

QUASI-TÖPLITZ FUNCTIONS IN KAM THEOREM*

MICHELA PROCESI[†] AND XINDONG XU[‡]

Abstract. We define and describe the class of quasi-Töplitz functions. We then prove an abstract KAM theorem where the perturbation is in this class. We apply this theorem to a nonlinear Schrödinger equation on the torus \mathbb{T}^d , thus proving existence and stability of quasi-periodic solutions.

Key words. Schrödinger equation, KAM theory, quasi-Töplitz functions

AMS subject classifications. 37K55, 35Q55, 37J40, 70H08, 70K43

DOI. 10.1137/110833014

1. Introduction. In this paper, we study a model nonlinear Schrödinger equation (NLS) with external parameters on the torus \mathbb{T}^d , and prove existence and stability of quasi-periodic solutions. In order to do this we introduce a new class of functions, which we denote as quasi-Töplitz. We focus on the equation

$$(1.1) \quad iu_t - \Delta u + \mathbb{M}_\xi u + f(|u|^2)u = 0, \quad x \in \mathbb{T}^d, t \in \mathbb{R},$$

where $f(y)$ is a real analytic function with $f(0) = 0$, while \mathbb{M}_ξ is a Fourier multiplier, namely, a linear operator which commutes with the Laplacian, and whose role is to introduce b parameters in order to guarantee that (1.1), linearized at $u = 0$, admits a quasi-periodic solution with b frequencies. More precisely we choose a finite set $\{\mathbf{n}^{(1)} = 0, \mathbf{n}^{(2)}, \dots, \mathbf{n}^{(b)}\}$ with $\mathbf{n}^{(i)} \in \mathbb{Z}^d$, and define \mathbb{M}_ξ so that the eigenvalues of the operator $\Delta + \mathbb{M}_\xi$ are

$$(1.2) \quad \begin{cases} \omega_j = |\mathbf{n}^{(j)}|^2 + \xi_j, & 1 \leq j \leq b, \\ \Omega_n = |n|^2, & n \notin \{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(b)}\}. \end{cases}$$

Equation (1.1) is a well-known model for the natural NLS, in which the Fourier multiplier is substituted by a multiplicative potential V . Existence and stability of quasi-periodic solutions of (1.1) via a KAM algorithm was proved in [11] for the more general case where $f(y)$ is substituted for with $f(y, x)$, $x \in \mathbb{T}^d$. With respect to that paper we use a different approach to prove measure estimates, based essentially on two ingredients: the fact that the equation has the total momentum $M = \int_{\mathbb{T}^d} \bar{u} \nabla u$ as an integral of motion, and the use of the properties of the quasi-Töplitz functions. These two ideas induce some significant simplifications which we think are interesting, in particular the conservation of momentum enables us to prove a stronger result, namely, our solutions are analytic while in [11] only Gevrey class is proven. Our dynamical result for the NLS (1.1) is the following theorem.

THEOREM 1.1. *There exists a positive-measure Cantor set \mathcal{C} such that for any $\xi = (\xi_1, \dots, \xi_b) \in \mathcal{C}$, the NLS equation (1.1) admits small amplitude analytic quasi-*

*Received by the editors May 5, 2011; accepted for publication (in revised form) March 21, 2013; published electronically July 3, 2013. This research was supported by the European Research Council under FP7, project Hamiltonian PDEs.

<http://www.siam.org/journals/sima/45-4/83301.html>

[†]Università di Roma La Sapienza, 00185 Roma, Italy (mprocesi@mat.uniroma1.it).

[‡]Università degli Studi di Napoli “Federico II,” Napoli 80126, Italy, and Department of Mathematics, Southeast University, Nanjing 210096 China (xindong.xu@seu.edu.cn, xuxdnju@gmail.com).

periodic solutions. The solutions are linearly stable, and we give a reducible normal form close to them.

This is obtained by proving that the NLS Hamiltonian fits the hypotheses of an abstract KAM theorem; see Theorem 5.1.

Before describing our results, and techniques in more detail, let us make a very brief excursus on the literature on quasi-periodic solutions for PDEs on \mathbb{T}^d , and on the general strategy of a KAM algorithm.

The existence of quasi-periodic solutions for (1.1) (as well as for the nonlinear wave equation) was first proved by Bourgain (see [6], [8], and [7]), by applying a combination of Lyapunov–Schmidt reduction and Nash–Moser generalized implicit function theorem in order to solve the small divisor problem. This method is very flexible, and may be effectively applied in various contexts, for instance in the case where $f(y)$ has only finite regularity; see [2] and [3]. As a drawback this method only establishes existence of the solutions but does not give information on the linear stability. In order to achieve this stronger result it is natural to extend to (1.1) the, by now classical, KAM techniques which were developed to study (1.1) with Dirichlet boundary conditions on the segment $[0, \pi]$, see [17]–[24]. A fundamental hypothesis in the aforementioned algorithms is that the eigenvalues Ω_n are simple, and this is clearly not satisfied in the case of (1.1) on \mathbb{T}^1 , where the eigenvalues are double. We mention that this hypothesis was weakened for the nonlinear wave equation by Chierchia and You in [9], by only requiring that the eigenvalues have finite and uniformly bounded multiplicity. Their method however does not extend trivially to the NLS on \mathbb{T}^1 , and surely may not be applied to the NLS in higher dimensions, where the multiplicity of Ω_n is of order $\Omega_n^{(d-1)/2}$. The first result on KAM theory on the torus \mathbb{T}^d was given in [12] for the *nonlocal* NLS:

$$iu_t - \Delta u + M_\xi u + f(|\Psi_s(u)|^2)\Psi_s(u) = 0, \quad x \in \mathbb{T}^d, t \in \mathbb{R},$$

where Ψ_s is a linear operator, diagonal in the Fourier basis, and such that $\Psi_s(e^{i\langle n, x \rangle}) = |n|^{-2s} e^{i\langle n, x \rangle}$ for some $s > 0$. The key points of that paper are (1) the use of conservation of total momentum to avoid the problems arising from multiplicity of the Ω_n , and (2) the fact that the presence of the nonlocal operator Ψ_s simplifies the proof of the *Melnikov nonresonance conditions* throughout the KAM algorithm. As we mentioned before, the more complicated problem of a KAM algorithm for the local NLS without momentum conservation was solved by Eliasson and Kuksin in [11].

Let us briefly describe the general strategy in the KAM algorithm for (1.1).

We expand the solution in Fourier series as $u = \sum_{n \in \mathbb{Z}^d} u_n \phi_n(x)$; here $\phi_n(x) = \sqrt{\frac{1}{(2\pi)^d}} e^{i\langle n, x \rangle}$ with $n \in \mathbb{Z}^d$ is the standard Fourier basis. Then we introduce standard action-angle coordinates for the modes \mathbf{n}_j by setting $u_{\mathbf{n}_j} = \sqrt{I_j^{(0)} + I_j} e^{i\vartheta_j}$, $j = 1, \dots, b$, where the $I_j^{(0)}$ are arbitrary sufficiently small numbers. Finally we set $u_n = z_n = z_n^+$, $\bar{u}_n = \bar{z}_n = z_n^-$ for all $n \neq \{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(b)}\}$. We get

$$(1.3) \quad H = \sum_{1 \leq j \leq b} \omega_j(\xi) I_j + \sum_{n \in \mathbb{Z}_1^d} \Omega_n |z_n|^2 + P(I, \vartheta, z, \bar{z}), \quad \mathbb{Z}_1^d := \mathbb{Z}^d \setminus \{\mathbf{n}_1, \dots, \mathbf{n}_b\}.$$

It is easily seen that H , and hence P , preserve the total momentum (see (2.5) below); moreover, P (and $\sum (\Omega_m - |m|^2) z_m \bar{z}_m$) are Töplitz/anti-Töplitz functions, namely, the Hessian matrix $\partial_{z_m^\sigma} \partial_{z_n^{\sigma'}} P$ depends on $z_m^\sigma, z_n^{\sigma'}$ only through $\sigma m + \sigma' n$.

Informally speaking the KAM algorithm consists in constructing a convergent sequence of symplectic transformations Φ_ν such that

$$(1.4) \quad \Phi_\nu \circ H := H_\nu = \sum_{1 \leq j \leq b} \omega_j^{(\nu)}(\xi) I_j + \sum_{n \in \mathbb{Z}_1^d} \Omega_n^{(\nu)}(\xi) |z_n|^2 + P_\nu(\xi, I, \vartheta, z, \bar{z}),$$

where $P_\nu \rightarrow 0$ in some appropriate norm. The symplectic transformation is well defined for all ξ which satisfy the *Melnikov nonresonance conditions*

$$(1.5) \quad |\langle \omega^{(\nu)}, k \rangle + \Omega^{(\nu)} \cdot l| \geq \gamma K_\nu^{-\varrho}$$

for all $k \in \mathbb{Z}^b$, $l \in \mathbb{Z}_1^d$ such that $(k, l) \neq (0, 0)$, $|l| \leq 2$, and $|k| < K_\nu$. Here ϱ, γ are appropriate constants. With these conditions in mind it is clear that a degeneracy $\Omega_n^{(\nu)} = \Omega_m^{(\nu)}$ poses problems since the left-hand side in (1.5) is identically zero for $k = 0, l = e_m - e_n$ (e_m with $m \in \mathbb{Z}_1^d$ is the standard basis vector). To avoid this problem we use the fact that all the H_ν have M as the constant of motion. This in turn implies that some of the Fourier coefficients of P_ν are identically zero so that the conditions (1.5) need to be imposed only on those k, l such that $\sum_{i=1}^b \mathbf{n}_i k_i + \sum_{m \in \mathbb{Z}_1^d} m l_m = 0$. Then, in our example, $k = 0$ automatically implies $n = m$. This is the key argument used in [12]. However, once that one has proved that the left-hand side of (1.5) is never identically zero, one still has to show that the quantitative bounds of (1.5) may be imposed on some positive measure set of parameters ξ . This is an easy task when $|l| = 0, 1$ or $l = e_m + e_n$ but may pose serious problems in the case $l = e_m - e_n$, where the nonresonance condition is of the form

$$(1.6) \quad |\langle \omega^{(\nu)}, k \rangle + \Omega_m^{(\nu)} - \Omega_n^{(\nu)}| \geq \gamma K_\nu^{-\varrho} \quad \forall k \in \mathbb{Z}^b, n, m \in \mathbb{Z}_1^d : |k| < K_\nu,$$

where $n - m = \sum_{i=1}^b \mathbf{n}_i k_i$. Indeed in this case for every fixed value of k one should in principle impose infinitely many conditions, since the momentum conservation only fixes $n - m$. In [12], the presence of Ψ_s implies that $\Omega_m^{(\nu)} - |m|^2 \approx \frac{\varepsilon}{|m|^s}$ so that if $|m|^s > c|k|^\tau$ the variation of $\Omega^{(\nu)}$ is negligible. This implies in turn that one has to impose only finitely many conditions for each k . In the case of (1.1), however, $s = 0$, so that this argument may not be applied. One wishes to impose the nonresonance conditions by verifying only a finite number of bounds for each k . To do this one needs some control on $\Omega_m^{(\nu)} - |m|^2$, for $|m|$ large, throughout the KAM algorithm. The ideal setting is when $\Omega_m^{(\nu)} - |m|^2$ is m -independent. This holds true for the first step of the KAM algorithm due to the fact that P is a Töplitz function. However it is easily seen that already P_1 is not a Töplitz function, and some wider class of functions must be defined.

In order to control the shift of the normal frequency, Eliasson and Kuksin in [11] define a *Töplitz-Lipschitz property*, which they show is satisfied by the NLS Hamiltonian, and preserved through the KAM iteration. With this property, they prove the existence of KAM tori. As a further difficulty they consider an NLS which does not have M as a constant of motion. This implies that some of the Melnikov nonresonance conditions (1.6) may not be imposed. At each step of the KAM algorithm they thus obtain a more complicated normal form.

In order to describe the Töplitz-Lipschitz property, given an analytic function $A(z, \bar{z})$, let $A_m^n(\pm) = \partial_{z_m} \partial_{z_n^\pm} A$ be its Hessian matrix. For all $n, m, c \in \mathbb{Z}^d$, one requires that the limit $A_m^n(\pm, c) := \lim_{t \rightarrow \infty} A_{m+tc}^{n+tc}(\pm)$ exists, and is attained with speed of order $\frac{1}{t}$. In dimension $d > 2$ one also requires similar conditions on the

limits $\lim_{s \rightarrow \infty} A_{m+sc'}^{n \mp sc'}(\pm, c)$ with c' orthogonal to c . In [13] an understanding of this property in \mathbb{T}^2 is given. A key step is to divide the region $\{|n - m| \leq N\} \subset \mathbb{Z}^d \times \mathbb{Z}^d$ into a finite number of Lipschitz domains.

In our paper we use a similar—but in our opinion more natural—approach. We define a class of functions, the *quasi-Töplitz functions* whose main properties are

1. the Poisson bracket of two quasi-Töplitz functions is quasi-Töplitz (Proposition 6.8),
2. the Hamiltonian flow generated by a quasi-Töplitz function preserves the quasi-Töplitz property (Proposition 6.8),
3. the solution of the homological equation with a quasi-Töplitz perturbation is quasi-Töplitz (Proposition 6.7).

Note that the Töplitz–Lipschitz property of [11] is closed only with respect to Poisson brackets when one of the functions is quadratic; this makes our definitions more flexible.

In this paper we strongly rely on the conservation of momentum for our definitions; however, this condition is not necessary in order to define the quasi-Töplitz functions; see, for instance, [4]. In the next paragraph we give a brief informal description of our method.

1.1. Brief description of the strategy. We start by fixing two Diophantine exponents $\tau_0 \ll \tau_1$. All our definitions, and constructions are based on some parameters $N \gg 1$, $\frac{1}{2} < \theta, \mu < 4$, and $\tau_0 \leq \tau \leq \tau_1/4d$ which are needed in order to ensure that the quasi-Töplitz functions are closed with respect to Poisson brackets (with slightly different parameters).

The first step in our construction is an intrinsic (and unique) description of affine subspaces described by equations with integer coefficients. We consider the equations $v_i \cdot x = p_i$, $i = 1, \dots, \ell$, $x, v_i \in \mathbb{Z}^d$, $p_i \in \mathbb{Z}$ describing the set of integral points x in an affine subspace; we then denote this set by $[v_i; p_i]_\ell$ and, by abuse of notation, call it an affine subspace. Given $N \gg 1$, an N -optimal presentation of an affine subspace of codimension ℓ is a (uniquely fixed if it exists) list $[v_i; p_i]_\ell$ such that $|v_i| < C_1 N$, and the p_i are *positive, ordered, and as small as possible* (see Definition 3.4).

This decomposition holds also for a single point (when $\ell = d$, in this case an N -optimal presentation will surely exist). Then we use the parameters $\frac{1}{2} < \theta, \mu < 4$, $\tau_0 \leq \tau \leq \tau_1/4d$ to define the notion of ℓ cut for a point m , and of *good points* of an affine subspace with respect to the parameters (N, θ, μ, τ) . Namely, if $[v_i; p_i]_\ell$ is the N -optimal presentation of m , then m has a cut at ℓ if $p_\ell < \mu N^\tau$ and $p_{\ell+1} > \theta N^{4d\tau}$. In the same way the (N, θ, μ, τ) -good points of an affine subspace $[v_i; p_i]_\ell$, with $p_\ell < \mu N^\tau$ are those points of $[v_i; p_i]_\ell$ which have a cut at ℓ with parameters (N, θ, μ, τ) (see Definition 3.6).

We then define the (N, θ, μ, τ) -bilinear functions, i.e., functions which are bilinear in the *high variables* $z_m^\sigma, z_n^{\sigma'}$ such that $|m|, |n| > \theta N^{\tau_1}$, and both m , and n have a cut with parameters (N, θ, μ, τ) . These functions may depend on I, ϑ , and on the *small variables* z_j^σ with $|j| < \mu N^3$ in a possibly complicated way (see Definition 4.1 for a precise statement).

Finally we define the *piecewise Töplitz functions* as those (N, θ, μ, τ) -bilinear functions which are Töplitz when restricted to the (N, θ, μ, τ) -good points of any affine subspace (see Definition 4.2, and Remark 4.1).

We can now define the (K, θ, μ) -quasi-Töplitz functions. Informally speaking given a function f , for all $N > K$, $\tau_0 \leq \tau \leq \tau_1/4d$, we project it on the (N, θ, μ, τ) -bilinear functions, and we say that f is quasi-Töplitz if all these projections are *well*

approximated by a piecewise Töplitz function. To be more precise, τ controls the size of the error function, namely, the (N, θ, μ, τ) -bilinear part of f is approximated by a piecewise Töplitz function with an error of the order $N^{-4d\tau}$, for all $N \geq K$ (see (4.5), and Definition 4.3).

The role of the parameters K, θ, μ is to ensure that if f, g are quasi-Töplitz with parameters K, θ, μ then $\{f, g\}$ is quasi-Töplitz for all $\theta' > \theta$ and $\mu' < \mu$ provided $K' > K$ is large enough (see Proposition 6.8).

We proceed by induction supposing that we have been able to perform ν KAM iterative steps, and that we have a Hamiltonian of the form (1.4), where $\sum_m (\Omega_m^{(\nu)} - |m|^2) |z_m|^2$ is quasi-Töplitz with parameters $(K_\nu, \theta_\nu, \mu_\nu)$ (note that K_ν is the ultraviolet cut-off at step ν). In order to solve the homological equation (and hence pass to step $\nu + 1$) we restrict to the subset of ξ for which (1.5) holds for all k, m, n (satisfying momentum conservation) for some $\varrho := \varrho(k, m, n) < 2d\tau_1$. The main point is to show that this restriction on the parameters only removes a small measure set.

For all natural $N \geq K_\nu$ we introduce a decomposition of \mathbb{Z}_1^d as

$$(1.7) \quad \mathbb{Z}_1^d := A_0 \cup \left(\bigcup_{\ell=1}^{d-1} A_\ell \right) \cup \{|m| \leq 4N^{\tau_1}\};$$

here, $A_0 \equiv A_0(N)$ is \mathbb{Z}_1^d minus a finite number of affine hyperplanes while $A_\ell := A_\ell(N)$ is the union of a finite number of affine spaces of codimension ℓ minus a finite number of affine spaces of codimension $\ell + 1$ (see Figure 3.1 for a picture in $d = 2$).

This decomposition is constructed as follows:

A_0 (defined in (3.8)) is chosen so that for all $|k| < N, m \in A_0$ the Melnikov denominators (1.6) are *not small*.

For all $0 < \ell < d$ we may write

$$A_\ell := \bigcup_{\substack{v_1, \dots, v_\ell \in \mathbb{Z}_1^d, p_1, \dots, p_\ell \in \mathbb{Z} \\ |v_i| < C_1 N, p_i < 4N^{\tau_1/4d}}} [v_i; p_i]_\ell^g,$$

where the $[v_i; p_i]_\ell^g \subset [v_i; p_i]_\ell$ (see Definition 3.9) are defined in order to ensure the following property: fix $\tau(p_\ell)$ by setting $N^\tau = \max(2p_\ell, N^{\tau_0})$; we have that all $m \in [v_i; p_i]_\ell^g$ are $(N, \theta, \mu, \tau(p_\ell))$ -good points for $[v_i; p_i]_\ell$ for all choices of $\frac{1}{2} < \theta, \mu < 4$ —this is the content of Lemma 3.10. Finally the fact that these sets provide a decomposition of \mathbb{Z}_1^d is the content of Proposition 3.11.

