CONSTRUCTION OF INVARIANT WHISKERED TORI BY A PARAMETERIZATION METHOD. PART II: QUASI-PERIODIC AND ALMOST PERIODIC BREATHERS IN COUPLED MAP LATTICES.

ERNEST FONTICH, RAFAEL DE LA LLAVE, AND YANNICK SIRE

ABSTRACT. We construct quasi-periodic and almost periodic solutions for coupled Hamiltonian systems on an infinite lattice which are translation invariant. The couplings can be long range, provided that they decay moderately fast with respect to the distance.

For the solutions we construct, most of the sites are moving in a neighborhood of a hyperbolic fixed point, but there are oscillating sites clustered around a sequence of nodes. The amplitude of these oscillations does not need to tend to zero. In particular, the almost periodic solutions do not decay at infinity.

The main result is an a-posteriori theorem. We formulate an invariance equation. Solutions of these equation are embeddings of an invariant torus. We show that, if there is an approximate solution of the invariance equation that satisfies some non-degeneracy conditions, there is a true solution close by.

This does not require that the system is close to integrable, hence it can be used to validate numerical calculations or formal expansions.

The proof of this a-posteriori theorem is based on a Nash-Moser iteration, which does not use transformation theory. Simpler versions of the scheme were developed in E. Fontich, R. de la Llave,Y. Sire *Jour. Diff. Eq.* **246**, 3136 (2009).

One technical tool, important for our purposes, is the use of weighted spaces that capture the idea that the maps under consideration are local interactions. Using these weighted spaces, the estimates of iterative steps are similar to those in finite dimensional spaces. In particular, the estimates are independent of the number of nodes that get excited. Using these techniques, given two breathers, we can place them apart and obtain an approximate solution, which leads to a true solution nearby. By repeating the process infinitely often, we can get solutions with infinitely many frequencies which do not tend to zero at infinity.

CONTENTS

1. Introduction

The goal of this paper is to prove theorems on persistence of invariant tori in some lattice systems. These models describe copies of identical systems placed on the nodes of a lattice and interacting with all the other systems in the lattice. The interaction can be of infinite range, but it has to decay sufficiently fast with the distance. We will assume that the dynamics is Hamiltonian and, for simplicity, we will also assume that the dynamics is analytic. We will consider *"whiskered tori"*. These are invariant tori such that the motion on them is a rotation and which are as hyperbolic as possible compatible with the fact that the motion is an irrational rotation (it is well known that the directions symplectically conjugate to the tangent of the tori have to be neutral). See Definition 3.1.

The main technical tool we will develop is a theorem of persistence of finite dimensional whiskered tori, namely Theorem 3.6 below, which has sufficiently good properties to allow us to use it recursively to construct tori with infinite frequencies.

The tori we consider in Theorem 3.6 are finite-dimensional whiskered tori with some local character. The motion on the torus is a rigid rotation with a Diophantine frequency. The preservation of the symplectic structure and the rotational motion on the tori force that there are some neutral directions in the normal directions. We will assume that, except for these directions, the normal directions are hyperbolic (they expand at exponential rates either in the future or in the past). In particular, the hyperbolic spaces are infinite dimensional.

The main technique to prove Theorem 3.6 is to derive an equation that implies invariance of the torus and that the motion on it is a rotation and to develop a theory for solutions of the equation.

Given a map F on a phase space $\mathcal{M}$ and a frequency $\omega \in \mathbb{R}^l$, it is easy to see that $K : \mathbb{T}^l \to \mathcal{M}$ is a parameterization of a torus with a rotation ω, if and only if

$$(1) \qquad\qquad F \circ K = K \circ T_\omega \,,$$

where T_ω denotes the rotation on the torus by ω. Similarly, for a vector field X, we seek parameterizations K satisfying

$$(2) \qquad\qquad X \circ K = \partial_\omega K$$

where ∂_ω is the derivative along the direction ω.

Our main result will be Theorem 3.6, which shows that if we have an approximate solution of the invariance equation which is also not too degenerate, there is a true solution which is close to the approximate one. Theorems of this form, that validate an approximate solution, will be called *a posteriori*, following the language in numerical analysis.

We emphasize that Theorem 3.6 does not assume that the system is close to integrable, so that the approximate solution could be produced in any way. Of course, when the system is close to integrable, we can take as approximate solutions the solutions of the integrable system, so that we recover the standard formulations of KAM theorems for quasi-integrable systems. The approximate solutions can be produced by a variety of methods, including Lindstedt series or numerical computations. In finite dimensions, some whiskered tori are generated by resonant averaging [dlLW04, Tre94] or by homoclinic tangencies [Dua08]. In such cases, Theorem 3.6 leads to justifications of the expansions or the numerical computations. We also note that Theorem 3.6 does not assume that the system is translation invariant (it assumes only the existence of some uniform bounds).

The *a posteriori* approach to KAM theorem was emphasized in [Mos66b, Mos66a, Zeh75, Zeh76a, Zeh76b]. There, it was pointed out that this *a posteriori* approach automatically allows to deduce results for finitely differentiable systems as well as to prove smooth dependence on parameters or analiticity of perturbative series. We refer the reader to [dlL01] for a comparison of different KAM methods.

In this paper, we use the *a posteriori* format to construct more complicated quasi-periodic solutions by juxtaposing two simpler solutions separated by a sufficiently long distance. The *a posteriori* format of Theorem 3.6, allows us to control the limit of the solutions, which will be an almost periodic solution. The ability to superimpose solutions far apart is greatly facilitated by assuming translation invariance, which will be an assumption in the second part of this work. One could

assume significantly less (e.g. some amount of uniformity). Nevertheless, this seems a natural assumption.

One important technical tool in this paper is the use of spaces of decay functions following [JdlL00, FdlLM11a]. These are spaces of functions whose norms quantify the effect that the motion of one particle does not affect much the motion of particles far apart. Besides that, they also enjoy certain Banach algebra properties so that the hyperbolic directions can be dealt with in the same manner than in the finite dimensional spaces.

Using spaces of decay functions, we can make quantitative the observation that, since the oscillations at one site do not affect those further apart, superimposing oscillations centered around sites far apart produces a very approximate solution. The error in the invariance equation (measured in the sense of an appropriate space of decay functions) is arbitrarily small if the centers are placed far enough. A rather simple calculation shows that the non-degeneracy conditions deteriorate also by an arbitrarily small amount. In summary, if the frequencies of the oscillations are jointly Diophantine (even if the constant is bad), we can satisfy all the requirements of the theorem by displacing the breathers far apart. If one makes appropriate choices – placing the subsequent centers of oscillation far enough apart – we will show that the process converges in a sense which is strong enough to justify that the limit is a solution of the system. This solution contains infinitely many frequencies.

The process of coupling the breathers does not require any smallness conditions in the coupling (it suffices to place the breathers far enough apart). On the other hand, establishing the existence of breathers by perturbing from those of the uncoupled system, does require some smallness conditions (we also require some mild smallness conditions on the perturbations to ensure that the system remains non-degenerate).

In the solutions that we construct most of the sites are near a hyperbolic equilibrium. These solutions, therefore, have an average energy close to that of the equilibrium solutions and are at the border of chaos (in particular, they are dynamically unstable). There are indications that these solutions play an important role in instability.

The results of the paper were summarized in [FdlLS09a], which perhaps can be used as a reading guide to the present paper.

To provide some motivation, we now mention several models found in the literature for which our method applies. These models can be described by the following formal Hamiltonian

$$(3) \qquad H(p,q) = \sum_{i \in \mathbb{Z}^N} \left(\frac{1}{2} |p_i|^2 + W(q_i) \right) + \sum_{k \in \mathbb{Z}^N} \sum_{i \in \mathbb{Z}^N} V_k(q_i - q_{i+k})$$

under some assumptions on the potentials W and V_k. Here the formal Hamiltonian structure is

$$\Omega_\infty = \sum_{i \in \mathbb{Z}^N} dp_i \wedge dq_i.$$

Note that, even if the sum defining the Hamiltonian and the symplectic form are formal and not meant to converge, Hamilton's equations are a well behaved system of differential equations (if the V_k decay fast enough, e.g. if they are finite range).

In fact the equations of motion are

$$\dot{q}_i = p_i$$

$$\dot{p}_i = -\nabla W(q_i) - \sum_{k \in \mathbb{Z}^N} \left\{ \nabla V_k(q_i - q_{i+k}) - \nabla V_k(q_{i-k} - q_i) \right\}.$$

The model (3) involves a local potential W for each particle and interaction potentials among pairs of particles. Of course, the interaction potentials are assumed to decay with $|k|$ fast enough. The method of proof also accommodates many body interactions. One important feature of the method is that, in some appropriate weighted spaces, the estimates we obtain are independent of the number and the position of the centers of oscillation.

If we take the lattice to be with one degree of freedom, the potential V to be just nearest neighbor (i.e. $V_k = 0$ for $|k| > 1$) and set $V_1(s) = \frac{\gamma}{2}s^2$, we obtain the so called 1-D Klein-Gordon system described by the formal Hamiltonian

$$H(q,p) = \sum_{n=-\infty}^{+\infty} \left(\frac{1}{2}p_n^2 + W(q_n) + \frac{\gamma}{2}(q_{n+1} - q_n)^2 \right),$$

and whose equations of motion are

$$(4) \qquad \ddot{q}_n + W'(q_n) = \gamma(q_{n+1} + q_{n-1} - 2q_n), \quad n \in \mathbb{Z}.$$

We note that the method we present applies to higher dimensional lattices and higher dimensional systems. We also do not need to assume that the symplectic form is the standard one. This is convenient when the symplectic form is degenerate. Changes in the symplectic form correspond to magnetic fields [Thi97]. Note that the systems with magnetic fields are not reversible.

For a review of the physical relevance of these models we refer the reader to [FW98]. Concerning the existence proof of periodic breathers, we refer to [MA94, AGT96, AG96, AKK01]. In the latter papers, the technique is based on a variational argument whereas in [MA94], the authors use an implicit function theorem. For quasi-periodic breathers in finite – but arbitrarily large systems, we mention [BV02, GY07, GVY08, CY07b]. The paper [Yua02] proves the existence of quasi-periodic breathers in the Fermi-Pasta-Ulam lattice. In all the cases above, the breathers are normally elliptic or dissipative. Quasi-periodic and almost periodic breathers for lattices of reversible systems with dissipation are considered in [CY07a].

Remark 1.1. There is a variety of results showing that for hyperbolic PDE's there are no quasi-periodic solutions of finite energy [Pyk96, SW99, KK08, KK10]. Since some of the models we consider are obtained as discretizations of nonlinear wave equations, it is interesting to understand why the results above do not apply to the discretized model, even if they apply to the PDE.

The reason is that the mechanism behind the proofs in the above papers that quasi-periodic solutions of non-linear PDE's have to radiate and send energy to infinity.

In the models we consider, there is no radiation because most of the media is near the hyperbolic regime.

We can understand the lack of radiation in the model but representative problem

$$(5) \qquad \ddot{q}_n + Aq_n = \gamma(q_{n+1} + q_{n-1} - 2q_n), \quad n \in \mathbb{Z}, \gamma > 0.$$

where $A < 0$. The equation (5) is a linearization of (4) near the point $q = 0$, which is a maximum of the potential (or a mimimum of W in the notation of (4)).

We see that if we substitute solutions of the form $q_n = \exp(\imath(\omega t + kn))$ in (5), we are lead to the dispersion relation

$$-\omega^2 + A = \gamma(2\cos(k) - 2).$$

If $|A|$ is small enough, this dispersion relation does not have any real solutions for ω and the only square roots are imaginary.

In this model, near the hyperbolic fixed points the equations do not propagate waves, so that there is no radiation and the arguments excluding quasi-periodic solutions do not apply.

Note, however that for PDE's, the dispersion relation would be $-\omega^2 + A = -k^2$. The unboudedness of the k^2 term makes it possible to have propagating waves always.

Note also that in the model in [FSW86], there is no propagation either because of the random nature of the media.

2. Basic setup and preliminaries

The method introduced in [FdlLS09a] will be extended to systems on infinite-dimensional manifolds provided that the coupling decays fast enough with respect of the distance among the sites. An important tool will be appropriate function spaces for these interactions. There are many other methods to establish the existence of whiskered tori [Gra74, Zeh76a, You99]. The present method has the advantage that it depends much less on the subtle geometric properties, so that it applies easily to infinite dimensional contexts.

Since we are interested in translation invariant problems and want to produce solutions that do not go to zero at infinity, it is natural to model the functional analysis in ℓ^∞, which has some subtle points that require attention.

The goal of this section is to set up the functional analysis spaces modelled after ℓ^∞ and capturing the idea that changes in one site have very small effect in sites that are far away. We anticipate that we need two types of spaces. One family of spaces for the mappings from the infinite dimensional space to itself and another kind of spaces for the mappings from a finite-dimensional torus to the infinite dimensional phase space. This corresponds to the F, K in (1) or the X, K in (2). The spaces we choose are patterned after the choices in [JdlL00, FdlLM11a]. Other spaces of functions in lattice systems are in [Rug02, BK95].

2.1. **Phase spaces.** In this paper we will assume that the phase space at each node is given by an Euclidean exact symplectic manifold $(M = \mathbb{T}^l \times \mathbb{R}^{2d-l}, \Omega = d\alpha)$ where $\mathbb{T} = \mathbb{R}\backslash\mathbb{Z}$. We will not assume that the symplectic form is given in the standard form of action-angle variables. For some calculations we consider $\tilde{M} = \mathbb{R}^l \times \mathbb{R}^{2d-l}$, the universal covering of M or the complex extensions of the above. This is natural since the KAM method requires to consider Fourier series.

It is possible to adapt our method to the case of a non-Euclidean manifold M using connectors and exponential mappings. This just requires some typographical effort. See the discussion in [FdlLS09a].

Then, the phase space of the lattice system will be a subset of

$$(6) \qquad M^{\mathbb{Z}^N} = \prod_{j \in \mathbb{Z}^N} M.$$

Since M is unbounded we will take as phase space

$$(7) \qquad \mathcal{M} = \ell^\infty(\mathbb{Z}^N) = \left\{ x \in M^{\mathbb{Z}^N} \mid \sup_{i \in \mathbb{Z}^N} |x_i| < \infty \right\}$$

which is a strict subset of $M^{\mathbb{Z}^N}$. We will endow $\mathcal{M}$ with the distance,

$$d(x, y) = \sup_{i \in \mathbb{Z}^N} d(x_i, y_i)$$

where $d(x_i, y_i)$ is the distance on the finite dimensional manifold M.

When M is $\mathbb{R}^{2d}$, $\mathcal{M}$ is a Banach space with the norm

$$\|x\|_\infty = \sup_{i \in \mathbb{Z}^N} |x_i|.$$

When $M = \mathbb{T}^l \times \mathbb{R}^{2d-l}$, $\mathcal{M}$ is a Banach manifold modelled on $\ell^\infty(\mathbb{Z}^N)$.

Notice that because M is an Euclidean space, the tangent space of $\mathcal{M}$ is trivial and can be identified with ℓ^∞. Given $x \in \mathcal{M}$ and $\xi \in \ell^\infty$, we can define $x + \xi$ by just adding the components. We see that if $\|\xi\|_\infty < 1/2$, the mapping $\xi \to x + \xi$ is injective, so that this defines a chart in $\mathcal{M}$.

Remark 2.1. The fact that we assume that the manifold has the structure $M = \mathbb{T}^l \times \mathbb{R}^{2d-l}$ is important here since it implies that $H^1(M) \sim H^1(\mathbb{T}^l)$ is non-trivial. This allows us to perform the construction in Appendix 12. If the manifold was such that its first de-Rham cohomology group were trivial, all the symplectic maps would be exact symplectic and the construction would not work. To deal with manifolds such that $H^1(M)$ is trivial ($M = \mathbb{R}^{2d}$ for instance), one can use the method developed by the authors in the finite-dimensional case (see [FdlLS09a, FdlLS09b]). This consists in perturbing the invariance equation for the tori by a translation term and prove, at the end of the convergence scheme, that the geometry implies that this term is zero. The method of [FdlLS09a] allows to deal with secondary tori (i.e. tori which are contractible to tori of lower dimension) directly. The present method would require to make some preliminary changes of variables.

The choice of ℓ^∞ is dictated by the fact that we want to deal with solutions that neither grow nor decrease at ∞. This, however, will lead to some complications, the functional analysis in ℓ^∞ being rather delicate. On the other hand, obtaining estimates in ℓ^∞ for several of our objects will be relatively easy.

Since we are going to deal with analytic functions, one has to define what is the complex extension of the manifold $\mathcal{M}$. By assumption, the manifold M is an Euclidean manifold, hence, it admits a complex extension $M^\mathbb{C}$. We define the complex extension of $\mathcal{M}$ as a subspace of the product of the complex extensions of M, i.e.

$$\mathcal{M}^\mathbb{C} = \left\{ z \in \prod_{j \in \mathbb{Z}^N} M^\mathbb{C} \mid \sup_{i \in \mathbb{Z}^N} |z_i| < \infty \right\}.$$

In the following, we will be considering mostly $\mathcal{M}^\mathbb{C}$ but to simplify the notation we will not write the superscript $\mathbb{C}$, if it does not lead to confusion.

2.2. **Some functional analysis in $\ell^\infty(\mathbb{Z}^N)$ and the spaces of decay functions.**
As emphasized in [FdlLM11a], $\ell^\infty(\mathbb{Z}^N)$ has a very complicated dual space which
cannot be identified with a space of sequences since there is no Riesz-representation
theorem. As a consequence, we have that the matrix elements of an operator do not
characterize the operator and, relatedly, the differential of a map is not represented
by its partial derivatives. The physical meaning is that one has to take into account
"boundary conditions at infinity".

For example, consider the functional $\mathcal{T}$ defined on a closed subspace of $\ell^\infty(\mathbb{Z})$
by the formula

$$\mathcal{T}(u) = \lim_{n \to +\infty} u_n.$$

By the Hahn-Banach theorem, $\mathcal{T}$ extends to $\ell^\infty(\mathbb{Z})$. The extended functional $\mathcal{T}$ is
non-trivial but we have

$$\partial_{u_i}\mathcal{T}(u) = 0$$

since the limit does not depend on u_i. Of course, the functional $\mathcal{T}$ is linear but it
is not represented by a matrix.

This phenomenon can be eliminated by assuming that we are dealing with func-
tions whose derivative is given by the matrix of partial derivatives. We will develop
some technology that allows to verify this assumption rather comfortably in the
cases of interest. A much more thorough treatment can be found in [FdlLM11a].

To formulate quantitatively the locality of the maps we will consider Banach
spaces whose norm makes precise that changing one coordinate affects little the
outcome of other coordinates far away.

We will make use of the so-called decay functions introduced in [JdlL00].

Definition 2.2. *We say that a function $\Gamma : \mathbb{Z}^N \to \mathbb{R}_+$ is a decay function when it
satisfies*

(1) $\displaystyle\sum_{j \in \mathbb{Z}^N} \Gamma(j) \le 1,$

(2) $\displaystyle\sum_{j \in \mathbb{Z}^N} \Gamma(i-j)\Gamma(j-k) \le \Gamma(i-k), \ \ i,k \in \mathbb{Z}^N$

The algebraic property (2) in definition 2.2 is important since it is the one that
allows us to construct Banach algebras.

The following elementary proposition is proved in detail in [JdlL00] and provides
an example of a decay function.

Proposition 2.3. *Given $\alpha > N$, $\theta \ge 0$, there exists $a > 0$, depending on α, θ, N
such that the function defined by*

$$\Gamma(i) = \begin{cases} a|i|^{-\alpha}e^{-\theta|i|} & \text{if } i \ne 0, \\ a & \text{if } i = 0 \end{cases}$$

is a decay function on $\mathbb{Z}^N$.

We note , as it is easily verified in [JdlL00], that $\Gamma(i) = C\exp(-\beta|i|)$ is not a
decay function for any $\beta, C > 0$.

If one considers other sets Λ in place of $\mathbb{Z}^N$, such as the Bethe lattice which also
admit decay functions (see [JdlL00]), many of the results of the present paper can
be adapted.

Definition 2.4. *Given two decay functions* Γ, Γ' *we say that* Γ *dominates* Γ' *and write* $\Gamma' \ll \Gamma$ *when*

$$\lim_{k \to \infty} \Gamma'(k)/\Gamma(k) = 0.$$

We say that a family of decay functions Γ_β, $\beta \in [0,1]$ *is an ordered family when* $\tilde{\beta} < \beta$ *implies* $\Gamma_{\tilde{\beta}} \ll \Gamma_\beta$.

Of course, the examples in Proposition 2.3 enjoy several other nice properties for example that $\Gamma(i)$ is a decreasing function of $|i|$. They also provide us with examples of families satisfying Definition 2.4. These properties play a role in the proof of Theorem 3.11. This is the reason why Theorem 3.11 is only stated for these functions. We refer the reader to Appendix 10 where the study of the spaces of decay functions is performed. In the following, we just give the definitions needed to state our main results.

Remark 2.5. Prof. L. Sadun pointed out that there is a very natural physical interpretation of the definition of decay functions. We note that a site i can affect another site j either directly or by affecting another site k which in turn affects the site j. Of course, more complicated effects involving longer chains of intermediate sites are also possible. If the direct interaction between two sites is bounded by a decay function, it follows that the effect mediated through intermediate sites is bounded by the same function. This makes it possible to comfortably carry out perturbation calculations.

We now introduce the functional spaces needed for our purposes. We introduce:
- The Banach space of decay linear operators

(8)
$$\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N)) = \left\{ \begin{array}{c} A \in \mathcal{L}(\ell^\infty(\mathbb{Z}^N)) \mid \exists \ \{A_{ij}\}_{i,j \in \mathbb{Z}^N}, \\ (Au)_i = \sum_{j \in \mathbb{Z}^N} A_{ij} u_j \ i \in \mathbb{Z}^N, \\ \sup_{i,j \in \mathbb{Z}^N} \Gamma(i-j)^{-1} |A_{ij}| < \infty \end{array} \right\}$$

where $\mathcal{L}(\ell^\infty(\mathbb{Z}^N))$ denotes the space of continuous linear maps from $\ell^\infty(\mathbb{Z}^N)$ into itself. We endow $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$ with the norm

(9)
$$\|A\|_\Gamma = \sup_{i,j \in \mathbb{Z}^N} \Gamma(i-j)^{-1} |A_{ij}|$$

- The space of C^1 functions on an open set $\mathcal{B} \subset \mathcal{M}$

$$C_\Gamma^1(\mathcal{B}) = \left\{ \begin{array}{c} F : \mathcal{B} \to \mathcal{M} \mid F \in C^1(\mathcal{B}), \ DF(x) \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N)), \\ \sup_{x \in \mathcal{B}} \|F(x)\| < \infty, \ \sup_{x \in \mathcal{B}} \|DF(x)\|_\Gamma < \infty \end{array} \right\}$$

with the norm

$$\|F\|_{C_\Gamma^1} = \max \big(\sup_{x \in \mathcal{B}} \|F(x)\|, \sup_{x \in \mathcal{B}} \|DF(x)\|_\Gamma \big).$$

For $r \in \mathbb{N}$, we have

$$C_\Gamma^r(\mathcal{B}) = \left\{ \begin{array}{c} F : \mathcal{B} \to \mathcal{M} \mid F \in C^r(\mathcal{B}), \ D^j F \in C_\Gamma^1(\mathcal{B}), \\ 0 \le j \le r-1 \end{array} \right\}.$$

We define a notion of analyticity for maps on lattices.

Definition 2.6. *Let* $\mathcal{B}$ *be an open set of* $\mathcal{M}$*. We say that* $F : \mathcal{B} \to \mathcal{M}$ *is analytic if it is in* $C_\Gamma^1(\mathcal{B})$ *with the derivatives understood in the complex sense.*

- The space of analytic embeddings on a strip

$$D_\rho = \left\{ z \in \mathbb{C}^l/\mathbb{Z}^l \mid |\mathrm{Im}\, z_i| < \rho,\ i = 1, \ldots, l \right\}.$$

Let $R \geq 1$ be an integer and consider $\underline{c} \in (\mathbb{Z}^N)^R$, i.e.

$$\underline{c} = (c_1, \ldots, c_R).$$

We introduce the following quantity

$$\|f\|_{\rho,\underline{c},\Gamma} = \sup_{i \in \mathbb{Z}^N} \min_{j=1,\ldots,R} \Gamma^{-1}(i - c_j)\|f_i\|_\rho,$$

where

$$\|f_i\|_\rho = \sup_{\theta \in D_\rho} |f_i(\theta)|.$$

We denote

(10) $$\mathcal{A}_{\rho,\underline{c},\Gamma} = \left\{ \begin{array}{c} f : D_\rho \to \mathcal{M} \mid f \in C^0(\overline{D}_\rho),\ f \text{ analytic in } D_\rho, \\ \|f\|_{\rho,\underline{c},\Gamma} < \infty \end{array} \right\}.$$

This space, with the norm $\|\cdot\|_{\rho,\underline{c},\Gamma}$, is a Banach space. If we consider a map A from D_ρ into the set of linear maps $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$, the associated norm is

$$\|A\|_{\rho,\Gamma} = \sup_{i,j \in \mathbb{Z}^N} \sup_{\theta \in D_\rho} \Gamma^{-1}(i - j)|A_{ij}(\theta)| = \sup_{\theta \in D_\rho} \|A(\theta)\|_\Gamma$$

2.3. Symplectic geometry on lattices.

In this section, we introduce the little geometry we need on the manifold $\mathcal{M}$ to be able to perform the iteration. We refer the reader to Appendix 11 where a more systematic description and properties of the objects are performed. We basically need symplectic geometry for the vanishing lemma 6.1 in Section 6, as it is usual in KAM theory.

Consider our finite-dimensional exact symplectic manifold $(M, \Omega = d\alpha)$ and the associated lattice

$$\mathcal{M} = \ell^\infty(\mathbb{Z}^N, M).$$

Define α_∞ and Ω_∞ to be the formal sums

$$\alpha_\infty = \sum_{j \in \mathbb{Z}^N} \pi_j^* \alpha, \qquad \Omega_\infty = \sum_{j \in \mathbb{Z}^N} \pi_j^* \Omega.$$

where π_j are the standard projections from $\mathcal{M}$ to M at the node $j \in \mathbb{Z}^N$. Let J be the symplectic matrix associated to the symplectic two-form Ω on M. We denote J_∞ the operator defined on $T\mathcal{M}$ by

$$J_\infty(z) = \mathrm{diag}\big(\ldots, J(\pi_i z), \ldots\big), \qquad z \in \mathcal{M}.$$

We introduce the following definitions.

Definition 2.7. *We say that a C_Γ^1 function $F : \mathcal{M} \to \mathcal{M}$ is symplectic if the following identity holds for any $z \in \mathcal{M}$*

$$DF^\top(z) J_\infty(F(z)) DF(z) = J_\infty(z)$$

where the product of two operators A and B in C_Γ^1 is given component-wise by $(AB)_{i,j \in \mathbb{Z}^N} = \sum_{k \in \mathbb{Z}^N} A_{ik} B_{kj}$.

Similarly, we have the following definition. Let $\hat{A}$ be the linear operator associated to the Liouville form α on M. We denote $\hat{A}_\infty$ the operator defined on $T\mathcal{M}$ by

$$\hat{A}_\infty(z) = \text{diag}\big(\ldots, \hat{A}(\pi_i z), \ldots\big), \qquad z \in \mathcal{M}.$$

Definition 2.8. *We say that a C^1_Γ function $F : \mathcal{M} \to \mathcal{M}$ is exact symplectic on $\mathcal{M}$ if there exists a one-form $\tilde{\alpha}$ defined on $T\mathcal{M}$ with matrix $\tilde{A}$ such that*

- *For every $j \in \mathbb{Z}^N$, there exists a smooth function W_j on M such that*

$$\tilde{\alpha}_j = dW_j$$

  where d is the exterior differentiation on M.
- *The following formula holds component-wise on the lattice*

$$DF(z)^\top \hat{A}_\infty(F(z)) = \hat{A}_\infty(z) + \tilde{A}(z),$$

The previous definitions are completely equivalent to the standard definitions of symplectic and exact symplectic maps in the finite-dimensional case, but they are among the mildest ones that we can imagine in infinite dimensions.

We anticipate that the symplectic structure, will only enter in this paper in two places: 1) The automatic reducibility in the center directions, 2) The vanishing lemma to show that for exact symplectic mappings several averages vanish. These applications are very finite dimensional.

The following lemma will be usefull for us (see Appendix 11).

Lemma 2.9. *Consider a function ψ defined on $\mathbb{T}^l$ (or a subset of it) with values in $\mathcal{M}$ and belonging to $\mathcal{A}_{\rho,\underline{c},\Gamma}$ for some $\rho > 0$. Then the bilinear form*

$$\psi^*\Omega_\infty = \sum_{j \in Z^N} \psi^*\Omega(\pi_j)$$

is a $2-$form on the torus $\mathbb{T}^l$.

