $f^* = \Pi_{N,\theta,\mu,\tau}(\mathfrak{m}_{\alpha,\beta}\circ\Phi_\xi)$, see (213). We claim that $\mathfrak{m}_{\alpha,\beta}\circ\Phi_\xi$ is not (N,θ,μ,τ)–bilinear, and so $f^* = \Pi_{N,\theta,\mu,\tau}(\mathfrak{m}_{\alpha,\beta}\circ\Phi_\xi) = 0$. Indeed,

$$\mathfrak{m}_{\alpha,\beta}\circ\Phi_\xi = (\xi+y)^{\frac{\alpha^{(1)}+\beta^{(1)}}{2}} e^{\mathrm{i}((\alpha^{(1)}-\beta^{(1)}),x)} z^{\alpha^{(2)}}\bar{z}^{\beta^{(2)}} \tag{215}$$

is (N,θ,μ,τ)–bilinear if and only if (see Definitions 8.3 and 8.2)

$$z^{\alpha^{(2)}}\bar{z}^{\beta^{(2)}} = z^{\tilde\alpha^{(2)}}\bar{z}^{\tilde\beta^{(2)}} z_m^\sigma z_n^{\sigma'}\,,$$

$$\sum_{j\in\mathbb{Z}^d\setminus S} |j|(\tilde\alpha_j^{(2)} + \tilde\beta_j^{(2)}) < \mu N^3\,, \quad |m|,|n| > \theta N^{\tau_1}\,, \quad |\alpha^{(1)} - \beta^{(1)}| < N\,, \tag{216}$$

and n, m have a cut at ℓ with parameters θ,μ,τ.

We deduce the contradiction that $\mathfrak{m}_{\alpha,\beta} = (u^{(1)})^{\alpha^{(1)}}(\bar{u}^{(1)})^{\beta^{(1)}}(u^{(2)})^{\tilde\alpha^{(2)}}(\bar{u}^{(2)})^{\tilde\beta^{(2)}}u_m^\sigma u_n^{\sigma'}$ is (N,θ,μ',τ)-bilinear because (recall that we suppose $\mathfrak{p}(\mathfrak{m}_{\alpha,\beta}) < N$)

$$\sum_{l=1}^n |\mathtt{j}_l|(\alpha_{\mathtt{j}_l}^{(1)} + \beta_{\mathtt{j}_l}^{(1)}) + \sum_{j\in\mathbb{Z}^d\setminus S}|j|(\tilde\alpha_j^{(2)} + \tilde\beta_j^{(2)}) \overset{(210),(216)}{<} \mathfrak{p}(\mathfrak{m}_{\alpha,\beta}) + \mu N^3 < N + \mu N^3 \overset{(203)}{<} \mu' N^3\,.$$

For the general case, we divide $F = \Pi_{\mathfrak{p}<N}F + \Pi_{\mathfrak{p}\ge N}F$. By the above claim

$$f^* = \Pi_{N,\theta,\mu,\tau}\Big(\big((Id - \Pi_{N,\theta,\mu',\tau})\Pi_{\mathfrak{p}\ge N}F\big)\circ\Phi_\xi\Big) = \Pi_{N,\theta,\mu,\tau}\Big(\big(\Pi_{\mathfrak{p}\ge N}(Id - \Pi_{N,\theta,\mu',\tau})F\big)\circ\Phi_\xi\Big)\,.$$

Finally, (214) follows by applying Lemma 11.9 to $\big(\Pi_{\mathfrak{p}\ge N}(Id - \Pi_{N,\theta,\mu'})F\big)\circ\Phi_\xi$ and using the fact that projections may only reduce the norm. □

Lemma 11.11. *Let $F \in \mathcal{T}_{R/2,N,\theta,\mu',\tau}$ with $\Pi_{\mathfrak{p}\ge N}F = 0$. Then $F\circ\Phi_\xi(\cdot;\xi) \in \mathcal{T}_{s,2r,N,\theta,\mu',\tau}$, $\forall\,\xi \in \varepsilon^2\mathfrak{K}_\alpha \cup 2\varepsilon^2\mathfrak{K}_\alpha$.*

Proof. Recalling Definition 8.8 we have

$$F = \sum_{A\in\mathcal{H}_N}\ \sum^{*}_{|m|,|n|>\theta N^{\tau_1},\sigma,\sigma'=\pm} F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)u_m^\sigma u_n^{\sigma'} \text{ with } F^{\sigma,\sigma'}(A,h) \in \mathcal{L}_{R/2}(N,\mu',h)\,.$$

denoting $A = [v_i;p_i]_\ell$ the apex $*$ means the sum restricted to those n, m which have a cut at ℓ with parameters (θ,μ,τ) and m has associated space A.