To prove the measure estimates we use the above decomposition with $N = K_\nu$. Then the quasi-Töplitz property with $N = K_\nu$ implies that for each $m \in [v_i; p_i]_\ell^g$,

$$(1.8) \quad \Omega_m^{(\nu)} = |m|^2 + \hat{\Omega}^{(\nu)}([v_i; p_i]_\ell) + \bar{\Omega}_m^{(\nu)} K_\nu^{-4d\tau(p_\ell)},$$

where $\hat{\Omega}^{(\nu)}$ is constant on all the points of $[v_i; p_i]_\ell$ while $\bar{\Omega}_m^{(\nu)}$ is bounded by ε_0 (see Lemma 4.4). We stress that here¹ $\tau = \tau(p_\ell)$ is fixed by the positive integer p_ℓ .

Roughly speaking, we fix k , choose *one point* m^g on each $[v_i; p_i]_\ell^g$, and impose the Melnikov conditions (1.6) with $\varrho = 2d\tau(p_\ell)$, $\gamma \rightsquigarrow 2\gamma$, and $m = m^g$ (see Definition 6.1(iv) for the precise formulation). This condition, and (1.8) ensure the second Melnikov condition for all $m \in [v_i; p_i]_\ell^g$ with $\varrho = 2d\tau(p_\ell)$ (see Lemma 6.2). This shows

¹Note that in the definition of quasi-Töplitz functions and of cuts, however, τ is left as a free parameter with the only restriction $\tau_0 \leq \tau \leq \tau_1/4d$.

that the infinitely many conditions (1.6) can be imposed by only requiring a finite subset of them.

In order to check the measure estimates we remark that to impose *one* Melnikov condition (i.e., with fixed $k, m \in [v_i; p_i]_\ell^g$, and $\varrho = 2d\tau(p_\ell)$) we need to remove a region of parameter sets of order $K_\nu^{-2d\tau(p_\ell)}$ (see Lemma 6.4). Thus we need to estimate the number of affine spaces $[v_i; p_i]_\ell$ with $p_\ell = p$; using Remark 3.2 it follows that this bound is proportional to $K_\nu^{d\tau(p)} = (2p)^d$. This concludes the problem of measure estimates, and we exclude a set of ξ of measure $\sum_{p \in \mathbb{N}; p > K_\nu^{\tau_0}} (2p)^{-d}$ (here we are giving only an informal argument; see Lemma 6.3 for the complete proof). In order to pass to the step $\nu + 1$ we need F_ν (the solution of the homological equation) to be quasi-Töplitz: this requires a further restriction of the parameter set (see Definition 6.1(iv), Remark 6.3, and Proposition 6.7).

Recalling that quasi-Töplitz functions are closed with respect to Poisson brackets we conclude that the new Hamiltonian is still quasi-Töplitz for some new parameters $\theta_{\nu+1}, \mu_{\nu+1}$ for all $N \geq K_{\nu+1}$.

2. Relevant notation and definitions.

2.1. Function spaces and norms. We start by introducing some notation. We fix b vectors $\{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(b)}\}$ in \mathbb{Z}^d called the *tangential sites*. We denote by $\mathbb{Z}_1^d := \mathbb{Z}^d \setminus \{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(b)}\}$ the complement, called the *normal sites*. Let $z = (\dots, z_n, \dots)_{n \in \mathbb{Z}_1^d}$ and its complex conjugate $\bar{z} = (\dots, \bar{z}_n, \dots)_{n \in \mathbb{Z}_1^d}$. We introduce the weighted norm

$$\|z\|_\rho = \sum_{n \in \mathbb{Z}_1^d} |z_n| e^{|\mathbf{n}|\rho} |n|^{d+1},$$

where $|n| = \sqrt{n_1^2 + n_2^2 + \dots + n_d^2}$, $n = (n_1, n_2, \dots, n_d)$, and $\rho > 0$. We denote by ℓ_ρ the Hilbert space of lists $\{w_j = (z_j, \bar{z}_j)\}_{j \in \mathbb{Z}_1^d}$ with $\|z\|_\rho < \infty$.

We consider the real torus $\mathbb{T}^b := \mathbb{R}^b / \mathbb{Z}^b$ naturally contained in the space $\mathbb{C}^b / \mathbb{Z}^b \times \ell_\rho$ as the subset where $I = z = \bar{z} = 0$. We then consider in this space the neighborhood of \mathbb{T}^b :

$$D(r, s) := \{(I, \vartheta, z, \bar{z}) : |\text{Im } \vartheta| < s, |I| < r^2, \|z\|_\rho < r, \|\bar{z}\|_\rho < r\},$$

where $|\cdot|$ denotes the sup-norm of complex vectors. Denote by \mathcal{O} an open and bounded parameter set in \mathbb{R}^b , and let $D = \max_{\xi, \eta \in \mathcal{O}} |\xi - \eta|$.

We consider functions $F(I, \vartheta, z, \bar{z}; \xi) : D(r, s) \times \mathcal{O} \rightarrow \mathbb{C}$ analytic in I, ϑ, z , and of class C_W^1 in ξ . We expand in Taylor–Fourier series as:

$$(2.1) \quad F(\vartheta, I, z, \bar{z}; \xi) = \sum_{l, k, \alpha, \beta} F_{lk\alpha\beta}(\xi) I^l e^{i\langle k, \vartheta \rangle} z^\alpha \bar{z}^\beta,$$

where the coefficients $F_{lk\alpha\beta}(\xi)$ are of class C_W^1 (in the sense of Whitney), the vectors $\alpha \equiv (\dots, \alpha_n, \dots)_{n \in \mathbb{Z}_1^d}$, $\beta \equiv (\dots, \beta_n, \dots)_{n \in \mathbb{Z}_1^d}$ have finitely many nonzero components $\alpha_n, \beta_n \in \mathbb{N}$, $z^\alpha \bar{z}^\beta$ denotes $\prod_n z_n^{\alpha_n} \bar{z}_n^{\beta_n}$, and finally $\langle \cdot, \cdot \rangle$ is the standard inner product in \mathbb{C}^b .

We use the following weighted norm for F :

$$(2.2) \quad \|F\|_{r,s} = \|F\|_{D(r,s), \mathcal{O}} \equiv \sup_{\substack{\|z\|_\rho < r \\ \|\bar{z}\|_\rho < r}} \sum_{\alpha, \beta, k, l} |F_{kl\alpha\beta}|_{\mathcal{O}} r^{2|l|} e^{|\mathbf{k}|s} |z^\alpha| |\bar{z}^\beta|,$$

$$(2.3) \quad |F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi \in \mathcal{O}} \left(|F_{kl\alpha\beta}| + \left| \frac{\partial F_{kl\alpha\beta}}{\partial \xi} \right| \right).$$

(The derivatives with respect to ξ are in the sense of Whitney.) To an analytic function F , we associate a Hamiltonian vector field with coordinates

$$X_F = (F_I, -F_{\vartheta}, \{iF_{z_n}\}_{n \in \mathbb{Z}_1^d}, \{-iF_{\bar{z}_n}\}_{n \in \mathbb{Z}_1^d}).$$

Consider a vector function $G : D(r, s) \times \mathcal{O} \rightarrow \ell_\rho$ with

$$G = \sum_{kl\alpha\beta} G_{kl\alpha\beta}(\xi) I^l e^{i(k, \vartheta)} z^\alpha \bar{z}^\beta,$$

where $G_{kl\alpha\beta} = (\dots, G_{kl\alpha\beta}^{(i)}, \dots)_{i \in \mathbb{Z}_1^d}$. Its norm is similarly defined as

$$\|G\|_{D(r,s), \mathcal{O}} = \sup_{\substack{\|z\|_\rho < r \\ \|\bar{z}\|_\rho < r}} \|\mathcal{M}G\|_\rho,$$

where

$$\mathcal{M}G = (\dots, \mathcal{M}G^{(i)}, \dots)_{i \in \mathbb{Z}_1^d}, \quad \mathcal{M}G^{(i)} = \sum_{\alpha, \beta, k, l} |G_{kl\alpha\beta}^{(i)}|_{\mathcal{O}} r^{2|l|} e^{|k|s} z^\alpha \bar{z}^\beta$$

is a majorant of $G^{(i)}$. We say that an analytic function F is regular if the function $(z, \bar{z}) \rightarrow \mathcal{M}X_F$ is analytic from $B_r \rightarrow \ell_\rho$. Its weighted norm is defined by²

$$\begin{aligned} \|X_F\|_{r,s} = \|X_F\|_{D(r,s), \mathcal{O}} &\equiv \sum_{j=1}^b \|F_{I_j}\|_{D(r,s), \mathcal{O}} + \frac{1}{r^2} \sum_{j=1}^b \|F_{\vartheta_j}\|_{D(r,s), \mathcal{O}} \\ (2.4) \quad &+ \frac{1}{r} (\|\partial_z F\|_{D(r,s), \mathcal{O}} + \|\partial_{\bar{z}} F\|_{D(r,s), \mathcal{O}}). \end{aligned}$$

A function F is said to satisfy momentum conservation if $\{F, M\} = 0$ with $M = \sum_{i=1}^b \mathbf{n}^{(i)} I_i + \sum_{m \in \mathbb{Z}_1^d} j |z_m|^2$. This implies that

$$(2.5) \quad F_{k,l,\alpha,\beta} = 0 \quad \text{if } \pi(k, \alpha, \beta) := \sum_{i=1}^b \mathbf{n}^{(i)} k_i + \sum_{m \in \mathbb{Z}_1^d} m(\alpha_m - \beta_m) \neq 0.$$

By Jacobi's identity momentum conservation is preserved by the Poisson bracket.

Remark 2.1. It will be useful to envision the conservation of momentum at fixed k as a relation between α, β ; to make this more evident we write

$$(2.6) \quad \pi(k, \alpha, \beta) = 0 \quad \text{as} \quad - \sum_{m \in \mathbb{Z}_1^d} m(\alpha_m - \beta_m) = \sum_{i=1}^b \mathbf{n}^{(i)} k_i := \pi(k).$$

DEFINITION 2.1. We denote by $\mathcal{A}_{r,s}$ the space of regular analytic functions in $D(r, s)$, and C_W^1 in \mathcal{O} which satisfy momentum conservation (2.5), and with finite semi-norm (2.4).

If \mathcal{S} is a set of monomials in $I_j, e^{i\vartheta_j}, z_m, \bar{z}_n$, we define the projection operator $\Pi_{\mathcal{S}}$ which to a given analytic function F associates the part of the series only relative to the monomials in \mathcal{S} .

²The norm $\|\cdot\|_{D_\rho(r,s), \mathcal{O}}$ for scalar functions is defined in (2.2).

We have the following useful result.

LEMMA 2.2. (i) *The majorant norm is closed under projections, namely, $\|\Pi_S f\|_{r,s} \leq \|f\|_{r,s}$, and $\|X_{\Pi_S f}\|_{r,s} \leq \|X_f\|_{r,s}$.*

(ii) *$\mathcal{A}_{r,s}$ is closed under Poisson brackets, with respect to the symplectic form $dI \wedge d\vartheta + idz \wedge d\bar{z}$; moreover by Cauchy estimates, if we denote $\delta = (\frac{r'}{r})^2 \min(s - s', 1 - \frac{r'}{r})$,*

$$\begin{aligned} \|[X_f, X_g]\|_{r',s'} &\leq 2^{2d+1} \delta^{-1} \|X_f\|_{r,s} \|X_g\|_{r,s}, \\ \|X_{\{f,g\}}\|_{r',s'} &\leq 2^{2d+1} \delta^{-1} \|X_f\|_{r,s} \|X_g\|_{r,s}. \end{aligned}$$

Proof. Item (i) is obvious. Item (ii) is proved in [5, respectively, Lemmas 2.15, and 2.16]. In [5] the interested reader can find an analysis of the properties of the majorant norm. Note that in [5] there is the restriction $r/2 < r' < r$ (the same for s), hence the term $(\frac{r'}{r})^2$ is substituted for by 4. \square

3. Affine subspaces. An affine space A of codimension ℓ in \mathbb{R}^d can be defined by a list of ℓ equations $A := \{x \mid v_i \cdot x = p_i\}$, 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*, and the vectors v_i lie in a prescribed ball B_N of radius some constant N . We set $C_1 := \max_i |n_i|$, and we denote by

$$\langle v_i \rangle_\ell = \text{Span}(v_1, \dots, v_\ell; \mathbb{R}) \cap \mathbb{Z}^d, \quad B_N := \{x \in \mathbb{Z}^d \setminus \{0\} : |x| < C_1 N\};$$

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

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

DEFINITION 3.1. *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 lexicographical³ order (in \mathbb{N}^s) or if $(|a|) = (|b|)$, and $a > b$ in the lexicographical order in \mathbb{Z}^s .*

For instance in \mathbb{Z}^2 , $(\pm 1, \pm 5) \prec (\pm 2, \pm 4)$ since $(1, 5) < (2, 4)$; on the other hand, we have $(1, 4) \prec (1, -4) \prec (-1, 4) \prec (-1, -4)$. This is due to the fact that these last vectors have the same components apart from the sign, and $(1, 4) > (1, -4) > (-1, 4) > (-1, -4)$ in the lexicographic ordering of \mathbb{Z}^2 .

LEMMA 3.2. *Every nonempty set of elements in $L \subset \mathbb{Z}^s$ has a unique minimum.*

Proof. We first consider the list of vectors $|L| \subset \mathbb{N}^s$ consisting of the vectors $(|a|)$ with $a \in L$. This list has a minimum with respect to the lexicographic ordering of \mathbb{N}^s . Naturally there may more than one vector, say $a \neq b \in L$ with $(|a|) = (|b|)$, which attain the minimum of $|L|$. This vectors are at most 2^s , and among them we choose the unique maximum in the lexicographical order in \mathbb{Z}^s . \square

Consider a fixed but large enough N .

DEFINITION 3.3. *We set \mathcal{H}_N the set of all affine spaces A which can be presented as $A = [v_i; p_i]_\ell$ for some $0 < \ell \leq d$ so that $v_i \in B_N$.*

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 $(p_1, \dots, p_\ell; v_1, \dots, v_\ell) \prec (q_1, \dots, q_\ell; w_1, \dots, w_\ell)$.

DEFINITION 3.4. *The N -optimal presentation $[l_i; q_i]_\ell$ of $A \in \mathcal{H}_N$ is the minimum in the sign lexicographical order of the presentations of A which satisfy the bound $v_i \in B_N$.*

³Recall that given two partially ordered sets A and B , the lexicographical order on the Cartesian product $A \times B$ is defined as $(a, b) < (a', b')$ if and only if either $a < a'$ or $a = a'$, and $b < b'$.

Given an affine subspace $A := \{x | v_i \cdot x = p_i, i = 1, \dots, \ell\}$ by the notation $A \xrightarrow{N} [v_i; p_i]_\ell$ we mean that the given presentation is N optimal.

Remark 3.1. (i) Note that each point $m = (m_1, \dots, m_d) \in \mathbb{Z}_1^d$ has an 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) 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.

Example 3.1. We now give an example of the N -optimal presentation of a point, and of an affine subspace. One may easily verify that for any affine subspace A there exists $\bar{N}(A)$ such that for all $N \geq \bar{N}(A)$ the N -optimal presentation is N independent.

Let us start with the case $m_0 = (-11, 15, 3, 27) \in \mathbb{Z}^4$. We have that for all $N > C_1^{-1}\sqrt{82}$ (recall that $C_1 = \max_i |n_i|$)

$$m_0 \xrightarrow{N} [0, 0, 0, 1; (0, 0, 9, -1), (0, 1, 4, -1), (3, 0, 2, 1), (1, 0, -5, 1)].$$

In general, given any point m_0 we will always find $\bar{N}(m_0)$ such that for all $N > \bar{N}(m_0)$ the N -optimal presentation is fixed, say $[p_i^{(0)}; v_i^{(0)}]_d$ and $p_i^{(0)} = 0$ for $i = 1, \dots, d-1$, while $p_d^{(0)} = \text{mcd}(m_1^{(0)}, \dots, m_d^{(0)})$.

Let us now study some affine subspaces.

If $d = 2$ consider the line $A := \{m \in \mathbb{Z}^2 : m = m_0 + tc, t \in \mathbb{R}\}$, with m_0 orthogonal to c (suppose also that the components of m_0 are coprime). Then $A \xrightarrow{N} [|m_0|^2; m^{(0)}]_1$ provided that $N \geq C_1^{-1}|m_0|$.

If $d = 4$ and $A := \{m \in \mathbb{Z}^4 : m = (-11, 15, 3, 27) + (1, 0, 0, 0)t, t \in \mathbb{R}\}$ we have that

$$A \xrightarrow{N} [0, 0, 3; (0, 0, 9, -1), (0, 1, 4, -1), (0, 0, 1, 0)]_3 \quad \forall N > C_1^{-1}\sqrt{82}.$$

If $B := \{m \in \mathbb{Z}^4 : m = (-11, 15, 3, 27) + (1, 0, 0, 0)t + (0, 1, 0, 0)s, t, s \in \mathbb{R}\}$ we have that

$$B \xrightarrow{N} [0, 3; (0, 0, 9, -1), (0, 0, 1, 0)]_2 \quad \forall N > C_1^{-1}\sqrt{82}.$$

LEMMA 3.5. (i) *If the presentation $A = [v_i; p_i]_\ell$ is N -optimal, we have*

$$(3.1) \quad 0 \leq p_1 \leq p_2 \leq \dots \leq p_\ell.$$

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

$$(3.2) \quad |(v, r)| \geq p_{j+1} \quad \forall r \in A.$$

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

(iv) *Finally, $-A$ has an N -optimal presentation $-A = [v'_i, p_i]_\ell$ with the same constants p_i and $(|v'_i|) = (|v_i|)$.*

Proof. (i) If $p_i < 0$ we can change the presentation changing p_i into $-p_i$ and v_i into $-v_i$. By definition this is a lower presentation lexicographically; we obtain a contradiction. Suppose now that (3.1) is false—say, for instance, that $p_1 > p_2 \geq 0$ —then by definition $\{p_2, p_1, \dots, p_\ell; v_2, v_1, \dots, v_\ell\}$ is a presentation of A , and it is lexicographically lower than $\{p_1, p_2, \dots, p_\ell; v_1, v_2, \dots, v_\ell\}$.

(ii) Take $v \in \langle v_1, \dots, v_\ell \rangle \cap B_N \setminus \langle v_1, \dots, v_j \rangle$, and any $r \in A$. We note that $\langle v, r \rangle$ is constant on A . There exists an $h > j$ such that if we substitute v_h , $h > j$, with v we obtain a new presentation. Again we deduce by minimality in the lexicographical order that $|\langle v, r \rangle| \geq p_h \geq p_{j+1}$.

(iii) Any presentation $A_j = [w_i, q_i]_j$ can be completed to a presentation $[w_i, q_i]_\ell$ of A so if $[q_1, \dots, q_j, w_1, \dots, w_j] \prec [p_1, \dots, p_j; v_1, \dots, v_j]$ we also have $[q_1, \dots, q_\ell; w_1, \dots, w_\ell] \prec [p_1, \dots, p_\ell; v_1, \dots, v_\ell]$ by the definition of lexicographical order, a contradiction.