2.4. **Diophantine properties.** KAM relies on approximation properties of the frequencies by rational numbers. In this section, we recall some well known notions. For diffeomorphisms, the relevant notion of Diophantine properties is given by the following

Definition 2.10. *Given $\kappa > 0$ and $\nu \geq l$, we define $D(\kappa, \nu)$ as the set of frequency vectors $\omega \in \mathbb{R}^l$ satisfying the Diophantine condition:*

$$|\omega \cdot k - n|^{-1} \leq \kappa |k|^\nu, \quad \text{for all } k \in \mathbb{Z}^l - \{0\} \text{ and } n \in \mathbb{Z}$$

with $|k| = |k_1| + \cdots + |k_l|$, where k_i are the coordinates of k.

For vector fields, one uses the following

Definition 2.11. *Given $\kappa > 0$ and $\nu \geq l - 1$, we define $D_h(\kappa, \nu)$ as the set of frequency vectors $\omega \in \mathbb{R}^l$ satisfying the Diophantine condition:*

$$|\omega \cdot k|^{-1} \leq \kappa |k|^\nu, \quad \text{for all } k \in \mathbb{Z}^l - \{0\},$$

where $|k| = |k_1| + \cdots + |k_l|$.

Given $f \in L^1(\mathbb{T}^l)$, we denote

$$\mathrm{avg}\,(f) = \int_{\mathbb{T}^l} f(\theta) d\theta.$$

We also denote by T_ω the rotation on $\mathbb{T}^l$ by ω :

$$T_\omega(\theta) = \theta + \omega.$$

In Section 9.2 we will discuss extensions of these definitions to infinite dimensional vectors which are well adapted to our applications.

3. Formulation of the results

We will first obtain a translated tori result, i.e. a KAM theorem for parameterized families of maps F_λ which are symplectic for all λ and such that F_0 is *exact* symplectic. This will allow us to avoid the considerations of vanishing of averages at each stage of the iteration. Then, we will prove a simple vanishing lemma (see Section 6) that shows that the added extra parameter vanishes. This yields to the desired invariant tori theorem. Going through translated curve theorems has become quite standard in KAM theory (see [Mos67, Rüs76a]) especially since [Sev99] pointed out that it deals with very degenerate situation. In our case, it is particularly advantageous since the parameters we need are finite dimensional and it avoids many infinite dimensional considerations.

The problem is the following: given an exact symplectic map F and a vector of frequencies $\omega \in D(\kappa, \nu)$ we wish to construct an invariant torus for F such that the dynamics of F restricted on it is conjugated to the translation T_ω. To this end, we search for an embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ in $\mathcal{A}_{\rho,\underline{c},\Gamma}$ such that for all $\theta \in D_\rho$, K satisfies the functional equation (1).

Notice that if (1) is satisfied, the image under F of a point in the range of K will also be in the range of K. If the range of $DK(\theta)$ is l-dimensional for all θ, then $K(\mathbb{T}^l)$ is an l-dimensional invariant torus. (Similarly, the geometric interpretation of (2) is that the vector field X at a point in the range of K is tangent to the range of K.)

The assumptions are that we are given a mapping K that satisfies (1) up to a very small error and that fullfills some non-degeneracy assumptions. We prove that the embedding K exists and also that the solution is unique up to composition on the right with translations.

Actually we are going to prove a more general result which works for parameterized families of symplectic maps F_λ, such that F_0 is exact symplectic, but only provides translated (and not invariant) tori. That is, given $\omega \in D(\kappa, \nu)$ and an approximate solution K of $F_{\lambda_0} \circ K - K \circ T_\omega = 0$ satisfying a set of non-degeneracy conditions, we search for an embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ in $\mathcal{A}_{\rho,\underline{c},\Gamma}$ such that

$$(11) \qquad\qquad\qquad F_\lambda \circ K = K \circ T_\omega$$

for some λ close to λ_0. The geometric interpretation of the invariance equations is illustrated in Figure 1.

We go through a Newton scheme to prove the existence of such a pair (λ, K). To this end, we introduce the operator $\mathcal{F}_\omega$

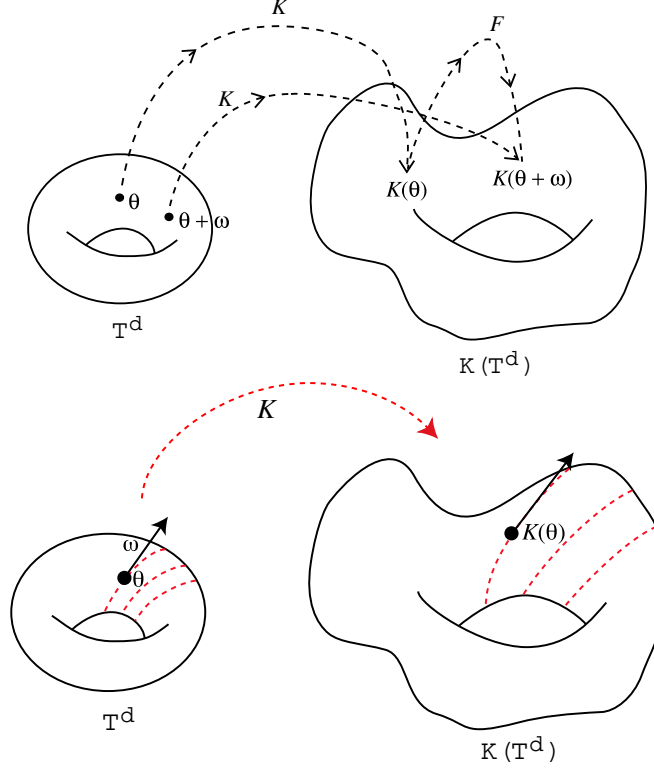$$\mathcal{F}_\omega(\lambda, K) = F_\lambda \circ K - K \circ T_\omega.$$

FIGURE 1. Illustration of the invariance equations (1), (2).

In the paper [FdlLS09a], the authors constructed invariant tori using a *posteriori* KAM theorems in finite dimensional systems. The general principles of this method remain valid in some infinite-dimensional systems such as lattices. We first introduce some notations and several non-degeneracy conditions.

Definition 3.1. *Consider $\rho > 0$, $\omega \in \mathbb{R}^l$, $\underline{c} = (c_1, \ldots, c_R) \in (\mathbb{Z}^N)^R$, $R \geq 1$, a decay function Γ, $\lambda \in \mathbb{R}^l$ and $F_\lambda : \mathcal{M} \to \mathcal{M}$ be a C_Γ^2 map.*

We say that $K : D_\rho \to \mathcal{M} \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ is a whiskered embedding for F_λ when we have:

The tangent space $T_{K(\theta)}\mathcal{M}$ has an invariant analytic splitting for all $\theta \in D_\rho$

$$(12) \qquad T_{K(\theta)}\mathcal{M} = \mathcal{E}_{K(\theta)}^s \oplus \mathcal{E}_{K(\theta)}^c \oplus \mathcal{E}_{K(\theta)}^u ,$$

where $\mathcal{E}_{K(\theta)}^s$, $\mathcal{E}_{K(\theta)}^c$ and $\mathcal{E}_{K(\theta)}^u$ are the stable, center and unstable invariant spaces respectively, which satisfy:

- *The projections $\Pi_{K(\theta)}^s$, $\Pi_{K(\theta)}^c$ and $\Pi_{K(\theta)}^u$ associated to this splitting are analytic with respect to θ considered as operators in $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$.*
- *The splitting (12) is characterized by asymptotic growth conditions (cocycles over T_ω): there exist $0 < \mu_1, \mu_2 < 1$, $\mu_3 > 1$ such that $\mu_1\mu_3 < 1$,*

$\mu_2\mu_3 < 1$ and $C_h > 0$ such that for all $n \geq 1$, $\theta \in D_\rho$ and $\lambda \in \mathbb{R}^l$

$$(13) \quad \|DF_\lambda \circ K \circ T_\omega^{n-1} \times \cdots \times DF_\lambda \circ Kv\|_{\rho,\underline{c},\Gamma} \leq C_h\mu_1^n\|v\|_{\rho,\underline{c},\Gamma}$$
$$\Longleftrightarrow \ v \in \mathcal{E}_{K(\theta)}^s$$

and

$$(14) \quad \|DF_\lambda^{-1} \circ K \circ T_\omega^{n-1} \times \cdots \times DF_\lambda^{-1} \circ Kv\|_{\rho,\underline{c},\Gamma} \leq C_h\mu_2^n\|v\|_{\rho,\underline{c},\Gamma}$$
$$\Longleftrightarrow \ v \in \mathcal{E}_{K(\theta)}^u.$$

- The center subspace $\mathcal{E}_{K(\theta)}^c$ is finite-dimensional, has dimension $2l$ and it is characterized by:

$$\|DF_\lambda \circ K \circ T_\omega^{n-1}(\theta) \times \cdots \times DF_\lambda \circ K(\theta)v\|_{\rho,\underline{c},\Gamma} \leq C_h\mu_3^n\|v\|_{\rho,\underline{c},\Gamma}$$
$$(15) \quad \|DF_\lambda^{-1} \circ K \circ T_\omega^{-(n-1)}(\theta) \times \cdots \times DF_\lambda^{-1} \circ K(\theta)v\|_{\rho,\underline{c},\Gamma} \leq C_h\mu_3^n\|v\|_{\rho,\underline{c},\Gamma}$$
$$\Longleftrightarrow \ v \in \mathcal{E}_{K(\theta)}^c.$$

It is important for applications that the spectral condition in Definition 3.1 is implied by a condition that can be verified by a finite calculation (see Definition 3.2 below) . Approximate invariance of the splitting is sufficient (see Proposition 4.2 below) to ensure there is a truly invariant splitting. So, that the final version of our results will have as a hypothesis the existence of approximately invariant tori with approximately invariant splitting (Definition 3.2) but will have as a conclusion the existence of exactly invariant tori with exactly invariant splittings (Definition 3.1).

Definition 3.2. *Consider* $\rho > 0$, $\omega \in \mathbb{R}^l$, $\underline{c} = (c_1, \ldots, c_R) \in (\mathbb{Z}^N)^R$, $R \geq 1$, *a decay function* Γ, $\lambda \in \mathbb{R}^l$ *and* $F_\lambda : \mathcal{M} \to \mathcal{M}$ *be a* C_Γ^2 *map.*

We say that $\tilde{K} : D_\rho \to \mathcal{M} \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ *satisfies the* η-*hyperbolic condition (or has an* $\eta-$*invariant splitting), if there exists an analytic splitting of* $T_{\tilde{K}(\mathbb{T}^l)}\mathcal{M}$,

$$(16) \qquad\qquad T_{\tilde{K}(\theta)}\mathcal{M} = \mathcal{E}_{\tilde{K}(\theta)}^s \oplus \mathcal{E}_{\tilde{K}(\theta)}^c \oplus \mathcal{E}_{\tilde{K}(\theta)}^u$$

such that, denoting $\Pi_{\tilde{K}(\theta)}^{s,c,u}$ *be the corresponding projections, we have*

(1) *The splitting is approximately invariant under the co-cycle* $DF \circ K$ *over* T_ω *in the sense that*

$$dist\Big(DF_\lambda(\tilde{K}(\theta))\mathcal{E}_{\tilde{K}(\theta)}^{s,c,u}, \mathcal{E}_{\tilde{K}(\theta)}^{s,c,u}\Big) < \eta.$$

(2) *There exists* $N \in \mathbb{N}, 0 < \tilde{\mu}_1, \tilde{\mu}_2 < 1$ *and* $\tilde{\mu}_3 > 1$ *such that* $\tilde{\mu}_1, \tilde{\mu}_3 < 1, \tilde{\mu}_2\tilde{\mu}_3 < 1$ *and*

$$(17) \quad \|DF_\lambda \circ \tilde{K} \circ T_\omega^{N-1}(\theta) \times \cdots \times DF_\lambda \circ \tilde{K}(\theta)v\|_{\rho,\underline{c},\Gamma} \leq \tilde{\mu}_1^N\|v\|_{\rho,\underline{c},\Gamma},$$
$$\forall v \in \mathcal{E}_{\tilde{K}(\theta)}^s,$$

$$(18) \quad \|DF_\lambda^{-1} \circ \tilde{K} \circ T_\omega^{-(N-1)}(\theta) \times \cdots \times DF_\lambda^{-1} \circ \tilde{K}(\theta)v\|_{\rho,\underline{c},\Gamma} \leq \tilde{\mu}_2^N\|v\|_{\rho,\underline{c},\Gamma},$$
$$\forall v \in \mathcal{E}_{\tilde{K}(\theta)}^u$$

and

$$\|DF_\lambda \circ \tilde{K} \circ T_\omega^{N-1}(\theta) \times \cdots \times DF_\lambda \circ \tilde{K}(\theta)v\|_{\rho,\underline{c},\Gamma} \leq \tilde{\mu}_3^N\|v\|_{\rho,\underline{c},\Gamma}$$
$$(19) \quad \|DF_\lambda^{-1} \circ \tilde{K} \circ T_\omega^{-(N-1)}(\theta) \times \cdots \times DF_\lambda^{-1} \circ \tilde{K}(\theta)v\|_{\rho,\underline{c},\Gamma} \leq \tilde{\mu}_3^N\|v\|_{\rho,\underline{c},\Gamma}$$
$$\forall v \in \mathcal{E}_{\tilde{K}(\theta)}^c.$$

Remark 3.3. Note that in Definition 3.2 we are using that the phase space is Euclidean. On a general manifold, the products used in (17), (18), (19) cannot be defined because, in general, $DF(x) : T_x\mathcal{M} \to T_{F(x)}\mathcal{M}$. Hence, in a general manifold, if $F \circ K(\theta) \neq K(\theta + \omega)$, we cannot define $DF \circ K(\theta + \omega)DF \circ K(\theta)$. In [FdlLS09a] one can find a definition of approximately invariant cocycles for general manifolds. In this paper, we will not consider such generality.

Definition 3.4. Given $\rho > 0$, $\omega \in \mathbb{R}^l$, $\underline{c} = (c_1, \ldots, c_R) \in (\mathbb{Z}^N)^R$, $R \geq 1$, a decay function Γ, $\lambda \in \mathbb{R}^l$ and an embedding $K : D_\rho \to \mathcal{M} \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, a pair (λ, K) is said to be non-degenerate (and we denote $(\lambda, K) \in ND_{loc}(\rho, \Gamma)$) if it satisfies the following conditions

- *Non degeneracy of the embedding: We have that the $l \times l$ matrix $DK^\top(\theta)DK(\theta)$ is invertible for all θ in D_ρ. We denote $N(\theta) = \left(DK^\top(\theta)DK(\theta)\right)^{-1}$ and we assume that*
$$\|N\|_{\rho,\Gamma} < \infty$$
- *Twist condition: let $P(\theta) = DK(\theta)N(\theta)$.*
  The average on $\mathbb{T}^l$ of the $l \times l-$matrix

(20) $$A_\lambda(\theta) = P(\theta + \omega)^\top \left([DF_\lambda(K)\Pi^c_{K(\theta)}J_\infty^{-1}P](\theta) - [\Pi^c_{K(\theta)}J_\infty^{-1}(K)P](\theta + \omega)\right)$$

  is non-singular.
- *Parameter cohomological non-degeneracy: The average on $\mathbb{T}^l$ of the $l \times l-$matrix*

(21) $$Q_\lambda(\theta) = \left((DK^\top(\omega + \theta)\Pi^c_{K(\theta)}J_\infty(K(\omega + \theta))\frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\right)$$

  is non-singular.

It is clear that the meaning of $\|N\|_\rho$ is a measure of the quality of the embedding. It grows if the embedding comes close to having a singularity. During the proof it will become clear that the meaning of A_λ is the change of the rotation when we move in the direction transversal to the torus. As we will see in calculations, the meaning of the invertibility of the average of Q_λ is that, by changing λ, we can adjust the obstructions to the cohomology equations.

For applications, it is important to note that the non-degeneracy hypothesis only depend on the approximate solution considered and that they are readily computable algebraic expressions. They are quite analogous to the condition numbers in numerical analysis.

First we state our main theorem, which provides the existence of a solution (λ, K) to the functional equation (11). This is the translated tori KAM theorem.

Theorem 3.5. *Let $F_\lambda : \mathcal{M} \to \mathcal{M}$ be a family of symplectic maps parameterized by $\lambda \in \mathbb{R}^l$, $\omega \in D(\kappa, \nu)$ for some $\kappa > 0, \nu \geq l$, $\rho_0 > 0$, Γ a decay function and $\underline{c} = (c_1, \ldots, c_R) \in (\mathbb{Z}^N)^R$. Assume we have $\lambda_0 \in \mathbb{R}^l$ and $K_0 : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ satisfying the following hypotheses*

- *For all $\lambda \in \mathbb{R}^l$, the maps F_λ belong to C^2_Γ and satisfy $\sup_{i \in \mathbb{Z}^N} \Gamma^{-1}(i)(F_\lambda(0))_i < \infty$.*
- *The map F_λ is real analytic and it can be extended holomorphically to some complex neighborhood of the image under K_0 of D_{ρ_0}:*
$$B_r = \{z \in \mathcal{M}| \exists \theta \, s.t. |\text{Im } \theta| < \rho_0 \, , \, |z - K_0(\theta)| < r\},$$

for some $r > 0$ and such that $\|DF_\lambda\|_{C^2_\Gamma(B_r)}$ is finite.

- $(\lambda_0, K_0) \in ND_{loc}(\rho_0, \Gamma)$ *i.e , the embedding K_0 is non-degenerate in the sense of Definition3.4.*
- *The embedding K_0 is η_0-hyperbolic in the sense of Definition 3.2 with η_0 sufficiently small (depending on $\|\Pi^{s,c,u}\|_{\rho_0,\underline{c},\Gamma}$, $\mu_{s,c,u}$, N, $\|F\|_{C^2_\Gamma(B_r)}$).*

Define the error E_0 by

$$E_0 = F_{\lambda_0} \circ K_0 - K_0 \circ T_\omega.$$

Denote also

$$\tilde{\varepsilon} = \max(\|E_0\|_{\rho_0,\underline{c},\Gamma}, \eta_0).$$

There exists a constant $C > 0$ depending on l, κ, ν, ρ_0, $\|DF_\lambda\|_{C^2_\Gamma(B_r)}$, $\|DK_0\|_{\rho_0,\underline{c},\Gamma}$, $\|N_0\|_{\rho_0,\Gamma}$, $\|\frac{\partial F_\lambda(K_0)}{\partial \lambda}\|_{\rho_0,\underline{c},\Gamma}$, $\|A^0_{\lambda_0}\|_{\rho_0,\Gamma}$, $|\text{avg}\,(A^0_{\lambda_0})|^{-1}$, $|\text{avg}\,(Q_{\lambda_0})|^{-1}$ (where $A^0_{\lambda_0}$, $Q^0_{\lambda_0}$ and N_0 are as in Definition 3.4,3.2 replacing K with K_0) and on $\|\Pi^{c,s,u}_{K_0(\theta)}\|_{\rho_0,\underline{c},\Gamma}$ such that, if for some δ, $0 < \delta < \min(1, \rho_0/12)$, we have the following conditions satisfied

$$C\kappa^4\delta^{-4\nu}\tilde{\varepsilon} < 1$$

and

$$C\kappa^2\delta^{-2\nu}\tilde{\varepsilon} < r$$

then, there exist an embedding $K_\infty \in ND_{loc}(\rho_\infty = \rho_0 - 6\delta, \Gamma)$ and a vector $\lambda_\infty \in \mathbb{R}^l$ such that

$$(22) \qquad\qquad F_{\lambda_\infty} \circ K_\infty = K_\infty \circ T_\omega.$$

Furthermore, we have the following estimates

$$(23) \qquad \begin{aligned} \|K_\infty - K_0\|_{\rho_\infty,\underline{c},\Gamma} &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon}, \\ |\lambda_0 - \lambda_\infty| &< C\kappa^2\delta^{-2\nu}\tilde{\varepsilon}. \end{aligned}$$

Additionnally, we have that the invariant embedding K_∞ admits invariant splittings, satisfying Definition (3.4).

Denoting the non-degeneracy constants corresponding to K_∞ by index ∞, we have:

$$(24) \qquad \begin{aligned} \|\Pi^{s,c,u}_\infty \circ K_\infty - \Pi^{s,c,u} \circ K_0\|_{\rho_\infty,c,\Gamma} &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon} \\ |\mu^\infty_{s,c,u} - \mu_{s,c,u}| &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon} \end{aligned}$$

and

$$(25) \qquad \begin{aligned} |\|N_0\|_{\rho_0,\Gamma} - \|N_\infty\|_{\rho_\infty,\Gamma}| &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon}, \\ |\|A^0_{\lambda_0}\|_{\rho_0,\Gamma} - \|A^\infty_{\lambda_\infty}\|_{\rho_\infty,\Gamma}| &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon}, \\ |\|Q^0_\lambda\|_{\rho_0,\Gamma} - \|Q^\infty_{\lambda_\infty}\|_{\rho_\infty,,\Gamma}| &\leq C\kappa^2\delta^{-2\nu}\tilde{\varepsilon}. \end{aligned}$$

The previous theorem will allow us to construct increasingly complicated solutions, the solutions of one stage being an approximate solution for the next stage. We will however be able to maintain enough control of the non-degeneracy conditions.

Of course, (25) is an easy consequence of (23) since the objects that enter in the degeneracy estimates are algebraic expressions of K.

We now come to the result on the existence of invariant tori. They correspond to localized quasi-periodic orbits on the manifold $\mathcal{M}$. These orbits are known as "breathers".

Theorem 3.6. *Let $F_\lambda : \mathcal{M} \to \mathcal{M}$ be a family of symplectic maps parameterized by $\lambda \in \mathbb{R}^l$, $\omega \in D(\kappa, \nu)$ for some $\kappa > 0, \nu \geq l$, $\rho_0 > 0$, Γ a decay function and $\underline{c} = (c_1, \ldots, c_R) \in (\mathbb{Z}^N)^R$. Assume we have $\lambda_0 \in \mathbb{R}^l$ and $K_0 : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ satisfying the following hypotheses*

- *The map F_0 is exact symplectic and $F_0(0) = 0$.*
- *For all $\lambda \in \mathbb{R}^l$, the maps F_λ belong to C_Γ^2 and satisfy $\sup_{i \in \mathbb{Z}^N} \Gamma^{-1}(i)(F_\lambda(0))_i < \infty$.*
- *The map F_λ is real analytic and it can be extended holomorphically to some complex neighborhood of the image under K_0 of D_{ρ_0}:*

$$B_r = \{z \in \mathcal{M} | \exists \theta \, s.t. |\mathrm{Im} \, \theta| < \rho_0 \,, \, |z - \mathrm{K}_0(\theta)| < \mathrm{r}\},$$

  for some $r > 0$ and such that $\|DF_\lambda\|_{C_\Gamma^2(B_r)}$ is finite.
- *$(0, K_0) \in ND_{loc}(\rho_0, \Gamma)$ i.e , the embedding K_0 is non-degenerate in the sense of Definition3.4.*
- *The embedding K_0 is η_0-hyperbolic in the sense of Definition 3.2 with η_0 sufficiently small (depending on $\|\Pi^{s,c,u}\|_{\rho_0,\underline{c},\Gamma}$, $\mu_{s,c,u}$, N, $\|F\|_{C_\Gamma^2(B_r)}$).*

Define the error E_0 by

$$E_0 = F_0 \circ K_0 - K_0 \circ T_\omega.$$

Denote also

$$\tilde{\varepsilon} = \max(\|E_0\|_{\rho_0,\underline{c},\Gamma}, \eta_0).$$

There exists a constant $C > 0$ depending on l, κ, ν, ρ_0, $\|DF_\lambda\|_{C_\Gamma^2(B_r)}$, $\|DK_0\|_{\rho_0,\underline{c},\Gamma}$, $\|N_0\|_{\rho_0,\Gamma}$, $\|\frac{\partial F_\lambda(K_0)}{\partial \lambda}\|_{\rho_0,\underline{c},\Gamma}$, $\|A_0^0\|_{\rho_0,\Gamma}$, $|\mathrm{avg}\,(A_0^0)|^{-1}$, $|\mathrm{avg}\,(Q_0)|^{-1}$ (where A_0^0, Q_0^0 and N_0 are as in Definition 3.4,3.2 replacing K with K_0) and on $\|\Pi_{K_0(\theta)}^{c,s,u}\|_{\rho_0,\underline{c},\Gamma}$ such that, if for some δ, $0 < \delta < \min(1, \rho_0/12)$, we have the following conditions satisfied

$$C\kappa^4\delta^{-4\nu}\tilde{\varepsilon} < 1$$

and

$$C\kappa^2\delta^{-2\nu}\tilde{\varepsilon} < r$$

Then, we have in (22)

$$\lambda_\infty = 0,$$

i.e. the torus K_∞ is actually an invariant torus for F_0 and we have

$$F_0 \circ K_\infty = K_\infty \circ T_\omega.$$

Remark 3.7. It is important to mention that we do not assume that the symplectic forms are the standard forms. This allows to consider the existence of external magnetic fields and magnetic interactions among the sites since the effect of a magnetic field is just a change of the symplectic form [Thi97]. Alternatively, if (p, q) are the conjugated coordinates, one can change $p \to p - A$ where A is the vector potential. Note that the introduction of a magnetic field destroys the reversibility under the usual involution $S(p, q) = (-p, q)$.

Remark 3.8. The whiskered tori that satisfy the spectral hypothesis have invariant manifolds that make them important in problems of stability. However, the proof of the stable manifold is not completely straightforward since the space ℓ^∞ does not have smooth cut-off functions. It is possible to show that these invariant manifolds have also some decay properties. This has been established in [FdlLM11b].

We have also the following result which provides local uniqueness.

Theorem 3.9. *Let* $\omega \in D(\kappa, \nu)$ *for some* $\kappa > 0, \nu > l$ *and* $K_1 \in ND(\rho)$ *and* $K_2 \in ND(\rho)$ *be two solutions of equation* (1) *such that* $K_1(D_\rho) \subset B_r, K_2(D_\rho) \subset B_r$. *There exists a constant* $C > 0$ *depending on* l, κ, ν, ρ, ρ^{-1}, $\|F\|_{C_\Gamma^2}$, $\|K_1\|_{\rho, \underline{c}, \Gamma}$, $\|N_1\|_{\rho, \Gamma}$, $\|A_1\|_{\rho, \Gamma}$, $|\mathrm{avg}\,(A_1)|^{-1}$ *such that if* $\|K_1 \circ T_\tau - K_2\|_{\rho, \underline{c}, \Gamma}$ *satisfies for some* $\tau \in \mathbb{R}^l$

$$C\kappa^2 \delta^{-2\nu} \|K_1 \circ T_\tau - K_2\|_{\rho_0, \underline{c}, \Gamma} \le 1,$$

where $\delta = \rho/4$, *then there exists a phase* $\tilde{\tau} \in \mathbb{R}^l$ *such that* $K_1 \circ T_{\tilde{\tau}} = K_2$ *in* D_ρ. *Moreover,*

$$|\tau - \tilde{\tau}| \le C\kappa^2 \rho^{-2\nu} \|K_1 \circ T_\tau - K_2\|_{\rho_0, \underline{c}\Gamma}.$$

Remark 3.10. It is important to remark that ALL constants in the previous theorems are independent of $\underline{c}$. This fact is crucial for the next result, which provides an existence theorem for almost-periodic functions.

The idea of the construction follows the one in the finite-dimensional case (see [FdlLS09a]. Notice here that the decay properties are on the hyperbolic subspace, the center one being finite-dimensional. Some small differences between the scheme of the present paper and [FdlLS09a] are detailed in Remark 4.20.

We also have the analogous result to Theorem 3.6 for vector-fields, Theorem 8.4. We will postpone the statement of Theorem 8.4 till Section 8 where we also present a proof.

As an application of Theorem 8.4, we will present a result on existence of solutions with infinitely many frequencies (also called *almost periodic* solutions). We note that, as indicated before, we will establish the theorem in two stages. In a first stage, we will continue the breathers from the uncoupled system to the whole system. In the second stage, we will couple infinitely many of these breathers so that we obtain solutions with infinitely many frequencies. We note that the smallness conditions and the elimination of a positive set of frequencies only occurs in the first stage.

The models we consider (26) have been considered in the Physics and Mathematics literature. They are models of many microscopic processes. See [BK04, BK98, DRAW02, FBGGn05, CF05, Gal08, BEMW07] and references there among many others.

A large variety of solutions for equations of this type have been constructed: space-localized periodic in time solutions, known as breathers (see [MA94, Jam01]), solitary waves ([Ioo00, IK00, FW94], [FP99, FP02, FP04a, FP04b]), pulsating traveling waves (see [JS05, Sir05]). The relevance of these solutions in biological phenomena has also been discussed (see [DPW92, PS04, Pey04]).