Composing with the map Φ_ξ in (197), since $m, n \notin S$, we get

$$F\circ\Phi_\xi = \sum_{\sigma,\sigma'=\pm\,;|m|,|n|>\theta N^{\tau_1}} F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)\circ\Phi_\xi\, z_m^\sigma z_n^{\sigma'}\,.$$

Each coefficient $F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)\circ\Phi_\xi$ depends on n, m, σ, σ' only through $A, \sigma m + \sigma' n, \sigma, \sigma'$. Hence, in order to conclude that $F\circ\Phi_\xi \in \mathcal{T}_{s,2r,N,\theta,\mu',\tau}$ it remains only to prove that $F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)\circ\Phi_\xi\in\mathcal{L}_{s,2r}(N,\mu',\sigma m+\sigma' n)$, see Definition 8.2. Each monomial $\mathfrak{m}_{\alpha,\beta}$ of $F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)\in\mathcal{L}_{R/2}(N,\mu',\sigma m+\sigma' n)$ satisfies

$$\sum_{l=1}^{n}(\alpha_{\mathtt{j}_l}+\beta_{\mathtt{j}_l})|\mathtt{j}_l|+\sum_{j\in\mathbb{Z}^d\setminus S}(\alpha_j+\beta_j)|j|<\mu' N^3 \quad\text{and}\quad \mathfrak{p}(\mathfrak{m}_{\alpha,\beta})<N$$

by the hypothesis $\Pi_{\mathfrak{p}\geq N}F=0$. Hence $\mathfrak{m}_{\alpha,\beta}\circ\Phi_\xi$ (see (215)) is (N,μ')-low momentum, in particular $|\alpha^{(1)}-\beta^{(1)}|\leq\mathfrak{p}(\mathfrak{m}_{\alpha,\beta})<N$. □

Proof. OF PROPOSITION 11.5. Since $F\in\mathcal{Q}^T_{R/2,K,\theta,\mu'}$ (see Definition 8.14), for all $N\geq K$, there is a Töplitz approximation $\tilde F\in\mathcal{T}_{R/2,N,\theta,\mu',\tau}$ of F, namely

$$\Pi_{N,\theta,\mu',\tau}F=\tilde F+N^{-4d\tau}\hat F \quad\text{with}\quad \|X_F\|_{R/2},\|X_{\tilde F}\|_{R/2},\|X_{\hat F}\|_{R/2}<2\|F\|^T_{R/2,K,\theta,\mu'}\,. \tag{217}$$

In order to prove that $f:=F\circ\Phi_\xi\in\mathcal{Q}^T_{s,r,K,\theta,\mu}$ we define its candidate Töplitz approximation

$$\tilde f:=\Pi_{N,\theta,\mu,\tau}((\Pi_{\mathfrak{p}<N}\tilde F)\circ\Phi_\xi)\,, \tag{218}$$

see (211). Lemma 11.11 applied to $\Pi_{\mathfrak{p}<N}\tilde F\in\mathcal{T}_{R/2,N,\theta,\mu',\tau}$ implies that $(\Pi_{\mathfrak{p}<N}\tilde F)\circ\Phi_\xi\in\mathcal{T}_{s,2r,N,\theta,\mu',\tau}$ and then, applying the projection $\Pi_{N,\theta,\mu,\tau}$ we get $\tilde f\in\mathcal{T}_{s,2r,N,\theta,\mu,\tau}\subset\mathcal{T}_{s,r,N,\theta,\mu,\tau}$. Moreover, by (218) and applying Lemma 11.7 to $\Pi_{\mathfrak{p}<N}\tilde F$ (note that $\Pi_{\mathfrak{p}<N}\tilde F$ is either zero or it is in $\mathcal{H}^D_{R/2}$ with $D\geq 2$ because it is bilinear), we get

$$\begin{aligned}\|X_{\tilde f}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha}\leq\|X_{(\Pi_{\mathfrak{p}<N}\tilde F)\circ\Phi_\xi)}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha} &\overset{(209)}{\lessdot} (8r/R)^{D-2}\|X_{\Pi_{\mathfrak{p}<N}\tilde F}\|_{R/2}\\ &\overset{(217)}{\lessdot} (8r/R)^{D-2}\|F\|^T_{R/2,K,\theta,\mu',\tau}\,.\end{aligned} \tag{219}$$

Moreover the Töplitz defect is

$$\begin{aligned}\hat f\quad &:= N^{4d\tau}(\Pi_{N,\theta,\mu,\tau}f-\tilde f)\overset{(218)}{=}N^{4d\tau}\,\Pi_{N,\theta,\mu,\tau}\big((F-\Pi_{\mathfrak{p}<N}\tilde F)\circ\Phi_\xi\big)\\ &= N^{4d\tau}\Pi_{N,\theta,\mu,\tau}\big((F-\tilde F)\circ\Phi_\xi\big)+N^{4d\tau}\Pi_{N,\theta,\mu,\tau}\big((\tilde F-\Pi_{\mathfrak{p}<N}\tilde F)\circ\Phi_\xi\big)\\ &\overset{(217),(211)}{=} \Pi_{N,\theta,\mu,\tau}(\hat F\circ\Phi_\xi)+N^{4d\tau}\Pi_{N,\theta,\mu,\tau}\Big(\big(F-\Pi_{N,\theta,\mu'}F\big)\circ\Phi_\xi\Big)\\ &\qquad +N^{4d\tau}\Pi_{N,\theta,\mu,\tau}\big((\Pi_{\mathfrak{p}\geq N}\tilde F)\circ\Phi_\xi\big)\\ &\overset{(213)}{=} \Pi_{N,\theta,\mu,\tau}(\hat F\circ\Phi_\xi)+N^{4d\tau}f^*+N^{4d\tau}\Pi_{N,\theta,\mu,\tau}\big((\Pi_{\mathfrak{p}\geq N}\tilde F)\circ\Phi_\xi\big)\,.\end{aligned}$$