(iv) As for the last statement it is enough to observe that there is a 1 to 1 correspondence between presentations $A = [w_j, q_j]$ of A and $-A$ with the constants $q_i \geq 0$; if $A = [w_j, q_j]$ we have $-A = [-w_j, q_j]$. The absolute value vectors of the two presentations are the same, the statement follows. \square

Remark 3.2. For fixed N, ℓ, p the number of affine spaces in \mathcal{H}_N , of codimension ℓ and such that $p_\ell \leq p$, is bounded by $(2C_1N)^{\ell d}(2p)^\ell$.

3.1. Parameters, and cuts. We shall need several auxiliary parameters in the course of our proof. We start by fixing some numbers:

$$(3.3) \quad \begin{aligned} \tau_0 &> \max(d + b, 12), \quad \tau_1 := (4d)^{d+1}(\tau_0 + 1), \\ c &\leq \frac{1}{2}, \quad C \geq 4, \quad N_0 \geq d!C_1^d C c^{-1}. \end{aligned}$$

In what follows N will always denote some large number, in particular $N > N_0$. For the purpose of this paper we may fix $c = \frac{1}{2}$ and $C = 4$; however, we give the definitions in the more general setting so that they are more flexible.

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 is always $\ell = d$. Set by convention $p_0 = 0$ and $p_{d+1} = \infty$.

We then give a definition involving the parameters θ, μ, τ which we call *allowable* if

$$\tau_0 \leq \tau \leq \tau_1 / (4d) = (4d)^d (\tau_0 + 1), \quad c < \theta, \mu < C.$$

We need to analyze certain *cuts*, for the values p_i associated with 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.

DEFINITION 3.6. *The point $m \xrightarrow{N} [v_i; p_i]$ has a cut $\ell \in \{0, 1, \dots, d\}$ with the parameters (N, θ, μ, τ) if ℓ is such that $p_\ell < \mu N^\tau$, $p_{\ell+1} > \theta N^{4d\tau}$ (recall that $p_0 = 0, p_{d+1} = \infty$).*

The space $A := \{x \mid v_i \cdot x = p_i, i = 1, \dots, \ell\}$ is denoted by $[v_i; p_i]_\ell$ and called the affine space associated with the cut of m .

In turn for every affine subspace $A \xrightarrow{N} [v_i; p_i]_\ell$ with $p_\ell < \mu N^\tau$, the set of points $m \in A$ with $|m| > \theta K^{\tau_1}$ which have ℓ as a cut with the parameters (N, θ, μ, τ) are called the (N, θ, μ, τ) -good points of A .

Notice that $\theta N^{4d\tau} > \mu N^\tau$ (since $N^{(4d-1)\tau} \geq N^{(4d-1)\tau_0} > Cc^{-1} > \theta\mu^{-1}$), so for any given $m \in \mathbb{Z}_1^d$ there is at most one choice of ℓ such that m has an ℓ cut with parameters (N, θ, μ, τ) . Note moreover that the affine subspace associated with a (N, θ, μ, τ) -good point of A is A .

Remark 3.3. 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.

Example 3.2. Fix $N > C_1^{-1}\sqrt{82}$, θ, μ, τ , and consider the affine subspace $A \xrightarrow{N} [0, 0, 3; (0, 0, 9, -1), (0, 1, 4, -1), (0, 0, 1, 0)]_3$ of Example 3.1. For all t large enough (i.e., $t > 66C_1N$), setting

$$m(t) = (-11, 15, 3, 27) + (1, 0, 0, 0)t$$

we have

$$m(t) \xrightarrow{N} [0, 0, 3, p_4(N, t); (0, 0, 9, -1), (0, 1, 4, -1), (0, 0, 1, 0), v_4(N)],$$

where $v_4(N) = (v_4^{(1)}(N), \dots, v_4^{(4)}(N))$ is a vector such that $|v_4(N)| < C_1N$, the first component $v_4^{(1)}(N) = 1$; finally $p_4(N, t) = t - P(N)$ with $|P(N)| < 33C_1N$. Hence m is a (N, θ, μ, τ) -good point of A provided that $t > \theta N^{4d\tau} - 33C_1N$.

Remark 3.4. (1) If ℓ is a cut for the point $m \xrightarrow{N} [v_i; p_i]$, with allowable parameters (N, θ', μ', τ) it is also so for all parameters (N, θ, μ, τ) with $c < \theta \leq \theta' < C$, $c < \mu' \leq \mu < C$.

(2) If for a given $\ell, \tau_0 \leq \tau \leq \tau_1/4d$ we have $p_\ell \leq cN^\tau$, $p_{\ell+1} \geq cN^{4d\tau}$, then ℓ is a cut with parameters (N, θ, μ, τ) for every choice of $c < \theta, \mu < C$.

LEMMA 3.7. Consider $m, r \in \mathbb{Z}_1^d$ with $m \xrightarrow{N} [v_i; p_i]$, $r \xrightarrow{N} [w_i; q_i]$; suppose that ℓ is a cut for m with the allowable parameters N, θ', μ', τ and suppose there exist parameters $c < \theta < \theta' < C$, $c < \mu' < \mu < C$:

$$(3.4) \quad |r - m| < C_1^{-1}(\mu - \mu')N^{\tau-1}, \quad C_1^{-1}(\theta' - \theta)N^{4d\tau-1},$$

then

(1) ℓ is a cut for the point r , for all allowable parameters (N, θ, μ, τ) for which (3.4) holds,

- (2) $\langle w_1, \dots, w_\ell \rangle = \langle v_1, \dots, v_\ell \rangle$,
- (3) $[w_i; q_i]_\ell = [v_i; p_i]_\ell + r - m$.

Proof. Fix θ, μ satisfying (3.4). Write $(v_i, r) = (v_i, r - m) + p_i$. For $i \leq \ell$, since $|v_i| \leq C_1N$ we have

$$(3.5) \quad |(v_i, r)| \leq p_i + |v_i||r - m| < \mu'N^\tau + (\mu - \mu')N^\tau = \mu N^\tau.$$

From (3.1) by the definition of N -optimal, for all $v \in B_N \setminus \langle v_1, \dots, v_\ell \rangle$ one has

$$(3.6) \quad |(v, r)| = |(v, m) + (v, r - m)| \geq p_{\ell+1} - |v||r - m| > \theta'N^{4d\tau} - C_1N|r - m| = \theta N^{4d\tau}.$$

(1), (2) By induction on i we wish to show that $q_i < \mu N^\tau$ and $w_i \in \langle v_1, \dots, v_\ell \rangle$ for all $i \leq \ell$. For $i = 0$ this is trivial, so assume that for $0 \leq i < \ell$, we have $\langle w_1, \dots, w_i \rangle \subset \langle v_1, \dots, v_\ell \rangle$. Since the v_i are independent, there exists $h \leq \ell$ such that $v_h \notin \langle w_1, \dots, w_i \rangle$. By (3.5) $q_{i+1} \leq |(v_h, r)| < \mu N^\tau$.

By contradiction suppose that $w_{i+1} \in B_N \setminus \langle v_1, \dots, v_\ell \rangle$; applying (3.6) we would get $(w_{i+1}, r) := q_{i+1} > \theta N^{4d\tau} > \mu N^\tau$, a contradiction.

Since the w_i (as well as the v_i) are linearly independent, clearly $\langle v_1, \dots, v_\ell \rangle = \langle w_1, \dots, w_\ell \rangle$. This proves (2). As a consequence for $s > \ell$, we again apply (3.6) to $w_s \in B_N \setminus \langle v_1, \dots, v_\ell \rangle$; we obtain $q_{j+1} > \theta N^{4d\tau}$. This completes the proof of (1).

(3) By (2) the space $[w_i; q_i]_\ell$ is the one parallel to $[v_i; p_i]_\ell$ and passing through r . The result follows. \square

Remark 3.5. Note that if we know that m, r both have an ℓ cut with parameters N, θ, μ, τ then we can deduce that the subspace $[w_i; q_i]_\ell$ is the one parallel to $[v_i; p_i]_\ell$ and passing through r provided that

$$(3.7) \quad |r - m| < C_1^{-1}c(N^{4d\tau-1} - Cc^{-1}N^{\tau-1}).$$

Notice that $C_1^{-1}c(N^{4d\tau-1} - Cc^{-1}N^{\tau-1}) \geq N^\tau$; actually in our computations we will have $|r - m| < N^3$.

Remark 3.6. With the above lemma we are stating that if m has an ℓ cut with parameters θ', μ', τ then, for all choices of $\theta < \theta', \mu' < \mu$, for which θ, μ are allowable parameters, there exists a spherical neighborhood B of m such that all points $r \in B$ have an ℓ cut with parameters N, θ, μ, τ . The radius of B is determined by (3.4). Note moreover that if r has a cut ℓ for some parameters then so has $-r$, and with the same parameters. Then Lemma 3.7 holds verbatim if in (3.4) we substitute $|m - r|$ with $|m + r|$.

The definitions which we have given are sufficient to define and analyze the quasi-Töplitz functions, which are introduced in section 4. 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 .

3.2. Standard cuts. The following construction will be useful: we divide

$$[CN^{4d\tau_0}, cN^{\tau_1/4d}) = \cup_{i=1}^{d-1} [N^{S_i}, N^{S_{i+1}}) \cup [N^{S_d}, cN^{\tau_1/4d})$$

by setting $N^{S_1} := CN^{4d\tau_0}$ and defining recursively

$$c^{-1}N^{S_{i+1}} = c^{-1}C \cdot (c^{-1}N^{S_i})^{4d}, \quad i = 1, \dots, d - 1.$$

By definition we get

$$c^{-1}N^{S_j} = (c^{-1}C)^{\sum_{i=0}^{j-1} (4d)^i} N^{(4d)^j \tau_0}.$$

Recalling that $N > N_0 = Cc^{-1}$, and $\tau_1 = (4d)^{d+1}(\tau_0 + 1)$, we get

$$c^{-1}N^{S_d} \leq N^{d(4d)^{d-1} + (4d)^d \tau_0} \leq N^{\tau_1/4d}.$$

We set

$$\varrho_0 := \tau_0, \quad \varrho_d := \frac{\tau_1}{4d}, \quad cN^{\varrho_i} := N^{S_i}, \quad 0 < i < d.$$

LEMMA 3.8. *For all allowable parameters $c < \theta, \mu < C$ and for each point $m \xrightarrow{N} [v_i; p_i]$ we construct a standard cut $\ell, 0 \leq \ell \leq d$, for m for which the parameter τ is one of the previously defined numbers $\varrho_i, i = 0, \dots, d$.*

If $|m| \geq N^{\tau_1}$, then $\ell < d$, if $p_1 < CN^{4d\tau_0}$ then $\ell > 0$.

Proof. Let $m \xrightarrow{N} [v_i; p_i]$. If $p_d \leq cN^{\tau_1/4d}$ then we set $\ell = d$ and $\tau = \varrho_d = \tau_1/4d$. If $p_1 \geq CN^{4d\tau_0}$ then we set $\ell = 0$ and $\tau = \varrho_0 = \tau_0$.

Otherwise if $p_1 < CN^{4d\tau_0}$, and $p_d > cN^{\tau_1/4d}$ then at least one of the $d-1$ intervals $(N^{S_i}, N^{S_{i+1}})$ with $i = 1, \dots, d-1$ does not contain any element of the ordered list $\{p_2, \dots, p_{d-1}\}$. The parameters ℓ, τ are fixed by setting $\tau = \varrho_{\bar{i}}$, where $cN^{\varrho_{\bar{i}}} = N^{S_{\bar{i}}}$, and \bar{i} is the smallest among the indices i such that the interval $(N^{S_i}, N^{S_{i+1}})$ does not contain any points of the list $\{p_2, \dots, p_{d-1}\}$; finally $\ell < d$ is the index for which $p_\ell \leq N^{S_\ell} = cN^\tau$, and $p_{\ell+1} \geq N^{S_{\ell+1}} = C(c^{-1}N^{S_\ell})^{4d} = CN^{4d\tau}$.

If $p_d \leq cN^{\tau_1/4d}$, we apply Cramer's rule to the equations $Vm = p$ given by the presentation. We have $|m| = |V^{-1}p| \leq cd!N^{\tau_1/4d}(C_1N)^{d-1} < N^{\tau_1}$ since $\frac{\tau_1}{4d} + d < \tau_1$ and as soon as $N > cd!C_1^{d-1}$. \square

3.3. Cuts, and good points. As shown in the introduction we need a decomposition of \mathbb{Z}_1^d as in (1.7). For any given N we set

$$(3.8) \quad A_0 = A_0(N) := \left\{ m \in \mathbb{Z}_1^d : m \xrightarrow{N} [v_i; p_i] \text{ with } p_1 > CK^{4d\tau_0} \right\}.$$

In order to define A_ℓ we set the following.

DEFINITION 3.9. For all $[v_i; p_i]_\ell \in \mathcal{H}_N$ with $1 \leq \ell < d$ and $p_\ell \leq cN^{\frac{\tau_1}{4d}}$, the set

$$(3.9) \quad [v_i; p_i]_\ell^g := \{x \in [v_i, p_i]_\ell \mid |x| > N^{\tau_1}, |(v, x)| \geq C \max(N^{4d\tau_0}, c^{-4d} p_\ell^{4d}), \forall v \in B_N \setminus \langle v_i \rangle_\ell\}$$

will be called the N -good portion of the subspace $A = [v_i; p_i]_\ell$.

Remark 3.7. Notice that every $v \in B_N \setminus \langle v_i \rangle_\ell$ gives a non constant linear function $v \cdot x$ on A . Thus the good points of A form a nonempty 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 form a finite union of affine subspaces of codimension one in A .

LEMMA 3.10. Given $p \leq cN^{\tau_1/(4d)}$ we fix $\tau(p)$ so that $N^{\tau(p)} = \max(N^{\tau_0}, c^{-1}p)$ (note that $\tau_0 \leq \tau \leq \tau_1/(4d)$). The following holds: for all $c < \theta, \mu < C$, and for all affine subspaces $[v_i; p_i]_\ell \in \mathcal{H}_N$ such that $p_\ell = p$, we have that every point $m \in [v_i; p_i]_\ell^g$ is an $(N, \theta, \mu, \tau(p))$ -good point for $[v_i; p_i]_\ell$.

Proof. By hypothesis (3.9)

$$p_{\ell+1} = (v_{\ell+1}, m) \geq C \max(N^{4d\tau_0}, c^{-4d} p^{4d});$$

recall that $p_\ell = p$. If $p \leq cN^{\tau_0}$ then $\tau(p) = \tau_0$ by definition. Since $p_{\ell+1} \geq cN^{4d\tau_0}$ m has the cut ℓ for all choices of $c < \theta, \mu < C$. Otherwise $cN^{\tau_1/(4d)} \geq p > cN^{\tau_0}$, and $p_{\ell+1} \geq Cc^{-4d}p^{4d}$. So in conclusion for all $c < \theta, \mu < C$ we have $p_\ell = p = cN^{\tau(p)} < \mu N^{\tau(p)}$, and $p_{\ell+1} \geq cN^{4d\tau(p)} > \theta N^{4d\tau(p)}$, hence the cut. \square

We now show that (1.7) provides a decomposition of \mathbb{Z}_1^d ; see Figure 3.1.

PROPOSITION 3.11. Each point $m \xrightarrow{N} [v_i; p_i]$ with $|m| > N^{\tau_1}$ and $p_1 < cN^{4d\tau_0}$ belongs to the set $[v_i; p_i]_\ell^g$ for some choice $0 < \ell < d$.

Proof. According to Lemma 3.8, each point m has a normalized cut $0 < \ell < d$ for all allowable θ, μ , and for some $\tau_0 < \tau < \tau_1/4d$ with τ in the finite list $\{\varrho_1, \dots, \varrho_d\}$. Thus for all $w \in B_N \setminus \langle v_i \rangle_\ell$ we have $|(m, w)| > \theta N^{4d\tau}$ for all $\theta < C$, moreover $p_\ell < \mu N^\tau$ for all $\mu > c$. Hence $|(m, w)| \geq cN^{4d\tau} > cN^{4d\tau_0}$ and $p_\ell \leq cN^\tau$. Combining these relations we obtain

$$|(m, w)| \geq C \max(N^{4d\tau_0}, c^{-4d} p_\ell^{4d}),$$

hence, $m \in [v_i; p_i]_\ell^g$ by Definition 3.9. \square

LEMMA 3.12. Given $p \leq cN^{\tau_1/4d}$, fix $\tau(p)$, as in Lemma 3.10, then the following holds. Given $m \in \mathbb{Z}_1^d$ with $m \in [v_i; p_i]_\ell^g$ and $p_\ell = p$, then for all $r \in \mathbb{Z}_1^d$, and for all parameters $c < \theta, \mu < C$ such that

$$(3.10) \quad |r - m| < C_1^{-1}(\mu - c)N^{\tau_0-1}, C_1^{-1}(C - \theta)N^{4d\tau_0-1},$$

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

Proof. We can apply Lemma 3.10 to m , obtaining the cut ℓ with parameters (N, θ', μ', τ) for all $c < \theta', \mu' < C$. Then, we may apply Lemma 3.7 obtaining the required cut for r for any choice of θ, μ satisfying (3.4) with respect to θ', μ' . Since θ', μ' can be taken arbitrarily close to c, C (3.4) follows from (3.10). \square

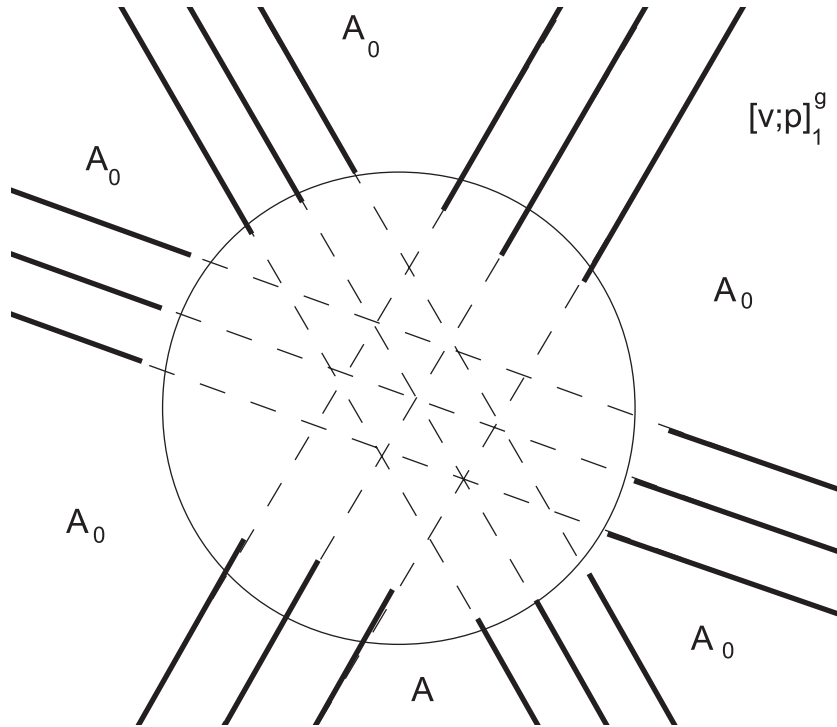


FIG. 3.1. A drawing of the standard decomposition in \mathbb{Z}_1^2 . A_0 is \mathbb{Z}_1^2 minus the dashed lines (each dashed line is described by an equation $[v; p]_1$). On each dashed line the set $[v; p]_1^g$ is signed in solid boldface. Note that $[v; p]_1^g$ is $[v; p]_1 \cap \mathbb{Z}_1^2$ minus a finite number of subspaces of codimension two, i.e., points.