There are already several other papers that have produced solutions with infinitely many frequencies. The paper [FSW86] produced such solutions by introducing some random terms and making the excitation of each oscillator goes to zero, so that its effect on the others was small. Frölich, Spencer and Wayne also assume that the coupling is high order in terms of the amplitude. The paper [CP95, Per03] considered oscillators but made the natural frequencies increase very fast so that there were no resonances in each of them. The paper [Pös90] proved a very abstract theorem that applies to perturbations of integrable systems and managed to recover several results as applications of this theorem. The paper [GY07] also considers coupled systems in one dimension, but produces tori with finitely many frequencies.

The solutions we construct are based on a different principle. We use that the solutions which are far apart interact very weakly even if they are large. Therefore, by placing solutions far apart, we will be able to make them interact weakly and we can satisfy the smallness conditions assumed by the general theorem. Notice that we are assuming that most of the sites are close to a hyperbolic orbit. Hence, the system will be very hyperbolic. This will allow us to deal with most of the normal directions using the methods of hyperbolic splittings and we will not need to consider the resonances that appear in the normally elliptic modes, which require more delicate estimates. We emphasize that we do not assume that the system is close to integrable.

Theorem 3.11. *Consider a lattice $\mathcal{M} = M^{\mathbb{Z}^N}$ with the symplectic form given by $\Omega_\infty = \sum_{n \in \mathbb{Z}^N} dq_n \wedge dp_n$. Consider the following Hamiltonian with respect to Ω_∞ given by*

$$(26) \qquad H(q,p) = \sum_{n \in \mathbb{Z}^N} \left(\frac{1}{2} p_n^2 + W(q_n) \right) + \varepsilon \sum_{j \in \mathbb{Z}^N} \sum_{n \in \mathbb{Z}^N} V_j(q_n - q_{n+j}).$$

Let Γ, Γ' be decay functions as in Proposition 2.3.
Denote by X the vector filed associated to the Hamiltonian (26).
Assume:

H1 *The system $\ddot{q} + W'(q) = 0$ admits a hyperbolic fixed point, which we will set without loss of generality at $q = 0$.*

H2 *There exists a set $\Xi_0 \subset \mathbb{R}^l$ of positive Lebesgue measure such that for all $\omega \in \Xi_0$, there exists a KAM torus invariant under the flow of $\ddot{q} + W'(q) = 0$ and non-degenerate in the sense of the standard KAM theory (twist condition).*

H3 *The potentials V_j and W are real analytic. Moreover, we assume that there exists a constant C_V such that*

$$\|V_k\|_{C^2,\rho} \le C_V \Gamma(k).$$

and also that $\nabla V_k(0) = 0$ for every k.

Fix $\rho' < \rho$. Then,

- *A)For all ε^* sufficiently small, we can find a set $\Xi_1(\varepsilon^*) \subset \Xi_0$, such that if $\omega \in \Xi_1(\varepsilon^*)$ and $|\varepsilon| < \varepsilon^*$, then the system (26) has a localized breather of frequency ω. There exists $K : \mathbb{T}^l \to \mathcal{M}$, $K \in A_{\rho', \underline{c}, \Gamma}$ such that*

$$X \circ K = \partial_\omega K$$

The embedding satisfies Definition 3.1 and we can choose the hyperbolicity and non-degeneracy constants uniformly.
  Furthermore we have

$$\mathrm{meas}(\Xi_0 \setminus \Xi_1(\varepsilon^*)) \to 0$$

as $\varepsilon^* \to 0$.

- *B) Consider now $\Xi_\infty = \Xi_1(\varepsilon^*)^{\mathbb{N}}$ endowed with the probability measure $\left(\frac{\mathrm{meas}(\cdot)}{\mathrm{meas}(\Xi_1(\varepsilon^*))} \right)^{\mathbb{N}}$. Then, there exists a set $\Xi_\infty^* \subset \Xi_\infty$, $\mathrm{meas}(\Xi_\infty^*) = 1$, such that if $\underline{\omega} \in \Xi_\infty^*$, there exist a sequence of centers $\underline{c}$ and a K analytic in some strip of $(\mathbb{T}^l)^{\mathbb{N}}$ so that*

$$J_\infty \nabla H \circ K = \partial_{\underline{\omega}} K$$

Note that we have stated Theorem 3.11 only for decay functions of the form given in Proposition 2.3. It is clear that the proof only uses a few properties of the function (e.g. monotonicity in the modulus of the argument). We have refrained from reformulating the theorem in more abstract terms.

We note that the only smallness conditions in ε enter just in the first stage of creating individual breathers around each site and in the preservation of the hyperbolic structure and other non-degeneracy conditions. The second stage, on the other hand does not require any other smallness conditions.

In the construction of infinite dimensional breathers out of single breathers we just need to exclude a few sequences of frequencies which are very resonant (they have measure zero in the probability measure indicated above). See Section 9.3. The smallness assumptions that we need to couple the sequences can be adjusted just by placing the different breathers far apart and we do not need any further smallness conditions in ε.

Note also that the only hypothesis on the one site system is the existence of positive measure of KAM tori (and the existence of a hyperbolic fixed point). This is implied if the system is close to a non-degenerate integrable system. Nevertheless, there are other arguments to show existence of KAM tori in systems very far from integrable [Dua94, Dua08]. Any of these systems could be taken as the basis for Theorem 3.11.

4. The Newton step

The sketch of the proof of Theorem 3.5 is roughly the same as for finite dimensional systems, with some minor changes detailed in Remark 4.20. Of course, even if the strategy is similar to that in finite dimensions, all the details need to be different since the situation is very different and we need to pay attention to the decay properties. With a view to applications to almost periodic solutions of Theorem 3.11, we also need to pay attention to the change in the non-degeneracy conditions and in the hyperbolicity properties and establish that many of the smallness assumptions are independent of the number and the geometry of the centers of oscillation.

The proof of Theorem 3.5 is based on a Newton iteration of Nash-Moser type. The estimates of the Newton step – including uniqueness – are summarized in Section 4.1 (See, Lemma 4.1). The fact that the inductive step can be iterated is more or less standard in KAM theory and it is done in Section 5.

The proof of the estimates of the Newton step are obtained in different stages

(1) We show that the approximate invariant hyperbolic splitting can be transformed in an invariant splitting. See Section 4.2.

(2) The equations for the Newton step can be divided into equations along the hyperbolic spaces (studied in Section 4.4) and the center space (studied in Section 4.3).

As usual in the study of cohomology equations, the equations in the center direction are much more subtle. In particular, the Diophantine properties and the geometric properties are only used in the equations in the center space.

(3) Once we have the estimates for the approximate solutions of the linearized equation, we show that, using the linearized equation, they improve the

solutions of the translated equation. Furthermore, we estimate the changes in the hyperbolicity constants and the non-degeneracy estimates.

(4) The passage from Theorem 3.5 to Theorem 3.6 is a geometric argument (a vanishing lemma) undertaken in Section 6.

4.1. **Estimates for the inductive step.** In this section we describe the inductive step of the proof of Theorem 3.5.

By Taylor's theorem we can write

$$\mathcal{F}_\omega(\lambda + \Lambda, K + \Delta) = \mathcal{F}_\omega(\lambda, K) + D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) + O(|(\Lambda, \Delta)|^2).$$

Assuming that (λ, K) is a pair that satisfies $\mathcal{F}_\omega(\lambda, K) = 0$ approximately with an error $E(\theta) = \mathcal{F}_\omega(\lambda, K)(\theta)$ we look for (Λ, Δ) such that $\mathcal{F}_\omega(\lambda + \Lambda, K + \Delta)$ is as small as possible. Then we are lead to consider the following Newton equation

$$(27) \qquad\qquad D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E,$$

where

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta)(\theta) = \frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\Lambda + DF_\lambda(K(\theta))\Delta(\theta) - \Delta(\theta + \omega).$$

To solve (27) we project the equation on both the center and the hyperbolic subspaces, taking advantage of the invariant splitting. Then we try to solve the projected equations. The one on the center subspace is reduced to two small divisors equations, essentially one on the tangent of the torus and the other on its conjugated directions. Taking advantage of the extra variable we can solve these equations up to a quadratic error. Using the conditions on the co-cycles over T_ω, we solve the projection on the stable and unstable subspaces.

The next result gives an approximate solution of (27) with precise estimates.

Lemma 4.1. *Under the hypotheses of Theorem 3.5 the equation*

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E$$

has an approximate solution (Λ, Δ) in the following sense: let

$$\tilde{E} = D_{\lambda,\kappa}\,\mathcal{F}_\omega(\lambda, K)\,(\Lambda, \Delta) + E\,.$$

For $0 < \delta < \rho$ we have the following estimates

$$\|\Delta\|_{\rho-\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-2\nu}\|E\|_{\rho,\underline{c},\Gamma},$$

$$|\Lambda| \leq C\|E\|_{\rho,\underline{c},\Gamma},$$

$$\|\tilde{E}\|_{\rho-\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-(2\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}\|\mathcal{F}_\omega(\lambda, K)\|_{\rho,\underline{c},\Gamma}.$$

Moreover, if Δ and $\tilde{\Delta}$ are solutions of (27) as above, i.e. solutions with quadratic error bounded by $C\kappa^2\delta^{-(2\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}\,\|\mathcal{F}_\omega\|_{\rho,\underline{c},\Gamma}$,

$$\|\Delta - \tilde{\Delta} - DK(\theta)\alpha\|_{\rho-\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-(2\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}\|\mathcal{F}_\omega(\lambda, K)\|_{\rho,\underline{c},\Gamma}.$$

In the previous estimates, the constant C depends on $\rho, l, \|DK\|_{\rho,\underline{c},\Gamma}, \|\Pi^{s,c,u}_{K(\theta)}\|_{\rho,\Gamma},$ $\|\frac{\partial F_\lambda}{\partial \lambda}\|_{\rho,\underline{c},\Gamma}$, the hyperbolicity constants and the decay function Γ but it does not depend on $\underline{c}$.

4.2. **Construction of invariant splittings out of approximately invariant ones.** The main result of this section will be Proposition 4.2, which establishes that given an approximately invariant splitting satisfying Definition 3.1, there is a truly invariant splitting nearby. Furthermore, we can estimate the distance between the true invariant splitting and the approximately invariant one. This, of course, implies the usual formulation of persistence of splittings under small perturbations.

The way that this result fits into the Newton scheme is that this will allow us to split the equation into different components. Compared to other estimates in the Newton step, the construction of invariant circles requires much less sophisticated analysis (it suffices to use contractions) and it does not require inductive assumptions nor making choices (e.g. the domain loss). The subtlety of the results comes because we have to choose appropriate spaces so that the estimates are uniform in the domains, the arrangement of the centers, etc. This uniformity of the results will be used when we consider the limit of infinitely many frequencies.

The method of proof we use is very similar to the standard proof using graph transforms [HP70, HPS77], which adapts very well to infinite dimensions [PS99]. Of course, there are several subtleties due to the infinite dimensional nature of the problem. In particular, we make essential use of the Banach algebra properties of the decay functions to make sure that we obtain estimates in the same spaces of functions (it is interesting to compare this with previous results in lattice dynamical systems). We emphasize that, in particular, the smallness conditions are independent of the centers of the embedding. This will be crucial when we consider the limit of a large number of centers.

Proposition 4.2. *Assume that the embedding $\tilde{K}$ has an δ-invariant hyperbolic splitting with respect to a map F (See Definition 3.2).*

There exists $\delta_0 > 0$ depending on $\|N\|_{\rho,\Gamma}, \|DF \circ \tilde{K}\|_{\rho,\underline{c},\Gamma}, \|DF^{-1} \circ \tilde{K}\|_{\rho,\underline{c},\Gamma}$ and $\|\Pi^{s,c,u}\|_{\rho,\underline{c},\Gamma}, \mu_{s,c,u}$ such that if $0 < \delta < \delta_0$ there is an analytic splitting

$$(28) \qquad\qquad T_{\tilde{K}(\theta)}\mathcal{M} = \tilde{\mathcal{E}}^s_{\tilde{K}(\theta)} \oplus \tilde{\mathcal{E}}^c_{\tilde{K}(\theta)} \oplus \tilde{\mathcal{E}}^\mu_{\tilde{K}(\theta)}$$

which is invariant under the co-cycle $DF \circ \tilde{K}$ over T_ω.

Let $\tilde{\Pi}^{s,c,\mu}_{\tilde{K}(\theta)}$ be the projections corresponding to the splitting (28). We furthermore have that there exist $C_h, 0 < \mu_1, \mu_2 < 1, \mu_3 > 1$ such that $\mu_1\mu_3 < 1, \mu_2\mu_3 < 1$ and the characterizations (13), (14), (15) of the splitting hold.

Moreover, there exists $C > 0$, depending on the same quantities as δ_0 does, such that for $0 < \delta < \delta_0$

$$\|\Pi^{s,c,\mu}_{\tilde{K}(\theta)} - \tilde{\Pi}^{s,c,\mu}_{\tilde{K}(\theta)}\|_{\rho,\underline{c},\Gamma} \le C\delta,$$

$$|\mu_{1,2,3} - \tilde{\mu}_{1,2,3}| < C\delta.$$

Proof. The ideas in this proof follow the ones in [FdlLS09a]. They have been taken from [HPPS70]. We make sure that the estimates are uniform with respect to δ and $\underline{c}$. We divide the proof into several steps.

Step 1: Construction of the invariant spaces. The existence of the invariant splitting will be done through the Banach fixed point principle applied to a graph transform operator.

We begin with the case of the stable bundle $\tilde{\mathcal{E}}^s_{\tilde{K}(\theta)}$. We describe the stable space $\tilde{\mathcal{E}}^s_{\tilde{K}(\theta)}$ as the graph of a linear map, i.e. $\tilde{\mathcal{E}}^s_{\tilde{K}(\theta)} = \text{Graph}\,(u \circ \tilde{K})$, where $u \circ \tilde{K}$ maps $\mathcal{E}^s_{\tilde{K}(\theta)}$ linearly into $\mathcal{E}^c_{\tilde{K}(\theta)} \oplus \mathcal{E}^u_{\tilde{K}(\theta)}$.

Since the splitting (16) is approximately invariant we can write the matrix $DF\big(\tilde{K}(\theta)\big)$ with respect to this decomposition as

$$DF\big(\tilde{K}(\theta)\big) = \begin{pmatrix} a_{11}(\theta) & a_{12}(\theta) \\ a_{21}(\theta) & a_{22}(\theta) \end{pmatrix}$$

with $\|a_{12}\|_{\rho,\underline{c},\Gamma} < C\delta$, $\|a_{21}\|_{\rho,\underline{c},\Gamma} < C\delta$. We also write

$$DF\big(\tilde{K}(\theta + (N-1)\,\omega)\big) \times \cdots \times DF\big(\tilde{K}(\theta)\big) = \begin{pmatrix} a_{11}^N(\theta) & a_{12}^N(\theta) \\ a_{21}^N(\theta) & a_{22}^N(\theta) \end{pmatrix}.$$

Note that by (17), (18) and (19) we have

(29) $$\|a_{11}^N\|_{\rho,\underline{c},\Gamma} \le \mu_1^N\,, \qquad \|a_{22}^{-N}\|_{\rho,\underline{c},\Gamma} \le \mu_3^N\,,$$

and

(30) $$\|a_{12}^N\|_{\rho,\underline{c},\Gamma} \le C\delta\,, \qquad \|a_{21}^N\|_{\rho,\underline{c},\Gamma} \le C\delta\,.$$

The graph condition over the co-cycle is

$$DF \circ \tilde{K}(\theta) \begin{pmatrix} \mathrm{Id} \\ u \circ \tilde{K}(\theta) \end{pmatrix} \in \mathrm{graph}\Big(u \circ \tilde{K}\big(T_\omega(\theta)\big)\Big).$$

This gives the functional equation for the map u

(31) $$u \circ \tilde{K}\big(T_\omega(\theta)\big)(a_{11} + a_{12}\ u \circ \tilde{K})(\theta) = (a_{21} + a_{22}\ u \circ \tilde{K})\,(\theta)\,.$$

Denoting $\tilde{u} = u \circ \tilde{K}$, (31) can be rewritten as

(32) $$\tilde{u} = a_{22}^{-1}\big[\tilde{u} \circ T_\omega(a_{11} + a_{12}\,\tilde{u}) - a_{21}\big].$$

Let $\mathcal{L}_\eta$ be the ball of radius η in the space of linear operators from $\mathcal{E}_{\tilde{K}(\theta)}^s$ into $\mathcal{E}_{\tilde{K}(\theta)}^c \oplus \mathcal{E}_{\tilde{K}(\theta)}^u$ with the norm $\|\cdot\|_\Gamma$.

Let $\mathcal{S}_\eta$ be the space of analytic sections from D_ρ to $\mathcal{L}_\eta$, i.e. the space of u : $D_\rho \to \mathcal{L}_\eta$ such that $u(\theta) : \mathcal{E}_{\tilde{K}(\theta)}^s \to \mathcal{E}_{\tilde{K}(\theta)}^c \oplus \mathcal{E}_{\tilde{K}(\theta)}^u$ with the norm $\|.\|_{\rho,\underline{c},\Gamma}$.

We take the operator $\mathcal{T} : \mathcal{S}_\eta \to \mathcal{S}_\eta$ defined as the right-hand side of (32). $\mathcal{T}$ is approximated by $\mathcal{T}_0 : \mathcal{S}_\eta \to \mathcal{S}_\eta$ defined by

$$\mathcal{T}_0\,\tilde{u} = a_{22}^{-1}\ \tilde{u} \circ T_\omega\ a_{11}.$$

We now consider $\mathcal{T}^N$ and $\mathcal{T}_0^N$. An elementary computation gives

$$\mathcal{T}_0^N\,\tilde{u} = a_{22}^{-1} \ldots a_{22}^{-1} \circ T_\omega^{N-1}\,\tilde{u} \circ T_\omega^N\ a_{11} \circ T_\omega^{N-1} \ldots a_{11}\,.$$

Moreover, taking into account that $\mathcal{T}$ is a degree two polynomial operator, we obtain by simple algebraic manipulations

(33) $$\|\mathcal{T}^N - \mathcal{T}_0^N\| < C\delta$$

and

(34) $$\mathrm{Lip}(\mathcal{T}^N - \mathcal{T}_0^N) < C\delta\,.$$

Using the Banach algebra properties of the decay norms, we have that:

(35) $$\|a_{11}^N(\theta) - a_{11}\big(T_\omega^{N-1}(\theta)\big) \ldots a_{11}(\theta)\|_{\rho,\underline{c},\Gamma} < C\delta\,,$$

(36) $$\|a_{22}^{-N}(\theta) - a_{22}^{-1}(\theta) \ldots a_{22}^{-1}\big(T_\omega^{N-1}(\theta)\big)\|_{\rho,\underline{c},\Gamma} < C\delta\,.$$

By (29), (35) and (36), if δ is small $\mathcal{T}_0^N$ sends $\mathcal{S}_\eta$ into $\mathcal{S}_\eta$ for all $\eta \in (0,1]$ and is a contraction in this domain. By (33) and (34), if δ is small $\mathcal{T}^N$ sends $\mathcal{S}_\eta$ into $\mathcal{S}_\eta$ for $\eta \in (C\delta, 1]$ and it is also a contraction.

Therefore $\mathcal{T}^N$ has a unique fixed point u^* in $\mathcal{S}_1$ which belongs to $\mathcal{S}_{C\delta}$. It is clear that $\mathcal{T}u^*$ is also a fixed point of $\mathcal{T}^N$, which belongs to $\mathcal{S}_{C'\delta} \subset \mathcal{S}_1$ for some $C' \geq C$. By uniqueness $\mathcal{T}u^* = u^*$.

A similar method can be applied for the center-unstable subspace. In this case the graph condition reads

$$DF_\lambda \circ \tilde{K}(\theta) \begin{pmatrix} v \circ \tilde{K}(\theta) \\ \mathrm{Id} \end{pmatrix} \in \mathrm{Graph}\Big(v \circ \tilde{K}\big(T_\omega(\theta)\big)\Big),$$

and the resulting operator $\mathcal{T} : \mathcal{S} \to \mathcal{S}$ is

(37) $$\mathcal{T}\tilde{v} = \big[(\tilde{a}_{11}\tilde{v} + \tilde{a}_{12})(\tilde{a}_{21}\tilde{v} + \tilde{a}_{22})^{-1}\big] \circ T_\omega^{-1},$$

where $\tilde{v} = v \circ \tilde{K}$. Repeating the same procedure as above, we construct the center-unstable space and obtain similar bounds.

Now let us consider $G = F^{-1}$. We observe that the stable space associated to the map G is the unstable space associated to F. Hence, applying the above procedure to G, we construct the unstable space $\tilde{\mathcal{E}}^u_{\tilde{K}(\theta)}$ and center-stable space $\tilde{\mathcal{E}}^{c,s}_{\tilde{K}(\theta)}$ with similar bounds. Finally we note that $\tilde{\mathcal{E}}^c_{\tilde{K}(\theta)} = \tilde{\mathcal{E}}^{c,s}_{\tilde{K}(\theta)} \bigcap \tilde{\mathcal{E}}^{c,u}_{\tilde{K}(\theta)}$.

Step 2: Estimates on the projections. We want to estimate the norm of the projection $\tilde{\Pi}^s_{\tilde{K}(\theta)}$ compared to the one of $\Pi^s_{\tilde{K}(\theta)}$.

Let $\xi \in T_{\tilde{K}(\theta)}\mathcal{M}$. Using the decomposition $\xi = (\xi^s, \xi^{cu}) \in \mathcal{E}^s_{\tilde{K}(\theta)} \oplus (\mathcal{E}^c_{\tilde{K}(\theta)} \oplus \mathcal{E}^u_{\tilde{K}(\theta)})$ we have the following representations

$$\Pi^s_{\tilde{K}(\theta)}\xi = (\xi^s, 0), \qquad \tilde{\Pi}^s_{\tilde{K}(\theta)}\xi = (\tilde{\xi}^s, \tilde{u}(\theta)\tilde{\xi}^s),$$

$$\Pi^{cu}_{\tilde{K}(\theta)}\xi = (0, \xi^{cu}), \qquad \tilde{\Pi}^{cu}_{\tilde{K}(\theta)}\xi = (\tilde{v}(\theta)\tilde{\xi}^{cu}, \tilde{\xi}^{cu}).$$

Then

$$\xi^s = \tilde{\xi}^s + \tilde{v}(\theta)\tilde{\xi}^{cu},$$

$$\xi^{cu} = \tilde{u}(\theta)\tilde{\xi}^s + \tilde{\xi}^{cu}$$

or equivalently

$$\begin{pmatrix} \tilde{\xi}^s \\ \tilde{\xi}^{cu} \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & \tilde{v}(\theta) \\ \tilde{u}(\theta) & \mathrm{Id} \end{pmatrix}^{-1} \begin{pmatrix} \xi^s \\ \xi^{cu} \end{pmatrix}$$

since the matrix $B = \begin{pmatrix} \mathrm{Id} & \tilde{v}(\theta) \\ \tilde{u}(\theta) & \mathrm{Id} \end{pmatrix}$ is invertible because is $O(\delta)$-close to the identity and moreover, by the Neumann series theorem, we can write

$$B^{-1} = \begin{pmatrix} \mathrm{Id} + w_{11}(\theta) & w_{12}(\theta) \\ w_{21}(\theta) & \mathrm{Id} + w_{22}(\theta) \end{pmatrix}$$

with $\|w_{ij}\|_{\rho,\Gamma} < C\delta$. Therefore using

$$\Big(\tilde{\Pi}^s_{\tilde{K}(\theta)} - \Pi^s_{\tilde{K}(\theta)}\Big) \begin{pmatrix} \xi^s \\ \xi^{cu} \end{pmatrix} = \begin{pmatrix} \tilde{\xi}^s - \xi^s \\ \tilde{u}(\theta)\,\tilde{\xi}^s \end{pmatrix},$$

this gives

$$\big\|\tilde{\Pi}^s_{\tilde{K}(\theta)} - \Pi^s_{\tilde{K}(\theta)}\big\|_{\rho,\underline{c},\Gamma} \leq C\delta.$$

Analogously one has $\big\|\tilde{\Pi}^{cu}_{\tilde{K}(\theta)} - \Pi^{cu}_{\tilde{K}(\theta)}\big\|_{\rho,\underline{c},\Gamma} < C\delta$.

The estimates for the projections Π^u, Π^{sc} are obtained in a similar way. From those we deduce readily the ones for Π^c by noting that $\Pi^c = \Pi^{cu} - \Pi^u$.

Step 3: Existence of μ_1, μ_2, C_h and estimates. Since the distance between the spaces $\mathcal{E}^{s,c,u}_{\tilde{K}(\theta)}$ and $\tilde{\mathcal{E}}^{s,c,u}_{\tilde{K}(\theta)}$ is bounded by $C\delta$, the restriction of $DF \circ \tilde{K} \circ T^{N-1}_\omega \times \cdots \times DF \circ \tilde{K}$ to them is separated by a distance less than $C\delta$. Therefore there exist μ_1, μ_2, μ_3 with μ_1, $\mu_2 < 1$, $\mu_3 > 1$, $\mu_1\mu_3 < 1$, $\mu_2\mu_3 < 1$ such that $|\mu_{1,2,3} - \tilde{\mu}_{1,2,3}| < C\delta$ and (17) holds for $\tilde{\mu}_1 = \mu_1$ and $v \in \tilde{\mathcal{E}}^s_{\tilde{K}(\theta)}$. Similarly for (18) and (19).

Once we have these last properties we deduce that there exists C_h such that (13), (14) and (15) hold for all $n \geq 1$. □

As a consequence of Proposition 4.2 we have the following result.

Proposition 4.3. *Under the hypotheses of Proposition 5.1, assume that $\|K - \tilde{K}\|_{\rho,\underline{c},\Gamma}$ is small enough. Then there exists an analytic splitting*

$$T_{\tilde{K}(\theta)}\mathcal{M} = \mathcal{E}^s_{\tilde{K}(\theta)} \oplus \mathcal{E}^c_{\tilde{K}(\theta)} \oplus \mathcal{E}^u_{\tilde{K}(\theta)}$$

invariant under the co-cycle $DF_\lambda \circ \tilde{K}$ over T_w.

Furthermore, there exists $C > 0$ such that

$$\left\|\Pi^{s,c,u}_{K(\theta)} - \Pi^{s,c,u}_{\tilde{K}(\theta)}\right\|_{\rho,\Gamma} \leq C\|K - \tilde{K}\|_{\rho,\underline{c},\Gamma},$$

$$|\mu_i - \tilde{\mu}_i| \leq C\|K - \tilde{K}\|_{\rho,\underline{c},\Gamma}, \qquad i = 1, 2, 3.$$

Proof. The invariant splitting

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$$

for $DF_\lambda \circ K$ is an approximate invariant splitting for $DF_\lambda \circ \tilde{K}$. Here we identify $T_{\tilde{K}(\theta)}\mathcal{M}$ with $T_{K(\theta)}\mathcal{M}$ since M is Euclidean. We then can take $\delta = C\|K - \tilde{K}\|_{\rho,\underline{c},\Gamma}$. □

4.3. **Solution of the linearized equation on the center subspace.** In this section we solve approximately the projection of equation (27) on the invariant center subspace provided by Proposition 4.3 and establish estimates. Projecting with $\Pi^c_{K(\theta+\omega)}$ and using the notation

$$\Delta^c(\theta) = \Pi^c_{K(\theta)}\Delta(\theta), \qquad E^c(\theta) = \Pi^c_{K(\theta+\omega)}E(\theta),$$

we obtain

$$(38) \qquad \Pi^c_{K(\theta+\omega)}\frac{\partial F_\lambda(K(\theta))}{\partial\lambda}\Lambda + DF_\lambda(K(\theta))\Delta^c(\theta) - \Delta^c(\theta + \omega) = -E^c(\theta).$$

4.3.1. *Estimates on cohomology equations.* We recall the well-known small divisors lemma (see [Rüs76a], [Rüs76b], [Rüs75], [dlL01]).

Proposition 4.4. *Let M be a finite dimensional Euclidean manifold and $\omega \in D(\kappa, \nu)$. Assume the mapping $h : D_\rho \to M$ is analytic on D_ρ and has zero average. Then for any $0 < \sigma < \rho$ the difference equation*

$$v(\theta + \omega) - v(\theta) = h(\theta)$$

has a unique zero average solution $v : \mathbb{T}^l \to M$, real analytic on $D_{\rho-\sigma}$ for any $0 < \sigma < \rho$. Moreover, we have the estimate

$$(39) \qquad \qquad \|v\|_{\rho-\sigma} \leq C\kappa\sigma^{-\nu}\|h\|_\rho,$$

where C only depends on ν and the dimension of the torus l.

We have the following corollary of Rüssmann's result in our context.