Lemmata 11.9 and 11.10 imply that, since $N^{4d\tau}2^{-\frac{N}{2\kappa}+1}\leq 1\ \forall N\geq K$ by (203),

$$\begin{aligned}\|X_{\hat f}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha} &\leq \|X_{\hat F\circ\Phi_\xi}\|_{s,r,\varepsilon^2\mathfrak{K}_\alpha}+N^{4d\tau}2^{-\frac{N}{2\kappa}+1}(\|X_{F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha}+\|X_{\tilde F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha})\\ &\lessdot \|X_{\hat F\circ\Phi_\xi}\|_{s,2r,\varepsilon^2\mathfrak{K}_\alpha}+\|X_{F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha}+\|X_{\tilde F\circ\Phi_\xi}\|_{s,2r,2\varepsilon^2\mathfrak{K}_\alpha}\\ &\overset{(209)}{\lessdot} (8r/R)^{D-2}(\|X_{\hat F}\|_{R/2}+\|X_F\|_{R/2}+\|X_{\tilde F}\|_{R/2}) \qquad (220)\\ &\overset{(217)}{\lessdot} (8r/R)^{D-2}\|F\|^T_{R/2,K,\theta,\mu'} \qquad (221)\end{aligned}$$

(to get (220) we also note that we can choose $\hat F,\tilde F$ so that they belong to the same $\mathcal{H}^D_{R/2}$ as F. The bound (204) follows by (209), (219), (221). □

11.12. **Reduction to constant coefficients.** For all $k \in S^c$ set $\mathtt{r}(k) := \mathtt{r}(A)$ to be the root of the component A of Γ_S to which k belongs (this is chosen once in one of the graphs in the same translation class). We have thus associated to each k an element $L(k) \in \mathbb{Z}^n$, see Theorem 2 and formula (36):

$$z_k = e^{-\mathrm{i}L(k).x} z'_k,\ y = y' + \sum_{k\in S^c} L(k)|z'_k|^2,\ x = x'. \tag{222}$$

to define a symplectic change of variables $\Psi : D(s, r/2) \to D(s, r)$ in which the normal form has constant coefficients. One may trivially check that

$$\|X_{F\circ\Psi}\|^\lambda_{s,r} \le 4e^{2d\kappa s}\|X_F\|^\lambda_{s,2r}.$$

We need to see what happens to (N, θ, μ, τ)–bilinear monomials first. Note that the momentum of z'_k is $\mathtt{r}(k)$. Take a monomial

$$\mathfrak{m} = \mathfrak{m}_{\alpha,\beta,k} = e^{\mathrm{i}(k,x)} y^l z^\alpha \bar z^\beta.$$

We have

$$\mathfrak{m}\circ\Psi = e^{\mathrm{i}(k',x)}(y' + \sum_{k\in S^c} L(k)|z'_k|^2)^l {z'}^\alpha \bar{z}'^\beta\,, \quad k' = k - \sum_j L(j)(\alpha_j - \beta_j).$$

Hence we obtain a sum of monomials

$$e^{\mathrm{i}(k',x)}(y')^h (z')^\alpha (\bar z')^\beta |z'|^{2g}\,, \quad h \le l\,, \quad |h| + |g| = |l| \tag{223}$$

all with momentum:

$$\pi_{\mathtt{r}}(k', \alpha, \beta) = \pi(k') + \sum_j (\alpha_j - \beta_j)\mathtt{r}(j) = \pi(k', \alpha, \beta) + \sum_j (\alpha_j - \beta_j)(\mathtt{r}(j) - j). \tag{224}$$

As in the previous section we define a cut off parameter

$$\mathfrak{p}(\mathfrak{m}_{\alpha,\beta,k}) := |k| + 2d\kappa(|\alpha| + |\beta|),$$

and set

$$\Pi_{\mathfrak{p}\ge K} F := \sum_{\mathfrak{p}(\mathfrak{m}_{\alpha,\beta,k})\ge K} F_{\alpha,\beta,k}\mathfrak{m}_{\alpha,\beta,k}\,, \quad \Pi_{\mathfrak{p}<K} := I - \Pi_{\mathfrak{p}\ge K}\,. \tag{225}$$

In the following Lemmas we assume that $s > (2d\kappa)^{-1}$.

Lemma 11.13. *For all $F \in \mathcal{H}_{s,r}$ we have*

$$\|X_{(\Pi_{\mathfrak{p}\ge K}F)\circ\Psi}\|_{s,r} \le 2^{-\frac{K}{8d\kappa}+2}\|X_{F\circ\Psi}\|_{2s,2r}\,. \tag{226}$$

Proof. When $\mathfrak{p}(\mathfrak{m}_{\alpha,\beta,k}) := |k| + 2d\kappa(|\alpha| + |\beta|) > K$ we distinguish two cases:

1. $2d\kappa(|\alpha| + |\beta|) > K/4$. We note that Ψ may only increase the degree in the normal variables of monomials so the total degree in the new variables is $> K/(8d\kappa)$ and the bound follows by the degree bounds 28.