4. Quasi-Töplitz functions. Now, and in the following, we fix $c = \frac{1}{2}$, $C = 4$.

DEFINITION 4.1. Given N, θ, μ, τ such that $1/2 < \theta, \mu < 4$, $\tau_0 \leq \tau \leq \tau_1/4d$, and $4N^3 < \frac{1}{2}N^{\tau_1}$ we say that a monomial

$$e^{i(k, \vartheta)} I^l z^\alpha \bar{z}^\beta z_m^\sigma z_n^{\sigma'}$$

is (N, θ, μ, τ) -bilinear if it satisfies momentum conservation (2.5), i.e.,

$$(4.1) \quad \sigma m + \sigma' n = -\pi(k, \alpha, \beta),$$

$$|k| < N, \quad |n|, |m| > \theta N^{\tau_1}, \quad \sum_j |j|(\alpha_j + \beta_j) < \mu N^3,$$

and moreover there exists $0 < \ell < d$ such that both n, m have an ℓ cut with parameters N, θ, μ, τ . By convention if $m \xrightarrow{N} [v; p_i]$ and $n \xrightarrow{N} [w; q_i]$ with $(p_1, \dots, p_\ell, v_1, \dots, v_\ell) \preceq (q_1, \dots, q_\ell, w_1, \dots, w_\ell)$ we say that the monomial has the cut $[v; p_i]_\ell$ (this defines univocally an affine subspace associated with the monomial). Note that by Lemma 3.8 we are sure that $\ell < d$. In $\mathcal{A}_{r,s}$ we consider the subspace of (N, θ, μ, τ) -bilinear functions, and call $\Pi_{(N, \theta, \mu, \tau)}$ the projection onto this subspace.

Notice that by Remark 3.5 the cut $[w; q_i]_\ell$ is completely fixed by $[v; p_i]_\ell$ and $\sigma m + \sigma' n$.

Having chosen $1/2, 4$ as bounds for the parameters θ, μ we will call *low momentum variables*, denoted by w^L , and spanning the space ℓ_ρ^L , the z_j^σ such that $|j| < 4N^3$.

Similarly we call *high momentum variables*, denoted by w^H , and spanning the space ℓ_ρ^H , the z_j^σ such that $|j| > N^{\tau_1}/2$. Notice that the low and high variables are separated. We may write uniquely

$$(4.2) \quad \Pi_{(N,\theta,\mu,\tau)} f = \sum_{\sigma,\sigma'=\pm} \sum_{\substack{|m|,|n|>\theta N^{\tau_1} \\ \exists \ell: m,n \text{ have an } \ell \text{ cut} \\ \text{with parameters } N,\theta,\mu,\tau}} f_{m,n}^{\sigma,\sigma'}(I, \vartheta, w^L) z_m^\sigma z_n^{\sigma'}$$

where

$$f_{m,n}^{\sigma,\sigma'}(I, \vartheta, w^L) = \sum_{\substack{|k|<N, |\alpha|+|\beta|<\mu N^3, \\ -\pi(k,\alpha,\beta)=\sigma m+\sigma' n}} f_{m,n,k,\alpha,\beta}^{\sigma,\sigma'}(I) e^{i(k,\vartheta)} z^\alpha \bar{z}^\beta;$$

finally $f_{m,n,k,\alpha,\beta}^{\sigma,\sigma'}(I)$ is an analytic function of I for $|I| < r^2$.

Given an affine subspace $A \xrightarrow{N} [v_i; p_i]_\ell$, we construct $(N, \theta, \mu, \tau, A)$ -restricted *Töplitz functions* by setting:

$$(4.3) \quad g(A, I, \vartheta, z) := \sum_{n,m,\sigma,\sigma',k,\alpha,\beta}^{(N,\theta,\mu,\tau,A)} g_{k,\alpha,\beta}^{\sigma,\sigma'}(\sigma m + \sigma' n, A; I) e^{i(k,\vartheta)} z^\alpha \bar{z}^\beta z_m^\sigma z_n^{\sigma'}$$

here the sum $\sum^{(N,\theta,\mu,\tau,A)}$ means the sum over those $n, m, \sigma, \sigma', k, \alpha, \beta$ such that $e^{i(k,\vartheta)} z^\alpha \bar{z}^\beta z_m^\sigma z_n^{\sigma'}$ is a (N, θ, μ, τ) -bilinear monomial with cut given by A . Finally $g_{k,\alpha,\beta}^{\sigma,\sigma'}(h, B; I)$ is an analytic function of I , for $|I| < r^2$, which is well defined for all $\sigma, \sigma' = \pm 1, k \in \mathbb{Z}^b, h \in \mathbb{Z}_1^d, \alpha, \beta \in \mathbb{N}^{\mathbb{Z}_1^d}$, and $B \xrightarrow{N} [w_i; q_i]_\ell \in \mathcal{H}_N$ such that $|k| < N, h = -\pi(k, \alpha, \beta), \sum_{j \in \mathbb{Z}_1^d} |j|(\alpha_j + \beta_j) < \mu N^3$, and $|q_\ell| < 4N^{\tau_1/4d}$.

Notice that the coefficient $g_{k,\alpha,\beta}^{\sigma,\sigma'}(\sigma m + \sigma' n, A; I)$ depends on m, n only through $\sigma m + \sigma' n, A; I$. The sum $\sum_{n,m,\sigma,\sigma',k,\alpha,\beta}^{(N,\theta,\mu,\tau,A)}$ instead selects those m, n such that $|m|, |n| > \theta N^{\tau_1}, |\sigma m + \sigma' n| < \mu N^3 + N, m, n$ have a cut ℓ, τ , and the cut of m is A .

DEFINITION 4.2. A function g is called piecewise Töplitz if it is of the form

$$g = \sum_{\substack{A \in \mathcal{H}_N \\ A \xrightarrow{N} [v_i; p_i]_\ell : |p_\ell| < \mu N^\tau}} g(A, I, \vartheta, z).$$

We denote the space of piecewise Töplitz functions as $\mathbb{F}(N, \theta, \mu, \tau) = \mathbb{F} \subset \mathcal{A}_{r,s}$

Remark 4.1. Notice that $\mathbb{F}(N, \theta, \mu, \tau)$ is a subset of the (N, θ, μ, τ) bilinear functions. Hence given $g \in \mathbb{F}(N, \theta, \mu, \tau)$ we may write it in the form (4.2)

$$g = \sum_{\sigma,\sigma'=\pm} \sum_{\substack{|m|,|n|>\theta N^{\tau_1} \\ \exists \ell: m,n \text{ have an } \ell \text{ cut} \\ \text{with parameters } N,\theta,\mu,\tau}} g_{m,n}^{\sigma,\sigma'}(I, \vartheta, w^L) z_m^\sigma z_n^{\sigma'}$$

and one has that

$$(4.4) \quad \begin{aligned} g_{m,n}^{\sigma,\sigma'}(I, \vartheta, w^L) &= g^{\sigma,\sigma'}(\sigma m + \sigma' n, [v_i; p_i]_\ell, I, \vartheta, w^L) \\ &:= \sum_{\substack{|k|<N, |\alpha|+|\beta|<\mu N^3, \\ -\pi(k,\alpha,\beta)=\sigma m+\sigma' n}} g_{k,\alpha,\beta}^{\sigma,\sigma'}(\sigma m + \sigma' n, [v_i; p_i]_\ell; I) e^{i(k,\vartheta)} z^\alpha \bar{z}^\beta \end{aligned}$$

if $|n|, |m| > \theta N^{\tau_1}$, $m \xrightarrow{N} [v_i; p_i]$, and there exists ℓ such that m, n have a cut at ℓ with parameters (N, θ, μ, τ) . Otherwise $g_{m,n}^{\sigma, \sigma'} = 0$.

Notice that $g^{\sigma, \sigma'}(\sigma m + \sigma' n, [v_i; p_i]_\ell, I, \vartheta, w^L)$ depends on m, n only through the subspace $[v_i; p_i]_\ell$, and $\sigma m + \sigma' n$. In other words the quadratic form representation (4.2) of an (N, θ, μ, τ) -piecewise Töplitz function has translation invariance in the sense that $g_{m,n}^{\sigma, \sigma'} = g_{m_1, n_1}^{\sigma, \sigma'}$ provided that: $\sigma m + \sigma' n = \sigma m_1 + \sigma' n_1$, there exists ℓ such that m, n, m_1, n_1 all have an ℓ, τ cut, and both m, m_1 have the same associated subspace $[v_i; p_i]_\ell$.

Given $f \in \mathcal{A}_{r,s}$ and $\mathcal{F} \in \mathbb{F}$, we define

$$(4.5) \quad \bar{f} := N^{4d\tau} (\Pi_{(N, \theta, \mu, \tau)} f - \mathcal{F}).$$

Finally set

$$(4.6) \quad \|X_f\|_{r,s}^T := \sup_{\substack{N \geq K, N \in \mathbb{N}, \\ \tau_0 \leq \tau \leq \tau_1/4d}} [\inf_{\mathcal{F} \in \mathbb{F}} (\max(\|X_f\|_{r,s}, \|X_{\mathcal{F}}\|_{r,s}, \|X_{\bar{f}}\|_{r,s}))].$$

DEFINITION 4.3. We say that $f \in \mathcal{A}_{r,s}$ is quasi-Töplitz of parameters (K, θ, μ) if $\|X_f\|_{r,s}^T < \infty$. We call $\|X_f\|_{r,s}^T$ the quasi-Töplitz norm of f .

Remark 4.2. Notice that our definition includes the Töplitz and anti-Töplitz functions by setting, for any N, θ, μ, τ , $\mathcal{F} = \Pi_{(N, \theta, \mu, \tau)} f$, and hence $\bar{f} = 0$. In the case of Töplitz functions one trivially has $\|X_f\|_{r,s}^T = \|X_f\|_{r,s}$.

Remark 4.3. Intuitively a quasi-Töplitz function is a function whose bilinear part is “well approximated” by a piecewise Töplitz function.

Given K, θ, μ , and a function $f \in \mathcal{A}_{r,s}$ we proceed as follows. For any choice of $N > K$ and $\tau_0 \leq \tau \leq \tau_1/4d$ we compute a “weighted distance” between $\Pi_{N, \theta, \mu, \tau} f$ and the subspace \mathbb{F} . First, for any $\mathcal{F} \in \mathbb{F}$, we define $\bar{f} := N^{4d\tau} (\Pi_{N, \theta, \mu, \tau} f - \mathcal{F})$, and compute $\|X_{\bar{f}}\|_{r,s}$ (since f , and \mathcal{F} are in $\mathcal{A}_{r,s}$ all this quantities are finite); then, in order to obtain a “distance,” we perform the infimum over $\mathcal{F} \in \mathbb{F}$. Essentially a function f is quasi-Töplitz if this weighted distance stays bounded as $N \rightarrow \infty$. Note that one could probably prove that the inf in our definition is actually a min, thus associating to f a “canonical choice” \mathcal{F} (depending on N, θ, μ, τ); this however is not needed in our construction, we only need a weaker decomposition as follows.

If f is quasi-Töplitz with parameters (K, θ, μ) then for any $N \geq K$, and $\tau_0 \leq \tau \leq \tau_1/4d$ there exist functions $\mathcal{F} \in \mathbb{F}(N, \theta, \mu, \tau)$, such that setting

$$\bar{f} := N^{4d\tau} (\Pi_{N, \theta, \mu, \tau} f - \mathcal{F}) \quad \text{we have} \quad \|X_{\mathcal{F}}\|_{r,s}, \|X_{\bar{f}}\|_{r,s} < 2\|X_f\|_{r,s}^T.$$

We now concentrate on the very special case of *diagonal quadratic functions* $Q(z) := \sum_{m \in \mathbb{Z}_1^d} Q_m z_m \bar{z}_m$. We notice that in this case we may reformulate the projection on (N, θ, μ, τ) -bilinear functions as:

$$\Pi_{(N, \theta, \mu, \tau)} Q(z) = \sum_{\substack{A \xrightarrow{N} [v_i; p_i]_\ell \in \mathcal{H}(N) \\ |p_\ell| \leq \mu N^\tau}} \sum_{m \in \mathbb{Z}_1^d}^{(N, \theta, \mu, \tau, A)} Q_m z_m \bar{z}_m,$$

where $\sum_m^{(N, \theta, \mu, \tau, A)}$ coincides with $\sum_{m, m_+, -, 0, 0, 0}^{(N, \theta, \mu, \tau, A)}$ of (4.3), namely, it is the sum over those m with $|m| > \theta N^{\tau_1}$ which have an ℓ cut with parameters (N, θ, μ, τ) associated with the affine space A .

LEMMA 4.4. *Let $Q(z)$ be a quasi-Töplitz diagonal quadratic function. There exist two diagonal quadratic functions $\mathcal{Q}(z) \in \mathbb{F}$, $\bar{Q}(z)$,*

$$(4.7) \quad \mathcal{Q}(z) = \sum_{\substack{A \stackrel{N}{\sim} [v_i; p_i]_{\ell} \in \mathcal{H}(N) \\ |p_{\ell}| \leq \mu N^{\tau}}} \sum_{m \in \mathbb{Z}_1^d}^{(N, \theta, \mu, \tau, A)} \mathcal{Q}(A) z_m \bar{z}_m,$$

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

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

$$(4.8) \quad Q_m = \mathcal{Q}(A) + N^{-4d\tau} \bar{Q}_m.$$

Moreover one has

$$(4.9) \quad |Q_m|, |\mathcal{Q}(A)|, |\bar{Q}_m| \leq 2 \|X_Q\|_r^T.$$

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

Hence we can fix quadratic, and diagonal functions $\mathcal{Q} \in \mathbb{F}$ and $\bar{Q} = N^{4d\tau} (\Pi_{N, \theta, \mu, \tau} \mathcal{Q} - \mathcal{Q})$ so that $\|X_{\mathcal{Q}}\|_r, \|X_{\bar{Q}}\|_r \leq 2 \|X_Q\|_r^T$. To conclude we need to show that a quadratic, diagonal, and piecewise Töplitz \mathcal{Q} is of the form (4.7). Indeed by (4.3) an (N, θ, μ, A) -restricted Töplitz function which is quadratic and diagonal is of the form

$$g(A, z) = g(A) \sum_m^{(N, \theta, \mu, \tau, A)} z_m \bar{z}_m.$$

Our last statement is proved by noting that

$$\|X_Q\|_r = 2 \sup_{\|z\|_{\rho} < r} \sum_{h \in \mathbb{Z}_1^d} |Q_h| \frac{|z_h|}{r} e^{\rho|h|} \geq |Q_j|$$

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

Remark 4.4. It is interesting to compare the set of quasi-Töplitz functions with the Töplitz–Lipschitz functions of [11]. The first observation is that the set of quasi-Töplitz functions is closed with respect to Poisson brackets, while the Töplitz–Lipschitz functions are closed only with respect to Poisson brackets when one of the functions is quadratic. This is due to the fact that the property of being quasi-Töplitz depends on the idea of (N, θ, μ, τ) -bilinear projection, and not on the Hessian of the function. Indeed one may easily produce functions which are quasi-Töplitz but not Töplitz–Lipschitz (even in the class of functions which preserve momentum).

A second more subtle point is whether the class of *quadratic* quasi-Töplitz and Töplitz–Lipschitz functions coincide; this should be true at least for $d \leq 2$, and we expect some inclusions to hold even in higher dimensions.

5. An abstract KAM theorem. The starting point for our KAM theorem is a family of Hamiltonians

$$(5.1) \quad H = \mathcal{N} + P, \quad \mathcal{N} = \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^d} \Omega_n(\xi) z_n \bar{z}_n, \quad P = P(I, \vartheta, z, \bar{z}, \xi),$$

defined in $D(r, s) \times \mathcal{O}$, where $\mathcal{O} \subset \mathbb{R}^b$ is open and bounded, say it is contained in a set of diameter D . The functions $\omega(\xi), \Omega_n(\xi)$ are well defined for $\xi \in \mathcal{O}$.

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 $(\vartheta, 0, 0, 0) \rightarrow (\vartheta + \omega(\xi)t, 0, 0, 0)$ 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.

We require the following hypotheses on \mathcal{N} , and P .

(A1) *Nondegeneracy*: The map $\xi \rightarrow \omega(\xi)$ is a C_W^1 diffeomorphism between \mathcal{O} and its image with $|\omega|_{C_W^1}, |\nabla\omega^{-1}|_{\mathcal{O}} \leq M$.

(A2) *Asymptotics of normal frequency*:

$$(5.2) \quad \Omega_n(\xi) = |n|^2 + \tilde{\Omega}_n(\xi),$$

where $\tilde{\Omega}_n$'s are C_W^1 functions of ξ with C_W^1 -norm uniformly bounded by some positive constant L with $LM < \frac{1}{2}$.

(A3) *Momentum conservation*: The perturbation P satisfies momentum conservation, it is real analytic, and C_W^1 in $\xi \in \mathcal{O}$; namely, $P \in \mathcal{A}_{r,s}$.

(A4) *Quasi-Töplitz property and regularity*: the functions P and $\sum_j \tilde{\Omega}_j |z_j|^2$ are quasi-Töplitz with parameters (K, θ, μ) , where

$$\frac{1}{2} < \theta, \mu < 4, \quad \left(\mu - \frac{1}{2}\right)K^{\tau_0}, (4 - \theta)K^{4d\tau_0} > 5K^4.$$

One has the bounds

$$\|X_P\|_{D(r,s),\mathcal{O}}^T < \infty, \|\langle \tilde{\Omega}z, z \rangle\|_{D(r,s),\mathcal{O}}^T < L.$$

Now we state our infinite dimensional KAM theorem.

THEOREM 5.1. *Assume Hamiltonian $\mathcal{N} + P$ in (5.1) satisfies (A1)–(A4). Then, for all $\gamma > 0$ small enough there exists a positive constant $\varepsilon = \varepsilon(\gamma, b, d, L, M, K, \theta, \mu)$ such that if $\|X_P\|_{D(r,s),\mathcal{O}}^T \leq \varepsilon$, then there exists a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$ with $\text{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma)$, and two maps (analytic in ϑ , and C_W^1 in ξ)*

$$\Psi : \mathbb{T}^b \times \mathcal{O}_\gamma \rightarrow D(r, s), \quad \tilde{\omega} : \mathcal{O}_\gamma \rightarrow \mathbb{R}^b,$$

where Ψ is $\frac{\varepsilon}{\gamma^2}$ -close to the trivial embedding $\Psi_0 : \mathbb{T}^b \times \mathcal{O} \rightarrow \mathbb{T}^b \times \{0, 0, 0\}$, and $\tilde{\omega}$ is ε -close to the unperturbed frequency ω , such that for any $\xi \in \mathcal{O}_\gamma$, and $\vartheta \in \mathbb{T}^b$, the curve $t \rightarrow \Psi(\vartheta + \tilde{\omega}(\xi)t, \xi)$ is a linearly stable quasi-periodic solution of the Hamiltonian system governed by $H = \mathcal{N} + P$.