Corollary 4.5. *Let $\mathcal{M} = \ell^\infty(\mathbb{Z}^N)$ and $\omega \in D(\kappa, \nu)$ and assume the mapping $h : \mathbb{T}^l \to \mathcal{M}$ belongs to $\mathcal{A}_{\rho,\underline{c},\Gamma}$ and has zero average. Then for any $0 < \sigma < \rho$ the difference equation*

$$v(\theta + \omega) - v(\theta) = h(\theta)$$

has a unique zero average solution $v : \mathbb{T}^l \to \mathcal{M}$, belonging to $\mathcal{A}_{\rho-\sigma,\underline{c},\Gamma}$ for any $0 < \sigma < \rho$. Moreover, we have the estimate

$$(40) \qquad \|v\|_{\rho-\sigma,\underline{c},\Gamma} \leq C\kappa\sigma^{-\nu}\|h\|_{\rho,\underline{c},\Gamma},$$

where C only depends on ν and linearly on the dimension of the torus l.

Remark 4.6. An important fact of the previous statement is that, since we consider the supremum norm on M and the equation is solved component by component, the estimates are independent on the dimension of $\mathcal{M}$.

Proof. We write the equation in coordinates $i \in \mathbb{Z}^N$ to get

$$v_i(\theta + \omega) - v_i(\theta) = h_i(\theta)$$

with v_i and h_i mapping $\mathbb{T}^l$ into M. We then apply the previous finite-dimensional result Proposition 4.4 to each of the components.

We also observe that the partial derivatives of the functions also satisfy

$$\partial_{\theta_k} v_i(\theta + \omega) - \partial_{\theta_k} v_i(\theta) = \partial_{\theta_k} h_i(\theta)$$

and that $\partial_{\theta_k} h_i$ has zero average. Then, $\partial_{\theta_k} v_i(\theta)$ will be the zero average solution obtained applying Proposition 4.4. Multiplying the estimates afforded by Proposition 4.4 by $\Gamma^{-1}(i - c_j)$ and taking the supremum in i and the infimum in j we get the desired result. $\square$

4.3.2. *Isotropic character of the torus.* One important issue is the approximate isotropic character of the approximate torus $K(\mathbb{T}^l)$. In our context the two-form Ω_∞ is formal but $K(\mathbb{T}^l)$ is finite-dimensional, therefore asking the forms to be isotropic amounts to ask that the pull-back $K^*\Omega_\infty$ vanishes. By the decay properties of K (see Lemma 2.9), this last object is a true form on $\mathbb{T}^l$ and we can write

$$K^*\Omega_\infty(\theta)(\xi, \eta) = \langle \xi, L(\theta)\eta \rangle, \qquad \xi, \eta \in \mathbb{R}^l.$$

The isotropic character of the torus is then equivalent to

$$L(\theta) \equiv DK(\theta)^\top J_\infty\big(K(\theta)\big) DK(\theta) = 0$$

for all $\theta \in \mathbb{T}^l$. Notice that L is an $l \times l-$ matrix.

We first consider the case when K is a solution of (11).

Lemma 4.7. *Let $(\mathcal{M}, \Omega_\infty = d\alpha_\infty)$ be the lattice manifold. Assume that $F_{\lambda_0} \circ K = K \circ T_\omega$, F_{λ_0} is symplectic, ω is rationally independent and $K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$. Then $L(\theta)$ is identically zero.*

Proof. Since $K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, using that F_{λ_0} is symplectic (see Appendix 11), we have

$$K^*\Omega_\infty = K^* F_{\lambda_0}^* \Omega_\infty = (K \circ T_\omega)^*\Omega_\infty.$$

By the condition on ω, T_ω is ergodic and therefore $K^*\Omega_\infty$ is constant. Hence L is also constant. Moreover, the fact that $\mathcal{M}$ is formally exact symplectic shows that $K^*\Omega_\infty = d(K^*\alpha_\infty)$, where now, d is the differential on the torus (see Appendix 11) . In coordinates this means that $L(\theta)$ has the form $DL_1(\theta)^\top - DL_1(\theta)$ for some finite-dimensional vector $L_1(\theta)$. Since the average of derivatives is zero we get that L is zero on $\mathbb{T}^l$. $\square$

4.3.3. *Geometric considerations on the center bundle $\mathcal{E}^c_{K(\theta)}$ in the exact case.* In this section, we show how to construct geometrically a very natural basis of the center subspace $\mathcal{E}^c_{K(\theta)}$ when K satisfies (11). Recall first that we are assuming that $\mathcal{E}^c_{K(\theta)}$ is finite-dimensional with dimension $2l$. We start with the following lemma.

Lemma 4.8. *The restriction $\Omega^c_{K(\theta)}$ of Ω_∞ to the finite dimensional space $\mathcal{E}^c_{K(\theta)}$ is a symplectic form on $\mathcal{E}^c_{K(\theta)}$.*

Proof. To prove the claim, it is enough to show that the form $\Omega^c_{K(\theta)}$ is non degenerate. Assume that $v \in \mathcal{E}^c_{K(\theta)}$ and $w \in \mathcal{E}^u_{K(\theta)}$. Then we have

$$\Omega^c_{K(\theta)}(v,w) = \Omega^c_{F \circ K(\theta)}(DF^n \circ Kv, DF^n \circ Kw) \quad n \in \mathbb{Z}.$$

Since the torus is invariant, we have that

$$DF \circ K \circ T^{n-1}_\omega(\theta) \times \cdots \times DF \circ K(\theta) = DF^n \circ K(\theta).$$

We deduce, sending $n \to +\infty$ and using the hyperbolic conditions (expansion/contraction properties), that $Omega^c_{K(\theta)}(u,v) = 0$ in the following cases

- $u, v \in \mathcal{E}^s_{K(\theta)}$,
- $u, v \in \mathcal{E}^u_{K(\theta)}$,
- $u \in \mathcal{E}^s_{K(\theta)} \cup \mathcal{E}^u_{K(\theta)}$ and $v \in \mathcal{E}^c_{K(\theta)}$,
- $v \in \mathcal{E}^c_{K(\theta)}$ and $v \in \mathcal{E}^s_{K(\theta)} \cup \mathcal{E}^u_{K(\theta)}$

Assume now that

$$\Omega^c_{K(\theta)}(c, \tilde{c}) = 0, \quad \forall \tilde{c} \in \mathcal{E}^c_{K(\theta)}.$$

By the previous argument, we have that for $u \in \mathcal{E}^u_{K(\theta)}$ and $s \in \mathcal{E}^s_{K(\theta)}$

$$\Omega^c_{K(\theta)}(\tilde{c}, u) = \Omega^c_{K(\theta)}(\tilde{c}, s) = 0.$$

Since Ω is non-degenerate, this leads to the desired result. $\square$

Since range $DK(\theta)$ is the tangent space of the torus $K(\mathbb{T}^l)$ and the dynamics on the torus is conjugated to a rotation, $DK(\theta)\mathbb{R}^l$ is contained in $\mathcal{E}^c_{K(\theta)}$. Moreover we have that $\Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta)\mathbb{R}^l$ also is contained in $\mathcal{E}^c_{K(\theta)}$. Instead of $\Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta)$ we will consider the matrix $\Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta) N(\theta)$ where $N(\theta)$ is the normalization $l \times l$-matrix $N(\theta) = [DK(\theta)^\top DK(\theta)]^{-1}$ previously introduced. Both have the same range because $N(\theta)$ is non-singular. The role of N is to provide some normalization for the symplectic conjugate.

Now we check that the range of $[\Pi^c_{K(\theta)} DK(\theta), \Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta) N(\theta)]$ is $2l$-dimensional. Indeed, assume that there is a linear combination

$$f = \sum_{j=1}^l \alpha_j \Pi^c_{K(\theta)} DK(\theta) e_j + \sum_{j=1}^l \beta_j \Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta) N(\theta) e_j = 0.$$

Then, for $1 \le k \le l$, using the isotropic character of $T_{K(\theta)} K(\mathbb{T}^l)$

$$0 = \sum_{j=1}^l \beta_j e_k^\top DK(\theta)^\top \Pi^c_{K(\theta)} J_\infty(K(\theta)) \Pi^c_{K(\theta)} J_\infty(K(\theta))^{-1} DK(\theta) N(\theta) e_j$$

$$= \sum_{j=1}^l \beta_j \langle e_k, e_j \rangle = \beta_k.$$

This calculation shows that f reduces to $\sum_{j=1}^{l} \alpha_j \Pi_{K(\theta)}^c DK(\theta) e_j$. Moreover, for $1 \le k \le l$

$$0 = \sum_{j=1}^{l} \alpha_j e_k^\top N(\theta)^\top DK(\theta)^\top \Pi_{K(\theta)}^c J_\infty(K(\theta))^{-\top} \Pi_{K(\theta)}^c J_\infty(K(\theta)) DK(\theta) e_j$$

$$= -\sum_{j=1}^{l} \alpha_j \langle e_k, e_j \rangle = \alpha_k.$$

Hence $\alpha_j = \beta_j = 0$ for all $j = 1, ..., l$. We conclude that

$$range[\Pi_{K(\theta)}^c DK(\theta), \Pi_{K(\theta)}^c J_\infty(K(\theta))^{-1} DK(\theta) N(\theta)] = \mathcal{E}_{K(\theta)}^c.$$

The treatment of equation (27) on the center subspace is greatly facilitated by observing that the preservation of the symplectic structure imposes that the system is close to a diagonal structure.

In our context, the dimension of the center subspace is $2l$ and $\mathcal{E}_{K(\theta)}^c \sim \mathbb{R}^{2l}$. In [FdlLS09a], the authors studied the case when $\mathcal{M}$ is finite-dimensional (with symplectic structure J). They used the set of vectors

$$\left\{ \frac{\partial K(\theta)}{\partial \theta_j}, J^{-1}(K(\theta)) \frac{\partial K(\theta)}{\partial \theta_j} \right\}_{j=1,...,l}$$

to perform a transformation which allows to approximately solve up to quadratic error the projected equation on the center subspace.

We consider the map $\tilde{M}(\theta)$ given in matrix notation by

(41) $$\tilde{M} = \left[\Pi_{K(\theta)}^c DK(\theta) \ \Pi_{K(\theta)}^c J_\infty(K(\theta))^{-1} DK(\theta) N(\theta) \right]$$

We will see that this map is a very convenient change of coordinates in the linearized equations which makes them solvable.

Remark 4.9. It is worth noting that all the quantities defined above (such as N, for instance) make perfect sense even if we are manipulating "infinite-dimensional" matrices. This is due to the fact that we are considering maps in $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$. For instance, since K is assumed to be in $\mathcal{A}_{\rho,\underline{c},\Gamma}$, one has for $i, j = 1, \ldots, l$

$$N(\theta)_{ij}^{-1} = \sum_{k \in \mathbb{Z}^N} \frac{\partial K_k}{\partial \theta_i} \frac{\partial K_k}{\partial \theta_j}.$$

Therefore, for $\theta \in D_\rho$ this leads to

$$|N(\theta)_{ij}^{-1}| \le \|DK\|_{\rho,\underline{c},\Gamma}^2 \min_{1 \le p,q \le R} \sum_{k \in \mathbb{Z}^N} \Gamma(k - c_p) \Gamma(k - c_q)$$

$$\le \|DK\|_{\rho,\underline{c},\Gamma}^2 \min_{1 \le p,q \le R} \Gamma(c_p - c_q) \le \|DK\|_{\rho,\underline{c},\Gamma}^2.$$

This gives that $\|N\|_{\rho,\Gamma} < \infty$.

Remark 4.10. It is also worth noticing that we have the following estimates for the generalized symplectic matrix $\tilde{J}_\infty$:

$$\|J_\infty\|_\Gamma = \Gamma^{-1}(0) \|J\|, \ \|J_\infty^{-1}\|_\Gamma = \Gamma^{-1}(0) \|J^{-1}\|$$

since $(J_\infty)_{ij} = J\delta_{ij}$, where δ_{ij} is the Krönecker symbol.

4.3.4. *Representation of DF_λ on the center subspace.* In this section we study a suitable representation of DF_λ applied to the basis of the center subspace given by the columns of $\tilde{M}(\theta)$. We begin by considering the case when K is a solution of (11).

Lemma 4.11. *Let K be a solution of equation* (11). *Then there exists a $2l \times 2l$ matrix $\mathcal{S}_\lambda(\theta)$ such that*

$$(42) \qquad DF_\lambda(K(\theta))\tilde{M}(\theta) = \tilde{M}(\theta + \omega)\mathcal{S}_\lambda(\theta),$$

with

$$(43) \qquad \mathcal{S}_\lambda(\theta) = \begin{pmatrix} \mathrm{Id}_l & A_\lambda(\theta) \\ 0_l & \mathrm{Id}_l \end{pmatrix}.$$

The matrix $A_\lambda(\theta)$ is

$$A_\lambda(\theta) = P(\theta + \omega)^\top[(DF_\lambda\Pi^c_{K(\theta)}J_\infty^{-1})(K(\theta))P(\theta) - [\Pi^c_{K(.)}J_\infty^{-1}(K)P](\theta + \omega)],$$

where $P(\theta) = DK(\theta)N(\theta)$.

Proof. Differentiating equation (11) with respect to θ, we get

$$DF_\lambda(K(\theta))DK(\theta) = DK(\theta + \omega).$$

This shows that $\mathcal{S}_\lambda(\theta)$ has the form

$$\begin{pmatrix} \mathrm{Id}_l & A_\lambda(\theta) \\ 0_l & B(\theta) \end{pmatrix},$$

where $A_\lambda(\theta), B(\theta)$ are $l \times l$ matrices. We will now show that $B = \mathrm{Id}_l$ via geometric properties. We should have

$$(44)$$
$$[DF_\lambda(K)\Pi^c_{K(\theta)}J_\infty^{-1}(K)DK\,N](\theta) = \Pi^c_{K(\theta+\omega)}DK(\theta+\omega)\,A_\lambda(\theta) + [\Pi^c_{K(\theta)}J_\infty^{-1}(K)DK\,N](\theta+\omega)B(\theta).$$

By the isotropic character of $K(\mathbb{T}^l)$ we have $DK^\top J_\infty(K)DK = 0$ and the definition of N, we have

$$(45) \qquad [DK^\top J_\infty(K)](\theta + \omega)[DF(K)\Pi^c_{K(\theta)}J_\infty(K)^{-1}DK\,N](\theta) = B(\theta).$$

Also by the symplecticness of F_λ

$$J_\infty(K(\theta + \omega))DF_\lambda(K(\theta)) = J_\infty(F_\lambda(K(\theta)))DF_\lambda(K(\theta)) =$$
$$[DF_\lambda(K)^{-\top}J_\infty(K)](\theta).$$

Then equation (45) becomes

$$B(\theta) = DK^\top(\theta + \omega)[DF_\lambda(K)^{-\top}DK\,N](\theta) = [DK^\top DK\,N](\theta) = \mathrm{Id}_l.$$

To obtain the expression of $A_\lambda(\theta)$ we multiply (44) by $(DK\,N)(\theta + \omega)^\top$ to get

$$(46)$$
$$A_\lambda(\theta) = P(\theta + \omega)^\top\left[[DF_\lambda(K)\Pi^c_{K(\theta)}J_\infty(K)^{-1}P](\theta) - [\Pi^c_{K(\theta)}J_\infty(K)^{-1}P](\theta + \omega)\right].$$

$\square$

The matrix $\tilde{M}(\theta)$ is not invertible since it is not square. However we can derive a generalized inverse for $\tilde{M}(\theta)$. As a motivation for subsequent developments, we first present Lemma 4.12 which deals with the geometric cancellations in the case of an exactly invariant torus. The case of interest for a KAM algorithm — when the torus is only approximately invariant — will be studied in Lemma 4.14 as a perturbation of Lemma 4.12.
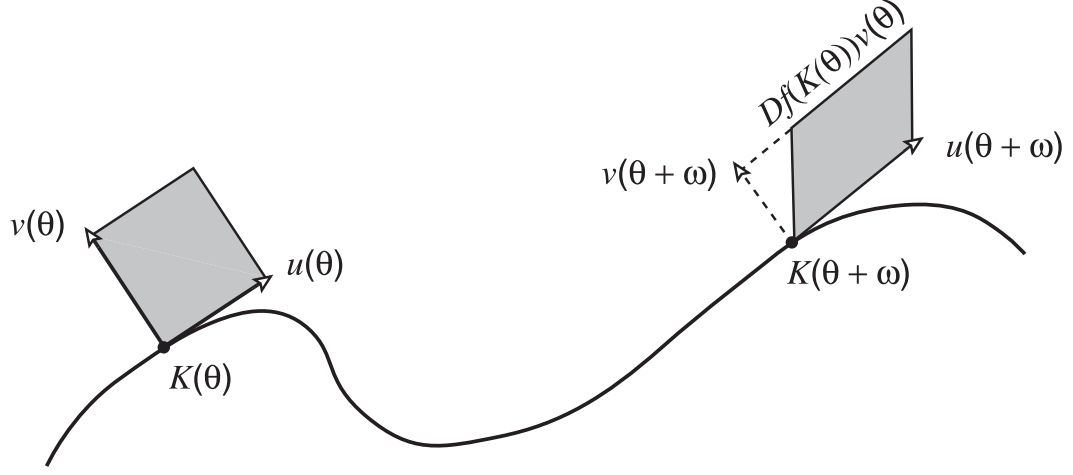
FIGURE 2. Illustration of the geometric reason for reducibility of the equations in the center. Note that the base of a rectangle gets mapped into the base of the other (differentiating the invariance). The preservation of the symplectic form implies that the symplectic area of both rectangles is the same.

A straightforward calculation shows that

$$(47) \quad \tilde{M}^\top \Pi_{K(\theta)}^c J_\infty(K) \tilde{M} = \begin{pmatrix} L & \mathrm{Id}_l \\ -\mathrm{Id}_l & (\Pi_{K(\theta)}^c)^\top N^\top DK^\top \Pi_{K(\theta)}^c J_\infty(K)^{-\top} DK \, N \end{pmatrix}.$$

Lemma 4.12. *Let K be a solution of* (11). *Then the matrix* $\tilde{M}^\top \Pi_{K(\theta)}^c J_\infty(K) \tilde{M}$ *is invertible and*

$$(\tilde{M}^\top \Pi_{K(\theta)}^c J_\infty(K) \tilde{M})^{-1} = \begin{pmatrix} (\Pi_{K(\theta)}^c)^\top N^\top DK^\top \Pi_{K(\theta)}^c J_\infty(K)^{-\top} DK \, N & -\mathrm{Id}_l \\ \mathrm{Id}_l & 0 \end{pmatrix}.$$

Proof. It follows from (47) and the isotropic character of the invariant torus, i.e. $L = 0$. □

Now we consider the case we are interested in, that is when K is an approximate solution with an error $E(\theta) = \mathcal{F}_\omega(\lambda, K)(\theta)$, assumed to be small. We will need the invertibility of $\tilde{M}(\theta)^\top J_\infty(K) \tilde{M}(\theta)$ in this case.

More precisely, we introduce

$$(48) \qquad\qquad e(\theta) = DF_\lambda(K(\theta)) \tilde{M}(\theta) - \tilde{M}(\theta + \omega) \mathcal{S}_\lambda(\theta),$$

where $\mathcal{S}_\lambda$ is given by (43). If we denote $e(\theta) = (e_1(\theta), e_2(\theta))$, a simple algebraic computation yields

$$e_1(\theta) = DE^c(\theta)$$
$$e_2(\theta) = [(DF_\lambda \Pi_{K(\theta)}^c J_\infty^{-1})(K) DK \, N](\theta) - \Pi_{K(\theta+\omega)}^c DK(\theta + \omega) A_\lambda(\theta)$$
$$- [\Pi_{K(.)}^c J_\infty^{-1} DK \, N](\theta + \omega) = O(E, DE)$$

by the choice of A_λ.

We first prove the approximate isotropic character of the torus

Lemma 4.13. *Under the previous conditions, let $K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ be a function which solves (11) approximately and let $E = F_\lambda \circ K - K \circ T_\omega$ be the error in equation (11). Then for $0 < \delta < \rho/2$,*

$$(49) \qquad \|L\|_{\rho-2\delta,\Gamma} \le C\kappa\,\delta^{-(\nu+1)}\,\|E\|_{\rho,\underline{c},\Gamma},$$

where C depends on $l, \nu, \rho,\ \ \|DK\|_{\rho,\underline{c},\Gamma}, \|F_{\lambda_0}\|_{C^1_\Gamma(B_r)}, \|J/infty\|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\underline{c},\Gamma}$.

Proof. We define the two-form on the torus $\mathbb{T}^l$

$$\Omega_e = K^*\Omega_\infty - (K \circ T_\omega)^*\Omega_\infty.$$

The corresponding matrix is $L - L \circ T_\omega$. Using that F_λ is symplectic we have that for any $(\xi, \eta) \in \mathbb{R}^{2l}$

$$\begin{aligned}
\Omega_e(\theta)(\xi, \eta) &= \big((F_\lambda \circ K)^*\Omega_\infty - (K \circ T_\omega)^*\Omega_\infty\big)(\theta)(\xi, \eta) \\
&= \sum_{i \in \mathbb{Z}^N} \Big[\Omega\big((F_\lambda)_i\big(K(\theta)\big)\big)\big(D((F_\lambda)_i \circ K)(\theta)\xi,\ D((F_\lambda)_i \circ K)(\theta)\eta\big) \\
&\quad - \Omega\big(K \circ T_\omega(\theta)\big)\big(D(K_i \circ T)(\theta)\xi,\ D(K_i \circ T)(\theta)\eta\big) \Big].
\end{aligned}$$

Since $D((F_\lambda)_i \circ K) - D(K_i \circ T) = DE_i$ and $\|DE\|_{\rho-\delta,\underline{c},\Gamma} \le \frac{c}{\delta}\|E\|_{\rho,\underline{c},\Gamma}$, using the decay properties of $F \circ K$ and $K \circ T_\omega$ to sum the series, we obtain that

$$(50) \qquad L - L \circ T_\omega = g$$

with $\|g\|_{\rho-\delta} \le C\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}$ for some C as in the statement. Now we use Proposition 4.4 to finish the proof. $\qquad\square$

The next step is to ensure the invertibility of the $2l \times 2l$-matrix $\tilde{M}^\top\Pi^c_{K(\theta)}J_\infty(K)\tilde{M}$. According to expression (47), we can write

$$\tilde{M}(\theta)^\top\Pi^c_{K(\theta)}J(K(\theta))_\infty\tilde{M}(\theta) = V(\theta) + R(\theta),$$

where

$$V = \begin{pmatrix} 0 & \mathrm{Id}_l \\ -\mathrm{Id}_l & N^\top DK^\top\Pi^c_{K(\theta)}J(K)^{-\top}_\infty DK\,N \end{pmatrix}$$

and

$$R = \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix}.$$

We have the following lemma, providing the desired invertibility result under a smallness assumption on E, namely (51) in the next lemma.

Lemma 4.14. *There exists a constant $C > 0$ such that if*

$$(51) \qquad C\kappa\delta^{-(\nu+1)}\|E\|_{\rho,\underline{c},\Gamma} \le 1/2$$

for some $0 < \delta < \rho/2$ then the matrix $\tilde{M}^\top(\theta)\Pi^c_{K(\theta)}J_\infty(K(\theta))\tilde{M}(\theta)$ is invertible for $\theta \in D_{\rho-2\delta}$ and there exists a matrix $\tilde{V}(\theta)$ such that

$$(\tilde{M}(\theta)^\top J_\infty(K(\theta))\tilde{M}(\theta))^{-1} = V(\theta)^{-1} + \tilde{V}(\theta)$$

with

$$\tilde{V}(\theta) = \Big(\sum_{k=1}^\infty (V(\theta)^{-1}R(\theta))^k\Big)V(\theta)^{-1},$$

where the series is absolutely convergent. Furthermore, we have the estimate

$$(52) \qquad \|\tilde{V}\|_{\rho-2\delta,\Gamma} \le C'\kappa\delta^{-(\nu+1)}\|E\|_{\rho,\underline{c},\Gamma},$$

where the constant $C' > 0$ depends on l, ν, $\|F\|_{C_\Gamma^1(B_r)}$, $\|J_\infty\|_{C_\Gamma^1(B_r)}$, $\|DK\|_{\rho,\underline{c},\Gamma}$, $\|N\|_{\rho,\Gamma}$ and $\|\Pi_{K(\theta)}^c\|_{\rho,\underline{c},\Gamma}$.

Proof. The matrix $V(\theta)$ is invertible with

$$V^{-1} = \begin{pmatrix} N^\top DK^\top \Pi_{K(\theta)}^c J_\infty(K)^{-\top} DK\, N & -\mathrm{Id}_l \\ \mathrm{Id}_l & 0 \end{pmatrix}.$$

We can write

$$\tilde{M}(\theta)^\top \Pi_{K(\theta)}^c J_\infty(K(\theta))\tilde{M}(\theta) = V(\theta)(\mathrm{Id}_{2l} + V(\theta)^{-1}R(\theta)).$$

To apply the Neumann series (and consequently justify the existence of the inverse of $\mathrm{Id}_{2l} + V^{-1}R$ as well as the estimates for its size), we have to estimate the term $V^{-1}R$. According to Lemma 4.13, we have the estimate for L

$$\|L\|_{\rho-2\delta,\Gamma} \le C\kappa\delta^{-(\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}$$

for all $\delta \in (0, \rho/2)$. This leads to the estimate

$$\|V^{-1}R\|_{\rho-2\delta,\Gamma} \le C\kappa\delta^{-(\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}$$

for $0 < \delta < \rho/2$, where $C > 0$ depends on l, ν, $\|F\|_{C_\Gamma^1(B_r)}$, $\|DK\|_{\rho,\underline{c},\Gamma}$, $\|N\|_{\rho,\Gamma}$ and $\|\Pi_{K(\theta)}^c\|_{\rho,\underline{c},\Gamma}$. Because of assumption (51), we have that the right-hand side of the last equation is less than $1/2$.

Then the matrix $\mathrm{Id}_{2l} + V(\theta)^{-1}R(\theta)$ is invertible with

$$\|(\mathrm{Id}_{2l} + V^{-1}R)^{-1}\|_{\rho-2\delta,\Gamma} \le \frac{1}{1 - \|V^{-1}R\|_{\rho-2\delta,\Gamma}} \le 2.$$

This ends the proof of Lemma 4.14. ∎

4.3.5. *Identification of the center subspace.* In this section, we identify the center space as being very close (up to terms that can be bounded by the error) to the range of the matrix $\tilde{M}$. This will allow us to use the range of $\tilde{M}$ in place of $\mathcal{E}_{K(\theta)}^c$ without changing the quadratic character of the method.

Proposition 4.15. *Denote by $\Gamma_{K(\theta)}$ the range of $\tilde{M}(\theta)$ and by $\Pi_{K(\theta)}^\Gamma$ the projection onto $\Gamma_{K(\theta)}$ according to the splitting $\mathcal{E}_{K(\theta)}^s \oplus \Gamma_{K(\theta)} \oplus \mathcal{E}_{K(\theta)}^u$.*
Then there exists a constant $C > 0$ such that if

$$\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma} \le C$$

we have the estimate

$$(53) \qquad \|\Pi_{K(\theta)}^c - \Pi_{K(\theta)}^\Gamma\|_{\rho-2\delta,\underline{c},\Gamma} \le C\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}$$

for every $\delta \in (0, \rho/2)$ and where C, as usual, depends on the non-degeneracy constants of the problem.

Proof. From (48) and Cauchy estimates (see Lemma 10.8 in Appendix 10), we have:

$$\mathrm{dist}_{\rho-\delta,\underline{c},\Gamma}((DF_\lambda \circ K)\Gamma_{K(\theta)}, \Gamma_{K(\theta)} \circ T_\omega) \le C\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma},$$

where *dist* stands for the distance between two spaces at the Grassmannian level. Using again equation (48) and iterating it, we obtain for $n \ge 1$

$$DF_\lambda(K(\theta + n\omega)) \times \cdots \times DF_\lambda(K(\theta))\tilde{M}(\theta) =$$

$$\tilde{M}(\theta + n\omega)\mathcal{S}_\lambda(\theta + (n-1)\omega) \times \cdots \times \mathcal{S}_\lambda(\theta) + R_n,$$

where

$$\|R_n\|_{\rho-\delta,\underline{c},\Gamma} \leq C_n\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}$$

and C_n depends on n.

Since $\mathcal{S}_\lambda(\theta)$ is upper triangular with Id_l on the diagonal, we have:

$$\mathcal{S}_\lambda(\theta + (n-1)\omega) \times \cdots \times \mathcal{S}_\lambda(\theta) = \begin{pmatrix} \mathrm{Id}_l & A_\lambda(\theta + (n-1)\omega) + \cdots + A_\lambda(\theta) \\ 0 & \mathrm{Id}_l \end{pmatrix}.$$

Therefore, by induction, we have for every $n \in \mathbb{N}$

$$\|DF_\lambda(K(\theta + n\omega)) \cdots DF_\lambda(K(\theta))\tilde{M}(\theta)\|_{\rho-\delta,\underline{c},\Gamma} \leq Cn + C_n\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}.$$

Identical calculations give that

$$\|DF_\lambda^{-1}(K(\theta - n\omega)) \cdots DF_\lambda^{-1}(K(\theta))\tilde{M}(\theta)\|_{\rho-\delta,\underline{c},\Gamma} \leq Cn + C_n\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}.$$

Note that, given any $\mu_3 > 1$ (as in Definition 3.1), there exists an integer $n_{\mu_3} \geq 0$ such that for all $n \geq n_{\mu_3}$, we have $Cn < \mu_3^n$. Consequently, choosing such n_{μ_3} there exists a constant C such that if the error satisfies

$$\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma} \leq C,$$

we have $Cn + C_n\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma} < \mu_3^n$. In other words, the above estimates hold for all sufficiently large n, provided that we impose a suitable smallness condition on $\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}$.