2. Otherwise $|k'| > K/2$ and the bound follows by the ultraviolet bounds 27. ☐

Fix parameters

$$(\mu' - \mu)N^3 > N\,, (\theta - \theta')N^{\tau_1} > 2d\kappa.$$

Lemma 11.14. *Take a function $F \in \mathcal{H}_{s,r}$, assume that*

$$\Pi_{\mathfrak{p}\ge N}F = \Pi^L_{N,\mu'}F = 0.$$

Then we have

$$f^* := \Pi_{N,\theta,\mu,\tau}(F - \Pi_{N,\theta',\mu',\tau}F)\circ\Psi = 0$$

Proof. We may assume that $F = \mathfrak{m}_{\alpha,\beta,k}$ is a monomial. If F is (N, θ', μ', τ)–bilinear the statement is clear. Otherwise f^* is a sum a monomials described by formula 223. If one of these monomials is bilinear its high variables either come from one of the new exponents g or already appear in α, β. In the first case this is possible only if $\mathfrak{m}$ is (N, μ')–low contrary to our hypothesis. In fact suppose that $g = \bar{g} + e_m$, where $m = n$ is the high variable with $|\mathtt{r}(m)| > \theta N^{\tau_1}$, and that

$$\sum_j |\mathtt{r}(j)|(\alpha_j + \beta_j + \bar{g}_j) < \mu N^3.$$

Then since $\bar{g}_j \geq 0$ and $|j - \mathtt{r}(j)| < 2d\kappa$ we have

$$\sum_j |j|(\alpha_j + \beta_j) < \mu N^3 + 2d\kappa(|\alpha| + |\beta|) < \mu N^3 + N < \mu' N^3.$$

Finally since $\mathfrak{p} < N$ we have $|k| < N$, we deduce that $\mathfrak{m}$ is low.

In the other case the two high variables m, n such that $|\mathtt{r}(m)|, |\mathtt{r}(n)| > \theta N^{\tau_1}$ already appear in $\mathfrak{m}_{\alpha,\beta,k}$. We claim that this implies $\mathfrak{m}_{\alpha,\beta,k}$ (N, θ', μ', τ)-bilinear contrary to the hypothesis. In fact write $\alpha = \bar{\alpha} + e_m, \beta = \bar{\beta} + e_n$. Applying Ψ the monomials appearing in f^* are of the form $\mathfrak{m}_{\bar{\alpha}+g,\bar{\beta}+g,k'} z_m \bar{z}_n$ with $|k'| < N$ and $\sum_j |\mathtt{r}(j)|(\bar{\alpha}_j + \bar{\beta}_j + 2g_j) < \mu N^3$. Then $\sum_j |j|(\bar{\alpha}_j + \bar{\beta}_j) < \mu N^3 + 2d\kappa(|\alpha| + |\beta|) < \mu' N^3$. Since $|j - \mathtt{r}(j)| < 2d\kappa$ we have

$$|m|, |n| > \theta N^{\tau_1} - 2d\kappa > \theta' N^{\tau_1}.$$

The fact that m, n have the correct cut is trivial, see Remark 8.7. Finally we are assuming that $\mathfrak{p}(m_{\alpha,\beta,k}) := |k| + 2d\kappa(|\alpha| + |\beta|) \leq N$ hence $|k| < N$ and we have that $\mathfrak{m}_{\alpha,\beta,k}$ is (N, θ', μ', τ)-bilinear.

□

We next analyze a function F with $\Pi^L_{N,\mu'} F = F$ and again we may assume that it is a monomial $F = \mathfrak{m}_{\alpha,\beta,k}$, in this case $f^* := \Pi_{N,\theta,\mu,\tau}(F - \Pi_{N,\theta',\mu',\tau} F) \circ \Psi = \Pi_{N,\theta,\mu,\tau} F \circ \Psi$ is a sum of monomials $\mathfrak{m}_{\alpha+\bar{g},\beta+\bar{g},k'} |z_m|^2$ arising from the terms 223 with $g = \bar{g} + e_m$.

Lemma 11.15. *Given a function F with $\Pi^L_{N,\mu'} F = F$ then $f^* = \Pi_{N,\theta,\mu,\tau} F \circ \Psi$ is piecewise Töplitz and diagonal.*

Proof. By the previous remarks we may compute explicitly f^* as:

$$\Pi^L_{N,\mu}\Big(\nabla_y F \circ \Psi\Big) \cdot \sum_{\substack{|m| > \theta N^{\tau_1},\\ m \in (N,\theta,\mu,\tau)-\mathrm{cut}}} L(m)|z_m|^2\,,$$

we have that $f^* \in \mathcal{T}_{(N,\theta,\mu,\tau)}$ since $L(m)$ is fixed on all the (N, θ, μ, τ)–good points of any subspace (by Theorem 5). □

Lemma 11.16. *Given a function $F \in \mathcal{T}_{(N,\theta',\mu',\tau)}$, then $\Pi_{N,\theta,\mu,\tau} F \circ \Psi \in \mathcal{T}_{(N,\theta,\mu,\tau)}$*

Proof. Recalling Definition 8.8 we have

$$F = \sum_{A \in \mathcal{H}_N} \sum^{*}_{|m|,|n| > \theta' N^{\tau_1}, \sigma,\sigma'=\pm} F^{\sigma,\sigma'}(A, \sigma m + \sigma' n) z_m^{\sigma} z_n^{\sigma'} \;\text{ with }\; F^{\sigma,\sigma'}(A, h) \in \mathcal{L}_{r,s}(N, \mu', h)\,.$$

denoting $A = [v_i; p_i]_\ell$ the apex $*$ means the sum restricted to those n, m which have a cut at ℓ with parameters (N, θ', μ', τ) and m has associated space A.