5.1. Application to the NLS. The NLS (1.1) is a Hamiltonian equation. We expand the solution in Fourier series as $u = \sum_{m \in \mathbb{Z}^d} u_m \phi_m(x)$ and obtain that the $u_m(t)$ are the Hamiltonian flow of

$$(5.3) \quad N + P = \sum_{i=1}^b (|n_i|^2 + \xi_i) |u_{n_i}|^2 + \sum_{n \in \mathbb{Z}_1^d} |n|^2 u_n \bar{u}_n + \int_{\mathbb{T}^d} g \left(\left| \sum_{m \in \mathbb{Z}^d} u_m \phi_m(x) \right|^2 \right) dx$$

with respect to the symplectic form $i \sum_{m \in \mathbb{Z}^d} du_m \wedge d\bar{u}_m$. Here g is a primitive of the analytic function f so it has a zero of degree at least two. The conservation of momentum follows by translation invariance.

As an example, if $f(u) = |u|^2u$, then $P = \sum_{m_1 - m_2 + m_3 - m_4 = 0}^{m_i \in \mathbb{Z}^d} u_{m_1} \bar{u}_{m_2} u_{m_3} \bar{u}_{m_4}$, and the constraint $m_1 - m_2 + m_3 - m_4 = 0$ ensures that P satisfies momentum conservation. We introduce standard action-angle coordinates $u_{n_j} = \sqrt{I_j^{(0)} + I_j} e^{i\vartheta_j}$, $j = 1, \dots, b$, $u_n = z_n, n \neq \{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(b)}\}$, where $4r^2 > I_i^{(0)} > 2r^2$, and obtain (1.3), where P is the last summand of (5.3). Let us suppose without loss of generality that $g(y) = y^p + O(y^{p+1})$, so that P is regular, and X_P is of order $|I_0|^{2p}r^{-2}$. It is easily seen that P is Töplitz (hence by Remark 4.2 P is quasi-Töplitz for all choices of θ, μ). Conditions (A1)–(A4) hold with $M = 1$ and any L (since $\tilde{\Omega} = 0$).

In order to apply Theorem 5.1 we fix $r = c\varepsilon^{\frac{1}{4p-2}}$, with c small. We have $\|X_P\|_{r,s}^T \leq C|I_0|^{2p}r^{-2}$ so the smallness condition is achieved.

6. KAM step. Theorem 5.1 is proved by an iterative procedure. We produce a sequence of Hamiltonians $H_\nu = \mathcal{N}_\nu + P_\nu$, and a sequence of symplectic transformations $X_{F_{\nu-1}}^1 H_{\nu-1} := H_\nu$, well defined on a domain $D(r_\nu, s_\nu) \times \mathcal{O}_\nu$. At each step, the perturbation becomes smaller at the cost of reducing the analyticity, and parameter domain. More precisely, the perturbation should satisfy $\|X_{P_{\nu+1}}\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_\nu}^T \leq \varepsilon_\nu^\kappa, \kappa > 1$. The sequence $r_\nu \rightarrow 0$ while $s_\nu \rightarrow s/4$ and $\mathcal{O}_\nu \rightarrow \mathcal{O}_\infty$. For simplicity of notation, we denote the quantities in the ν th step without subscript, i.e., $\mathcal{O}_\nu = \mathcal{O}$, $\omega_\nu = \omega$, and so on. The quantities in the $(\nu + 1)$ th step are denoted with subscript “+.” Most of the KAM procedure is completely standard; see [12] for proofs. The new part is (1) to show that the quasi-Töplitz property (A4) for P and $\langle \tilde{\Omega}z, \bar{z} \rangle$ are kept by KAM iteration, and (2) prove the measure estimate using the quasi-Töplitz property.

For simplicity, below we always use the same symbol C to denote constants independent of the iteration.

One step. Suppose that the Hamiltonian (5.1), well defined in $D(r, s) \times \mathcal{O}$, satisfies (A1)–(A4). Moreover P and $\langle \tilde{\Omega}z, \bar{z} \rangle$ are quasi-Töplitz with parameters (K, θ, μ) , and we have

$$(6.1) \quad |\omega|_{C_W^1}, |\nabla\omega^{-1}|_{\mathcal{O}} \leq M, \quad |\tilde{\Omega}_n|_{C_W^1} \leq L, \\ \|\langle \tilde{\Omega}z, \bar{z} \rangle\|_{D(r,s), \mathcal{O}}^T \leq L, \quad \|X_P\|_{D(r,s), \mathcal{O}}^T \leq \varepsilon.$$

Our aim is to construct (1) an open set $\mathcal{O}_+ \subset \mathcal{O}$ of positive measure, (2) a one-parameter group of symplectic transformations Φ_F^t , well defined for all $\xi \in \mathcal{O}_+, t \leq 1$, such that $\Phi_F^1 H := H_+ = \mathcal{N}_+ + P_+$ still satisfies (A1)–(A4) in the domain $D(r_+, s_+)$. Finally P_+ , and $\langle \tilde{\Omega}^+ z, \bar{z} \rangle$ are quasi-Töplitz with new parameters (K_+, θ_+, μ_+) , and we have

$$|\omega_+|_{C_W^1}, |\nabla\omega_+^{-1}|_{\mathcal{O}} \leq M_+; |\tilde{\Omega}_n^+|_{C_W^1}, \|\langle \tilde{\Omega}^+ z, \bar{z} \rangle\|_{D(r_+, s_+), \mathcal{O}_+}^T \leq L_+, \\ \|X_{P_+}\|_{D(r_+, s_+), \mathcal{O}_+}^T \leq \varepsilon_+ = \varepsilon^\kappa.$$

Let us define

$$R := \sum_{k, 2|p|+|\alpha|+|\beta| \leq 2} P_{k,p,\alpha,\beta} e^{i\langle k, \vartheta \rangle} I^p z^\alpha \bar{z}^\beta, \quad \langle R \rangle := \sum_{i=1}^b P_{0,e_i,0,0} I_i + \sum_{j \in \mathbb{Z}_1^d} P_{0,0,e_j,e_j} |z_j|^2.$$

Remark 6.1. The quadratic function R is quasi-Töplitz, and satisfies the bounds $\|X_R\|_{r,s}^T \leq 2\|X_P\|_{r,s}^T$.

The generating function of our symplectic transformation, denoted by F , solves the “homological equation”:

$$(6.2) \quad \{\mathcal{N}, F\} = \Pi_{\leq K} R - \langle R \rangle,$$

where $\Pi_{\leq K}$ is the projection which collects all terms in R with $|k| \leq K$, and K is fixed to be the quasi-Töplitz parameter of $P, \tilde{\Omega}$. It's well known (and immediate) that F is uniquely defined by the homological equation for those ξ such that $\langle \omega(\xi), k \rangle + \Omega(\xi) \cdot l \neq 0$. In order to have quantitative bounds, we restrict to a set \mathcal{O}_+ where (see Lemma 6.2)

$$(6.3) \quad |\langle \omega(\xi), k \rangle + \Omega(\xi) \cdot l| \geq \gamma K^{-2d\tau_1}, \quad |k| \leq K, |l| \leq 2, (k, l) \neq 0,$$

where $k \in \mathbb{Z}^b, l \in \mathbb{Z}^d$, and $(k, l = \alpha - \beta)$ satisfy momentum conservation (2.5). Then H in the new variables is:

$$H_+ := e^{\{F, \cdot\}} H = \mathcal{N}_+ + P_+,$$

where $\mathcal{N}_+ = \mathcal{N} + \langle R \rangle$ and $P_+ = e^{\{F, \cdot\}} H - \mathcal{N}_+$.

6.1. The set \mathcal{O}_+ . The set of no-resonant parameter is defined.

DEFINITION 6.1. \mathcal{O}_+ is defined to be the open subset of \mathcal{O} such that

(i) for all $|k| < K, h \in \mathbb{Z}, (h, k) \neq (0, 0)$,

$$(6.4) \quad |\langle \omega, k \rangle + h| > 2\gamma K^{-\tau_0};$$

(ii) for all $|k| < K, l \in \mathbb{Z}^d$, such that $|l| = 1$, and l, k satisfy momentum conservation (i.e., $l = \pm e_m$ with $-\pi(k) = \pm m$)

$$(6.5) \quad |\langle \omega, k \rangle + \Omega \cdot l| > 2\gamma K^{-\tau_0};$$

(iii) for all $|k| < K, |l| = 2$, such that l, k satisfy momentum conservation, and moreover $l \neq e_m - e_n$ or $l = e_m - e_n$, and $\max(|m|, |n|) \leq 8K^{\tau_1}$, we set

$$(6.6) \quad |\langle \omega, k \rangle + \Omega \cdot l| > 2\gamma K^{-2d\tau_1};$$

(iv) for all N with $K \leq N \leq 2K^{\tau_1/\tau_0}$, for all affine spaces $[v_i, p_i]_\ell$ in \mathcal{H}_N ($1 \leq \ell < d$) with $|p_\ell| < cN^{\tau_1/4d}$ we choose a point $m^g \in [v_i; p_i]_\ell^g$. For each such m^g , and for all k such that $|k| \leq K$, we require

$$(6.7) \quad |\langle \omega, k \rangle + \Omega_{m^g} - \Omega_{n^g}| > 2\gamma \min(N^{-2d\tau_0}, 2^{-4d}|p_\ell|^{-2d}),$$

where $n^g = m^g + \pi(k)$ (see (2.6) for the definition of $\pi(k)$).

The set \mathcal{O}_+ is defined in order to ensure Lemma 6.2 below.

LEMMA 6.2. For all $\xi \in \mathcal{O}_+$, for all $k \in \mathbb{Z}^b, |k| \leq K$, and $l \in \mathbb{Z}^d, |l| \leq 2$ which satisfy momentum conservation, we have

$$(6.8) \quad |\langle \omega, k \rangle + l \cdot \Omega| \geq \gamma K^{-2d\tau_1}.$$

Before proving the lemma we give some relevant notation.

We know that $\tilde{\Omega}(z) := \sum_m \tilde{\Omega}_m |z_m|^2$ is quasi-Töplitz quadratic and diagonal, hence given θ, μ, τ , we apply Lemma 4.4 with $Q(z) = \tilde{\Omega}(z)$ to obtain the bounds (4.8), and (4.9) for all $m \xrightarrow{N} [v_i; p_i]$ which have a cut at ℓ with parameters (N, θ, μ, τ) :

$$(6.9) \quad \tilde{\Omega}_m = \hat{\Omega}([v_i; p_i]_\ell) + N^{-4d\tau} \bar{\Omega}_m.$$

Let us fix an affine subspace $A \xrightarrow{N} [v_i; p_i]_\ell$. By Lemma 3.10 there exists $\tau := \tau(p_\ell)$ (depending only on p_ℓ) such that every $m \in [v_i; p_i]_\ell^g$ has a cut at ℓ with parameters $(N, \theta, \mu, \tau(p_\ell))$ for all $\frac{1}{2} < \theta, \mu < 4$, hence

$$(6.10) \quad |\tilde{\Omega}_m - \hat{\Omega}([v_i; p_i]_\ell)| < 2LN^{-4d\tau(p_\ell)};$$

here $\hat{\Omega}([v_i; p_i]_\ell)$ plays the role of $\mathcal{Q}(A)$ while by (6.1) L dominates the Töplitz norm of $\tilde{\Omega}$. Note that in particular this relation holds for m^g .

Proof of Lemma 6.2. The cases with $|l| = 0, 1$ follow trivially from the definitions (6.4), and (6.5) since τ_1 is large with respect to τ_0 ; same for $\pm l = e_m + e_n$, and $l = e_m - e_n$ with $\max(|m|, |n|) < 8K^{\tau_1}$.

For the remaining cases we proceed in two steps: first we fix k , $N = K$, and one subspace $A \xrightarrow{K} [v_i; p_i]_\ell$, and we consider (6.7) with this choice of $k, [v_i; p_i]_\ell$. We show that this inequality implies that (6.8) holds for all $l = e_m - e_n$ such that $m \in [v_i; p_i]_\ell^g$ and $n = m + \pi(k)$. We prove this fact by using (6.10) with $N = K$. Finally Proposition 6.2 ensures that every point $m \notin A_0$ with $|m| > 4K^{\tau_1}$ must belong to some $[v_i; p_i]_\ell^g$.

Let m be any point in $[v_i; p_i]_\ell^g$. Let us first notice that

$$(6.11) \quad \langle \omega, k \rangle + |m|^2 - |n|^2 = \langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle,$$

hence (6.8) with $l = e_m - e_n$ is surely satisfied if $|\langle \pi(k), m \rangle| \geq 2K^3$ because in that case (6.11) is greater than $2K^3 - C_1^2 K^2 - |\omega|K > K^3$ provided that K is large with respect to C_1 and ω .

If on the other hand $|\langle \pi(k), m \rangle| < 2K^3$, then $\pi(k) \in B_K^g$ is in $\langle v_i \rangle_\ell$, otherwise we would have $|\langle \pi(k), m \rangle| > \frac{1}{2}K^{4d\tau_0}$ by definition of $[v_i; p_i]_\ell^g$, and recalling that $K^{4d\tau_0} > 4K^3$ by hypothesis. Thus for all $m \in [v_i; p_i]_\ell^g$ either (6.8) is trivially satisfied or

$$|m|^2 - |n|^2 = |\pi(k)|^2 - 2\langle \pi(k), m \rangle = |\pi(k)|^2 - 2\langle \pi(k), m^g \rangle;$$

recall that m^g is one fixed point in $[v_i; p_i]_\ell^g$ on which we have imposed the nonresonance conditions (6.7).

We apply (6.10) with $N = K$ to m, m^g and $n = m + \pi(k), n^g = \pi(k) + m^g$. We set $n \xrightarrow{K} [w_i; q_i]$, since $(\mu - \frac{1}{2})K^{\tau(p_\ell)}, (4 - \theta)K^{4d\tau(p_\ell)} > 5K^4$ we may apply Lemma 3.12 (with $r = n$) to conclude that n has an ℓ cut $[w_i; q_i]_\ell$ with parameters θ, μ, τ . Note moreover that, by Lemma 3.7(3) $[w_i; q_i]_\ell$ is completely fixed by $[v_i; p_i]_\ell$, and k . We have

$$|\tilde{\Omega}_n - \hat{\Omega}([w_i; q_i]_\ell)| < 2LK^{-4d\tau(p_\ell)},$$

and this relation holds also for $n^g = m^g + \pi(k)$. This implies that

$$|\tilde{\Omega}_m - \tilde{\Omega}_n - \tilde{\Omega}_{m^g} + \tilde{\Omega}_{n^g}| \leq 8LK^{-4d\tau(p_\ell)},$$

where by definition of τ , $K^{\tau(p_\ell)} = \max(K^{\tau_0}, 2|p_\ell|)$, and hence

$$(6.12) \quad \begin{aligned} |\langle \omega, k \rangle + \Omega_m - \Omega_n| &\geq |\langle \omega, k \rangle + \Omega_{m^g} - \Omega_{n^g}| - 8LK^{-4d\tau(p_\ell)} \\ &\geq \frac{\gamma}{2} \min(K^{-2d\tau_0}, 2^{-4d}|p_\ell|^{-2d}) \geq \gamma K^{-\tau_1}. \end{aligned}$$

Now we may apply Proposition 3.11 with $N = K$ to conclude that every point m with $|m| > 8K^{\tau_1}$ and $p_1 < CK^{4d\tau_0}$ belongs to some $[v_i; p_i]_\ell^g$. So the measure estimates for the points m which fall in this case are covered by (6.6).

Finally if $m \in A_0$ of (3.8), i.e., if we have $p_1 > CK^{4d\tau_0}$ then

$$|\pm \langle \omega, k \rangle + \Omega_m - \Omega_n| > |\pm \langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \tilde{\Omega}_m - \tilde{\Omega}_n| > K^{4d\tau_0} - 2K^2$$

since $\pi(k) \in B_K$, and hence $|\langle \pi(k), m \rangle| > p_1$.

We have shown that conditions (ii)–(iv) in \mathcal{O}^+ imply (6.8). \square

Remark 6.2. This lemma is essentially saying that by improving *only one* non-resonant condition (6.7), we impose *all* the conditions (6.8) with $l = e_m - e_n$ such that $m \in [v_i; p_i]_j^g$ and $n = m + \pi(k)$.

Remark 6.3. Notice that up to now we only use (6.7) and (6.10) with $N = K$. Indeed the other nonresonance conditions are only required in order to show that the quasi-Töplitz property is preserved in solving the homological equation.

LEMMA 6.3. *The set \mathcal{O}_+ is open, and has $|\mathcal{O} \setminus \mathcal{O}_+| \leq C\gamma K^{-\tau_0+b+d/2}$.*

For the measure estimates, given $\varrho > 0$ we define.

$$\mathcal{R}_{k,l}^\varrho := \{ \xi \in \mathcal{O} \mid |\langle \omega, k \rangle + \Omega \cdot l| < \gamma K^{-\varrho} \}.$$

LEMMA 6.4. *For all $(k, l) \neq (0, 0)$, $|k| \leq K$, and $|l| \leq 2$, which satisfy momentum conservation, one has $|\mathcal{R}_{k,l}^\varrho| \leq C\gamma K^{-\varrho}$.*

Proof. By assumption \mathcal{O} is contained in some open set of diameter D .

Choose a to be a vector such that $\langle k, a \rangle = |k|$; we have

$$\left| \frac{d}{dt} (\langle k, \omega(\xi + ta) \rangle + \Omega \cdot l) \right| \geq M(|k| - ML) \geq \frac{M}{2}$$

which leads to

$$\int_{\mathcal{R}_{k,l}^\varrho} d\xi \leq 2M^{-1}\gamma K^{-\varrho} \int_{\xi+ta \cap \mathcal{R}_{k,l}^\varrho} dt \int d\xi_2 \dots d\xi_b \leq 2M^{-1}D^{b-1}\gamma K^{-\varrho}. \quad \square$$

Proof of Lemma 6.3. The first statement is trivial, indeed (ii)–(iv) are a finite number of inequalities; notice that in (iv) for each $[v_i, p_i]_\ell^g$, and k we impose only one condition by choosing one couple m^g, n^g . Finally by Remark 3.2 there are a finite number of $[v_i, p_i]_\ell^g$. Item (i) apparently has infinitely many conditions since $h \in \mathbb{Z}$, however, we note that all but a finite number (i.e., $|h| < 2|\omega|K$) are trivially satisfied.