As a consequence, $\Gamma_{K(\theta)}$ is an approximately invariant bundle, and we also have bounds on the rate of growth of the co-cycle both in positive and negative times. Using Proposition 4.2, this shows that indeed one can find a true invariant subspace $\tilde{\mathcal{E}}_{K(\theta)}$ close to $\Gamma_{K(\theta)}$. Since this invariant subspace should be of the same dimension of the center space $\mathcal{E}_{K(\theta)}^c$, we deduce that

$$\tilde{\mathcal{E}}_{K(\theta)} = \mathcal{E}_{K(\theta)}^c.$$

$\square$

4.3.6. *Estimates on the center subspace.* We recall the projection into the center subspace of the linearized equation

$$(54) \qquad \Pi_{K(\theta+\omega)}^c \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda + DF_\lambda(K(\theta))\Delta^c(\theta) - \Delta^c(\theta + \omega) = -E^c(\theta).$$

To shorten the notation till the end of the section we will write $\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda$ instead of $\Pi_{K(\theta+\omega)}^c \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda$.

We introduce the new function $W(\theta)$ through

$$(55) \qquad \Delta^c(\theta) = \tilde{M}(\theta)W(\theta) + \hat{e}(\theta)W(\theta),$$

where

$$(56) \qquad \hat{e} = \Pi_{K(\theta+\omega)}^c - \Pi_{K(\theta+\omega)}^\Gamma$$

which was estimated in Proposition 4.15.

Substituting (55) into equation (54) we get

$$(57) \qquad DF(K(\theta))\tilde{M}(\theta)W(\theta) - \tilde{M}(\theta+\omega)W(\theta+\omega)$$

$$= -E^c(\theta) - \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda + \hat{e}(\theta+\omega)W(\theta+\omega) - DF(K(\theta))\hat{e}(\theta)W(\theta)$$

We anticipate that the term $\hat{e}W$ will be quadratic in the error. Similarly, writing

$$\frac{\partial F_\lambda}{\partial \lambda}(K(\theta)) = \Pi^\Gamma_{K(\theta+\omega)}\frac{\partial F_\lambda}{\partial \lambda}(K(\theta)) + \hat{e}\frac{\partial F_\lambda}{\partial \lambda}(K(\theta)),$$

we also anticipate that the term $\hat{e}\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda$ will be quadratic in the error. As a consequence, we will ignore these two terms and the equation for W is

$$(58) \qquad DF_\lambda(K(\theta))\tilde{M}(\theta)W(\theta) - \tilde{M}(\theta+\omega)W(\theta+\omega) = -E^c(\theta) - \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda.$$

Multiplying equation (58) by $\tilde{M}(\theta+\omega)^\top \Pi^c_{K(\theta)} J_\infty(K(\theta+\omega))$, using Lemma 4.14 (giving the invertibility of $\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M}$) and equation (48), we end up with

$$(59) \qquad \left[\begin{pmatrix} \mathrm{Id}_l & A_\lambda(\theta) \\ 0_l & \mathrm{Id}_l \end{pmatrix} + B(\theta) \right] W(\theta) - W(\theta+\omega) = p_1(\theta) + p_2(\theta)$$

$$-[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M}](\theta+\omega)^{-1}[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta+\omega)\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda,$$

where

$$(60) \qquad B(\theta) = [\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M}](\theta+\omega)^{-1}[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta+\omega)e(\theta),$$

$$(61) \qquad p_1(\theta) = -V(\theta+\omega)^{-1}[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta+\omega)E^c(\theta)$$

and

$$(62) \qquad p_2(\theta) = -\tilde{V}(\theta+\omega)[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta+\omega)E^c(\theta).$$

The next result provides the estimates of the previously introduced quantities.

Lemma 4.16. *Assume $\omega \in D(\kappa,\nu)$ and δ and $\|E\|_{\rho,\underline{c},\Gamma}$ satisfy (51). Then using the linear change of variables (55), equation (58) becomes*

$$[\mathcal{S}_\lambda(\theta) + B(\theta)]\, W(\theta) - W(\theta+\omega) = p_1(\theta) + p_2(\theta)$$

$$- [\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M}](\theta+\omega)^{-1}[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta+\omega)\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda,$$

where B, p_1 and p_2 are given by equations (60), (61) and (62) respectively.

Moreover the following estimates hold: for p_1 we have

$$(63) \qquad \|p_1\|_{\rho,\underline{c},\Gamma} \leq C\|E\|_{\rho,\underline{c},\Gamma},$$

where C only depends on $\|J(K)\|_{\rho,\underline{c},\Gamma}$, $\|N\|_\rho$, $\|DK\|_{\rho,\underline{c},\Gamma}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\underline{c},\Gamma}$. For p_2 and B, we have

$$(64) \qquad \|p_2\|_{\rho-2\delta,\underline{c},\Gamma} \leq C\kappa\delta^{-(\nu+1)}\|E\|^2_{\rho,\underline{c},\Gamma}$$

and

$$(65) \qquad \|B\|_{\rho-2\delta,\Gamma} \leq C\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma},$$

where C depends l, ν, ρ, R, $\|N\|_{\rho,\Gamma}$, $\|DK\|_{\rho,\underline{c},\Gamma}$, $\|F\|_{C^1_\Gamma(B_r)}$, $\|J_\infty\|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\underline{c},\Gamma}$.

Proof. We have basically to estimate $\tilde{M}(\theta+\omega)^\top J_\infty(K(\theta))E^c(\theta)$ and $\tilde{M}(\theta+\omega)^\top J_\infty(K)e(\theta)$ and then use Lemma 4.14. First we bound

$$
\begin{aligned}
|E_i^c(\theta)| &\leq \sum_{j\in\mathbb{Z}^N} |(\Pi_{K(\theta+\omega)}^c)_{ij}| \, |E_j(\theta)| \\
&\leq \sum_{j\in\mathbb{Z}^N} \|\Pi_{K(\theta+\omega)}^c\|_{\rho,\Gamma}\, \Gamma(i-j)\, \|E\|_{\rho,\underline{c},\Gamma}\, \max_k\, \Gamma(j-c_k) \\
&\leq \|\Pi_{K(\theta+\omega)}^c\|_{\rho,\Gamma}\, \|E\|_{\rho,\underline{c},\Gamma}\, \sum_{k=1}^R \Gamma(i-c_k)\,.
\end{aligned}
$$

For $1\leq i\leq l$ we have, taking into account that J_∞ is uncoupled,

$$
(66) \qquad \left|\left(\tilde{M}(\theta+\omega)^\top J_\infty(K)E^c(\theta)\right)_i\right| \leq C \sum_{j\in\mathbb{Z}^N} |\partial_{\theta_i} K_j(\theta+\omega)|\, |E_j^c(\theta)|.
$$

We estimate from above by

$$
\sum_{j\in\mathbb{Z}^N} |D_{\theta_i} K_j(\theta+\omega)|\, |E_j^c(\theta)|
$$

$$
\begin{aligned}
(67) \qquad &\leq \sum_j \|DK\|_{\rho,\underline{c},\Gamma}\, \max_m\, \Gamma(j-c_m) \|\Pi_{K(\theta+\omega)}^c\|_{\rho,\Gamma} \|E\|_{\rho,\underline{c},\Gamma} \sum_{k=1}^R \Gamma(j-c_k) \\
&\leq R\, \|DK\|_{\rho,\underline{c},\Gamma}\, \|\Pi_{K(\theta+\omega)}^c\|_{\rho,\Gamma}\, \|E\|_{\rho,\underline{c},\Gamma}\,.
\end{aligned}
$$

For $l+1\leq i\leq 2l$, one gets

$$
(68) \qquad \left|\left(\tilde{M}(\theta+\omega)^\top J_\infty(K(\theta))\, E^c(\theta)\right)_i\right| \leq
$$
$$
C\left|\left(N(\theta+\omega)^\top DK(\theta+\omega)^\top \tilde{J}\left(K(\theta+\omega)^\top\right) E^c(\theta)\right)_i\right|.
$$

We get a similar bound for (68) taking into account that N is a bounded finite-dimensional matrix. Now the bounds (63) and (64) follow immediately from Lemma 4.14. For the estimate on B we use Cauchy estimates for $e_1(\theta)$. From

$$
B(\theta) = (V(\theta+\omega)^{-1} + \tilde{V}(\theta+\omega))\tilde{M}(\theta+\omega)^\top e(\theta)
$$

we have

$$
\begin{aligned}
\|B\|_{\rho-2\delta} &\leq \|V(\theta+\omega)^{-1}\|_{\rho-2\delta}\|\tilde{M}(\theta+\omega)^\top e(\theta)\|_{\rho-2\delta} \\
&\quad + \|\tilde{V}(\theta+\omega)\tilde{M}(\theta+\omega)^\top e(\theta)\|_{\rho-2\delta}
\end{aligned}
$$

and using estimate (52) we end up with

$$
\|B\|_{\rho-2\delta,\Gamma} \leq C\delta^{-1}\|E\|_{\rho,\underline{c},\Gamma} + \kappa\delta^{-(\nu+1)}\|E\|_{\rho,\underline{c},\Gamma}\, \delta^{-1}\|E\|_{\rho,\underline{c},\Gamma}.
$$

This gives the desired result. $\qquad\qquad\square$

4.3.7. *Approximate solvability of the linearized equation on the center subspace.* In this section we find a solution of equation (59) up to quadratic error. The convergence of the Newton scheme is of course not affected (see [Zeh75]).

For that we introduce the following operator

$$
\mathcal{L}W(\theta) = \begin{pmatrix} \mathrm{Id}_l & A_\lambda(\theta) \\ 0_l & \mathrm{Id}_l \end{pmatrix} W(\theta) - W(\theta+\omega).
$$

Then equation (63) can be written as

(69)

$$\mathcal{L}W(\theta) + B(\theta)W(\theta) = p_1(\theta) + p_2(\theta)$$
$$- [\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M}](\theta + \omega)^{-1}[\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta + \omega)\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda,$$

We will reduce equation (69) to two small divisors equations. Generically their right-hand sides will not have zero average, but we will use the freedom in choosing Λ and in fixing the average of the solution to solve one after the other. By Lemma 4.14 we can write

$$[(\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)\tilde{M})^{-1}\tilde{M}^\top \Pi^c_{K(\theta)} J_\infty(K)](\theta + \omega)\frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\Lambda = Q(\theta)\Lambda + q(\theta)\Lambda,$$

where the $2l \times l$ matrix Q is

$$Q(\theta) = V(\theta + \omega)^{-1}\tilde{M}(\theta + \omega)^\top \Pi^c_{K(\theta)} J_\infty(K(\theta + \omega))\frac{\partial F_\lambda(K(\theta))}{\partial \lambda}$$

and q satisfies for all $\delta \in (0, \rho/2)$

$$\|q\|_{\rho-2\delta,\Gamma} \leq C\kappa\delta^{-(\nu+1)}\left\|\frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\right\|_{\rho,\underline{c},\Gamma}\|E\|_{\rho,\underline{c},\Gamma},$$

where the constant C depends on l, ν, ρ, R, $\|N\|_{\rho,\Gamma}$, $\|DK\|_{\rho,\underline{c},\Gamma}$, $\|F\|_{C^1_\Gamma(B_r)}$, $\|J_\infty\|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\Gamma}$.

We will take as an approximate solution the solution v of

(70) $$\mathcal{L}v(\theta) = p_1(\theta) - Q(\theta)\Lambda$$

obtained from (69) by removing the terms containing B, p_2 and q.

Proposition 4.17. *Assume $\omega \in D(\kappa, \nu)$ and (λ, K) is a non-degenerate pair (i.e. $(\lambda, K) \in ND_{loc}(\rho, \Gamma)$). If the error $\|E\|_{\rho,\underline{c},\Gamma}$ satisfies (51), there exist a mapping $v \in \mathcal{A}_{\rho-2\delta,\underline{c},\Gamma}$ for any $0 < \delta < \rho/2$ and a vector $\Lambda \in \mathbb{R}^l$ solving equation (70).*

Moreover there exists a constant $C > 0$ depending on $\nu, \rho, l, R, \|K\|_{\rho,\underline{c},\Gamma}, |\text{avg}(Q_\lambda)|^{-1}$, $|\text{avg}(A_\lambda)|^{-1}, \|N\|_{\rho,\Gamma}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\Gamma}$ such that

(71) $$\|v\|_{\rho-2\delta,,\underline{c},\Gamma} < C\kappa^2\delta^{-2\nu}\|E\|_{\rho,\underline{c},\Gamma}$$

and

$$|\Lambda| < C\|E\|_{\rho,\underline{c},\Gamma}.$$

Proof. We denote $T(\theta)$ the right-hand side of equation (70), i.e. we have to solve

(72) $$\mathcal{L}v(\theta) = T(\theta),$$

with

$$T = p_1 - Q\Lambda.$$

We now decompose equation (72) into two equations. Writing $v = (v_1, v_2)^\top$, $T(\theta) = (T_1(\theta), T_2(\theta))^\top$ equation (72) is equivalent to

(73) $$v_1(\theta) + A_\lambda(\theta)v_2(\theta) = v_1(\theta + \omega) + T_1(\theta),$$

(74) $$v_2(\theta) = v_2(\theta + \omega) + T_2(\theta).$$

A simple computation shows that

$$T_2(\theta) = -[DK^\top \Pi^c_{K(\theta)} J_\infty(K)] \circ T_\omega \circ (E^c_2 + \frac{\partial F_\lambda(K(\theta))}{\partial \lambda})$$

We begin by solving equation (74). To apply Proposition 4.4 we choose $\Lambda \in \mathbb{R}^l$ such that
$$\text{avg}\,(T_2) = 0.$$

This condition is equivalent to
$$\text{avg}\left(DK^\top(\omega + \theta)\Pi^c_{K(\theta)}J_\infty(K(\omega + \theta))(E^c(\theta) + \frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\Lambda)\right) = 0.$$

This leads to
$$\text{avg}\left((DK^\top(\omega + \theta)\Pi^c_{K(\theta)}J_\infty(K(\omega + \theta))\frac{\partial F_\lambda(K(\theta))}{\partial \lambda}\right)\Lambda$$
$$= -\text{avg}\left((DK^\top(\omega + \theta)\Pi^c_{K(\theta)}J_\infty(K(\omega + \theta))E^c(\theta)\right).$$

Note that the matrix which applies to Λ is the average of Q_λ which, by hypothesis, is invertible. This gives $|\Lambda| < C\,\|E\|_{\rho,\underline{c},\Gamma}$.

From the expression of T and the value of Λ obtained above, we have that there exists a constant C such that
$$\|T_i\|_{\rho,\underline{c},\Gamma} \le C\|E\|_{\rho,\underline{c},\Gamma},$$

for $i = 1, 2$.

Then Proposition 4.4 provides us with an analytic solution v_2 on $D_{\rho-\delta}$ with arbitrary average and

(75) $$\|v_2\|_{\rho-\delta,\underline{c},\Gamma} \le C\kappa\delta^{-\nu}\|T_2\|_{\rho,\underline{c},\Gamma} + |\text{avg}\,(v_2)|.$$

Now we come to equation (74). To apply Proposition 4.4 we choose $\text{avg}\,(v_2)$ such that $\text{avg}\,(R_1 - A_\lambda v_2) = 0$. This condition is equivalent to
$$\text{avg}\,(v_2) = \text{avg}\,(A_\lambda)^{-1}(\text{avg}\,(T_1) - \text{avg}\,(A_\lambda v_2^\perp)),$$

where $v_2 = v_2^\perp + \text{avg}\,(v_2)$. This is possible since by the twist condition $\text{avg}\,(A_\lambda)$ is invertible.

We have that
$$|\text{avg}\,(v_2)| \le C\kappa\delta^{-\nu}\,\|E\|_{\rho,\underline{c},\Gamma}.$$

Then we take v_1 as the unique analytic solution of (74) with zero average. Furthermore, we have the estimate
$$\|v_1\|_{\rho-2\delta,\underline{c},\Gamma} \le C\kappa\delta^{-\nu}\|T_1 - A_\lambda v_2\|_{\rho-\delta,\underline{c},\Gamma}.$$

Collecting the previous bounds we get the result. $\qquad\qquad\square$

We now come back to the solutions of (38). The above procedure allows us to prove the following proposition, providing an approximate solution of the projection of $D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E$ on the center subspace.

Proposition 4.18. *Let (Λ, W) be as in Proposition 4.17 and assume the hypotheses of that proposition hold. Define $\Delta^c(\theta) = \tilde{M}(\theta)W(\theta)$. Then, equation (38) is approximately solvable and we have the following estimates*

(76) $$\|\Delta^c\|_{\rho-2\delta,\underline{c},\Gamma} \le C\kappa^2\delta^{-2\nu}\|E\|_{\rho,\underline{c},\Gamma},$$

$$|\Lambda| \le C\|E\|_{\rho,\underline{c},\Gamma},$$

where the constant C depends on $\nu, \rho, l, R, |\text{avg}\,(Q_\lambda)|^{-1}, |\text{avg}\,(A_\lambda)|^{-1}, \|N\|_\rho, \|\frac{\partial F_\lambda(K)}{\partial\lambda}\|_{\rho,\underline{c},\Gamma}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\Gamma}$ and

(77) $$\|D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta^c) + E^c\|_{\rho-2\delta,\underline{c},\Gamma} \le C\kappa^2\delta^{-(2\nu+1)}(\|E\|^2_{\rho,\underline{c},\Gamma} + \|E\|_{\rho,\underline{c},\Gamma}|\Lambda|),$$

where the constant C depends on l, ν, ρ, R, $\|F\|_{C^1_\Gamma(B_r)}$, $\|DK\|_{\rho,\underline{c},\Gamma}$, $\|N\|_\rho$, $|\mathrm{avg}\,(A_\lambda)|^{-1}$, $|\mathrm{avg}\,(Q_\lambda)|^{-1}$ and $\|\frac{\partial F_\lambda(K)}{\partial \lambda}\|_{\rho,\underline{c},\Gamma}$.

Proof. For the first estimate we take $\theta \in D_{\rho-2\delta}$ and write $W = (W_1, W_2)$. Then

$$\Delta^c(\theta) = \Pi^c_{K(\theta)}DK(\theta)W_1(\theta) + \Pi^c_{K(\theta)}J^{-1}_\infty(K(\theta))DK(\theta)N(\theta)W_2(\theta).$$

We have

$$\left|(DK(\theta)W_1(\theta))_i\right| = \left|\sum_{j=1}^l D_{\theta_j}K_i(\theta)W_{1,j}(\theta)\right|$$
$$\leq \|DK\|_{\rho,\underline{c},\Gamma}\,\max_k \Gamma(i - c_k)\|W_1\|_{\rho-2\delta}.$$

Also, since J_∞ is uncoupled and N is finite-dimensional,

$$\left|\left(J^{-1}_\infty(K(\theta))DK(\theta)N(\theta)W_2(\theta)\right)_i\right| \leq \|J(K)\|_\rho\left|\left(DK(\theta)N(\theta)W_2(\theta)\right)_i\right|$$
$$\leq \|J(K)\|_\rho\|N\|_\rho\|DK\|_{\rho,\underline{c},\Gamma}\,\max_k \Gamma(i - c_k)\|W_2\|_{\rho-2\delta}.$$

Now using (71), we obtain (76).

For (77), using the previous notations and Lemma 4.16,

(78)
$$D_{\lambda,k}\,\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta_c)(\theta) + E^c(\theta)$$
$$= \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda + \tilde{M}(\theta + \omega)\big[\mathcal{S}_\lambda(\theta)W(\theta) - W(\theta + \omega)\big]$$
$$+ e(\theta)W(\theta) + E^c(\theta) = \frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda$$
$$+ \tilde{M}(\theta + \omega)\big[p_1(\theta) - Q(\theta)\Lambda\big] + e(\theta)W(\theta) + E^c(\theta).$$

Note that

$\tilde{M}(\theta + \omega)p_1(\theta) =$
$\tilde{M}(\theta + \omega)\big[-([\tilde{M}^\top \Pi^c_{K(\theta)}J_\infty(K)\tilde{M})^{-1}](\theta + \omega)[\tilde{M}^\top \Pi^c_{K(\theta)}J_\infty(K)](\theta + \omega)E^c(\theta) - p_2(\theta)\big]$

and also that $\tilde{M}(\tilde{M}^\top \Pi^c_{K(\theta)}J_\infty(K)\tilde{M})^{-1}\tilde{M}^\top J_\infty(K)$ is symmetric and

$$\tilde{M}\left\{\tilde{M}^\top \Pi^c_{K(\theta)}J_\infty(K)\tilde{M})^{-1}\tilde{M}^\top J_\infty(K) - \mathrm{Id}\right\}$$

maps the vectors of the center subspace to zero because it is generated by the columns of $\tilde{M}$.

Also we have

$\tilde{M}(\theta + \omega)Q(\theta)\Lambda =$
$\tilde{M}(\theta + \omega)\left[[\tilde{M}^\top J_\infty(K)\tilde{M})^{-1}\tilde{M}^\top \Pi^c_{K(\theta)}J_\infty(K)](\theta + \omega)\frac{\partial F_\lambda}{\partial \lambda}(K(\theta))\Lambda - q(\theta)\Lambda\right].$

We recall that here the derivative of F with respect to λ actually means the projection of it into the center subspace.

Therefore (78) becomes

$$\tilde{M}(\theta + \omega)[q(\theta)\Lambda - p_2(\theta)] + e(\theta)W(\theta)$$

and (77) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.4. **Solution of the equation in the hyperbolic subspaces.** In this section, we study the projection of the Newton equations (27) on the the hyperbolic spaces.

According to the splitting (12), there exist projections on the linear spaces $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$. The analytic regularity of the splitting implies the analytic dependence of these projections in θ. We denote $\Pi^s_{K(\theta+\omega)}$ (resp. $\Pi^u_{K(\theta+\omega)}$) the projections on the stable (resp. unstable) invariant subspace.

We project equation (27) on the stable and unstable spaces to obtain

$$(79) \quad \Pi^s_{K(\theta+\omega)}\Big(\frac{\partial F(K(\theta))}{\partial\lambda}\Lambda + DF_\lambda(K(\theta))\Delta(\theta) - \Delta(\theta+\omega)\Big) = -\Pi^s_{K(\theta+\omega)}E(\theta),$$

$$(80) \quad \Pi^u_{K(\theta+\omega)}\Big(\frac{\partial F(K(\theta))}{\partial\lambda}\Lambda + DF_\lambda(K(\theta))\Delta(\theta) - \Delta(\theta+\omega)\Big) = -\Pi^u_{K(\theta+\omega)}E(\theta).$$

The invariance of the splitting reads

$$\Pi^s_{K(\theta+\omega)}DF_\lambda(K(\theta))\Delta(\theta) = DF_\lambda(K(\theta))\Pi^s_{K(\theta)}\Delta(\theta)$$

for the stable part and

$$\Pi^u_{K(\theta+\omega)}DF_\lambda(K(\theta))\Delta(\theta) = DF_\lambda(K(\theta))\Pi^u_{K(\theta)}\Delta(\theta)$$

for the unstable one.

We define

$$\Delta^{s,u}(\theta) = \Pi^{s,u}_{K(\theta)}\Delta(\theta).$$

Using the notation $\theta' = \theta + \omega$, equations (79)-(80) become

$$(81) \qquad DF_\lambda(K) \circ T_{-\omega}(\theta')\Delta^s(T_{-\omega}(\theta')) - \Delta^s(\theta') = -\tilde{E}^s(\theta',\Lambda),$$

where

$$\tilde{E}^s(\theta',\Lambda) = \Pi^s_{K(\theta')}\Big(\frac{\partial F(K(T_{-\omega}(\theta')))}{\partial\lambda}\Lambda\Big) + \Pi^s_{K(\theta')}E \circ T_{-\omega}(\theta')$$

and

$$(82) \qquad DF_\lambda(K) \circ T_{-\omega}(\theta')\Delta^u(T_{-\omega}(\theta')) - \Delta^u(\theta') = -\tilde{E}^u(\theta',\Lambda),$$

where

$$\tilde{E}^u(\theta',\Lambda) = \Pi^u_{K(\theta')}\Big(\frac{\partial F(K(T_{-\omega}(\theta')))}{\partial\lambda}\Lambda\Big) + \Pi^u_{K(\theta')}E \circ T_{-\omega}(\theta').$$

Contrary to the projections on the center subspace, the projections on the hyperbolic subspaces can be solved exactly. The key point for the estimates is the Banach algebra property of the decay functions.

Proposition 4.19. *Fix $\rho > 0$ then equation (81) (resp. (82)) admits a unique analytic solution $\Delta^s : D_\rho \to \mathcal{E}^s_{K(\theta)}$ (resp. $\Delta^u : D_\rho \to \mathcal{E}^u_{K(\theta)}$). Furthermore there exists a constant C depending only on the hyperbolicity constant μ_1 (resp. μ_2), the norm of the projector $\|\Pi^s_{K(\theta)}\|_{\rho,\Gamma}$ (resp. $\|\Pi^u_{K(\theta)}\|_{\rho,\Gamma}$) and $\|\frac{\partial F_\lambda(K)}{\partial\lambda}\|_{\rho,\underline{c},\Gamma}$ such that*

$$(83) \qquad \|\Delta^{s,u}\|_{\rho,\underline{c},\Gamma} \leq C(\|E\|_{\rho,\underline{c},\Gamma} + |\Lambda|)$$

Proof. We only give the proof for the stable case, the unstable case being very similar. Using equation (81), we claim

$$(84) \quad \Delta^s(\theta') = \sum_{k=0}^{\infty}(DF_\lambda(K) \circ T_{-\omega}(\theta') \times \cdots \times DF_\lambda(K) \circ T_{-k\omega}(\theta'))\tilde{E}^s(T_{-k\omega}(\theta'),\Lambda).$$

We introduce the map

$$\mathcal{F}^{co}(k, \theta') = DF_\lambda(K) \circ T_{-\omega}(\theta') \times \cdots \times DF_\lambda(K) \circ T_{-k\omega}(\theta').$$

From the definition of $\tilde{E}^s$ we have $\|\tilde{E}^s(T_{-k\omega}(\theta'), \Lambda)\|_{\rho,\underline{c},\Gamma} \leq C(\|E\|_{\rho,\underline{c},\Gamma} + |\Lambda|)$. Using (13) we obtain that the k term in (84) is bounded by

$$\|\mathcal{F}^{co}(k, \theta')\tilde{E}^s(T_{-k\omega}(\theta'), \Lambda)\|_{\rho,\underline{c},\Gamma} \leq C_h \mu_1^k \|\tilde{E}^s(T_{-k\omega}(\cdot), \Lambda)\|_{\rho,\underline{c},\Gamma}$$

and therefore

$$\|\Delta^s\|_{\rho,\underline{c},\Gamma} \leq C_h \|\tilde{E}^s\|_{\rho,\underline{c},\Gamma} \sum_{k=0}^\infty \mu_1^k \leq C(\|E\|_{\rho,\underline{c},\Gamma} + |\Lambda|),$$

since $\mu_1 < 1$.   □

Remark 4.20. It is perhaps interesting to compare the method of proof of this paper with that of [FdlLS09a]. Both papers use the invariance equations and formulate a quasi-Newton method that can be solved using techniques from hyperbolic lore and some geometric identities. One of the strengths of the set-up based on decay functions is that we can obtain estimates independent on the number and positions of the centers of activity by methods that resemble the finite dimensional methods.

Both [FdlLS09a] and the present paper use a counterterm to adjust some of the constants and then prove a vanishing lemma. In [FdlLS09a], the counterterm is obtained adding $J^{-1} \circ K_0 DK_0 \lambda$. In this paper, we consider a family of symplectic maps. This allows us to use the vanishing lemma only once at the end of the proof, whereas in [FdlLS09a], the vanishing lemma had to be used at each iterative step.

We also deal in a different way with the invertibility of the linear change of variables M. In the present paper, we obtain the invertibility on the range using some geometric identities, whereas in [FdlLS09a], we used some easier argument based on finite dimensional arguments.

One geometric aspect that required several changes (due in part to the changes in the counterterm and to the infinite dimensional character) is the estimates on the difference between the center space and the range of the change of variables M.