Composing with the map Ψ in (222), since $m, n \notin S$ and $|\mathtt{r}(m)|, |\mathtt{r}(n)| > \theta N^{\tau_1}$ implies $|m|, |n| > \theta' N^{\tau_1}$, we get $\Pi_{(N,\theta,\mu,\tau)} F \circ \Psi =$

$$\sum_{A\in\mathcal{H}_N} \sideset{}{^*}\sum_{|\mathtt{r}(m)|,|\mathtt{r}(n)|>\theta N^{\tau_1},\sigma,\sigma'=\pm} \Pi^L_{N,\mu}\Big(F^{\sigma,\sigma'}(A,\sigma m+\sigma' n)\circ\Psi\, e^{-\mathrm{i}(\sigma L(m)+\sigma' L(n),x)}\Big)(z'_m)^{\sigma}(z'_n)^{\sigma'}\,.$$

Each coefficient $F^{\sigma,\sigma'}(A, \sigma m + \sigma' n) \circ \Phi$ depends on n, m, σ, σ' only through $A, \sigma m + \sigma' n, \sigma, \sigma'$, same for $\sigma L(m) + \sigma' L(n)$. □

Proposition 11.17. (Quasi–Töplitz) *Let*

$$\vec{p} = (r, s, K, \theta, \mu, \lambda, \varepsilon^2\mathfrak{K}_\alpha),\ \vec{p}\,' = (2r, 2s, K, \theta', \mu', \lambda, \varepsilon^2\mathfrak{K}_\alpha)$$

be admissible parameters and

$$(\mu' - \mu)K^3 > K\,, \qquad (\theta - \theta')K^{\tau_1} > 2d\kappa > s^{-1}\,, \qquad K^{\tau_1} 2^{-\frac{K}{8\kappa d}+2} < 1\,. \tag{227}$$

If $F \in \mathcal{Q}^T_{\vec{p}\,'}$, then $f := F \circ \Psi \in \mathcal{Q}^T_{\vec{p}}$ and

$$\|X_f\|^T_{\vec{p}} \lessdot e^{2d\kappa s}\|X_F\|^T_{\vec{p}\,'}\,. \tag{228}$$

Proof. Consider $N \geq K$ and suppose that F has no N, μ'–low terms. In this case the proof is identical to that of Proposition 11.5 provided we use the corresponding Lemmata of this section. We conclude the proof by noting that $\Pi_{(N,\theta,\mu,\tau)}(\Pi^L_{N,\mu'} F \circ \Psi) \in \mathcal{T}_{N,\theta,\mu,\tau}$ by Lemma 11.15. Hence in this case the Töplitz defect is zero. □

11.18. **The final step.** In the final step we diagonalize block by block the matrices following Theorem 4. The linear change Ξ has a *finite block structure* in the sense that the Hilbert space $\ell^{a,p}$ is decomposed into an orthogonal sum of subspaces $V_{\mathcal{A},a}$ indexed by the combinatorial pairs $\mathcal{A}, a$ and, if we write a vector as a finite vector with coordinates in these subspaces, the linear transformation Ξ is given by the finite matrix $U_{\mathcal{A}} = (U_{i,j}(\xi))$ with entries depending on ξ and uniformly bounded by some value U, see Remark5.2. Denote by $\Xi^* : F \mapsto F \circ \Xi$ the map induced on functions. One may trivially see that the majorant norm

$$\sup_{\xi\in\varepsilon^2\mathfrak{K}_\alpha}\|M\Xi z'\|^2_{s,r} \leq \sum_{i\in S^c}\Big(\sum_{j:\,\mathtt{r}(j)=\mathtt{r}(i)}\sup_{\xi\in\varepsilon^2\mathfrak{K}_\alpha}|U_{i,j}(\xi)||z'_j|\Big)^2 e^{2a|i|}|i|^{2p} \leq 2^{2p}(d{+}1)^2U^2e^{4d\kappa|a|}|z'|^2_{s,r}\,.$$

We now restrict to thet domain $\mathcal{O}_0 = \varepsilon^2\mathfrak{K}_\alpha$ of measure of order ε^{2n}, which is all contained in one of the connected components of Theorem 4. Recall that one of the domains $\mathfrak{K}_e$ is contained in the elliptic region.