Let us prove the measure estimates; to impose (6.4) with $h = 0$ we have to remove

$$(6.13) \quad \left| \bigcup_{|k| \leq K} \mathcal{R}_{k,0}^{\tau_0} \right| \leq C(b)\gamma K^{-\tau_0+b}.$$

For $h \in \mathbb{Z}$ we set

$$\tilde{\mathcal{R}}_{k,h}^\varrho := \{ \xi \in \mathcal{O} \mid |\langle \omega, k \rangle + h| < \gamma K^{-\varrho} \},$$

and note that $\tilde{\mathcal{R}}_{k,h}^\varrho$ is empty if $|h| > 2|\omega||k|$. As in Lemma 6.4 for fixed (k, h) we have $|\tilde{\mathcal{R}}_{k,h}^\varrho| \leq C\gamma K^{-\varrho}$. Then

$$(6.14) \quad \left| \bigcup_{|k| \leq K, |h| \leq 2|\omega||k|} \tilde{\mathcal{R}}_{k,h}^{\tau_0} \right| \leq C(b)\gamma K^{-\tau_0+b+1}.$$

In order to impose the first Melnikov condition (6.5) we note that by momentum conservation in $\mathcal{R}_{k,l}^{\tau_0}$ we have $l = \pm e_{\mp\pi(k)}$. Then we have to remove

$$(6.15) \quad \left| \bigcup_{|k| \leq K, l = \pm e_{\mp\pi(k)}} \mathcal{R}_{k,l}^{\tau_0} \right| \leq C(b)\gamma K^{-\tau_0+b}.$$

If $l = \pm(e_m + e_n)$ the momentum conservation fixes $n = \mp\pi(k) - m$; we notice that the condition

$$\left| \pm \langle \omega, k \rangle + |m|^2 + |n|^2 + \tilde{\Omega}_m + \tilde{\Omega}_n \right| < \frac{1}{2}$$

implies $|\pm \langle \omega, k \rangle + |m|^2 + |n|^2| < 1$, and hence $|m|^2 + |n|^2 < 2|\omega|K$, and we have to remove a set of parameters:

$$(6.16) \quad \left| \bigcup_{\substack{k \leq K, \\ l = \pm(e_m + e_n)}} \mathcal{R}_{k,l}^{\tau_0} \right| = \left| \bigcup_{k \leq K} \bigcup_{\substack{l = \pm(e_m + e_n) \\ |m| \leq C(b)\sqrt{K}, n = -\pi(k) - m}} \mathcal{R}_{k,l}^{\tau_0} \right| \leq C\gamma K^{-\tau_0 + b + d/2}.$$

In conclusion one gets (6.4), and (6.5) with $\tau_0 > b + d/2$ and $l \neq \pm(e_m - e_n)$ by removing an open set of measure $C\gamma K^{-\tau_0 + b + d/2}$.

One trivially has

$$(6.17) \quad \left| \bigcup_{k \leq K} \bigcup_{\substack{l = \pm(e_m - e_n), m - n = \mp \pi(k), \\ \max(|m|, |n|) \leq 8K^{\tau_1}}} \mathcal{R}_{k,l}^{2d\tau_1} \right| \leq C\gamma K^{-d\tau_1 + b},$$

so we have (6.6) by removing an open set of measure $C\gamma K^{-d\tau_1 + b}$.

In order to deal with the last case, for all natural N such that $K \leq N \leq 2K^{\tau_1/\tau_0}$, for all affine subspaces $[v_i; p_i]_\ell$, and for all $|k| \leq K$ we set

$$(6.18) \quad \mathcal{R}_{k, [v_i; p_i]_\ell}^N := \{ \xi \mid |\langle \omega, k \rangle + \Omega_{m^g} - \Omega_{n^g}| < 2\gamma \min(N^{-2d\tau_0}, 2^{-4d}|p_\ell|^{-2d}) \}.$$

Following Lemma 6.4, $|\mathcal{R}_{k, [v_i; p_i]_\ell}^N| < C\gamma \min(N^{-2d\tau_0}, 2^{-4d}|p_\ell|^{-2d})$. By Remark 3.2 we have

$$\begin{aligned} & \left| \bigcup_{K \leq N \leq K^{\tau_1/\tau_0}} \bigcup_{\ell=0, \dots, d-1} \bigcup_{\frac{1}{2}N^{\tau_0} \leq |p_\ell| \leq 4N^{\frac{\tau_1}{4d}}} \bigcup_{\substack{[v_i; p_i]_\ell^g \\ |k| < K}} \mathcal{R}_{k, [v_i; p_i]_\ell}^N \right| \\ & \leq C\gamma \sum_{N \geq K} \sum_{\ell=0}^{d-1} \sum_{|p_\ell| > \frac{1}{2}N^{\tau_0}} |p_\ell|^{-2d-1+d} N^{\ell d} K^b \leq 4^d C_2 \gamma K^{-d\tau_0 + b}, \end{aligned}$$

so that we have (6.7) by removing an open set of measure $C\gamma K^{-d\tau_0 + b}$. \square

6.2. Quasi-Töplitz property. The main proposition of our paper is the following proposition.

PROPOSITION 6.5. *The functions $P_+, \tilde{\Omega}^+|z|^2$ are quasi-Töplitz with parameters (K_+, θ_+, μ_+) such that*

$$4K_+ < \sqrt{(\mu - \mu_+)(K_+)^{3/2}}, \quad 4\mu_+ K_+^4 < (\theta_+ - \theta) K_+^{4d\tau_0 - 1}.$$

The key to our strategy is based on the following three propositions which are proved in the appendix.

PROPOSITION 6.6. *For any $N \geq K$, $k \in \mathbb{Z}^b$ with $|k| < K$, and for all $|m|, |n| \geq \theta N^{\tau_1}$ such that $m - n = -\pi(k)$, $m \xrightarrow{N} [v_i; p_i]$, $n \xrightarrow{N} [w_i; q_i]$, and m, n have an ℓ cut with parameters θ, μ, τ for some choice of ℓ, τ one has*

$$\begin{aligned} & |\langle \omega, k \rangle + |m|^2 - |n|^2 + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)| \\ & = |\langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)| \\ & \geq \begin{cases} \gamma K^{-2d\tau_1\tau/\tau_0}, & \pi(k) \in \langle v_i \rangle_\ell, \\ \frac{1}{2}N^{4d\tau}, & \text{otherwise,} \end{cases} \end{aligned}$$

where $\hat{\Omega}([v_i; p_i]_\ell)$, and $\hat{\Omega}([w_i; q_i]_\ell)$ are defined by (6.9).

PROPOSITION 6.7. *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*

$$(6.19) \quad \|X_F\|_{r,s}^T \leq C\gamma^{-2} K^{3\tau_1^2/\tau_0} \|X_P\|_{r,s}^T,$$

where C is some constant.

Analytic quasi-Töplitz functions are closed under the Poisson bracket. More precisely we have the following proposition.

PROPOSITION 6.8. (i) Given $f^{(1)}, f^{(2)} \in \mathcal{A}_{r,s}$, quasi-Töplitz with parameters (K, θ, μ) we have that $\{f^{(1)}, f^{(2)}\} \in \mathcal{A}_{r',s'}$ is quasi-Töplitz for all parameters (K', θ', μ') such that $K', \theta', \mu', r', s'$ satisfy

$$(6.20) \quad \frac{1}{(K')^2} \leq (\mu - \mu'), \quad \frac{2\mu'}{(K')^{4d\tau_0-4}} < (\theta' - \theta), \quad e^{-(s-s')K'}(K')^{\tau_1} < 1.$$

We have the bounds

$$(6.21) \quad \|X_{\{f^{(1)}, f^{(2)}\}}\|_{r',s'}^T \leq C_1 \delta^{-1} \|X_{f^{(1)}}\|_{r,s}^T \|X_{f^{(2)}}\|_{r,s}^T,$$

where $\delta = (\frac{r'}{r})^2 \min(s - s', 1 - \frac{r'}{r})$

(ii) Given $f^{(1)}, f^{(2)}$ as in item (i), with $C_1 e \|X_{f^{(1)}}\|_{r,s}^T \delta^{-1} \ll 1$, the function $f^{(2)} \circ \phi_{f^{(1)}}^t := e^{t\{f^{(1)}, \cdot\}} f^{(2)}$, for $t \leq 1$, is quasi-Töplitz in $\mathcal{D}(r', s')$ for all parameters (K', θ', μ') such that

$$(6.22) \quad \frac{(\ln K')^2}{(K')^2} \leq (\mu - \mu'), \quad \frac{2\mu'(\ln K')^2}{(K')^{4d\tau_0-4}} < (\theta' - \theta), \quad e^{-(s-s')\frac{K'}{(\ln K')^2}}(K')^{\tau_1} < 1;$$

we have the bounds

$$\|X_{f^{(2)} \circ \phi_{f^{(1)}}^t}\|_{r',s'}^T \leq (1 - C_1 e \delta^{-1} \|X_{f^{(1)}}\|_{r,s}^T)^{-1} \|X_{f^{(2)}}\|_{r,s}^T.$$

7. Estimate, and KAM iteration.

7.1. Estimate on the coordinate transformation. We estimate X_F and ϕ_F^1 where F is given by (6.2).

LEMMA 7.1. Let $D_i = D(\frac{i}{4}r, s_+ + \frac{i}{4}(s - s_+))$, $0 < i \leq 4$, then

$$(7.1) \quad \|X_F\|_{D_3 \times \mathcal{O}_+} \leq C\gamma^{-2}K^{4d\tau_1}\varepsilon, \quad \|X_F\|_{D_3 \times \mathcal{O}_+}^T \leq C\gamma^{-2}K^{3\tau_1^2/\tau_0}\varepsilon.$$

LEMMA 7.2. Let $\eta = \varepsilon^{\frac{1}{3}}$, $D_{i\eta} = D(\frac{i}{4}\eta r, s_+ + \frac{i}{4}(s - s_+))$, $0 < i \leq 4$. If $\varepsilon \ll (\frac{1}{2}\gamma^2 K^{-3\tau_1^2/\tau_0})^3$, we then have that

$$(7.2) \quad \phi_F^t : D_{2\eta} \rightarrow D_{3\eta}, \quad -1 \leq t \leq 1,$$

is an analytic map, moreover,

$$(7.3) \quad \|\phi_F^t(z) - (z)\|_{D_{1\eta} \times \mathcal{O}_+} \leq C\gamma^{-2}K^{4d\tau_1}\varepsilon^{1/3}.$$

Proof. We first notice that

$$\|X_F\|_{3\eta}^T \leq c'\eta^{-2} \|X_F\|_{D_3 \times \mathcal{O}_+}^T \leq C\varepsilon^{-2/3}\gamma^{-2}K^{3\tau_1^2/\tau_0}\varepsilon < 1$$

by our smallness assumption. Let us denote by $\mathcal{B}_{2\eta}$ the space of close to identity analytic symplectic maps $D_{2\eta} \rightarrow \mathbb{C}^{2b} \times \ell_\rho$ with finite norm (2.4). Similarly we call $C([0, 1], \mathcal{B}_{2\eta})$ the Banach space of all continuous functions $t \mapsto \phi^t$ from $[0, 1]$ to $\mathcal{B}_{2\eta}$ endowed with the norm $\sup_{t \in [0,1]} \|\cdot\|_{2\eta}$. Consider the ball of radius $\rho := 2\|X_F\|_{3\eta} < 1$, and centered in $\phi^0 = id$. For ϕ^t in such ball consider the map

$$(7.4) \quad P(\phi^t) := id + \int_0^t X_F \circ \phi^s ds.$$

It is simple to see that the above map is a contraction, in particular

$$\sup_{t \in [0,1]} \left\| \int_0^t X_F \circ \phi^s ds \right\|_{2\eta} \leq \sup_{t \in [0,1]} \|X_F \circ \phi^t\|_{2\eta} \leq (1 + \rho) \|X_F\|_{3\eta} \leq \rho.$$

The lemma follows since the Hamiltonian flow ϕ_F^t generated by F at time $t \in [0, 1]$ is found as the fixed point of P . \square

7.2. Estimate of the new perturbation. The symplectic map ϕ_F^1 defined above transforms H into $H_+ = \mathcal{N}_+ + P_+$, where $\mathcal{N}_+ = \mathcal{N} + \langle R \rangle$, and

$$(7.5) \quad \begin{aligned} P_+ &= \int_0^1 (1-t) \{ \{ \mathcal{N}, F \}, F \} \circ \phi_F^t dt + \int_0^1 \{ \Pi_{\leq K} R, F \} \circ \phi_F^t dt + (P - \Pi_{\leq K} R) \circ \phi_F^1 \\ &= \int_0^1 \{ R(t), F \} \circ \phi_F^t dt + (P - \Pi_{\leq K} R) \circ \phi_F^1, \end{aligned}$$

with $R(t) = (1-t)(\mathcal{N}_+ - \mathcal{N}) + t\Pi_{\leq K} R$. Hence

$$X_{P_+} = \int_0^1 (\phi_F^t)^* X_{\{R(t), F\}} dt + (\phi_F^1)^* X_{(P - \Pi_{\leq K} R)}.$$

LEMMA 7.3. *The new perturbation P_+ satisfies the estimate*

$$\|X_{P_+}\|_{D(r_+, s_+)} \leq C\gamma^{-2} K^{4d\tau_1} \varepsilon^{4/3}.$$

Proof. According to Lemma 7.2,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} \leq c\gamma^{-2} K^{4d\tau_1} \varepsilon^{1/3}, \quad -1 \leq t \leq 1,$$

thus

$$\begin{aligned} \|D\phi_F^t\|_{D_{1\eta}} &\leq 1 + \|D\phi_F^t - Id\|_{D_{1\eta}} \leq 2, \quad -1 \leq t \leq 1, \\ \|X_{\{R(t), F\}}\|_{D_{2\eta}} &\leq \eta^{-2} \|X_{\{R(t), F\}}\|_{D_2} \leq C\gamma^{-2} K^{4d\tau_1} \eta^{-2} \varepsilon^2, \end{aligned}$$

and

$$\|X_{(P - \Pi_{\leq K} R)}\|_{D_{2\eta}} \leq C\eta\varepsilon,$$

and we have

$$\|X_{P_+}\|_{D(r_+, s_+)} \leq C\eta\varepsilon + C(\gamma^{-2} K^{4d\tau_1}) \eta^{-2} \varepsilon^2 \leq C\gamma^{-2} K^{4d\tau_1} \varepsilon^{4/3}. \quad \square$$

We need to show that P_+ is quasi-Töplitz, and estimate its Töplitz norm. We notice that $R(t)$ and $P - \Pi_{\leq K} R$ in (7.5) are quasi-Töplitz, by hypothesis (A4). Then, by Proposition 6.8(ii), we have that $R(t) \circ \phi_F^t = e^{\{F, \cdot\}} R(t)$ and $(P - \Pi_{\leq K} R) \circ \phi_F^t$ are quasi-Töplitz as well. Recalling Proposition 6.8, and repeating the reasoning of Lemma 7.3 with quasi-Töplitz norm, one has the following lemma.

LEMMA 7.4. *Set $\varepsilon_+ := C\gamma^{-2} K^{3\tau_1^2/\tau_0} \varepsilon^{4/3}$, then*

$$\|X_{P_+}\|_{D(r_+, s_+)}^T \leq \varepsilon_+.$$

7.3. Iteration lemma. In order to make the KAM machine work fluently, for any given $s, \varepsilon, r, \gamma$, and for all $\nu \geq 1$, we define the following sequences:

$$\begin{aligned}
 (7.6) \quad s_\nu &= s \left(1 - \sum_{i=2}^{\nu+1} 2^{-i} \right), \\
 r_\nu &= \frac{1}{4} \eta_{\nu-1} r_{\nu-1} = 2^{-2\nu} \left(\prod_{i=0}^{\nu-1} \varepsilon_i \right)^{\frac{1}{3}} r_0, \\
 \varepsilon_\nu &= c\gamma^{-2} K_{\nu-1}^{3\tau_1^2/\tau_0} \varepsilon_{\nu-1}^{\frac{4}{3}}, \quad \eta_\nu = \varepsilon_\nu^{\frac{1}{3}}, \\
 M_\nu &= M_{\nu-1} + \varepsilon_{\nu-1}, \quad L_\nu = L_{\nu-1} + \varepsilon_{\nu-1}, \\
 \mu_\nu &= \mu - \sum_{i=1}^{\nu} (\chi)^{-i}, \quad \theta_\nu = \theta + \sum_{i=1}^{\nu} (\chi)^{-i}, \\
 K_\nu &= c(s_{\nu-1} - s_\nu)^{-1} \ln \varepsilon_\nu^{-1},
 \end{aligned}$$

where $c, 1 < \chi < \frac{4}{3}$ is a constant, and the parameters $r_0, \varepsilon_0, L_0, s_0$, and K_0 are defined to be r, ε, L, s , and bounded by $\ln \varepsilon^{-1}$, respectively.

We iterate the KAM step, and proceed by induction.

LEMMA 7.5. *Suppose at the ν -step of KAM iteration, the Hamiltonian*

$$H_\nu = \mathcal{N}_\nu + P_\nu$$

is well defined in $D(r_\nu, s_\nu) \times \mathcal{O}_\nu$, where \mathcal{N}_ν is usual “integrable normal form,” P_ν , and $\sum \tilde{\Omega}_n^\nu |z_n|^2$ satisfy (A4) for $(K_\nu, \theta_\nu, \mu_\nu)$, ω_ν , and Ω_n^ν are C_W^1 smooth:

$$\begin{aligned}
 |\omega_\nu|_{C_W^1}, |\nabla \omega_\nu^{-1}|_{\mathcal{O}} &\leq M_\nu, |\tilde{\Omega}_n^\nu|_{C_W^1} \leq L_\nu, \quad |\Omega_n^\nu - \Omega_n^{\nu-1}|_{\mathcal{O}_\nu} \leq \varepsilon_{\nu-1}, \\
 \|X_{P_\nu}\|_{D(r_\nu, s_\nu), \mathcal{O}_\nu}^T &\leq \varepsilon_\nu, \quad \|\langle \tilde{\Omega}^\nu z, \bar{z} \rangle\|_{D(r_\nu, s_\nu), \mathcal{O}_\nu}^T \leq L_\nu.
 \end{aligned}$$

Then there exists a symplectic and quasi-Töplitz change of variables for parameter $(K_{\nu+1}, \theta_\nu, \mu_\nu)$,

$$(7.7) \quad \Phi_\nu : D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_{\nu+1} \rightarrow D(r_\nu, s_\nu),$$

where $|\mathcal{O}_{\nu+1} \setminus \mathcal{O}_\nu| \leq \gamma K_{\nu+1}^{-\tau_0 + b + \frac{d}{2}}$, such that on $D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_{\nu+1}$ we have

$$H_{\nu+1} = H_\nu \circ \Phi_\nu = e_{\nu+1} + \mathcal{N}_{\nu+1} + P_{\nu+1} = e_{\nu+1} + \langle \omega_{\nu+1}, I \rangle + \langle \Omega^{\nu+1} z, \bar{z} \rangle + P_{\nu+1}$$

with $\omega_{\nu+1} = \omega_\nu + \sum_{|l|=1} l P_{0,l,0,0}$, $\Omega_n^{\nu+1} = \Omega_n^\nu + P_{0,0,e_n,e_n}^\nu$.