5. Iteration of the modified Newton method, convergence and proof of Theorem 3.5

This section is devoted to the iteration of the Newton method. We derive first the usual KAM estimates for convergence. We assume that we are under the assumptions of Theorem 3.5. We note that with the decay norms the estimates are very similar to the ones we obtained in the finite-dimensional case (see [FdlLS09a]).

5.1. **Iteration of the method.** Let (λ_0, K_0) be an approximate solution of (11) (i.e. a solution of the linearized equation with error E_0). Following the Newton scheme we define the following sequence of approximate solutions

$$K_m = K_{m-1} + \Delta K_{m-1}, \qquad m \geq 1,$$
$$\lambda_m = \lambda_{m-1} + \Lambda_{m-1}, \qquad m \geq 1,$$

where $(\lambda_{m-1}, \Delta K_{m-1})$ is a solution of

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\Lambda_{m-1}, \Delta K_{m-1}) = -E_{m-1}$$

with $E_{m-1}(\theta) = \mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\theta)$. The next lemma states a classical result in KAM theory: the approximation to the solution at step m has an error which is bounded in a smaller complex domain by the square of the norm of the error at step $m-1$.

Proposition 5.1. *Assume $(\lambda_{m-1}, K_{m-1}) \in ND_{loc}(\rho_{m-1}, \Gamma)$ is an approximate solution of equation (11) and that the following holds*

$$r_{m-1} = \|K_{m-1} - K_0\|_{\rho_{m-1}, \underline{c}, \Gamma} < r.$$

If E_{m-1} is small enough such that lemma 4.18 applies then there exists a function $\Delta K_{m-1} \in \mathcal{A}_{\rho_{m-1}-3\delta_{m-1}, \underline{c}, \Gamma}$ for any $0 < \delta_{m-1} < \rho_{m-1}/3$ and a vector $\Lambda_m \in \mathbb{R}^l$ such that

(85) $$\|\Delta K_{m-1}\|_{\rho_{m-1}-2\delta_{m-1}, \underline{c}, \Gamma} \leq (C^1_{m-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-2\nu})\|E_{m-1}\|_{\rho_{m-1}, \underline{c}, \Gamma},$$

(86) $$|\Lambda_m| \leq C\|E_{m-1}\|_{\rho_{m-1}, \underline{c}, \Gamma}$$

(87)
$$\|D\Delta K_{m-1}\|_{\rho_{m-1}-3\delta_{m-1}, \underline{c}, \Gamma} \leq (C^1_{m-1}\delta_{m-1}^{-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)})\|E_{m-1}\|_{\rho_{m-1}, \underline{c}, \Gamma},$$

where C^1_{m-1}, C^2_{m-1} depend only on ν, l, $|F_\lambda|_{C^1_\Gamma(B_r)}$, $\|DK_{m-1}\|_{\rho_{m-1}, \underline{c}, \Gamma}$, $\|\Pi^c_{K_{m-1}(\theta)}\|_{\rho_{m-1}, \Gamma}$, $\|\Pi^s_{K_{m-1}(\theta)}\|_{\rho_{m-1}, \Gamma}$, $\|\Pi^u_{K_{m-1}(\theta)}\|_{\rho_{m-1}, \Gamma}$, $|\text{avg}(Q_{\lambda_{m-1}})|^{-1}$ and $|\text{avg}(A_{\lambda_{m-1}})|^{-1}$. Moreover, if $K_m = K_{m-1} + \Delta K_{m-1}$ and

$$r_{m-1} + C^1_{m-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-2\nu} < r$$

then we can redefine C^1_{m-1} and C^2_{m-1} and all previous quantities such that the error $E_m(\theta) = \mathcal{F}_\omega(\lambda_m, K_m)(\theta)$ satisfies

(88) $$\|E_m\|_{\rho_m, \underline{c}, \Gamma} \leq C_{m-1}\kappa^4\delta_{m-1}^{-4\nu}\|E_{m-1}\|^2_{\rho_{m-1}, \underline{c}, \Gamma},$$

where we take $\rho_m = \rho_{m-1} - 3\delta_{m-1}$.

Remark 5.2. The estimate (88) showing that the norm of the error at step m is essentially bounded by the square of the norm of the error at step $m-1$ was already in [Zeh75, Zeh76a].

Proof. Taking into account that $\Delta K_{m-1}(\theta)$ is the sum of its three projections on the stable, center and unstable subspaces estimates (85) and (86) follow from Proposition 4.18 and Proposition 4.19. Estimate (87) follows from estimate (85) and Cauchy's inequalities. Define the remainder of the Taylor expansion

$$\mathcal{R}(\lambda, \lambda', K, K') = \mathcal{F}_\omega(\lambda, K) - \mathcal{F}_\omega(\lambda', K')$$
$$- D_{\lambda, K}\mathcal{F}_\omega(\lambda', K')(\lambda - \lambda', K - K').$$

Then putting $\lambda = \lambda_m$ and $\lambda' = \lambda_{m-1}$ we have

$$E_m(\theta) = E_{m-1}(\theta) + D_{\lambda, K}\mathcal{F}_\omega(\lambda_{m-1}, K_{m-1}(\theta))(\Lambda_{m-1}, \Delta K_{m-1}(\theta))$$
$$+ \mathcal{R}(\lambda_{m-1}, \lambda_m, K_{m-1}, K_m)(\theta).$$

According to estimate (77) and since the equations on the hyperbolic subspace are *exactly* solved, we have

$$\|E_{m-1} + D_{\lambda, K}\mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\Lambda_{m-1}, \Delta K_{m-1})\|_{\rho_m, \underline{c}, \Gamma}$$
$$\leq c_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|^2_{\rho_{m-1}, \underline{c}, \Gamma}.$$

Estimate (88) then follows from Taylor's remainder. $\qquad\square$

In the following, we derive the changes in the non-degeneracy conditions during the iterative step.

For the twist condition, we have the following lemma, which is prove easily noting that we are just perturbing finite dimensional matrices.

Lemma 5.3. *Assume that the hypothesis of Proposition 5.1 hold. If* $\|E_{m-1}\|_{\rho_{m-1}}$ *is small enough, then*

- *If* $DK_{m-1}^\top DK_{m-1}$ *is invertible with inverse* N_{m-1}
  then $DK_m^\top DK_m$ *is invertible with inverse* N_m *and we have*

$$\|N_m\|_{\rho_m} \leq \|N_{m-1}\|_{\rho_{m-1}} + C_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1},\underline{c},\Gamma}.$$

- *If* $\mathrm{avg}\,(A_{\lambda_{m-1}})$ *is non singular then,* $\mathrm{avg}\,(A_{\lambda_m})$ *is non-singular and we have the estimate*

$$|\mathrm{avg}\,(A_{\lambda_m})|^{-1} \leq |\mathrm{avg}\,(A_{\lambda_{m-1}})|^{-1} + C'_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1},\underline{c},\Gamma}.$$

- *If* $\mathrm{avg}\,(Q_{\lambda_{m-1}})$ *is non singular then,* $\mathrm{avg}\,(Q_{\lambda_m})$ *is non-singular and we have the estimate*

$$|\mathrm{avg}\,(Q_{\lambda_m})|^{-1} \leq |\mathrm{avg}\,(Q_{\lambda_{m-1}})|^{-1} + C''_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1},\underline{c},\Gamma}.$$

5.2. **Iteration of the Newton step and convergence.** Once we have the estimates for a step of the iterative method, following the standard scheme in KAM theory one can prove the convergence of the method. One takes $0 < \delta_0 < \min(1, \rho_0/12)$, $\delta_m = \delta_0/2$ and $\rho_m = \rho_{m-1} - 3\delta_{m-1}$.

From (88) we have that

$$(89) \qquad \|E_m\|_{\rho_m,\underline{c},\Gamma} \leq C_{m-1}\kappa^4\delta_0^{-4\nu}2^{4\nu(m-1)}\|E_{m-1}\|_{\rho_{m-1},\underline{c},\Gamma}^2.$$

Moreover the constants C_m are bounded uniformly in m by a constant C. Using (89) iteratively one obtains

$$\|E_m\|_{\rho_m,\underline{c},\Gamma} \leq \left(C\kappa^4\delta_0^{-4\nu}2^{4\nu}\|E_0\|_{\rho_0,\underline{c},\Gamma}\right)^{2^m}.$$

Then, if $\|E_0\|_{\rho_0,\underline{c},\Gamma}$ is small enough the iteration converges to a pair $(\lambda_\infty, K_\infty) \in ND_{loc}(\rho_\infty, \Gamma)$, with $K_\infty \in \mathcal{A}_{\rho_\infty,\underline{c},\Gamma}$ such that $F_{\lambda_\infty} \circ K_\infty = K_\infty \circ T_w$.

6. Vanishing lemma and existence of invariant tori. Proof of Theorem 3.6

In this context of infinite-dimensional lattices, one could ask for the existence of a vanishing lemma, which would ensure that the translated tori are actually invariant ones. The issue is that we do not have a true symplectic form on the whole manifold $M^{\mathbb{Z}^N}$ but just a formal one. The proof we give is inspired by the one in [FdlLS09a] but we have to take into account that formal forms make sense only via their pull-backs to $\mathbb{T}^l$.

The following lemma, called the vanishing lemma, is more or less equivalent to showing that some averages cancel, but is somewhat easier to implement.

Lemma 6.1. *Assume* F_λ *is analytic smooth in* $\lambda \in \mathbb{R}^l$ *and maps* $\mathcal{M}$ *into itself. Assume* $\omega \in D(\kappa, \nu)$ *and let* $(\lambda, K) \in ND_{loc}(\rho, \Gamma)$, *where* $K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ *is a solution of*

$$F_\lambda \circ K = K \circ T_\omega.$$

Assume furthermore:

- F_0 is exact symplectic, F_λ is symplectic for $\lambda \neq 0$ and F_λ is constructed as in Appendix 12.
- F_λ extends analytically to a neighborhood of $K(\mathbb{T}^l)$,
- We have $|\lambda| \leq \lambda^*$, where λ^* depends only on derivatives of F and the symplectic structure J.

Then

$$\lambda = 0.$$

Proof. We will write equation (11) as

$$(90) \qquad F_0 \circ K = R_\lambda \circ K \circ T_\omega,$$

where $R_\lambda = F_0 \circ F_\lambda^{-1}$. We denote

$$(91) \qquad \hat{\theta}_i = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_l) \in \mathbb{T}^{l-1}$$

and similarly $\hat{\omega}_i = (\omega_1, \ldots, \omega_{i-1}, \omega_{i+1}, \ldots, \omega_l) \in \mathbb{R}^{l-1}$. We also denote $\sigma_{i,\hat{\theta}_i} : \mathbb{T} \to \mathbb{T}^l$ the path given by

$$(92) \qquad \sigma_{i,\hat{\theta}_i}(\eta) = (\theta_1, \ldots, \theta_{i-1}, \eta, \theta_{i+1}, \ldots, \theta_l).$$

We will compute the integral $\int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* F_0^* \alpha_\infty$ in two different ways. Note that the quantity $\int_{\sigma_{i,\hat{\theta}_i}} K^* F_0^* \alpha_\infty$ is well-defined since K has decay on $\mathcal{M}$.

Using that F_0 is exact symplectic we have:

$$(93) \qquad \begin{aligned} \int_{\sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} K^* F_0^* \alpha_\infty &= \int_{\sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} (K^* \alpha_\infty + dW_K) \\ &= \int_{\sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} K^* \alpha_\infty = \int_{K \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} \alpha_\infty. \end{aligned}$$

Similarly, we have

$$(94) \qquad \begin{aligned} \int_{\sigma_{i,\hat{\theta}_i}} (R_\lambda \circ K \circ T_\omega)^* \alpha_\infty &= \int_{\sigma_{i,\hat{\theta}_i}} T_\omega^* (R_\lambda \circ K)^* \alpha_\infty \\ &= \int_{\sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} (R_\lambda \circ K)^* \alpha_\infty. \end{aligned}$$

Since the averages over $\mathbb{T}^{l-1}$ are the same, one gets

$$(95) \qquad \begin{aligned} 0 &= \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} [K^* \alpha_\infty - (R_\lambda \circ K)^* \alpha_\infty] \\ &= -\int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* (R_\lambda^* \alpha_\infty - \alpha_\infty). \end{aligned}$$

Now, we estimate $K^* \left(R_\lambda^* \alpha_\infty - \alpha_\infty \right)$. We know that

$$R_\lambda = F_0 \circ F_\lambda^{-1}.$$

Therefore, we have

$$K^* \left(R_\lambda^* \alpha_\infty - \alpha_\infty \right) = K^* \left((F_\lambda^{-1})^* (F_0^* \alpha_\infty - F_\lambda^* \alpha_\infty) \right).$$

This gives, using the exact symplecticness of F_0

$$0 = \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^*(F_\lambda^{-1})^*(\alpha_\infty - F_\lambda^* \alpha_\infty).$$

By the construction of the map F_λ, we have

$$(F_\lambda^{-1})^*(\alpha_\infty - F_\lambda^* \alpha_\infty) = \frac{1}{2}(F_\lambda^{-1})^* \Big\{ \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k^j F_0^* \delta_k^j + d\beta \Big\},$$

where β is a smooth function on the torus and $(\delta_k^j)_{k=1,\dots,l}$ is a basis of $H^1(\mathbb{T}^l)$. This gives

$$(F_\lambda^{-1})^*(\alpha_\infty - F_\lambda^* \alpha_\infty) = \frac{1}{2} \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k^j R_\lambda^* \delta_k^j + d\bar{\beta}.$$

Therefore, one gets

$$\frac{1}{2} \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k^j \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* R_\lambda^* \delta_k^j = 0.$$

By the smoothness of F_λ with respect to λ we can write

$$R_\lambda = \mathrm{Id} + O(|\lambda|).$$

This gives

$$0 = \frac{1}{2} \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k^j \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* R_\lambda^* \delta_k^j$$

$$= \frac{1}{2} \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k^j \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* \delta_k^j + (|\lambda|^2).$$

Since K is an embedding, $\{\delta_j^k\}_{1 \le j \le l}$ is a basis of $H^1(K(\mathbb{T}^l))$. Consequently, the map $v \mapsto \int_{\mathbb{T}^{l-1}} \int_{\sigma_{i,\hat{\theta}_i}} K^* \delta_k^j v$ is invertible and the smallness assumption on λ (which is satisfied in particular by the KAM theorem) ensures, by the Implicit Function Theorem, that $\lambda = 0$. □

Once we have the vanishing lemma, Theorem 3.6 follows directly from Theorem 3.5 by using the construction in Appendix 12. Indeed, starting with the exact symplectic map F we construct the family F_λ to which we apply Theorem 3.5, with an approximate solution $(\lambda_0 = 0, K_0)$, to obtain $(\lambda_\infty, K_\infty)$ such that

$$F_{\lambda_\infty} \circ K_\infty = K_\infty \circ T_w.$$

The vanishing lemma implies that $\lambda_\infty = 0$.

7. Uniqueness results

In this section, we prove Theorem 3.9. We closely follow the proof in [FdlLS09a]. It is based on showing that the operator $D\mathcal{F}_\omega(K)$ has an approximate left inverse (as in [Zeh75, Zeh76a]). Notice first that the composition on the right by every translation of a solution of (1) is also a solution. Therefore, one cannot expect a strict uniqueness result. Moreover, the second statement in Lemma 4.1 and the calculation on the hyperbolic directions show that, roughly speaking, two solutions
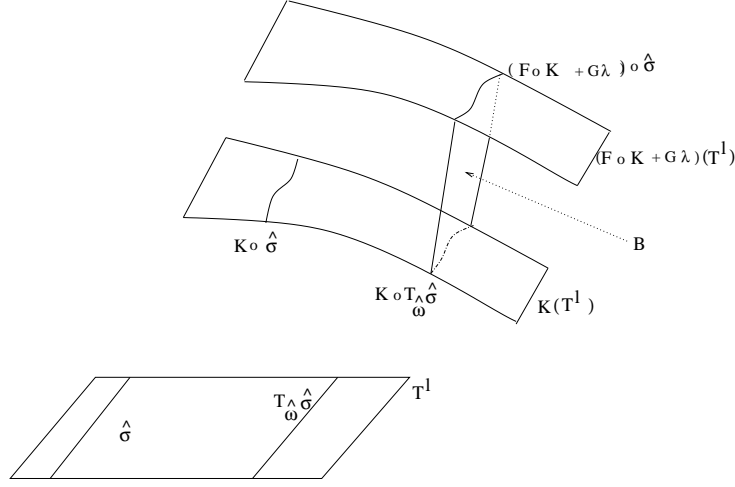
FIGURE 3. Illustration of the proof of the vanishing lemma, Lemma 6.1.

of the linearized equation differ by their average. Moreover this difference is in the direction of the tangent space of the torus. The idea behind the local uniqueness result is to prove that one can transfer the difference of the averages between two solutions to a difference of phase between the two solutions.

Now we assume that the embeddings K_1 and K_2 satisfy the hypotheses in Theorem 3.9, in particular K_1 and K_2 are solutions of (1), or (11) with $\lambda = 0$. If $\tau \neq 0$ we write K_1 for $K_1 \circ T_\tau$ which is also a solution. Therefore $\mathcal{F}_\omega(0, K_1) = \mathcal{F}_\omega(0, K_2) = 0$. By Taylor's theorem we can write

$$
\begin{aligned}
(96) \quad 0 = \mathcal{F}_\omega(0, K_1) - \mathcal{F}_\omega(0, K_2) = & D_{\lambda, K} \mathcal{F}_\omega(0, K_2)(0, K_1 - K_2) \\
& + \mathcal{R}(0, 0, K_1, K_2).
\end{aligned}
$$

Moreover, there exists $C > 0$ such that

$$
\|\mathcal{R}(0, 0, K_1, K_2)\|_{\rho, \underline{c}, \Gamma} \leq C \|K_1 - K_2\|_{\rho, \underline{c}, \Gamma}^2
$$

since $F \in C_\Gamma^2$. Hence we end up with the following linearized equation

$$
D_{\lambda, K} \mathcal{F}_\omega(0, K_2)(0, K_1 - K_2) = -\mathcal{R}(0, 0, K_1, K_2).
$$

We denote $\Delta = K_1 - K_2$.

Projecting this equation on the center subspace, writing $\Delta^c(\theta) = \Pi_{K_2(\theta)}^c \Delta(\theta)$ and making the change of function $\Delta^c(\theta) = \tilde{M}(\theta) W(\theta)$, where $\tilde{M}$ is defined in (41) with $K = K_2$, we obtain

$$
\begin{aligned}
(97) \quad DF(K_2(\theta)) \tilde{M}(\theta) W(\theta) - & \tilde{M}(\theta + \omega) W(\theta + \omega) \\
& = -\Pi_{K_2(\theta+\omega)}^c \mathcal{R}(0, 0, K_1, K_2).
\end{aligned}
$$

Applying the property $DF(K_2(\theta)) \tilde{M}(\theta) = \tilde{M}(\theta + \omega) \mathcal{S}_0(\theta)$ for solutions of (1), multiplying both sides by $\tilde{M}(\theta + \omega)^\top J_\infty(K)$ and using that $\tilde{M}^\top J_\infty(K) \tilde{M}$ is invertible we get

$$
\mathcal{S}_0(\theta) W(\theta) - W(\theta + \omega) =
$$
$$
-[(\tilde{M}^\top J_\infty(K) \tilde{M})^{-1} \tilde{M}^\top J_\infty(K)](\theta + \omega) \Pi_{K_2(\theta+\omega)}^c \mathcal{R}(0, 0, K_1, K_2).
$$

We get bounds for W, from the fact that it solves the previous equation, using the methods in Section 4.3.7. We write $W = (W_1, W_2)$. Since $\mathcal{S}_0$ is triangular we begin by looking for W_2. We search for it in the form $W_2 = W_2^\perp + \text{avg}\,(W_2)$. We have $\|W_2^\perp\|_{\rho-\delta,\underline{c},\Gamma} = C\kappa\delta^{-\nu}(\|K_1 - K_2\|_{\rho,\underline{c},\Gamma})^2$. For W_1 we have

$$W_1(\theta) - W_1(\theta+\omega) = [NDK_2^\top](\theta)(\Pi_{K_2(\theta+\omega)}^c \mathcal{R}(0,0,K_1,K_2))_1(\theta)$$

(98)
$$- A_0(\theta)W_2^\perp(\theta) - A_0(\theta)\text{avg}\,(W_2).$$

The condition that the right-hand side of (98) has zero average gives $|\text{avg}\,(W_2)| \leq C\kappa\delta^{-\nu}(\|K_1 - K_2\|_{\rho,\underline{c},\Gamma})^2$. Then

$$\|W_1 - \text{avg}\,(W_1)\|_{\rho-2\delta} \leq C\kappa^2\delta^{-2\nu}(\|K_1 - K_2\|_{\rho,\underline{c},\Gamma})^2$$

but $\text{avg}\,(W_1)$ is free. Then

$$\left\|\Delta^c - (\text{avg}\,(\Delta^c)_1, 0)^\top\right\|_{\rho-2\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-(2\nu+1)}\big(\|K_1 - K_2\|_{\rho,\underline{c},\Gamma}\big)^2.$$

The next step is done in the same way as in [dlLGJV05]. We quote Lemma 14 of that reference using our notation.

Lemma 7.1. *There exists a constant C such that if $C\|K_1 - K_2\|_{\rho,\underline{c},\Gamma} \leq 1$ then there exits a phase $\tau_1 \in \{\tau \in \mathbb{R}^l \mid |\tau| < \|K_1 - K_2\|_{\rho,\underline{c},\Gamma}\}$ such that*

$$\text{avg}\,\big(NDK_2^\top \Pi_{K(\theta)}^c(K_1 \circ T_{\tau_1} - K_2)(\theta)\big) = 0.$$

The proof is based on the application of the Banach fixed point theorem in $\mathbb{R}^l$.

As a consequence of Lemma 7.1, if τ_1 is as in the statement, then $K \circ T_{\tau_1}$ is a solution of (1) such that if

$$W = (\tilde{M}(\theta+\omega)^\top \tilde{M}(\theta+\omega))^{-1}\tilde{M}(\theta+\omega)^\top \Pi_{K(\theta)}^c(K_1 \circ T_{\tau_1} - K_2),$$

for all $\delta \in (0, \rho/2)$ and we have the estimate

$$\|W\|_{\rho-2\delta} < C\kappa^2\delta^{-2\nu}\|\mathcal{R}\|_{\rho,\underline{c},\Gamma}^2 \leq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,\underline{c},\Gamma}^2.$$

This leads to on the center subspace

$$\|\Pi_{K(\theta)}^c(K_1 \circ T_{\tau_1} - K_2)\|_{\rho-2\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,\underline{c},\Gamma}^2.$$

Furthermore, we can show that $\Delta^h = \Pi_{K(\theta)}^h(K_1 - K_2)$ satisfies the estimate

$$\|\Delta^h\|_{\rho-2\delta,\underline{c},\Gamma} < C\|\mathcal{R}\|_{\rho,\underline{c},\Gamma}.$$

All in all, we have proven the estimate for $K_1 \circ T_{\tau_1} - K_2$ (up to a change in the original constants)

$$\|K_1 \circ T_{\tau_1} - K_2\|_{\rho-2\delta,\underline{c},\Gamma} \leq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,\underline{c},\Gamma}^2.$$

We are now in position to perform the same scheme used in Section 5. We can take a sequence $\{\tau_m\}_{m\geq 1}$ such that $|\tau_1| \leq \|K_1 - K_2\|_{\rho,\underline{c},\Gamma}$ and

$$|\tau_m - \tau_{m-1}| \leq \|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1},\underline{c},\Gamma}, \qquad m \geq 2,$$

and

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m,\underline{c},\Gamma} \leq C\kappa^2\delta_m^{-2\nu}\|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1},\underline{c},\Gamma}^2,$$

where $\delta_1 = \rho/4$, $\delta_{m+1} = \delta_m/2$ for $m \geq 1$ and $\rho_0 = \rho$, $\rho_m = \rho_0 - \sum_{k=1}^m \delta_k$ for $m \geq 1$. By an induction argument we end up with

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m,\underline{c},\Gamma} \leq (C\kappa^2\delta_1^{-2\nu}2^{6\nu}\|K_1 - K_2\|_{\rho_0,\underline{c},\Gamma})^{2^m}2^{-2\nu(3+m)}.$$

Therefore, under the smallness assumptions on $\|K_1 - K_2\|_{\rho_0,\underline{c},\Gamma}$, the sequence $\{\tau_m\}_{m \geq 1}$ converges and one gets

$$\|K_1 \circ T_{\tau_\infty} - K_2\|_{\rho/2,\underline{c},\Gamma} = 0.$$

Since both $K_1 \circ T_{\tau_\infty}$ and K_2 are analytic in D_ρ and coincide in $D_{\rho/2}$ we obtain the result.

8. Results for flows on lattices

In this section, we study vector-fields defined on lattices. The models we consider – which have appeared naturally in solid state physics and in biophysics, see [CF05] for a review – consist of a sequence of copies of an individual system arranged in a lattice coupled with their neighbors.

The basic result we will prove is Theorem 8.4. The proof is based on applying the result for maps to the time-one map of the flow. This requires to study the decay properties of flows generated by Hamiltonian systems with decay, which may have some independent interest. Again, it is important to emphasize that the results we prove have an *a posteriori* format, showing that close to approximate solutions – which satisfy some hyperbolicity and twist conditions – there is a true solution. We will not need that the system is close to integrable.

8.1. **Integrating decay vector fields.** In this section, we study properties of vector-fields with decay. We will show in Proposition 8.1 that the flows they generate are families of diffeomorphisms with decay. As we will see, this is a consequence of the composition properties of spaces of functions with decay, which in turn is a property of the Banach algebra properties under multiplication. This, together with some more delicate study of the non-degeneracy conditions, will allow us to apply the existence Theorem 8.4 in the next section to the time-one map of the flow for a model problem given by a vector-field X.

We consider the equation on $\mathcal{M} = \ell^\infty(\mathbb{Z}^N)$

$$(99) \qquad\qquad \partial_\omega K(\theta) = X \circ K(\theta),$$

where X is a vector-field on $\mathcal{M}$ and K maps $\mathbb{T}^l$ into $\mathcal{M}$.

We note that we can deal with systems represented by formal Hamiltonians which are given by formal sums which do not need to converge and hence do not define a function, but however their partial derivatives and therefore the differential equations they determine re well-defined. Therefore, the invariance equations make sense. This is one of the reasons why the present method, based on the study of the invariance equation has advantages over the more classical methods [Zeh76a] based on transformations of the Hamiltonian function.

We prove that decay vector fields generate flows $\{S_t\}_{t \in \mathbb{R}}$ such that all S_t are decay diffeomorphisms.

Proposition 8.1. *Let X be a C^r vector-field, $r \geq 1$, on an open set $\mathcal{B} \subset \mathcal{M}$ (recall that we consider $\mathcal{M}$ endowed with the ℓ^∞ topology) and consider the differential equation*

$$(100) \qquad\qquad x' = X(x).$$

Let $\mathcal{B}_1 \subset \mathcal{B}$ be an open set such that $d(\mathcal{B}_1, \mathcal{B}^c) = \eta > 0$.

Then there exist $T > 0$ such that for all the initial conditions $x_0 \in \mathcal{B}_1$ there is a unique solution x_t of the Cauchy problem corresponding to (100) defined for

$|t| < T$. We denote by $S_t(x_0) = x_t$. Note that, by the uniqueness result, we have $S_{t+s} = S_t \circ S_s$ when all the maps are defined and the composition makes sense. Moreover

(1) For all $t \in (-T, T)$, $S_t : \mathcal{B}_1 \to \mathcal{B}$ is a diffeomorphism onto its image.
(2) If $X \in C_\Gamma^r(\mathcal{B})$ then $S_t \in C_\Gamma^r(\mathcal{B}_1)$ for all $t \in (-T, T)$. Moreover, there exist $C, \mu > 0$ such that

$$\|DS_t(x)\|_\Gamma \leq Ce^{\mu t}, \qquad x \in \mathcal{B}_1, \ \ t \in (-T, T).$$

Note also that, when $\mathcal{B} = \mathcal{M}$, we have $T = \infty$.

Remark. When M is the complexified manifold the derivatives can be considered as complex derivatives. Therefore in such a case S_t is analytic according to Definition 2.6.

Moreover if 0 is an equilibrium point of X, $S_t(0) = 0$ and hence if $\psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, by Lemma 10.15, $S_t \circ \psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$.

Proof. The first claim (1) follows from the standard proof of existence, uniqueness and regularity of solutions of ordinary differential equations in Banach spaces [Hal80]. Let $m = \|X\|_{C^1}$ and $T < \eta/m$. Then we have that $S_t(x)$ is C^r with respect to $(t, x) \in (-T, T) \times \mathcal{B}_1$. The uniqueness implies the flow property

$$S_{t+s}(x) = S_t(S_s(x))$$

and this property implies that $S_t^{-1} = S_{-t}$.