Using the bounds of Remark 5.2 and passing to the majorant norm for vector fields we have

$$\|X_{\Xi^* f}\|^\lambda_{s,r} \leq \mathtt{A}\|X_f\|^\lambda_{s,Cr}\,, \quad \mathtt{A} := (d+1)2^pU\,e^{2d\kappa|a|}\,. \tag{229}$$

Next we need to control the Töplitz norms. We remark that, since we are making linear transformations among variables z_k which have the same root, any monomial in these variables is replaced by a homogeneous sum of monomials in the new variables, all of which have the same root-momentum, so the space $\mathcal{L}_{s,r}(N, \mu, h)$ is mapped into itself. The bilinear functions $\mathcal{B}_{\underline{p}}$, with $\underline{p} := (N, \theta, \mu, \tau)$ are mapped in $\mathcal{B}_{\underline{p}^1}$, with $\underline{p}^1 := (N, \theta^1, \mu^1, \tau)$, provided that $\theta, \theta^1, \mu, \mu^1$ are parameters which satisfy the neighborhood lemma 7.22, so that if m has a $\underline{p}$–cut, also $m + u$ has a $\underline{p}^1$–cut for all possible types $u \in \mathcal{Z}$. The new estimate on parameters that we need is, using Formula (54) for $r - m \in \mathcal{Z}$ is:

$$2d\kappa < \min(\kappa^{-1}(\mu^1 - \mu)N^{\tau-1},\ \kappa^{-1}(\theta - \theta^1)N^{4d\tau-1}). \tag{230}$$

We now claim that for more restricted parameters p' we have that $\Xi^*\mathcal{Q}^T_{\underline{p}}$ is contained in $\mathcal{Q}^T_{\underline{p}'}$. it remains to understand what happens to the space $\mathcal{T}_{\underline{p}}$ we claim that $\Pi_{\underline{p}'}\Xi^*\mathcal{T}_{\underline{p}} \subset \mathcal{T}_{\underline{p}'}$. Take thus a function $g = \sum_{m,n}^{(A,\underline{p})} \mathtt{g}(\sigma m + \sigma' n) z^\sigma_m z^{\sigma'}_n$ as in Formula (70). We have that $A^g_{\underline{p}}$ is contained in some stratum $\Sigma_{\mathcal{A},a}$ for some combinatorial pair $\mathcal{A}, a$. For the space B associated to n we have that $B^g_{\underline{p}}$ is contained in a stratum $\Sigma_{\mathcal{B},b}$. Note that the pair $\mathcal{B}, b$ is determined by $\mathcal{A}, a$ and $\sigma m + \sigma' n$.

Now the change of variables Ξ acts on z_m giving a linear combination of z_{m-u_a+u} where u_a is the type of m and u runs over the types appearing in $\mathcal{A}$ similarly for $\mathcal{B}$.

Consider

(231)
$$\Pi_{\underline{p}'}\Xi^* g = \sum_{m,n}^{(A,\underline{p})} \Pi_{\underline{p}'}\Xi^*\mathtt{g}(\sigma m + \sigma' n) \sum_{v\in\mathcal{A},\ k_1 = u_v - u_a} U_{\mathcal{A}}(\xi)_{a,v} z^\sigma_{m+k_1} \sum_{w\in\mathcal{B}\ ,\ k_2 = u_w - u_b} U_{\mathcal{B}}(\xi)_{b,w} z^{\sigma'}_{n+k_2}.$$

We have already remarked that $\Xi^*\mathtt{g}(\sigma m + \sigma' n) \in \mathcal{L}_{s,r}(N,\mu,h)$. Formula (231) gives a sum $\sum_{m',n'} g^{\sigma,\sigma'}_{m',n'} z^\sigma_{m'} z^{\sigma'}_{n'}$ where both m', n' have a $\underline{p}'$ cut and either the associated space of m' precedes that of n' or the opposite case occurs, moreover $m' \in \Sigma_{\mathcal{A},v},\ n' \in \Sigma_{\mathcal{B},w}$. Reordering Formula (231) it is easily seen that the coefficient

$$g^{\sigma,\sigma'}_{m',n'} = U_{\mathcal{A}}(\xi)_{a,v} U_{\mathcal{B}}(\xi)_{b,w} \mathtt{g}(\sigma(m' - k_1) + \sigma'(n' - k_2)) =$$
$$U_{\mathcal{A}}(\xi)_{a,v} U_{\mathcal{B}}(\xi)_{b,w} \mathtt{g}(\sigma m' + \sigma' n' - \sigma k_1 - \sigma' k_2)$$

depends only upon $\sigma m' + \sigma' n'$ and v hence the claim. Indeed the only thing to make explicit is how to remove the restriction $m' - k_1 \in A^g_{\underline{p}}$. This follows from the estimate on the parameters $\underline{p}'$ which ensures that, if m', n' have a $\underline{p}'$ cut at ℓ then the vectors $m' - k_1, m' - k_2$ have a $\underline{p}$ cut at ℓ. This we do as usual by the neighborhood Lemma noticing that $|k_1|, |k_2| \le 2d\kappa$ since they are differences of two elements in $\mathcal{Z}$ (cf. Remark 7.30). So the requirement is by (54):

$$2d\kappa < \min(\kappa^{-1}(\mu - \mu')N^{\tau-1},\ \kappa^{-1}(\theta' - \theta)N^{4d\tau-1}.$$

Summarizing we have performed 4 changes of coordinates called $\Psi^{(1)}, \Phi_\xi, \Psi, \Xi$. The final Hamiltonian is thus $H \circ \Psi^{(1)} \circ \Phi_\xi \circ \Psi \circ \Xi$, this by definition is *The Hamiltonian of the NLS in the final coordinates.* Recall that the perturbation P refers to the Hamiltonian $H \circ \Psi^{(1)} \circ \Phi_\xi = \mathcal{N} + P$ (Definition 2.2) which by abuse of notation we have still called H.