$\mathcal{N}_{\nu+1}$ is an “integrable normal form.” $P_{\nu+1}$ and $\sum \tilde{\Omega}_n^{\nu+1} |z_n|^2$ satisfy (A4) for parameters $(K_{\nu+1}, \theta_{\nu+1}, \mu_{\nu+1})$. Functions $\omega_{\nu+1}$, and $\Omega_n^{\nu+1}$ are C_W^1 smooth:

$$\begin{aligned}
 |\omega_{\nu+1}|_{C_W^1}, |\nabla \omega_{\nu+1}^{-1}|_{\mathcal{O}} &\leq M_{\nu+1}, |\tilde{\Omega}_n^{\nu+1}|_{C_W^1} \leq L_{\nu+1}, \quad |\Omega_n^{\nu+1} - \Omega_n^\nu|_{\mathcal{O}_{\nu+1}} \leq \varepsilon_\nu, \\
 \|X_{P_{\nu+1}}\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_{\nu+1}}^T &\leq \varepsilon_{\nu+1}, \quad \|\langle \tilde{\Omega}^{\nu+1} z, \bar{z} \rangle\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_{\nu+1}}^T \leq L_{\nu+1}.
 \end{aligned}$$

- By Proposition 6.5, the new perturbations $P_{\nu+1}$, and $\langle \tilde{\Omega}^{\nu+1} z, z \rangle$ satisfy the quasi-Töplitz property for parameters $(K_{\nu+1}, \theta_{\nu+1}, \mu_{\nu+1})$. As we can see, when we require $\tau_1 > \tau_0 > 12$,

$$\forall N \geq K_{\nu+1} = c(s_{\nu-1} - s_\nu)^{-1} \ln \varepsilon_\nu^{-1} > K_0 2^\nu$$

implies the inequality

$$2N \leq \sqrt{(\mu_\nu - \mu_{\nu+1})} N^{3/2}, \quad 4\mu' N^4 < (\theta_{\nu+1} - \theta_\nu) N^{4d\tau_0 - 1}.$$

- Since the set of Hamiltonians which Poisson commute with M (the momentum) is closed under Poisson brackets (or by using Lemma 4.4 in [12]) $P_{\nu+1}$ satisfies momentum conservation (namely, it Poisson commutes with M).

7.4. Convergence. Suppose that the assumptions of Theorem 5.1 are satisfied. Recall

$$\varepsilon_0 = \varepsilon, r_0 = r, s_0 = s, M_0 = M, L_0 = L, \mathcal{N}_0 = \mathcal{N}, P_0 = P,$$

\mathcal{O} is an open set. The assumptions of the iteration lemma are satisfied when $\nu = 0$ if ε_0, γ are sufficiently small. Inductively, we obtain sequences

$$\begin{aligned} \mathcal{O}_{\nu+1} &\subset \mathcal{O}_\nu, \\ \Psi^\nu &= \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_{\nu+1} \rightarrow D(r_0, s_0), \nu \geq 0, \\ H \circ \Psi^\nu &= H_{\nu+1} = \mathcal{N}_{\nu+1} + P_{\nu+1}. \end{aligned}$$

Let $\tilde{\mathcal{O}} = \bigcap_{\nu=0}^\infty \mathcal{O}_\nu$, since at ν step the parameter we excluded is bounded by $C\gamma K_\nu^{-\tau_0+b+d/2}$, the total measure we excluded with the infinity step of KAM iteration is bounded by γ which guarantees $\tilde{\mathcal{O}}$ is a nonempty set; actually it has positive measure.

As in [20, 21], with Lemma 7.2, $\mathcal{N}_\nu, \Psi^\nu, D\Psi^\nu, \omega_\nu$ converge uniformly on $D(0, \frac{s}{2}) \times \tilde{\mathcal{O}}$ with

$$\mathcal{N}_\infty = e_\infty + \langle \omega_\infty, I \rangle + \sum_n \Omega_n^\infty z_n \bar{z}_n.$$

Since $K_\nu = c(s_{\nu-1} - s_\nu)^{-1} \ln \varepsilon_\nu^{-1}$, we have $\varepsilon_\nu = c\gamma^2 K_{\nu-1}^{\frac{3\tau_1^2}{\tau_0}} \varepsilon_{\nu-1}^{\frac{4}{3}} \rightarrow 0$ once ε is sufficiently small. And with this we have ω_∞ is slightly different from ω .

Let ϕ_H^t be the flow of X_H . Since $H \circ \Psi^\nu = H_{\nu+1}$, there is

$$(7.8) \quad \phi_H^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_{\nu+1}}^t.$$

The uniform convergence of $\Psi^\nu, D\Psi^\nu, \omega_\nu$, and X_{H_ν} implies that the limits can be taken on both sides of (7.8). Hence, on $D(0, \frac{s}{2}) \times \tilde{\mathcal{O}}$ we get

$$(7.9) \quad \phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t$$

and

$$\Psi^\infty : D\left(0, \frac{s}{2}\right) \times \tilde{\mathcal{O}} \rightarrow D(r, s) \times \mathcal{O}.$$

From 7.9, for $\xi \in \tilde{\mathcal{O}}$, $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ is an embedded torus which is invariant for the original perturbed Hamiltonian system at $\xi \in \tilde{\mathcal{O}}$. The normal behavior of this invariant tori is governed by normal frequency Ω_∞ .

Appendix A. Proof of Propositions 6.6, 6.7, and 6.8.

A.1. Proposition 6.6.

Proof. By hypothesis

$$(A.1) \quad \begin{aligned} |m|, |n| &\geq \theta N^{\tau_1}, & m \xrightarrow{N} [v_i; p_i], & n \xrightarrow{N} [w_i; q_i], \\ |q_\ell|, |p_\ell| &\leq \mu N^\tau, & |q_{\ell+1}|, |p_{\ell+1}| &\geq \theta N^{4d\tau}, & [v_i; p_i]_\ell \prec [w_i; q_i]_\ell. \end{aligned}$$

By definition of quasi-Töplitz (see (6.10)), one has

$$(A.2) \quad |\tilde{\Omega}_m - \hat{\Omega}([v_i; p_i]_\ell)|, |\tilde{\Omega}_n - \hat{\Omega}([w_i; q_i]_\ell)| \leq 2LN^{-4d\tau}.$$

Recall that $m - n = -\pi(k)$, so one has

$$|m|^2 - |n|^2 = \langle m + n, m - n \rangle = |\pi(k)|^2 - 2\langle \pi(k), m \rangle.$$

If $\pi(k) \notin \langle v_i \rangle_\ell$ then $|\langle \pi(k), m \rangle| > N^{4d\tau} > K^3$, and the denominator is not small:

$$|\langle \omega, k \rangle + |m|^2 - |n|^2 + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)| > \frac{1}{2}N^{4d\tau},$$

since (again by definition of quasi-Töplitz) $|\hat{\Omega}([v_i; p_i]_\ell)|, |\hat{\Omega}([w_i; q_i]_\ell)| \leq 2L$.

If $\pi(k) \in \langle v_i \rangle_\ell$ then the value of $\langle \pi(k), m \rangle$ is fixed for all $m \in [v_i; p_i]_\ell$.

We know that $m \xrightarrow{K} [v'_i; p'_i]$ has a standard cut, so that $m \in [v'_i; p'_i]_{\bar{\ell}}^g$ for some $\bar{\ell}$. If $2^{4d}K^{\tau_1} < N^{\tau_0}$ then

$$\begin{aligned} & |\langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)| \\ & \stackrel{(A.2)}{\geq} |\langle \omega, k \rangle + \Omega_m - \Omega_n| - 4LN^{-4d\tau} \\ & \stackrel{(6.12)}{\geq} \gamma \min(K^{-2d\tau_0}, 2^{-4d}|p'_\ell|^{-2d}) - 4L|N|^{-4d\tau} \geq \frac{\gamma}{2} \min(K^{-2d\tau_0}, |p'_\ell|^{-2d}), \end{aligned}$$

since $|p'_\ell| < 4K^{\tau_1/4d}$ by the definition of standard cut.

If on the other hand we have $2^{4d}K^{\tau_1} > N^{\tau_0}$ we proceed as follows. We have seen that we may restrict to the case $\pi(k) \in \langle v_i \rangle_j$, where

$$|m|^2 - |n|^2 = |\pi(k)|^2 - 2\langle \pi(k), m \rangle = |\pi(k)|^2 - 2\langle \pi(k), m^g \rangle,$$

where (notice that $N < 2K^{\tau_1/\tau_0}$), $m^g := m^g(N)$ is the point in $[v_i; p_i]_\ell^g$ chosen for the measure estimates (6.7).

We notice that m^g, n^g satisfy the conditions (A.1), so we apply (A.2) to m, n, m^g, n^g . We have

$$\begin{aligned} & |\langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)| \\ & \geq |\langle \omega, k \rangle + \Omega_{m^g} - \Omega_{n^g}| - 4LN^{-4d\tau} \\ & \geq \frac{\gamma}{2} \min(N^{-2d\tau_0}, 2^{-2d}|p_\ell|^{-2d}) - 4LN^{-4d\tau} \\ & \geq \frac{\gamma}{4} \min(N^{-2d\tau_0}, 2^{-2d}|p_\ell|^{-2d}) \geq \gamma K^{\frac{-2d\tau_1}{\tau_0}} \end{aligned}$$

since by definition $|p_j| < \mu'N^\tau < 4N^\tau$, $N \leq 2K^{\tau_1/\tau_0}$. \square

A.2. Proposition 6.7.

Proof. The quasi-Töplitz property is a condition on the (N, θ, μ, τ) -bilinear part of F , where F is at most quadratic. Hence we only need to consider the quadratic terms

$$(A.3) \quad \Pi_{(N, \theta, \mu, \tau)} F = \sum_{\substack{|k| < N, |m|, |n| > \theta N^{\tau_1} \\ \exists \ell: m, n \text{ have an } \ell \text{ cut} \\ \text{with parameters } N, \theta, \mu, \tau}} e^{i\langle k, \vartheta \rangle} (F_{k, 0, e_m, e_n} z_m \bar{z}_n + F_{k, 0, e_m + e_n, 0} z_m z_n) + \text{complex conjugate}.$$

Recall that

$$(A.4) \quad F_{k,0,e_m,e_n} = \frac{P_{k,0,e_m,e_n}}{\langle k, \omega \rangle + \tilde{\Omega}_m - \tilde{\Omega}_n}, \quad F_{k,0,e_m+e_n,0} = \frac{P_{k,0,e_m+e_n,0}}{\langle \omega, k \rangle + \tilde{\Omega}_m + \tilde{\Omega}_n}.$$

By hypothesis $|m|, |n| > \theta N^{\tau_1}$ so in the case of $F_{k,0,e_m+e_n,0}$ one has

$$|F_{k,0,e_m+e_n,0}| = \frac{|P_{k,0,e_m+e_n,0}|}{\langle k, \omega \rangle + |m|^2 + |n|^2 + \tilde{\Omega}_m + \tilde{\Omega}_n} \leq |P_{k,0,e_m+e_n,0}| N^{-\tau_1},$$

since

$$|\langle k, \omega \rangle + |m|^2 + |n|^2 + \tilde{\Omega}_m + \tilde{\Omega}_n| > 2N^{\tau_1} - cK - 2L.$$

We proceed in the same way for $\partial_\xi F_{k,0,e_m+e_n,0}$. This means that $F_{k,0,e_m+e_n,0}$ is quasi-Töplitz with the ‘‘Töplitz approximation’’ equal to zero. Recalling that P is quasi-Töplitz we deduce, by Remark 4.1, that if $m \xrightarrow{N} [v_i; p_i]$, $|m|, |n| > \theta N^{\tau_1}$, and m, n have a cut ℓ, τ , then we have:

$$P_{k,0,e_m,e_n} = \mathcal{P}_k(m - n, [v_i; p_i]_\ell) + N^{-4d\tau} \bar{P}_{k,0,e_m,e_n}.$$

Note that by definition (see (4.4)) for all m, n which have an ℓ, τ cut the Töplitz approximation $\mathcal{P}_k(m - n, [v_i; p_i]_\ell)$ must depend only on $m - n$ on the affine subspace $[v_i; p_i]_\ell$, and on k . Moreover the approximation (A.5) must hold for all $m \in [v_i; p_i]_\ell$ which have a cut ℓ, τ (naturally if we fix τ , and an affine subspace $[v_i; p_i]_\ell$ it may well be possible that no integer point $m \in [v_i; p_i]_\ell$ has a cut ℓ, τ).

Finally since $\sum_m \tilde{\Omega}_m z_m \bar{z}_m$ is quasi-Töplitz, diagonal, and quadratic we have

$$\tilde{\Omega}_m = \hat{\Omega}([v_i; p_i]_\ell) + N^{-4d\tau} \bar{\Omega}_m$$

for all $m \xrightarrow{N} [v_i; p_i]$ which have an ℓ, τ cut.

We wish to show that

$$(A.5) \quad F_{k,0,e_m,e_n} = \mathcal{F}_k(m - n, [v_i; p_i]_\ell) + N^{-4d\tau} \bar{F}_{k,0,e_m,e_n};$$

here \mathcal{F}_k is the k Fourier coefficient of the Töplitz approximation \mathcal{F} .

By hypothesis we have conditions (A.1), and $\langle v_1, \dots, v_\ell \rangle = \langle w_1, \dots, w_\ell \rangle$. This in turn implies that the subspace $[w_i, q_i]_\ell$ is obtained from $[v_i, p_i]_\ell$ by translation by $m - n = -\pi(k)$. If $\pi(k) \notin \langle v_i \rangle_\ell$ then the denominator in the first of (A.4) is

$$|\langle k, \omega \rangle + \Omega_m - \Omega_n| > |\langle k, \omega \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle| - 2L > \frac{1}{4} N^{4d\tau},$$

and we may again set $\mathcal{F}_k(m - n, [v_i, p_i]_\ell) = 0$. Otherwise we set

$$\mathcal{F}_k(m - n, [v_i, p_i]_\ell) = \frac{\mathcal{P}_k(m - n, [v_i, p_i]_\ell)}{\langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell)}.$$

We notice that $\langle \pi(k), m \rangle$ depends only on the subspace $[v_i, p_i]_\ell$, and on $\pi(k)$. Moreover by definition $\hat{\Omega}(\cdot)$ depends only on the affine subspace on which it is computed; finally $[w_i; q_i]_\ell$ depends only on $[v_i; p_i]_\ell$, and on k . Hence $\mathcal{F}_k(m - n, [v_i, p_i]_\ell)$ depends only on $k, m - n$, and $[v_i, p_i]_\ell$ as was our claim. Finally we apply Proposition 6.6 to bound

the denominator. In order to bound the derivatives in ξ of F we proceed in the same way, only the denominators may appear to the power of two.

Finally to bound \bar{F} we notice that

$$\bar{F}_{k,m,n} = \frac{\bar{P}_{k,m,n}}{\mathcal{D}} + N^{4d\tau} \mathcal{P}_k(m-n, [v_i; p_i]_\ell) \frac{\tilde{\Omega}_m - \hat{\Omega}([v_i; p_i]_\ell) - \tilde{\Omega}_n + \hat{\Omega}([w_i; q_i]_\ell)}{D\mathcal{D}},$$

where

$$\mathcal{D} = \langle \omega, k \rangle + |\pi(k)|^2 - 2\langle \pi(k), m \rangle + \hat{\Omega}([v_i; p_i]_\ell) - \hat{\Omega}([w_i; q_i]_\ell), \quad D = \langle \omega, k \rangle + \Omega_m - \Omega_n,$$

and $N^{4d\tau} |\tilde{\Omega}_m - \hat{\Omega}([v_i; p_i]_\ell)| \leq 2L$. In conclusion taking the $\sup_{N>K, \tau<\tau_1}$

$$\|X_F\|_{r,s}^T \leq C\gamma^{-2} N^{\frac{3\tau_1}{\tau_0}} \|X_P\|_{r,s}^T \quad \square$$

A.3. Proposition 6.8. Before proving Proposition 6.8, we discuss some technical lemmas, and set up some notation. We divide the Poisson bracket into four terms: $\{\cdot, \cdot\} = \{\cdot, \cdot\}^{I,\vartheta} + \{\cdot, \cdot\}^L + \{\cdot, \cdot\}^H + \{\cdot, \cdot\}^R$, where the superscripts L, H, R identify the variables in which we are performing the derivatives (the symbol R summarizes the derivatives in all the w_i which are neither low nor high momentum). We call a monomial

$$e^{i\langle k, \vartheta \rangle} I^l z^\alpha \bar{z}^\beta$$

1. of (N, μ) -low momentum if $|k| < N$ and $\sum_j |j|(\alpha_j + \beta_j) < \mu N^3$. Denote by $\Pi_{N,\mu}^L$ the projection on this subspace;

2. of N -high frequency if $|k| \geq N$. Denote by Π_N^U the projection on this subspace.

Recall that the projection symbol $\Pi_{N,\theta,\mu,\tau}$ is given in Definition 4.1. A function f then may be uniquely represented as $f = \Pi_{N,\theta,\mu,\tau} f + \Pi_{N,\mu}^L f + \Pi_N^U f + \Pi_R f$, where $\Pi_R f$ is by definition the projection on those monomials which are neither (N, θ, μ, τ) -bilinear nor of (N, μ) -low momentum nor of N -high frequency.

A technical lemma is given below.

LEMMA A.1. *The following splitting formula holds:*

$$\begin{aligned} \text{(A.6)} \quad & \Pi_{N,\theta',\mu',\tau} \{f^{(1)}, f^{(2)}\} \\ &= \Pi_{N,\theta',\mu',\tau} \left(\{ \Pi_{N,\theta,\mu,\tau} f^{(1)}, \Pi_{N,\theta,\mu,\tau} f^{(2)} \}^H + \{ \Pi_{N,\theta,\mu,\tau} f^{(1)}, \Pi_{N,2\mu}^L f^{(2)} \}^{I,\vartheta} \right. \\ & \quad + \{ \Pi_{N,\theta,\mu,\tau} f^{(1)}, \Pi_{N,2\mu}^L f^{(2)} \}^L + \{ \Pi_N^U f^{(1)}, f^{(2)} \} \\ & \quad + \{ \Pi_{N,2\mu}^L f^{(1)}, \Pi_{N,\theta,\mu,\tau} f^{(2)} \}^{I,\vartheta} + \{ \Pi_{N,2\mu}^L f^{(1)}, \Pi_{N,\theta,\mu,\tau} f^{(2)} \}^L \\ & \quad \left. + \{ f^{(1)}, \Pi_N^U f^{(2)} \} \right). \end{aligned}$$

Proof. We perform a case analysis: we replace each $f^{(i)}$ with a single monomial to show which terms may contribute nontrivially to the projection $\Pi_{N,\theta',\mu',\tau} \{f^{(1)}, f^{(2)}\}$.

Consider the expression

$$\Pi_{N,\theta',\mu',\tau} \{ e^{i\langle k^{(1)}, \vartheta \rangle} I^{l^{(1)}} z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}}, e^{i\langle k^{(2)}, \vartheta \rangle} I^{l^{(2)}} z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}} \}.$$

If one or both of the $|k^{(i)}| > N$ then one or both monomials are of high frequency, and we obtain the last term in the second and third lines of (A.6).