We note that the standard theory of existence of solutions, also gives that, when $\mathcal{B} = \mathcal{M}$, we have $T = \infty$ (recall that our definition of C^r implies that the derivatives are bounded, so that X is globally Lipschitz).

(2) By the general theory we also have that DS_t satisfies the first order variational equation

$$(DS_t(x))' = DX(S_t(x))DS_t(x), \qquad DS_0(x) = \mathrm{Id},$$

and by the theory of linear systems we know that $DS_t(x)$ is the limit of the sequence given by

(101) $$\Phi_t^0(x) = \mathrm{Id},$$

(102) $$\Phi_t^k(x) = \mathrm{Id} + \int_0^t DX(S_s(x))\Phi_s^{k-1}(x)\,ds, \qquad k \geq 1$$

for $t \in (-T, T), x \in \mathcal{B}_1$. To get that $DS_t(x) \in \mathcal{L}_\Gamma$ we make estimates of the Γ-norm of Φ_t^k.

Using (101) and (102) we obtain by induction that

$(1, k)$ $\Phi_t^k(x) \in \mathcal{L}_\Gamma, \quad \forall x \in \mathcal{B}_1, \quad \forall t \in (-T, T),$

$(2, k)$ $\displaystyle\sup_{x \in \mathcal{B}_1} \left\|\Phi_t^{k+1}(x) - \Phi_t^k(x)\right\|_\Gamma \leq \frac{1}{(k+1)!}\left(\|X\|_{C_\Gamma^1}|t|\right)^{k+1}, \forall t \in (-T, T),$

for $k \geq 0$.

In particular, we note that if $\Phi_t^k(x)$ is a linear operator represented by its matrix, then, so is $\Phi_t^{k+1}(x)$.

Writing

$$\Phi_t^k(x) = \mathrm{Id} + \sum_{j=1}^k \left(\Phi_t^j(x) - \Phi_t^{j-1}(x)\right),$$

by $(1,k),(2,k)$ we easily obtain that $\Phi_t^k(x)$ converges in $\mathcal{L}_\Gamma$. Hence $DS_t(x) \in \mathcal{L}_\Gamma$ and

$$\|DS_t(x)\|_\Gamma \leq \Gamma^{-1}(0) + \sum_{j=1}^{\infty} \frac{1}{j!} \big(\|X\|_{C_\Gamma^1}|t|\big)^j = \Gamma^{-1}(0) - 1 + \exp\big(\|X\|_{C_\Gamma^1}|t|\big).$$

The higher order derivatives of S_t satisfy higher order variational equations which are also linear equations. By an analogous argument we get that $D^k S_t(x) \in \mathcal{L}_\Gamma(\mathcal{M}, \mathcal{L}^{k-1}(\mathcal{M}, \mathcal{M}))$. We present the details for the case $k = 2$ and leave to the reader the adaptation of the typography for larger k. We have:

$$(D^2 S_t(x))' = DX(S_t(x))D^2 S_t(x)$$
$$+ D^2 X(S_t(x))(DS_t(x),\ DS_t(x)), \qquad D^2 S_0(x) = 0.$$

Let

$$G_t(x) = \int_0^t D^2 X(S_s(x))(DS_s(x),\ DS_s(x))\, ds.$$

By Lemma 10.5, $G_t(x) \in \mathcal{L}_\Gamma^2$. We can write

$$D^2 S_t(x) = \int_0^t DX(S_s(x))D^2 S_s(x)\, ds + G_t(x).$$

The sequence given by

$$\Psi_t^0(x) = 0$$

$$\Psi_t^k(x) = \int_0^t DX(S_s(x))\Psi_s^{k-1}(x)\, ds + G_t(x)$$

converges to $D^2 S_t(x)$. Similarly to the case $k = 1$ one proves by induction that $\Psi_t^k(x) \in \mathcal{L}_\Gamma^2$. Since $\mathcal{L}_\Gamma^2$ is complete we obtain that $D^2 S_t(x) \in \mathcal{L}_\Gamma^2$. $\qquad\square$

We also remark that, using the standard argument of adding extra equations [Hal80], we can obtain the smooth dependence on parameters.

8.2. **Invariant tori for flows.** The following result is our main KAM theorem for vector-fields on lattices. For the sake of simplicity, we have not formulated the most general result possible but have rather stated the result for models that appear in the Physics literature.

In the following, for sake of simplicity, we consider equations of the type

$$(103) \qquad\qquad\qquad\qquad \dot{u} = J_\infty \nabla H(u)$$

where

- The operator J_∞ is given by

$$J_\infty(z) = \mathrm{diag}\big(\ldots, J, \ldots\big),$$

  where J is the standard symplectic form.
- The ∇ operator is the standard operator induced by the $\ell^2(\mathbb{Z}^N)$ metric on the lattice.

The previous equation (103) arise in the context of statistical physics as described in the introduction. If one desires to consider more general vector-fields X, we refer the reader to the paper [FdlLS09a] where algorithms and proofs are provided in this more general context.

We first describe the non-degeneracy conditions. The linearized equation

$$\frac{d\Delta}{dt} = J_\infty D\nabla H(K(\theta + \omega t))\Delta$$

plays a crucial role. We denote $A(\theta) \equiv J_\infty D\nabla H(K(\theta))$ qnd we remark that since the vector field $J_\infty \nabla H$ has decay, by Proposition 8.1, the vector field $J_\infty \nabla H$ generates an evolution operator, denoted $U_\theta(t)$ with decay. We have

$$\frac{d}{dt}U_\theta(t) = A(\theta + \omega t)U_\theta(t),$$

and $U_\theta(0) = \mathrm{Id}$. We now have the following definitions.

Condition 8.2. *(Spectral non-degeneracy condition) An embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ we say that K is hyperbolic non-degenerate if there is an analytic splitting*

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$$

invariant under the linearized equation (106) *in the sense that*

$$U_\theta(t)\mathcal{E}^{s,c,u}_{K(\theta)} = \mathcal{E}^{s,c,u}_{K(\theta+\omega t)}.$$

Moreover the center subspace $\mathcal{E}^c_{K(\theta)}$ has dimension $2l$. We denote $\Pi^s_{K(\theta)}$, $\Pi^c_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ the projections associated to this splitting and we denote

$$U^{s,c,u}_\theta(t) = U_\theta(t)|_{\mathcal{E}^{s,c,u}_{K(\theta)}}.$$

Furthermore, we assume that there exist β_1, β_2, $\beta_3 > 0$ and $C_h > 0$ independent of θ satisfying $\beta_3 < \beta_1$, $\beta_3 < \beta_2$ and such that the splitting is characterized by the following rate conditions:

$$(104) \qquad \begin{aligned} \|U^s_\theta(t)U^s_\theta(\tau)^{-1}\|_{\rho,\underline{c}\Gamma} &\le C_h e^{-\beta_1(t-\tau)}, & t \ge \tau \ge 0, \\ \|U^u_\theta(t)U^u_\theta(\tau)^{-1}\|_{\rho,\underline{c},\Gamma} &\le C_h e^{\beta_2(t-\tau)}, & t \le \tau \le 0, \\ \|U^c_\theta(t)U^c_\theta(\tau)^{-1}\|_{\rho,\underline{c},\Gamma} &\le C_h e^{\beta_3|t-\tau|}, & t,\tau \in \mathbb{R}. \end{aligned}$$

Condition 8.3. *let $N(\theta) = [DK(\theta)^\top DK(\theta)]^{-1}$ and $P(\theta) = DK(\theta)N(\theta)$. The average on $\mathbb{T}^l$ of the matrix*

$$S(\theta) = N(\theta)DK(\theta)^\top[A(\theta)J_\infty - J_\infty A(\theta)]DK(\theta)N(\theta).$$

is non-singular. Here $A(\theta) = J_\infty D\nabla H(K(\theta))$.

We now state our theorem

Theorem 8.4. *Let H be a formal Hamiltonian function on T^*M such that the associated vector-field $X = J_\infty \nabla H$ is an exact symplectic C^2_Γ, analytic vector-field in $\mathcal{M}$. For some decay function Γ, let $\omega \in D_h(\kappa,\nu)$ for some $\kappa > 0$, $\nu \ge l-1$, $\rho_0 > 0$ and $\underline{c} = (c_1,\ldots,c_R) \in (\mathbb{Z}^N)^R$. Denote S_t the flow associated to X.*

Consider the equation

$$(105) \qquad \sum_{i=1}^{l} \omega_i \frac{\partial K}{\partial \theta_i}(\theta) = (X \circ K)(\theta).$$

Assume

(1) *X extends analytically to a complex neighborhood $\mathcal{U}$ of $K_0(D_{\rho_0})$:*

$$B_r = \big\{z \in \mathcal{M} | \exists \theta \in \mathbb{T}^l s.t |Im\,\theta| < \rho_0,\ |z - K_0(\theta)| < r\big\},$$

for some $r > 0$.

(2) *There exists $K_0 \in \mathcal{A}_{\rho_0, \Gamma, \underline{c}, \Gamma}$ such that $K_0 \in ND_{loc}(\rho_0, \Gamma)$ (the embedding K_0 is non-degenerate) in the sense that it satisfies non-degeneracy conditions 8.2 and 8.3.*

(3) *There exists a constant $C > 0$ depending on l, κ, ν, ρ_0, $\|H\|_{C^3(B_r)_\Gamma}$, $\|DK_0\|_{\rho_0, \underline{c}, \Gamma}$, $\|N_0\|_{\rho_0}$, $\|S_0\|_{\rho_0}$, $|\text{avg}\,(S_0)|^{-1}$ (where S_0 and N_0 are as in definitions 8.2-8.3 replacing K by K_0) and $\|\Pi^{c,s,u}_{K_0(\theta)}\|_{\rho_0, \Gamma}$ such that $E_0 = J_\infty \nabla H(K_0) - \partial_\omega K_0$ satisfies the following estimates*

$$C\kappa^4 \delta^{-4\nu} \|E_0\|_{\rho_0, \underline{c}, \Gamma} < 1$$

and

$$C\kappa^2 \delta^{-2\nu} \|E_0\|_{\rho_0, \underline{c}, \Gamma} < r,$$

where $0 < \delta < \min(1, \rho_0/12)$ is fixed.

Then there exists an analytic embedding $K \in \mathcal{A}_{\rho - 6\delta, \underline{c}, \Gamma}$ such that $K \in ND_{loc}(\rho - 6\delta, \Gamma)$ and satisfies equation (105) for all $t \in \mathbb{R}$.

Proof. We will only sketch the proof and refer the reader to [FdlLS09a] where the complete proofs are provided in the case of finite-dimensions. in the present framework, the Banach algebra properties of our spaces make the proofs in the infinite dimensional case very similar to the finite-dimensional context. We consider the following linearized equation:

$$(106) \qquad \frac{d\Delta}{dt} - A(\theta + \omega t)\Delta = -E(\theta).$$

We first project equation (106) on the center subspace and on the hyperbolic subspaces. On the center subspace, one has

$$(107) \qquad \partial_\omega \Delta^c(\theta) - A(\theta)\Delta^c(\theta) = -E^c(\theta),$$

Using the following proposition ([Rüs76a], [Rüs76b], [Rüs75], [dlL01]),

Proposition 8.5. *Assume that $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geq l - 1$, i.e.*

$$|\omega \cdot k|^{-1} \leq \kappa |k|^\nu, \qquad \text{for all } k \in \mathbb{Z}^l \setminus \{0\}.$$

Let $h : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ be a real analytic function with zero average. Then, for any $0 < \delta < \rho$ there exists a unique analytic solution $v : D_{\rho - \delta} \supset \mathbb{T}^l \to \mathcal{M}$ of the linear equation

$$\sum_{j=1}^{l} \omega_j \frac{\partial v}{\partial \theta_j} = h$$

having zero average. Moreover, if $h \in \mathcal{A}_{\rho, \underline{c}, \Gamma}$ then v satisfies the following estimate

$$\|v\|_{\rho - \delta, \underline{c}, \Gamma} \leq C\kappa \delta^{-\nu} \|h\|_{\rho, \underline{c}, \Gamma}, \qquad 0 < \delta < \rho.$$

The constant C depends on ν and the dimension of the torus l but is independent of $\underline{c}$.

one can prove the following reducibility property

Lemma 8.6. *Assume $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geq l - 1$ and $\|E\|_\rho$ is small enough. Then there exist a matrix $B(\theta)$ and vectors p_1 and p_2 such that equation*

$$(108) \qquad [\partial_\omega \tilde{M}(\theta) - A(\theta)\tilde{M}(\theta)]\xi(\theta) + \tilde{M}(\theta)\partial_\omega \xi(\theta) = -E^c(\theta),$$

can be written as

$$\left[\begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix} + B(\theta) \right] \xi(\theta) + \partial_\omega \xi(\theta) = p_1(\theta) + p_2(\theta).$$

Moreover, the following estimates hold:

(109) $$\|p_1\|_{\rho,\underline{c},\Gamma} \leq C \|E\|_{\rho,\underline{c},\Gamma},$$

where C just depends on $\|J_\infty(K)\|_{\rho,\Gamma}$, $\|N\|_\rho$, $\|DK\|_{\rho,\underline{c},\Gamma}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\underline{c},\Gamma}$. For p_2 and B we have

(110) $$\|p_2\|_{\rho-2\delta,\underline{c},\Gamma} \leq C\kappa\delta^{-(\nu+1)} \|E\|^2_{\rho,\underline{c},\Gamma}$$

and

(111) $$\|B\|_{\rho-2\delta,\Gamma} \leq C\kappa\delta^{-(\nu+1)} \|E\|_{\rho,\underline{c}\,Gamma}$$

for $\delta \in (0, \rho/2)$, where C depends on l, ν, $\|N\|_\rho$, $\|DK\|_{\rho,\Gamma}$, $|H|_{C^3(B_r)_\Gamma}$, $|J|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_{\rho,\Gamma}$.

The solution of the reduced equations works in the same way as in the case of maps. We sketch the procedure and we emphasize on the differences.

We write $\xi = (\xi_1, \xi_2)$. We introduce the operator

(112) $$\mathcal{L}\xi = \begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix} \xi + \partial_\omega \xi = p_1(\theta),$$

where $p_1 = (p_{11}, p_{12})$. Using this decomposition of $\mathcal{E}^c_{K(\theta)}$ we can write equation (112) in the form

$$S(\theta)\xi_2(\theta) + \partial_\omega \xi_1(\theta) = p_{11}(\theta),$$
$$\partial_\omega \xi_2(\theta) = p_{12}(\theta).$$

where

$$p_{12}(\theta) = DK(\theta)^\top J_\infty DE(\theta).$$

In order to be able to solve this small divisor equations, one has to ensure that the average on $\mathbb{T}^l$ of $DK(\theta)^\top J_\infty DE(\theta)$ is zero. This is the main difference with the finite-dimensional case and we perform now the computation. We use the fact that J_∞ has a special structure. Indeed, we have

$$(J_\infty)_{ij} = J\delta_{ij}$$

and then

$$(DK^\top J_\infty DE)_{ij} = \sum_{k \in \mathbb{Z}^N} (DK^\top)_{ik} (J_\infty DE)_{kj}.$$

But, we have

$$(J_\infty DE)_{kj} = J(DE)_{kj}$$

and then

$$(DK^\top J_\infty DE)_{ij} = \sum_{k \in \mathbb{Z}^N} (DK^\top)_{ik} J(DE)_{kj}.$$

Thereofore, the average on $\mathbb{T}^l$ of $DK^\top J_\infty DE^c$ amounts to compute the average on $\mathbb{T}^l$ of $(DK^\top)_{ik} J(DE)_{kj}$.

Remark 8.7. Here we have use the fact

$$E^c = \tilde{M}E + \hat{e}E$$

where $\hat{e} = \pi^c_{K(\theta+\omega)} - \pi^\Gamma_{K(\theta+\omega)}$ and the term $\hat{e}E$ being quadratic in the error, one can omit it.

By the computations in [dlLGJV05], one proves that thew average of $(DK^\top)_{ik}J(DE)_{kj}$ is zero. Hence this gives the desired result.

We now project the linearized equation (106) on the stable and unstable subspaces by using the projections $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ respectively. We denote $\Delta^s(\theta) = \Pi^s_{K(\theta)}\Delta(\theta)$, $\Delta^u(\theta) = \Pi^u_{K(\theta)}\Delta(\theta)$.

Using the previous notation, we obtain

$$(113) \qquad \partial_\omega \Delta^s(\theta) - A(\theta)\Delta^s(\theta) = -\Pi^s_{K(\theta)}E(\theta)$$

for the stable part and

$$(114) \qquad \partial_\omega \Delta^u(\theta) - A(\theta)\Delta^u(\theta) = -\Pi^u_{K(\theta)}E(\theta)$$

for the unstable one.

The following result provides the solution of the previous equations.

Proposition 8.8. *Given $\rho > 0$, equations (113) and (114) admit unique analytic solutions $\Delta^s : D_\rho \to \mathcal{E}^s$ and $\Delta^u : D_\rho \to \mathcal{E}^u$ respectively, such that $\Delta^{s,u}(\theta) \in \mathcal{E}^{s,u}_{K(\theta)}$. Furthermore there exist constants $C^{s,u}$ such that*

$$(115) \qquad \|\Delta^{s,u}\|_{\rho,\underline{c},\Gamma} \le C^{s,u}\|E\|_{\rho,\underline{c},\Gamma},$$

where $C^{s,u}$ depend on β_1, $\|\Pi^s_{K(\theta)}\|_{\rho,\Gamma}$ (resp. β_2, $\|\Pi^u_{K(\theta)}\|_{\rho,\Gamma}$) and C_h but is independent of $\underline{c}$.

The previous definition closes the proof of Theorem 8.4.

$\square$

9. Proof of Theorem 3.11

The goal of this section is to prove Theorem 3.11. We proceed in three stages:

(1) In the first stage, we construct quasiperiodic breathers around one site indexed by a frequency $\omega \in \Xi(\varepsilon*)$. This will be a straightforward application of Theorem 3.6. See Section 9.1. We will use as initial approximation the solutions in which one site is oscillating quasi-periodically and the others are at the fixed point. This is an exact solution when $\varepsilon = 0$ and will be an approximate solution when ε is sufficiently small. Note that, since the system is translation invariant, the center site can be chosen to be any point on the lattice.

(2) In a second stage, carried out in Section 9.3 we show that, given two solutions which are centered around two groups of sites, if we displace far enough these solutions and add them, we obtain an approximate solution (for a slightly slower decay function). Then we can conclude to the existence of a true solution close to them. The estimates of solutions displaced will be the content of the *coupling lemma* (Lemma 9.8), which is the centerpiece of the argument. This second stage of coupling different solutions requires several new techniques. In particular, a detailed discussion of Diophantine vectors in infinite dimensions. It will also be crucial that many of

the estimates that we have obtained before are uniform in the number and the geometry of the sites.

(3) Finally, in a third stage, we will show that there is a limit to this process of clustering breathers. We obtain a well defined limit if the centers are placed far enough apart.

We will need the following definition.

Definition 9.1. *Given $m \in \mathbb{Z}^N$, let $\tau^m : \mathcal{M} \to \mathcal{M}$ be defined by*

$$\left(\tau^m(x)\right)_i = x_{i+m}, \qquad i \in \mathbb{Z}^N.$$

In particular if $F : \mathcal{M} \to \mathcal{M}$ and $k : \mathbb{T}^p \to \mathcal{M}$

$$(\tau^m F)_i(x) = F_{i+m}(x),$$
$$(\tau^m k)_i(\theta) = k_{i+m}(\theta).$$

Let S_t be the flow of the system associated to the Hamiltonian in the statement of Theorem 3.11, and let $\tilde{S}_t = \tau^m S_t \, \tau^{-m}$, with $m \in \mathbb{Z}^N$. Both S_t and $\tilde{S}_t$ satisfy the same initial value problem, hence they coincide wherever they are defined. As a consequence we have, using $F = S_1$,

$$F = \tau^m \circ F \circ \tau^{-m}.$$

From this we deduce that if $K_\omega : \mathbb{T}^{rl} \to \mathcal{M}$ with $\omega \in \mathbb{R}^{rl}$ is a solution of $F \circ K_\omega = K_\omega \circ T_\omega$ then for all $m \in \mathbb{Z}^N$ we have that $\tau^m K_\omega$ is also a solution.

9.1. Existence of quasi-periodic breathers centered around one site (Part A of Theorem 3.11).

Since the problem is invariant under translations, we will choose, without loss of generality, to center the breather at the origin. We then consider the Hamiltonian

$$H_\varepsilon(q,p) = \sum_{n \in \mathbb{Z}^N} \left(\frac{1}{2} p_n^2 + W(q_n) \right) + \varepsilon \sum_{j \in \mathbb{Z}^N} \sum_{n \in \mathbb{Z}^N} V_j(q_n - q_{n+j}).$$

We note that, by Proposition 8.1 for ε small enough, we can obtain a time-1 map, which we will denote by F_ε. This map will be exact symplectic by Proposition 11.9 in Appendix 11.

We also note that, for $\varepsilon = 0$, F_0 is an uncoupled map

$$(F_0(x))_i = f_0(x_i), \quad i \in \mathbb{Z}^N$$

with f_0 the time-1 map of the the flow on M corresponding to the Hamiltonian $\frac{1}{2}p^2 + W(q)$.

Assumption **H2** of Theorem 3.11, implies that, for $\omega \in \Xi_0$ we can find an embedding $k_\omega : \mathbb{T}^l \to M$ such that

$$f_0 \circ k_\omega = k_\omega \circ T_\omega$$

We can then consider the embedding $K : \mathbb{T}^l \to \mathcal{M}$ defined by:

$$\left(K_\omega(\theta)\right)_i = \begin{cases} k_\omega(\theta) & i = 0 \\ 0 & i \neq 0 \end{cases}$$

Note that $F_0 \circ K_\omega = K_\omega \circ T_\omega$. What we want to do is to check that K_0 satisfies the hypothesis of Theorem 3.6.

We start by embedding F_ε into a family $F_{\varepsilon,\lambda}$ $\lambda \in \mathbb{R}^l$ constructed by setting

$$\left(F_{\varepsilon,\lambda}(\theta)\right)_i = \begin{cases} \left(F_\varepsilon(\theta)\right)_i + (0,\lambda) & i = 0 \\ \left(F_\varepsilon(\theta)\right)_i & i \neq 0 \end{cases}$$

We can think of $F_{\varepsilon,\lambda}$ as the composition of the map F_ε and a translation in the direction of the action in the $i = 0$ component. Both maps are symplectic but the translation is not exact symplectic.

To verify the quality of the embedding, we note that

$$\sum_{n \in \mathbb{Z}^N} \frac{\partial(K_\omega)_n}{\partial\theta_i} \frac{\partial(K_\omega)_n}{\partial\theta_j} = \frac{\partial(k_\omega)}{\partial\theta_i} \frac{\partial(k_\omega)}{\partial\theta_j}$$

and, by assumption the later is non-degenerate uniformly in ω.

We note that for the uncoupled map, the twist condition and the parameter nondegeneracy conditions in Definition 3.4 reduce to the conditions for the time-one map. Similarly, the hyperbolicity conditions (Definition 3.1) are satisfied whenever $\varepsilon = 0$. The stability results developed before show that these conditions remain true (with uniform values) for $|\varepsilon| \ll 1$.

If we choose $\Xi_0^*(\varepsilon)$ with uniform Diophantine constants, and chose $\varepsilon*$ accordingly, we obtain from Theorem 3.6 the existence of the KAM tori. The uniformity of the hyperbolicity and the non-degeneracy constants is a consequence of the perturbation results for the non-degeneracy conditions.

9.2. **Number-theoretic properties of infinite sequences of frequencies.**
This section is devoted to some results on infinite sequences of frequencies. We want to introduce the concept of Diophantine sequence (see Definition 9.2) and show that these sequences are very abundant in the sense that they have full probability with respect to several probability measures.

Let us consider $\Xi_0 \subset [-L, L]^l$ with $l \in \mathbb{N}$ and $L > 0$. Assume that Ξ_0 has positive Lebesgue measure. Later we will take as Ξ_0 to be a subset of $\mathcal{D}(\kappa_0, \nu_0)$ such that there are KAM tori in the uncoupled system with this Diophantine properties (See Assumption **H2** in Theorem 3.11.)

To discuss infinite products of measures, we consider the normalized probability measure

$$\text{meas}_*(.) = \frac{\text{meas}(.)}{\text{meas}(\Xi_0)}$$

where $\text{meas}(\cdot)$ is any measure absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^l$. By a theorem of Kolmogorov [Dur96], the product set $\Xi_0^\mathbb{N}$ with the product σ-algebra can be endowed with the product probability measure $\text{meas}_*^\mathbb{N}$.

Note that, there are different sets Ξ_0 which satisfy the assumption **H2** of Theorem 3.11. Each of these choices will lead to mutually singular measures in the infinite product. Nevertheless, we do not include Ξ_0 in the notation for the infinite measure. The result will, of course, be valid for all choices.

Now we introduce a notion of Diophantine sequences in $(\mathbb{R}^l)^\mathbb{N}$ which is well adapted for our needs. Basically, we just require that for every $r \geq 1$ the first r components are Diophantine, even if the exponent and the constant change with r. The Diophantine properties of the sequence are just the sequence of Diophantine properties of the truncations. This will be natural for us since at every stage of the argument we will be working with just a finite number of frequencies.

We introduce the following notation. We will consider

$$\underline{\omega} = (\omega_1, \omega_2, \dots) \in \Xi_0^{\mathbb{N}} \quad \text{and} \quad \underline{k} = (k_1, k_2, \dots) \in (\mathbb{Z}^l)^{\mathbb{N}} \setminus \{0\}.$$

We denote $\underline{\omega}^{(r)} = (\omega_1, \dots, \omega_r)$ and $\underline{k}^{(r)} = (k_1, \dots, k_r)$ the truncated sequences of length r. Hence

$$\underline{\omega}^{(r)} \cdot \underline{k}^{(r)} = \sum_{i=1}^{r} \omega_i \cdot k_i \quad \text{and} \quad \left| k^{(r)} \right| = \sum_{i=1}^{r} |k_i|,$$

where $k_i = (k_{i,1}, \dots, k_{i,l}) \in \mathbb{Z}^l$ and $|k_i| = |k_{i,1}| + \cdots + |k_{i,l}|$. Also, given $\omega_1 \in \mathbb{R}^{rl}$ and $\omega_2 \in \mathbb{R}^l$ we will write $(\omega_1, \omega_2) \in \mathbb{R}^{(r+1)l}$ the concatenation of the vectors ω_1, ω_2.

Definition 9.2. *We define*

$$\mathcal{D} = \bigcup_{(\underline{\kappa}, \underline{\nu}) \in (\mathbb{R}^+)^{\mathbb{N}} \times (\mathbb{R}^+)^{\mathbb{N}}} \mathcal{D}(\underline{\kappa}, \underline{\nu}),$$

where $\underline{\kappa} = (\kappa_1, \dots \kappa_r, \dots)$, $\underline{\nu} = (\nu_1, \dots \nu_r, \dots)$ *and*

$$\mathcal{D}(\underline{\kappa}, \underline{\nu}) = \left\{ \begin{array}{l} \underline{\omega} \in \Xi_0^{\mathbb{N}} \,|\, \forall r \geq 1, \ \left| \sum_{i=1}^{r} \omega_i \cdot k_i - m \right|^{-1} \leq \kappa_r |k^{(r)}|^{\nu_r}, \\ \forall \underline{k} \in (\mathbb{Z}^N)^{\mathbb{N}} \text{ s.t. } k^{(r)} \neq 0, \ \forall m \in \mathbb{Z} \end{array} \right\}.$$

Remark 9.3. Note that, since we are considering infinite dimensions, the notion of $|k|$ we are using in Definition 9.2 could matter. We note however that changing the norms only changes the sequence $\underline{\kappa}$, the sequence $\underline{\nu}$ remaining the same.

The next result ensures that there are many sequences of Diophantine vectors.

Lemma 9.4. *Let $\underline{\nu}$ be a given sequence such that $\nu_r > rl$. Then,*

$$\text{meas}(\cdot)_*^{\mathbb{N}} \Big(\Xi_0^{\mathbb{N}} \setminus \bigcup_{\underline{\kappa} \in (\mathbb{R}^+)^{\mathbb{N}}} \mathcal{D}(\underline{\kappa}, \underline{\nu}) \Big) = 0.$$

Proof. We follow the standard argument for the finite dimensional case (see [dlL01] for a pedagogical exposition).

Notice first that L – the size of the box in $\mathbb{R}^l$ containing our set Ξ_0 – is fixed. We start by considering r fixed.