Proposition 11.19. *The perturbation of the Hamiltonian of the NLS in the final coordinates is quasi-Töplitz for the parameters* $\vec{p}_0 = (r_0 = r/(2\mathtt{A}), s_0 = s/4, \theta_0 = \mathtt{C}/2, \mu_0 = 2\mathtt{c}, K_0 > N_0, \lambda = 2\varepsilon^2, \mathcal{O} = \varepsilon^2\mathfrak{K}_\alpha)$. *We have the bounds:*

$$\|X_{P\circ\Psi\circ\Xi}\|^T_{\vec{p}_0} \le C(\varepsilon r + \varepsilon^5 r^{-1})\,, \tag{232}$$

Proof. By Corollary 11.6 we have that P is quasi-Töplitz with parameters $s, r, K, \theta = \mathtt{C}\frac{3}{4}, \mu = \mathtt{c}\frac{5}{4}, 2\varepsilon^2, \varepsilon^2\mathfrak{K}_\alpha)$. Since, by Formula (50):

$$(\frac{3}{2}\mathtt{c} - \frac{5}{4}\mathtt{c})N_0^2 > 4d\kappa^2\,, \quad (\frac{3}{4} - \frac{3}{8})\mathtt{C}N_0^{\tau_1} > 4d\kappa^2\,, \quad N_0^{\tau_1} 2^{-\frac{N_0}{2\kappa}+1} < 1$$

we apply Proposition 11.17 and obtain the desired bounds for $P \circ \Psi$ with the parameters $(s/2, r/2, \theta = \frac{3}{8}\mathtt{C}, \mu = \frac{3}{2}\mathtt{c}, K > N_0, 2\varepsilon^2, \mathcal{O}_0)$. Then we apply the last change of variables Ξ we need to satisfy again the neighborhood Lemma and reduce the parameters to $\theta_0 =$

$\mathtt{C}/2, \mu_0 = 2\mathtt{c}$ moreover we reduce the analyticity radius by $\frac{1}{\mathtt{A}}$. We obtain the desired result. $\square$

11.20. **Final conclusions: solutions of the NLS.**

Proposition 11.21. *The Hamiltonian of the NLS in the final coordinates is a compatible Hamiltonian in the sense of Definition 10.2 and satisfies the hypotheses of Theorem 7 provided we choose $r = \varepsilon^2$ and ε small.*

Proof. The fact that it satisfies the smallness condition in Theorem 7 follows from (232). We need to verify all the conditions $(A1)-(A5)$.
$(A1)$ *Non–degeneracy:* The map $\xi \to \omega(\xi)$ is $\xi \mapsto \mathtt{v} - 2\xi$ so it is a lipeomorphism from $\mathcal{O}_0$ to its image with $|\omega^{-1}|_\infty^{lip} \le 1$. We have $|\omega(\xi) - \mathtt{v}| \le \varepsilon^2$ since by assumption $|\xi| \le \varepsilon^2$.

$(A2)$ *Asymptotics of normal frequency:* For all $n \in S^c$ we have a decomposition:

$$\Omega_n(\xi) = \sigma(n)(|\mathtt{r}(n)|^2 + 2\vartheta_n(\xi)). \tag{233}$$

In our case we start from $\tilde\Omega_n(\xi) = 0$. We know that the $\vartheta_n(\xi)$ are in a finite list of analytic functions which are homogeneous of degree one in ξ. As for (120), by homogeneity, we can fix $M \ge 1$ so that $2 + 2|\vartheta|_\infty^{lip} \le M\,, \quad 2|\vartheta|_\infty \le M\varepsilon^2$. This fixes the parameters M, L, then we fix K_0 large enough (independently form r_0) and choose $\gamma < \min(2M, B)\varepsilon^2$ with $B = B(K_0, \mathcal{O}_0)$ given in the iterative Lemma.

$(A3)$ *Regularity and Quasi–Töplitz property:* The function $\tilde\Omega(z) = 0$. The functions P, $\vartheta(z) := \sum_j \vartheta_j |z_j|^2$ are M–regular, preserve momentum as in (44), are Lipschitz in the parameters ξ. Then P is quasi-Töplitz with parameters $(s_0, r_0, K_0, \theta_0, \mu_0, \gamma/M, \varepsilon^2\mathfrak{K}_\alpha)$ by Proposition 11.19, with the bounds (232). Moreover we know that the functions ϑ_i are constant of the strata $\Sigma_{\mathcal{A},a}$ of §7.26 hence $\vartheta(z) := \sum_j \vartheta_j|z_j|^2$ is regular, preserve momentum and is quasi–Töplitz and for all $N \ge K_0$ $\tau_0 \le \tau \le \tau_1/4d$ we have $\Pi_{(N,\theta,\mu,\tau)} \sum_j \vartheta_j |z_j|^2 \in \mathcal{T}_{(N,\theta,\mu,\tau)}$.

$(A4)$ *Smallness condition*: We compute $\Theta, |\vec\epsilon| \le C\varepsilon^3\gamma^{-1}$ by (232). The condition

$$\Theta < 1\,,\ LM|\vec\epsilon| < 1\,,\ \gamma^{-1}\|X_{\tilde\Omega}\|_{\vec p}^{T} < 1\,,\ \kappa e|\vec\epsilon| K_0^{4d\tau_1} \ll 1. \tag{234}$$

This translates to the condition (188) in which $\alpha = 3$ by (205) with $r = \varepsilon^2$. Finally we note that the third condition in $(A4)$ is trivial since $\tilde\Omega = 0$.