Suppose now that $|k^{(1)}|, |k^{(2)}| < N$; we wish to understand under which conditions on the $\alpha^{(i)}, \beta^{(i)}$ this expression is not zero. By direct inspection, one of the following situations (apart from a trivial permutation of the indexes 1, 2) must hold:

1. one has $z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}} = z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z_m^\sigma z_j^{\sigma_1}$ and $z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}} = z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}} z_n^{\sigma'} z_j^{-\sigma_1}$, where $|m|, |n| \geq \theta' N^{\tau_1}$ have a cut for some ℓ with parameters (N, θ', μ', τ) , and $z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}}$ is of (N, μ') -low momentum. The derivative in the Poisson bracket is on w_j ;
2. one has $z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}} = z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z_m^\sigma z_n^{\sigma'}$ and $z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}} = z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}}$, where $|m|, |n| \geq \theta' N^{\tau_1}$ have a cut for some ℓ with parameters (N, θ', μ', τ) and $z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}}$ is of (N, μ') -low momentum. The derivative in the Poisson bracket is on I, ϑ ;
3. one has $z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}} = z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z_m^\sigma z_n^{\sigma'} z_j^{\sigma_1}$ and $z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}} = z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}} z_j^{-\sigma_1}$, where $|m|, |n| \geq \theta' N^{\tau_1}$ have a cut for some ℓ with parameters (N, θ', μ', τ) , and $z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}}$ is of (N, μ') -low momentum. The derivative in the Poisson bracket is on w_j ;
4. one has $z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}} = z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z_m^\sigma$ and $z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}} = z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}} z_n^{\sigma'}$, where $|m|, |n| \geq \theta' N^{\tau_1}$ have a cut for some ℓ with parameters (N, θ', μ', τ) , and $z^{\bar{\alpha}^{(1)}} \bar{z}^{\bar{\beta}^{(1)}} z^{\bar{\alpha}^{(2)}} \bar{z}^{\bar{\beta}^{(2)}}$ is of (N, μ') -low momentum. The derivative in the Poisson bracket is on I, ϑ .

Case 1. We apply momentum conservation to both monomials and obtain

$$\sigma_1 j = -\sigma m - \pi(k^{(1)}, \bar{\alpha}^{(1)}, \bar{\beta}^{(1)}) = \sigma' n + \pi(k^{(2)}, \bar{\alpha}^{(2)}, \bar{\beta}^{(2)}).$$

Recall that

$$\sum_{l \in \mathbb{Z}_1^d} |l|(\bar{\alpha}_l^{(1)} + \bar{\beta}_l^{(1)} + \bar{\alpha}_l^{(2)} + \bar{\beta}_l^{(2)}) \leq \mu' N^3 \longrightarrow \sum_{l \in \mathbb{Z}_1^d} |l|(\bar{\alpha}_l^{(i)} + \bar{\beta}_l^{(i)}) \leq \mu' N^{\tau_1}$$

and by hypothesis $|k^{(i)}| \leq N$; this implies that $|j| > \theta' N^{\tau_1} - \mu' N^3 - CN > \theta N^{\tau_1}$ for $N > K'$ respecting (6.20) (recall that C is a constant so that $|\pi(k)| \leq C|k|$). Hence $\min(|m|, |n|, |j|) > \theta N^{\tau_1}$. By momentum conservation $|\sigma m + \sigma_1 j|, |-\sigma_1 j + \sigma' n| \leq CN + \mu' N^3 \leq 5N^3$; by hypothesis n, m have a cut ℓ with parameters (N, θ', μ', τ) . By Lemma 3.7 also $j \xrightarrow{N} [w_i; q_i]$ has a cut ℓ with parameters (N, θ, μ, τ) . Then $e^{i(k^{(i)}, \vartheta)} z^{\alpha^{(i)}} \bar{z}^{\beta^{(i)}}$ are by definition (N, θ, μ, τ) -bilinear. The derivative in the Poisson bracket is on j which is a high momentum variable.

As m, n run over all possible vectors in \mathbb{Z}_1^d with $|m|, |n| \geq \theta' N$, we obtain the first term in (A.6).

Case 2. Following the same argument $e^{i(k^{(1)}, \vartheta)} z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}}$ is (N, θ', μ', τ) -bilinear and $e^{i(k^{(2)}, \vartheta)} z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}}$ is (N, μ') -low momentum. We obtain the second contribution in (A.6).

Case 3. We apply momentum conservation to the second monomial, and obtain $-\sigma_1 j = -\pi(k^{(2)}, \bar{\alpha}^{(2)}, \bar{\beta}^{(2)})$. This implies that

$$\begin{aligned} |j| + \sum_{l \in \mathbb{Z}_1^d} |l|(\bar{\alpha}_l^{(1)} + \bar{\beta}_l^{(1)}) &\leq |\pi(k^{(2)}, \bar{\alpha}^{(2)}, \bar{\beta}^{(2)})| + \sum_{l \in \mathbb{Z}_1^d} |l|(\bar{\alpha}_l^{(1)} + \bar{\beta}_l^{(1)}) \\ &\leq CN + \sum_{l \in \mathbb{Z}_1^d} |l|(\bar{\alpha}_l^{(1)} + \bar{\beta}_l^{(1)} + \bar{\alpha}_l^{(2)} + \bar{\beta}_l^{(2)}) \\ &\leq \mu' N^3 + CN \leq \mu N^3 \end{aligned}$$

if $N > K'$ with K' satisfying (6.20). Then $e^{i(k^{(1)}, \vartheta)} z^{\alpha^{(1)}} \bar{z}^{\beta^{(1)}}$ is, by definition, (N, θ, μ, τ) -bilinear, and $e^{i(k^{(2)}, \vartheta)} z^{\alpha^{(2)}} \bar{z}^{\beta^{(2)}}$ is $(N, 2\mu)$ -low momentum. The derivative in the Poisson bracket is on j which is a low momentum variable. We obtain the third contribution in (A.6).

Case 4. We apply momentum conservation to both monomials and we get

$$\min(|\sigma m|, |\sigma' n|) \leq \max_{i=1,2} (|-\pi(k^{(i)}, \bar{\alpha}^{(i)}, \bar{\beta}^{(i)})|) \leq CN + \mu' N^3,$$

which is in contradiction to the hypothesis $|m|, |n| \geq \theta' N^{\tau_1}$. Hence Case 4 does not give any contribution.

The third line in (A.6) is dealt with just as the second line is by exchanging the indexes 1, 2. \square

In order to show that $\{f^{(1)}, f^{(2)}\}$ is quasi-Töplitz, for all $N > K'$ and τ we have to provide a decomposition

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

so that $\mathcal{F}^{(1,2)} \in \mathbb{F}$, and

$$(A.7) \quad \|X_{\mathcal{F}^{(1,2)}}\|_{r', s'}, \|X_{\bar{f}^{(1,2)}}\|_{r', s'} < \delta^{-1} C \|X_{f^{(1)}}\|_{r, s}^T \|X_{f^{(1)}}\|_{r, s}^T$$

for some constant C .

Using Remark 4.3, we substitute in (A.6) $\Pi_{N, \theta', \mu', \tau} f^{(i)} = \mathcal{F}^{(i)} + N^{-4d\tau} \bar{f}^{(i)}$, with $\mathcal{F}^{(i)} \in \mathbb{F}$.

LEMMA A.2. Consider the function

$$\begin{aligned} \mathcal{F}^{(1,2)} = & \Pi_{N, \theta', \mu', \tau} \left(\{\mathcal{F}^{(1)}, \mathcal{F}^{(2)}\}^H + \{\mathcal{F}^{(1)}, \Pi_{N, 2\mu}^L f^{(2)}\}^{(I, \vartheta)+L} \right. \\ & \left. + \{\Pi_{N, 2\mu}^L f^{(1)}, \mathcal{F}^{(2)}\}^{(I, \vartheta)+L} \right), \end{aligned}$$

where we have denoted $\{\cdot, \cdot\}^{(I, \vartheta)+L} = \{\cdot, \cdot\}^{(I, \vartheta)} + \{\cdot, \cdot\}^L$. (i) One has $\mathcal{F}^{(1,2)} \in \mathbb{F}$. (ii) Setting $\bar{f}^{(1,2)} = N^{4d\tau} (\Pi_{N, \theta', \mu', \tau} \{f^{(1)}, f^{(2)}\} - \mathcal{F}^{(1,2)})$ one has that the bounds (A.7) hold.

Proof. In order to prove the first statement it is useful to write

$$\mathcal{F}^{(i)} = \sum_{\substack{A=[v_i; p_i]_\ell \in \mathcal{H}_N \\ |p_\ell| < \mu N^\tau}} \sum_{\sigma, \sigma' = \pm 1} \sum_{m, n}^{(N, \theta', \mu', \tau, A)} [\mathcal{F}^{(i)}]^{\sigma, \sigma'}(I, \vartheta, w^L; \sigma m + \sigma' n, [v_i; p_i]_\ell) z_m^\sigma z_n^{\sigma'}$$

where $\sum^{(N, \theta', \mu', \tau, A)}$ is the sum over those n, m which respect (4.1), and have the ℓ cut at $A = [v_i; p_i]_\ell$ with the parameters θ', μ', τ . For compactness of notation we will omit the dependence on (I, ϑ, w^L) .

The fact that $\{\mathcal{F}^{(1)}, \Pi_{N, 2\mu}^L f^{(2)}\}^{I, \vartheta+L} \in \mathbb{F}$ is obvious. Indeed the coefficient of $z_m^\sigma z_n^{\sigma'}$ is

$$\{\mathcal{F}^{(1)}(\sigma m + \sigma' n, [v_i; p_i]_\ell), \Pi_{N, 2\mu}^L f^{(2)}\}^{I, \vartheta+L},$$

the same for $\{\mathcal{F}^{(2)}, \Pi_{N, 2\mu}^L f^{(1)}\}^{I, \vartheta+L}$.

Suppose now that n, m respect (4.1), and have the ℓ cut $[v_i; p_i]_\ell$ with the parameters θ', μ', τ . By the rules of Poisson brackets the coefficient of $z_m^\sigma z_n^{\sigma'}$ in the expression $\{\mathcal{F}^{(1)}, \mathcal{F}^{(2)}\}^H$ is

$$(A.8) \quad \sum_{\substack{r \in \mathbb{Z}_1^d, \sigma_1 = \pm 1 \\ |r| \geq \theta N^{\tau_1} \\ |\sigma m + \sigma_1 r| \leq \mu N^3 \\ |-\sigma_1 r + \sigma' n| \leq \mu N^3}} -\sigma_1 [\mathcal{F}^{(1)}]^{\sigma, \sigma_1}(\sigma m + \sigma_1 r, [v_i; p_i]_\ell) [\mathcal{F}^{(2)}]^{-\sigma_1, \sigma'}(-\sigma_1 r + \sigma' n; [w_i; q_i]_\ell).$$

Since $|\sigma m + \sigma_1 r|, |\sigma' n - \sigma_1 r| \leq \mu N^3$ and $|m|, |n| > \theta' N^{\tau_1}$ we have that the condition $|r| > \theta N^{\tau_1}$ is automatically fulfilled. By Lemma 3.7 r, n, m all have an ℓ cut with parameters (θ, μ, τ) . We set $m \xrightarrow{N} [v_i; p_i], n \xrightarrow{N} [v'_i; p'_i], r \xrightarrow{N} [w_i; q_i]$. Again by Lemma 3.7 $\langle v_i \rangle_\ell = \langle v'_i \rangle_\ell = \langle w_i \rangle_\ell$, moreover $[w_i; q_i]_\ell$ is completely fixed by $[v_i; p_i]_\ell, \sigma, \sigma_1$, and by $\sigma m + \sigma_1 r := h$. We may suppose (the other cases are done in the same way) that

$$(p_1, \dots, p_\ell, v_1, \dots, v_\ell) \preceq (q_1, \dots, q_\ell, w_1, \dots, w_\ell) \preceq (p'_1, \dots, p'_\ell, v'_1, \dots, v'_\ell).$$

Note also that this order relation depends only on $\sigma, \sigma', \sigma_1, [v_i; p_i]_\ell, \sigma m + \sigma' n$, and $\sigma m + \sigma_1 r = h$. Then we may change variables in the sum over r in (A.8):

$$\sum_{\sigma_1 = \pm 1} \sum_{\substack{h: |h| < \mu N^3 \\ |\sigma m + \sigma' n - h| \leq \mu N^3}} -\sigma_1 [\mathcal{F}^{(1)}]^{\sigma, \sigma_1}(h, [v_i; p_i]_\ell) [\mathcal{F}^{(2)}]^{-\sigma_1, \sigma'}(\sigma m + \sigma' n - h; [w_i; q_i]_\ell);$$

this expression only depends on $[v_i; p_i]_\ell$. The estimate (A.7) for $\mathcal{F}^{(1,2)}$ follows by Cauchy estimates since

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

We now compute

$$\begin{aligned} \bar{f} &= \Pi_{N, \theta', \mu', \tau} \left(\{\Pi_{N, \theta, \mu, \tau} f^{(1)}, \bar{f}^{(2)}\}^H + \{\bar{f}^{(1)}, \mathcal{F}^{(2)}\}^H \right. \\ &\quad + \{\bar{f}^{(1)}, \Pi_{N, \mu}^L f^{(2)}\}^{J, \vartheta} + \{\bar{f}^{(1)}, \Pi_{N, \mu}^L f^{(2)}\}^L + N^{4d\tau} \{\Pi_N^U f^{(1)}, f^{(2)}\} \\ &\quad \left. + \{\Pi_{N, \mu}^L f^{(1)}, \bar{f}^{(2)}\}^{I, \vartheta} + \{\Pi_{N, \mu}^L f^{(1)}, \bar{f}^{(2)}\}^L + N^{4d\tau} \{f^{(1)}, \Pi_N^U f^{(2)}\} \right). \end{aligned}$$

Since $e^{-N(s-s')} < N^{-\tau_1}$, one has

$$\|X_{\{f^{(1)}, \Pi_N^U f^{(2)}\}}\|_{r', s'} \leq N^{-\tau_1} 2^{2d+1} \delta^{-1} \|X_{f^{(1)}}\|_{r, s} \|X_{f^{(2)}}\|_{r, s},$$

by the Cauchy and smoothing estimates. The estimate (A.7) follows. \square

Proof of Proposition 6.8. Proposition 6.8(i) follows Lemma A.2.

(ii) Given $f^{(i)}, i = 1, \dots, J$ as in item (i), and applying repeatedly (6.20), the nested Poisson bracket

$$\{f^{(1)}, \{f^{(2)}, \dots, \{f^{(J-1)}, f^{(J)}\} \dots\}$$

is quasi-Töplitz in $\mathcal{D}(r_+, s_+)$ with parameters (K_+, θ_+, μ_+) if

$$(A.9) \quad \frac{1}{N^2} \leq \frac{(\mu - \mu')}{J}, \quad \frac{2\mu'}{N^{4d\tau_0-4}} < \frac{\theta' - \theta}{J}, \quad e^{-\frac{s-s'}{J}} N(N)^{\tau_1} < 1$$

for all $N > K_+$. For given N we bound all the terms in $e^{\{F, \cdot\}} G$ containing $J > (\ln N)^2$ Poisson brackets by $N^{-\tau_1}$ by using the standard bound

$$\begin{aligned} \sum_{k > J} \frac{\|X_{ad(f^{(1)})^k f_2}\|_{r', s'}}{k!} &\leq (2e\delta^{-1} \|X_{f^{(1)}}\|_{r, s})^{J+1} \|X_{f^{(2)}}\|_{r, s} \\ &\leq CN^{-\tau_1} \|X_{f^{(1)}}\|_{r, s} \|X_{f^{(2)}}\|_{r, s} \end{aligned}$$

provided that $2e\delta^{-1} \|X_{f^{(1)}}\|_{r, s} < \frac{1}{2}$. We then apply (A.9) with $J = (\ln N)^2$, we get the restriction (6.22). So applying item (i) repeatedly we get, for all $k < J$,

$$\frac{1}{k!} \|X_{ad(f^{(1)})^k f_2}\|_{r', s'}^T \leq (Ce\delta^{-1} \|X_{f^{(1)}}\|_{r, s}^T)^k \|X_{f^{(2)}}\|_{r, s},$$

and the result follows. \square

Acknowledgments. We wish to thank Massimiliano Berti, Luca Biasco, and Claudio Procesi for their careful reading, and many helpful suggestions. Finally we wish to thank the anonymous referees whose remarks helped us to make the paper more readable.

REFERENCES

- [1] M. BERTI AND L. BIASCO, *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys., 305 (2011), pp. 741–796.
- [2] M. BERTI AND P. BOLLE, *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Math. J., 134 (2006), pp. 359–419.
- [3] M. BERTI AND P. BOLLE, *Quasi-periodic solutions with Sobolev regularity of NLS on T^d , and a multiplicative potential*, J. Eur. Math. Soc. (JEMS), 15 (2013), pp. 229–286.
- [4] M. BERTI, L. BIASCO, AND M. PROCESI, *Quasi-periodic solutions for the reversible derivative wave equations*, preprint, arXiv:1211.2136, 2012.
- [5] M. BERTI, L. BIASCO, AND M. PROCESI, *KAM theory for the Hamiltonian derivative wave equations*, Ann. Sci. École Norm. Sup. (4), 46 (2013), pp. 301–373.
- [6] J. BOURGAIN, *Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Ann. of Math. (2), 148 (1998), pp. 363–439.
- [7] J. BOURGAIN, *Global Solutions of Nonlinear Schrödinger Equations*, Colloquium Publications, AMS, Providence, RI, 1999.
- [8] J. BOURGAIN, *Green's Function Estimates for Lattice Schrödinger Operators, and Applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, NJ, 2005.
- [9] L. CHERCHIA AND J. YOU, *KAM tori for 1D nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys., 211 (2000), pp. 498–525.
- [10] W. CRAIG AND C.E. WAYNE, *Newton's method, and periodic solutions of nonlinear wave equations*, Comm. Pure. Appl. Math., 46 (1993), pp. 1409–1498.
- [11] L.H. ELIASSON AND S.B. KUKSIN, *KAM for the non-linear Schrödinger equation*, Ann. of Math. (2), 172 (2010), pp. 371–435.
- [12] J. GENG AND J. YOU, *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Comm. Math. Phys., 262 (2006), pp. 343–372.
- [13] J. GENG, X. XU, AND J. YOU, *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math., 226 (2011), pp. 5361–5402.
- [14] J. GENG AND X. XU, *Almost periodic solutions of one dimensional Schrödinger equation with the external parameters*, J. Dynam. Differential Equations, 25 (2013), pp. 435–450.
- [15] T. KAPPELER AND Z. LIANG, *A KAM theorem for the defocusing NLS equation*, J. Differential Equations, 252 (2012), pp. 4068–4113.
- [16] T. KAPPELER AND J. PÖSCHEL, *KdV & KAM*, Springer-Verlag, Berlin, 2003.
- [17] S.B. KUKSIN, *Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum*, Funct. Anal. Appl., 21 (1987), pp. 192–205.
- [18] S.B. KUKSIN, *Nearly Integrable Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics 1556, Springer-Verlag, Berlin, 1993.
- [19] S.B. KUKSIN AND J. PÖSCHEL, *Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math. (2), 143 (1996), pp. 149–179.
- [20] J. PÖSCHEL, *On elliptic lower dimensional tori in Hamiltonian systems*, Math. Z., 202 (1989), pp. 559–608.
- [21] J. PÖSCHEL, *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv., 71 (1996), pp. 269–296.
- [22] J. PÖSCHEL, *A KAM Theorem for some nonlinear partial differential equations*, Ann. Scvola. Norm. Sup. Pisa Cl. Sci. (4), 23 (1996), pp. 119–148.
- [23] J. PÖSCHEL AND E. TRUBOWITZ, *Inverse Spectral Theory*, Academic Press, New York, 1987.
- [24] C.E. WAYNE, *Periodic, and quasi-periodic solutions for nonlinear wave equations via KAM theory*, Comm. Math. Phys., 127 (1990), pp. 479–528.