$(A5)$ *Non–degeneracy (Melnikov conditions):* For all $(k, l) \ne 0$ compatible with momentum conservation the function $\langle\omega, k\rangle + (\Omega, l)$ is of the form $\langle \mathtt{v}, k\rangle + (\mathtt{V}, l) - 2\sum_i k_i\xi_i + 2\theta$ where θ can be $0, \pm\vartheta_j, \pm\vartheta_j \pm \vartheta_k$. From the main result of [20] we know that all the functions $\sum_i k_i\xi_i + \theta$ are analytic, homogeneous of degree 1 and different from 0 and give distinct eigenvalues on distinct blocks. Hence in each connected component $(\mathbb{R}_+)^n_\alpha$ of $(\mathbb{R}_+)^n \setminus \mathfrak{A}$ we can choose a compact domain $\mathfrak{K}_\alpha$ which does not intersect any of the zero curves of the functions $(\langle\omega, k\rangle + (\Omega, l))0$ for $|k| < 16\sqrt{n}$, this amunts, as explained in Remark 6.6, to taking the $\mathfrak{K}_\alpha$ disjoint from finitely many hypersurfaces describing the resultants given in that Remark. Now fix $\mathfrak{K}_\alpha$, for each k, l as above such that $\langle \mathtt{v}, k\rangle + (\mathtt{V}, l) = 0$ we have that $\langle\omega, k\rangle + (\Omega, l)$ is homogeneous of degree one and non zero. Then for all positive ρ such that $\rho\xi \in \mathfrak{K}_\alpha$ we have

$$\frac{|\langle\omega(\rho\xi), k\rangle + (\Omega(\rho\xi), l) - (\langle\omega(\xi), k\rangle + (\Omega(\xi), l))|}{(|1-\rho||\xi|} = \frac{\langle\omega, k\rangle + (\Omega, l)}{|\xi|} \ge a > 0$$

for some positive $a := a(\alpha)$. We repeat the same argument for all α and we choose $\mathcal{O}_0 = \varepsilon^2 \mathfrak{K}$.

$\square$

For compatible Hamitonians we have proved in the previous section a general Theorem 7 which ensures the existence of KAM–tori, now we can apply this Theorem to the NLS and have as final result Theorem **??** .

References

[1] M. Berti and L. Biasco. Branching of Cantor manifolds of elliptic tori and applications to PDEs. *Comm. Math. Phys.*, 305(3):741–796, 2011.

[2] M. Berti, L. Biasco, and M. Procesi. KAM theory for the Hamiltonian derivative wave equation. *Annales Scientifiques de l'ENS*, 46(2), 2013.

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

[4] M. Berti and Ph. Bolle. Quasi-periodic solutions Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential. to appear on Eur. Jour. Math.

[5] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math. (2)*, 148(2):363–439, 1998.

[6] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.*, 181(1):39–113, 2010.

[7] Walter Craig and C. Eugene Wayne. Newton's method and periodic solutions of nonlinear wave equations. *Comm. Pure Appl. Math.*, 46(11):1409–1498, 1993.

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

[9] J. Geng and Y. Yi. Quasi-periodic solutions in a nonlinear schrödinger equation. *J. Differential Equations*, 233:512–542, 2007.

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

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

[12] G. Gentile and M. Procesi. Periodic solutions for a class of nonlinear partial differential equations in higher dimension. *Comm. Math. Phys.*, 289(3):863–906, 2009.

[13] Guido Gentile and Michela Procesi. Periodic solutions for the Schrödinger equation with nonlocal smoothing nonlinearities in higher dimension. *J. Differential Equations*, 245(11):3253–3326, 2008.

[14] B. Grébert and L. Thomann. Resonant dynamics for the quintic nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(3):455–477, 2012.

[15] Sergej Kuksin and Jürgen Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math. (2)*, 143(1):149–179, 1996.

[16] Massimiliano Berti Pietro Baldi and Riccardo Montalto. KAM for quasi-linear and fully nonlinear forced KdV. arXiv:1211.6672, 2012.

[17] J. Pöschel. A KAM-theorem for some nonlinear partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 23(1):119–148, 1996.

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

[19] M. Procesi. A normal form for beam and non-local nonlinear Schrödinger equations. *Journal of Physics A: Mathematical and Theoretical*, 43(43):434028, 2010.

[20] M. Procesi, C. Procesi, and Nguyen Bich Van. The energy graph of the non–linear Schrödinger equation. *Rend. Lincei Mat. Appl.*, 24:1–73, 2013.

[21] W. M. Wang. Quasi-periodic solutions of the Schrödinger equation with arbitrary algebraic nonlinearities. Preprint, arXiv:0907.3409.

[22] W.-M. Wang. Supercritical Nonlinear Schrödinger equations I: Quasi-Periodic Solutions. arXiv:1007.0156, 2010.

[23] John Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *Amer. J. Math.*, 58 (1):141–163, 1936.

[24] X. Xu and M. Procesi. Quasi-Töplitz Functions in KAM Theorem. Preprint 2010, to appear in *SIAM J. of Math. Anal.*