For $k \in \mathbb{Z}^{rl}$, $\kappa_r \in \mathbb{R}^+$, $\nu_r \in \mathbb{R}^+$, we define

$$B_{k,m,\kappa_r,\nu_r} = \left\{ \omega \in ([-L,L]^l)^{\mathbb{N}} \,|\, |\omega^{(r)} \cdot k - m| < \kappa_r^{-1} |k|^{-\nu_r} \right\}.$$

We note that

(116) $$\Xi_0^{\mathbb{N}} \setminus \mathcal{D}(\underline{\kappa}, \underline{\nu}) = \bigcup_{r \geq 1} \bigcup_{k \in \mathbb{Z}^{rl} \setminus \{0\}} B_{k,m,\kappa_r,\nu_r}.$$

Geometrically, the sets B_{k,m,κ_r,ν_r} are slabs of width $2\kappa_r^{-1}|k|^{-\nu_r-1}$. As a consequence, we obtain that

$$\text{meas}^{\mathbb{N}}(B_{k,m,\kappa_r,\nu_r}) \cap \Xi_0^{\mathbb{N}} \leq \text{meas}^r(B_{k,m,\kappa_r,\nu_r}) \cap \Xi_0^r \leq C_r \kappa_r^{-1} |k|^{-\nu_r-1}.$$

Moreover, given k, the number of sets B_{k,m,κ_r,ν_r} intersecting Ξ_0^r is bounded by a constant depending on the dimension times $|k|$. Hence, we have that for $\nu_r > rl$,

$$\text{meas}^{\mathbb{N}}(\cup_{k \in \mathbb{Z}^{rl} \setminus \{0\}, m} B_{k,m,\kappa_r,\nu_r}) \leq \sum_{k \in \mathbb{Z}^{rl} \setminus \{0\}} \text{meas}^{\mathbb{N}}(B_{k,m,\kappa_r,\nu_r}) C_{rl} |k|$$

$$\leq \text{meas}(\Xi_0)^{-r} 2rl \ \kappa_r^{-1} \sum_{s=1}^{\infty} C'_{rl} \frac{s^{rl-1}}{s^{\nu_r}} \leq C''_{rl,\nu_r} \ \kappa_r^{-1},$$

where C_{rl}, C'_{rl} and C''_{rl,ν_r} are explicit constants. The right-hand side of the previous expression can be estimated from above by $\sum_{r \geq 1} C_{r,\nu_r} \kappa_r^{-1}$.

By choosing a suitable sequence $\underline{\kappa}$, the sum can be made as small as desired. $\square$

9.3. Constructing more complicated breathers out of simpler ones. The coupling lemma.

The main goal of this section is to prove Lemma 9.8 that shows that if we have two solutions of the invariant equation and put them in places separated sufficiently far apart, when we add them, we obtain a very approximate solution of the invariance equations.

In Lemma 9.14 we will show that these solutions obtained superimposing the two non-degenerate (in the sense of Definition 3.4) solutions centered around very far apart centers also satisfy the same non-degeneracy assumptions with only slight worse constants.

We will also show in Lemma 9.13 that if the approximate solutions are (up to a bounded error η) superpositions of one site breathers, then, they satisfy the hyperbolicity conditions of Theorem 3.5 with uniform bounds. The crucial point of Lemma 9.13 is that the estimates on the η allowed and the non-degeneracy constants are independent of $\underline{c}$, the finite set of sites that we are considering. This will be a relatively easy consequence of all the uniformity properties that we have developed so far.

9.3.1. *Some elementary calculations with the decay functions in Proposition 2.3.* Theorem 3.11 is only formulated with the the scale of decay functions Γ_β defined by

$$\Gamma_\beta(i) = \begin{cases} a \, |i|^{-\alpha} \, e^{-\beta|i|} & \text{if} \quad i \neq 0, \\ a & \text{if} \quad i = 0, \end{cases}$$

with $\alpha = \alpha_0$ fixed and $0 < \beta \leq \beta_0$.

In the proof of Proposition 2.3 in [JdlL00] it is shown that the value of a can be chosen as any value less than some a_0 independent of β. Actually we have

$$a_0(\alpha) < \left(2^{\alpha+1} K_{N,\alpha} + 2\right)^{-1}, \qquad \text{with} \quad K_{N,\alpha} = \sum_{j \in \mathbb{Z}^N \setminus \{0\}} |j|^{-\alpha}.$$

In the definition of both Γ and Γ_β we take the value of $a = \min\left(a_0(\alpha), a_0(2\alpha)\right)$.

With this choice we have the following properties:

(1) if $\tilde{\beta} < \beta$ then $\Gamma_\beta(i) \leq \Gamma_{\tilde{\beta}}(i)$ for all $i \in \mathbb{Z}^N$.

(2) if $\tilde{\beta} < \beta$ then $\displaystyle\lim_{|m| \to \infty} \frac{\Gamma_\beta(m)}{\Gamma_{\tilde{\beta}}(m)} = 0$.

(3) for any $\beta, \tilde{\beta} \leq \beta_0$

$$(117) \qquad\qquad \Gamma(i) \leq \frac{1}{a} \, \Gamma_\beta(i) \, \Gamma_{\tilde{\beta}}(i), \qquad i \in \mathbb{Z}^N.$$

We will encounter the quantity $\displaystyle\sum_{k=1}^{r} \Gamma_\beta(i - c_k)$. To be able to estimate it in a convenient way, independently on r, we will work with sequences of sites $\underline{c} = (c_1, c_2, \dots) \in (\mathbb{Z}^N)^{\mathbb{N}}$ satisfying the property

Definition 9.5. *We say that a sequence of sites $\underline{c}$ is spatially non-resonant when for all $i \in \mathbb{Z}^N$ there exist at most two different sites c_p, c_q in the sequence such that $|c_p - i| = |c_q - i|$.*
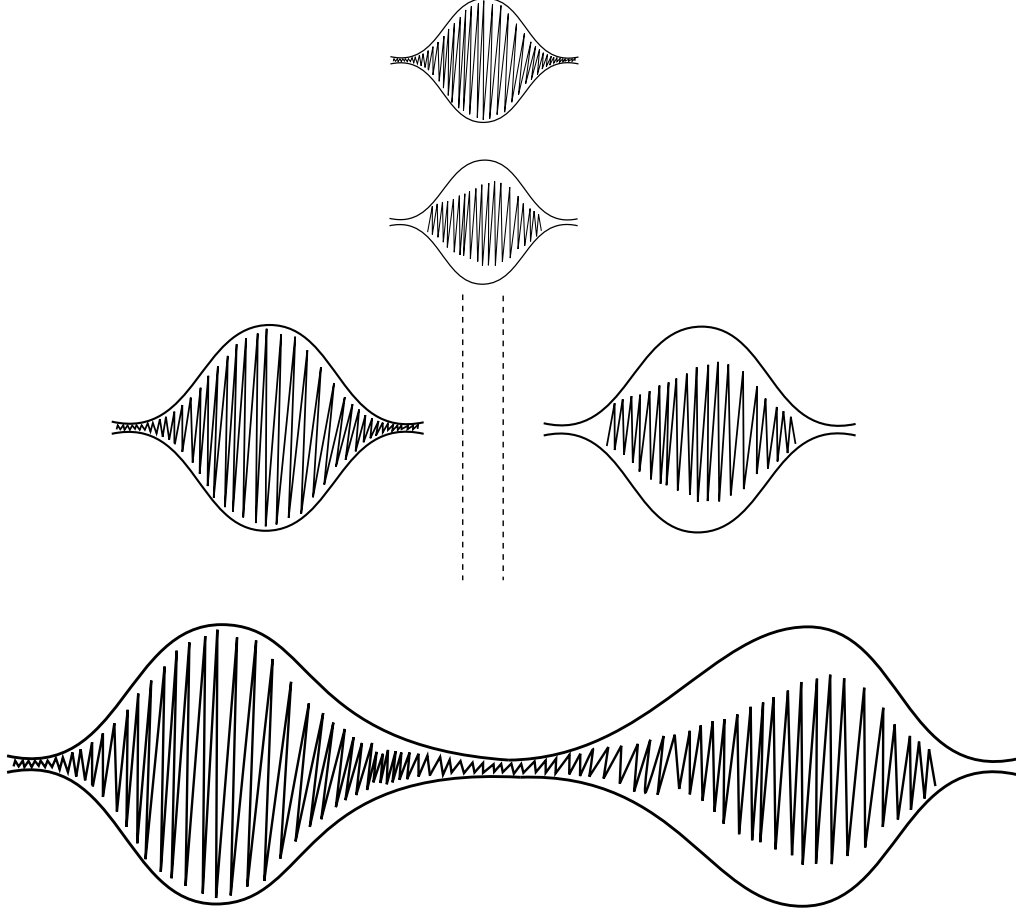
FIGURE 4. Given two breathers, placing them far apart, we obtain
an approximate solution. Using the a-posteriori Theorem 3.6, we
obtain that there is a true solution close to it. See Lemma 9.8

Remark 9.6. If we arrange the sites c_k in a coordinate plane of $\mathbb{Z}^N$, for instance $\mathbb{Z}^2 \times \{0\}^{N-2}$, and for all k we have $|(c_{k+1} - c_k)_2| < |(c_{k+1} - c_k)_1|$ of all k, then $\underline{c}$ is spatially non resonant according to Definition 9.5.

Lemma 9.7. *Let $\underline{c}$ be a spatially non-resonant sequence.*
Let $\beta \in (0, \beta_0)$. Then for every $i \in \mathbb{Z}^N$ and $r \geq 2$ we have

$$(118) \qquad \sum_{k=1}^{r} \Gamma_\beta(i - c_k) < \frac{2}{1 - e^{-\beta}} \; \max_k \; \Gamma_\beta(i - c_k).$$

Proof. Let $i \in \mathbb{Z}^N$ be fixed. Let k_0 be such that $|i - c_{k_0}| = \min_k |i - c_k|$. By the spatially non resonant propertym in the sum (118) for any value $\Gamma_\beta(i - c_k)$ there are at most two terms taking the same value. Then we can group the terms in pairs. Moreover if $|i - c_p| > |i - c_q|$ then

$$\Gamma_\beta(i - c_p) < \Gamma_\beta(i - c_q) \; e^{-\beta(|i-c_p| - |i-c_q|)}.$$

Therefore

$$\sum_{k=1}^{r} \Gamma_\beta(i - c_k) < 2\Gamma_\beta(i - c_{k_0}) + 2\sum_{m=1}^{\infty} \Gamma_\beta(i - c_{k_0})e^{-\beta m}$$

and (118) follows. □

9.3.2. *Statement and proof of the* Coupling Lemma.

Lemma 9.8. (Coupling lemma) *Let* $K_{\underline{\omega}_1} \in \mathcal{A}_{\rho,\underline{c}_1,\Gamma_\beta} \cap ND_{\mathrm{loc}}(\rho,\Gamma_\beta)$, $K_{\omega_2} \in \mathcal{A}_{\rho,c_2,\Gamma} \cap ND_{\mathrm{loc}}(\rho,\Gamma)$ *be the parameterizations of two invariant tori for* F, *localized around* $\underline{c}_1$ *and* c_2 *respectively, vibrating with frequencies* $\underline{\omega}_1 \in \mathbb{R}^{rl}$ *and* $\omega_2 \in \mathbb{R}^l$ *respectively.*

Then, if $|m|$ *is large enough,* $K_{\omega_{12}} : \mathbb{T}^{(r+1)l} \to \mathcal{M}$ *defined by*

$$K_{\omega_{12}} = K_{\underline{\omega}_1} + \tau^m K_{\omega_2}$$

is an approximate solution of

$$F \circ K = K \circ T_{\omega_{12}}$$

in the following sense: given $0 < \tilde{\beta} < \beta$ *we have the estimate*

$$\|F \circ K_{\omega_{12}} - K_{\omega_{12}} \circ T_{\omega_{12}}\|_{\rho,\underline{c}_{12},\Gamma_{\tilde{\beta}}}$$

(119)
$$\leq \max\left(\|K_{\omega_1}\|_{\rho,\underline{c}_1,\Gamma_\beta}, \|K_{\omega_2}\|_{\rho,c_2,\Gamma}\right)\Phi(m)$$

where $c_{12} = (\underline{c}_1, c_2 - m)$, *and* Φ *depends on* $F, \underline{c}_1, c_2, \beta, \tilde{\beta}$ *and*

$$\lim_{|m|\to\infty} \Phi(m) = 0.$$

Remark 9.9. Note that the approximate torus $K_{\omega_{12}}$ is in an space of slightly slower decay than the space of the invariant tori $K_{\underline{\omega}_1}$, K_{ω_2} since the decay estimate (119) involves the weight $\Gamma_{\tilde{\beta}}$ instead of the weight Γ_β.

It is important to emphasize that if we choose two solutions, fix a decay function slower than that of the two solutions, any target smallness for the error of the coupled solution can be accomplished by setting the translated solution far enough. In other words, we can choose all the parameters of the Lemma 9.8 and adjust all the requirements by setting the solutions far apart.

Remark 9.10. Notice also that in the Lemma 9.8, the approximate torus $K_{\omega_{12}}$ is defined on $\mathbb{T}^{rl} \times \mathbb{T}^l = \mathbb{T}^{(r+1)l}$. To make the notations coherent we embed the tori $\mathbb{T}^{rl}$ and $\mathbb{T}^l$ into $\mathbb{T}^{(r+1)l}$ identifying $\mathbb{T}^{rl}$ with $\mathbb{T}^{rl} \times \{0\}$ and $\mathbb{T}^l$ with $\{0\} \times \mathbb{T}^l$ respectively. Hence if $\theta = (\theta_1, \theta_2) \in \mathbb{T}^{(r+1)l}, K_{\underline{\omega}_1}(\theta) = K_{\underline{\omega}_1}(\theta_1)$ and $K_{\omega_2}(\theta) = K_{\omega_2}(\theta_2)$.

Remark 9.11. For simplicity, we have stated Lemma 9.8 as joining together a breather around one site to an already constructed solution. It is possible (and perhaps more natural) to prove a lemma that asserts that given two solutions (each containing oscillations around many sites) one can displace them and obtain a very approximate solution (in a slower decay space). We leave the precise formulation and the proof to the reader.

Before proving Lemma 9.8 we establish a lemma with two technical estimates. Given $\underline{c}_1 = (c_{1,1}, \ldots, c_{1,r}) \in (\mathbb{Z}^N)^r$ and $c_2, m \in \mathbb{Z}^N$ we introduce the sets of indices

$$\mathcal{I}_1 = \left\{i \in \mathbb{Z}^N \mid \min_k |i - c_{1,k}| < |i + m - c_2|\right\},$$

$$\mathcal{I}_2 = \mathbb{Z}^N \setminus \mathcal{I}_1,$$

and the functions

$$B_1(\beta, \tilde{\beta}, m) = \sup_{i \in \mathcal{I}_1} \frac{\Gamma_\beta(i + m - c_2)}{\max_k \Gamma_{\tilde{\beta}}(i - c_{1,k})},$$

$$B_2(\beta, \tilde{\beta}, m) = \sup_{i \in \mathcal{I}_2} \frac{\max_k \Gamma_\beta(i - c_{1,k})}{\Gamma_{\tilde{\beta}}(i + m - c_2)}.$$

Lemma 9.12. *If* $0 < \tilde{\beta} < \beta$ *we have*

$$\lim_{|m| \to \infty} B_1(\beta, \tilde{\beta}, m) = \lim_{|m| \to \infty} B_2(\beta, \tilde{\beta}, m) = 0.$$

Proof. First we note that if $i \in \mathcal{I}_1$ then

(120) $$|i + m - c_2| > \frac{1}{2} \min_k |c_{1,k} + m - c_2|.$$

Indeed, let k_0 be such that $|i - c_{1,k_0}| = \min_k |i - c_{1,k}|$. Then

$$|i + m - c_2| \geq |c_{1,k_0} + m - c_2| - |i - c_{1,k_0}|$$
$$> |c_{1,k_0} + m - c_2| - |i + m - c_2|$$

and hence

$$|i + m - c_2| > \frac{1}{2} |c_{1,k_0} + m - c_2| \geq \frac{1}{2} \min_k |c_{1,k} + m - c_2|.$$

Moreover, if $i \in \mathcal{I}_1$, by the monotonicity of $\Gamma_{\tilde{\beta}}$ we have $\max_k \Gamma_{\tilde{\beta}}(i - c_{1,k}) > \Gamma_{\tilde{\beta}}(i + m - c_2)$. Now

$$\frac{\Gamma_\beta(i + m - c_2)}{\max_k \Gamma_{\tilde{\beta}}(i - c_{1,k})} < \frac{\Gamma_\beta(i + m - c_2)}{\Gamma_{\tilde{\beta}}(i + m - c_2)}.$$

The bound (120) shows that when $|m| \to \infty$, $|i + m - c_2|$ goes to infinity uniformly in $i \in \mathcal{I}_1$. Hence by the second property of the scale Γ_β we obtain the first limit. The second limit is proved in an analogous way, checking first that if $i \in \mathcal{I}_2$

$$\min_k |i - c_{1,k}| \geq \frac{1}{2} \min_k |c_{1,k} + m - c_2|$$

and using that if $i \in \mathcal{I}_2$

$$\frac{\max_k \Gamma(i - c_{1,k})}{\Gamma_{\tilde{\beta}}(i + m - c_2)} \leq \frac{\max_k \Gamma_\beta(i - c_{1,k})}{\max_k \Gamma_{\tilde{\beta}}(i - c_{1,k})}.$$

$\square$

Proof of the coupling lemma Lemma 9.8. We denote $E(K) = F \circ K - K \circ T_{\omega_{12}}$ the error of the invariance equation for the coupled breather.

We are going to estimate the i-th component of $E = E(K_{\omega_{12}})$. Note that the torus $\tau^m K_{\omega_2}$ is localized around the site $c_2 - m$. We distinguish two cases: either $i \in \mathcal{I}_1$ or $i \in \mathcal{I}_2$. In the first case we write

$$E_i = F_i(K_{\underline{\omega}_1}) + \int_0^1 \left[DF\left(K_{\underline{\omega}_1} + s\tau^m K_{\omega_2}\right) \tau^m K\omega_2 \right]_i ds$$
$$- \left[K_{\underline{\omega}_1} \circ T_{\omega_{12}} \right]_i - \left[\tau^m K_{\omega_2} \circ T_{\omega_{12}} \right]_i.$$

We recall that $\theta = (\theta_1, \theta_2) \in \mathbb{T}^{rl} \times \mathbb{T}^l$. Since $K_{\underline{\omega}_1}$ does not depend on θ_2 then $F(K_{\underline{\omega}_1}) = K_{\underline{\omega}_1} \circ T_{\omega_{12}}$. Therefore

$$\|E_i\|_\rho \leq \sum_j \|F\|_{C^1_\Gamma} \Gamma(i-j) \|K_{\omega_2}\|_{\rho,c_2,\Gamma} \Gamma(j+m-c_2)$$

$$+ \|K_{\omega_2}\|_{\rho,c_2,\Gamma} \Gamma(i+m-c_2)$$

$$\leq \left(\|F\|_{C^1_\Gamma} + 1\right) \|K_{\omega_2}\|_{\rho,c_2,\Gamma} \Gamma(i+m-c_2).$$

Similarly, if $i \in \mathcal{I}_2$ we expand F around $\tau^m K_{\omega_2}$ and we obtain

$$\|E_i\|_\rho \leq \left(\frac{2}{1-e^{-\beta}} \|F\|_{C^1_\Gamma} + 1\right) \|K_{\underline{\omega}_1}\|_{\rho,\underline{c}_1,\Gamma_\beta} \max_k \Gamma_\beta(i - c_{1,k}).$$

We take $\tilde{\beta} < \beta$ and we compute

$$\|E\|_{\rho,\underline{c}_{12},\Gamma_{\tilde{\beta}}} = \max\left(\sup_{i\in\mathcal{I}_1} \min\left(\min_k \Gamma_{\tilde{\beta}}^{-1}(i-c_{1,k}), \Gamma_{\tilde{\beta}}^{-1}(i+m-c_2)\right) \|E_i\|_\rho,\right.$$

$$\left.\sup_{i\in\mathcal{I}_2} \min\left(\min_k \Gamma_{\tilde{\beta}}^{-1}(i-c_{1,k}), \Gamma_{\tilde{\beta}}^{-1}(i+m-c_2)\right) \|E_i\|_\rho\right)$$

$$\leq C \max\left(\sup_{i\in\mathcal{I}_1} \min_k \Gamma_{\tilde{\beta}}^{-1}(i-c_{1,k}) \Gamma(i+m-c_2),\right.$$

$$\left.\sup_{i\in\mathcal{I}_2} \Gamma_{\tilde{\beta}}^{-1}(i+m-c_2) \max_k \Gamma_\beta(i-c_{1,k})\right)$$

$$= C \max\left(B_1(2\beta_0, \tilde{\beta}, m), B_2(\beta, \tilde{\beta}, m)\right),$$

where

$$C = \left(\frac{2}{1-e^{-\beta}} \|F\|_{C^1_\Gamma} + 1\right) \max\left(\|K_{\underline{\omega}_1}\|_{\rho,\underline{c}_1,\Gamma_\beta}, \|K_{\omega_2}\|_{\rho,c_2,\Gamma}\right).$$

$\square$

9.3.3. Statement and proof of Lemma 9.13. Verifying the non-degeneracy condition of the coupled solutions.

In this section, we verify the nondegeneracy conditions provided that ε is small enough and that K is sufficiently close to the an uncoupled solution with all the sites far enough apart. That is, we consider situations when we are close to the completely uncoupled solution.

The result Lemma 9.13 will be clear because all the uncoupled solutions for the uncoupled dynamics satisfy the non-degeneracy assumptions. The change of the non-degeneracy assumptions between this uncoupled case can be controlled by elementary perturbation theories. Thanks to the systematic use of our framework, we have perturbation theories which are uniform on the excited sites.

Let $\underline{c}, \underline{\omega}$ be sequences of r sites and frequencies. We consider k_{ω_i}, parameterizations of invariant tori w.r.t. f_0, the time-one map of just one site. We denote

$$K^* = \sum_{i=1}^r \tau^{c_i} k_{\omega_i}$$

We note that $F_0 \circ K^* = K^* \circ T_{\underline{\omega}}$ and that K^* is uniformly non degenerate.

The hyperbolic splitting for $F_0 \circ K^*$ is

(121) $$\Pi^{s,c,u} = \oplus_{i\in\mathbb{Z}^N} \Pi_i^{s,c,u}$$

where $\Pi_i^{s,c,u}$ is the splitting corresponding to the i torus. If i is an indzex in $\underline{c}$, then we have $\Pi_i^c = \mathrm{Id}_{\mathbb{R}^l}$, $\Pi_i^s = 0$, $\Pi_i^u = 0$. Otherwise, one gets $\Pi_i^c = 0$, and Π_i^s, Π_i^s

are the projections corresponding to the stable and unstable directions at the fixed point.

Notice also that the $rl \times rl$ matrix $DK^\top DK$ is block diagonal. The diagonal has r $l \times l$ blocks $\{Dk_{\omega_i}^\top Dk_{\omega_i}\}_{i=1}^r$.

Lemma 9.13. (Hyperbolicity conditions) *Under the assumptions and the notation in Theorem 3.11. In particular, F_ε is an analytic family of exact symplectic maps in $C_\Gamma^2(\mathcal{B})$. Let $K \in \mathcal{A}_{\rho,\underline{c},\tilde{\Gamma}}$ with $\tilde{\Gamma} < \Gamma$.*

Assume that ε, $\|K - K^*\|_{\rho,\underline{c},\tilde{\Gamma}} \equiv \eta$ are small enough than a number that is independent of $\underline{c}$ and of Γ – it depends only on $\|F\|_{C_\Gamma^2(\mathcal{B}_1)}$, $\|\partial_\varepsilon F\|_{C_\Gamma^2(\mathcal{B}_1 \times \{|\varepsilon| \le \varepsilon^*\})}$, $\|\partial_\varepsilon^2 F\|_{C_\Gamma^2(\mathcal{B}_1 \times \{|\varepsilon| \le \varepsilon^*\})}$ and the hyperbolicity constants of the uncoupled splitting.

Then, K, F_ε satisfies the non-degeneracy conditions in Definition 3.1 with uniform constants.

Proof. We make the elementary remark

$$(122) \quad DF_\varepsilon \circ K = DF_0 \circ K^* + \big(DF_0 \circ K - DF_0 \circ K^*\big)\big(DF_\varepsilon \circ K - DF_0 \circ K\big)$$

and we will control the terms in parenthesis.

By the estimates in composition in Section 10.5, we obtain that:

$$\|DF_0 \circ K - DF_0 \circ K^*\|_{\mathcal{L}_{\rho,\tilde{\Gamma}}} \le C\|K - K^*\|_{\underline{c},\tilde{\Gamma}}$$
$$\|DF_\varepsilon \circ K - DF_0 \circ K\|_{\mathcal{L}_{\tilde{\Gamma}}} \le C|\varepsilon|$$

so we obtain that $\|DF_\varepsilon \circ K - DF_0 \circ K^*\|_{\mathcal{L}_{\rho,\tilde{\Gamma}}}$ is small.

The splitting indicated in (121) is invariant for DF_0. Hence, it is approximately invariant for $DF_\varepsilon \circ K$ and this satisfies the conditions for approximately invariant splittings Definition 3.2.

We note that Lemma 4.2 ensures that, if we place the solutions far apart, the approximately invariant solutions satisfy the hypothesis of Lemma 3.2, and, applying Lemma 4.2, we obtain that there are invariant spaces satisfying Definition (3.1). $\qquad\square$

9.3.4. *Statement and proof of Lemma 9.14. Verifying the non-degeneracy assumptions of coupled solutions.*

Lemma 9.14. (Twist conditions) *Assume that K_1, K_2 are embeddings in $\mathcal{A}_{\rho,\underline{c}_1,\Gamma_\beta}$ $\mathcal{A}_{\rho,\underline{c}_2,\Gamma_\beta}$ both $\underline{c}_1$, $\underline{c}_2$ being finite sequences, and Γ_β as before. Assume that K_1, K_2 satisfy the non-degeneracy conditions in Definition 3.4.*

Then, for m sufficiently large,

$$\tilde{K}(\theta_1, \theta_2) = K_1(\theta_1) + \tau_m K_2(\theta_2)$$

satisfies the non-degeneracy assumptions in Definition 3.4. Furthermore, the non-degeneracy constants of $\tilde{K}$ can be made as close to desired to the constants verified both by K_1, K_2 if we choose $|m|$ large enough.

We will be using the notation that n_1 is the number of sites in $\underline{c}_1$ and that θ_1 stands for all the $n_1 \times l$ variables corresponding to all the sites in $\underline{c}_1$. Similarly for K_2.

Proof. We start by estimating the non-degeneracy condition of the embedding.

We see that the $n_1 + n_2$ matrix $D\tilde{K}^\top DK$ splits naturally into blocks depending on whether we take derivatives with respect to variables in θ_1 or in θ_2.

We will also introduce the notation that $\Phi(m)$ stands for any quantity (vector, matrix, function, etc.) which can be made arbitrarily small by making m large.

$$(123) \qquad D\tilde{K}^\top DK = \begin{pmatrix} D_{\theta_1} K^\top D_{\theta_1} K_1 & D_{\theta_1} K_1^\top \tau^m D_{\theta_2} K_2 \\ \tau^m D_{\theta_2} K_2^\top D_{\theta_1} K_1 & D_{\theta_2} \tau^m K_2^\top D_{\theta_2} \tau^m K_2 \end{pmatrix}.$$

Since

$$D_{\theta_2} \tau^m K_2^\top D_{\theta_2} \tau^m K_2 = (D_{\theta_2} K_2^\top D_{\theta_2} K_2)$$

we see that the diagonal elements of $D\tilde{K}^\top DK$ are precisely those of the uncoupled system and are, therefore invertible.

We will show that the non-diagonal elements in (123) can be made arbitrarily small by choosing m large enough. Then, it will follow that $D\tilde{K}^\top DK$ is invertible and that

$$(124) \quad \tilde{N} = (D\tilde{K}^\top DK)^{-1} = \begin{pmatrix} (D_{\theta_1} K^\top D_{\theta_1} K_1)^{-1} & 0 \\ 0 & (D_{\theta_2} K_2^\top D_{\theta_2} K_2)^{-1} \end{pmatrix} + \Phi(m) =$$

$$\begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix} + \Phi(m)$$

We estimate the off-diagonal elements of (123). We observe that, we can estimate the entries of $n_1 l \times n_2 l$ upper right block as follows:

$$\left| \left(D_{\theta_1} K_1^\top \tau^m D_{\theta_2} K_2 \right)_{p,q} \right| \leq \sum_i \left| \frac{\partial K_{1,i}}{\partial \theta_{1,p}} \frac{\partial K_{2,i+m}}{\partial \theta_{2,q}} \right|$$

$$\leq \|DK_1\|_{\rho,\underline{c}_1,\Gamma_\beta} \|DK_2\|_{\rho,\underline{c}_2,\Gamma} \sum_i \max_k \Gamma_\beta(i - c_{1,k}) \max_l \Gamma(i + m - c_{2,l})$$

$$\leq \frac{2}{1 - e^{-\beta}} \|DK_1\|_{\rho,\underline{c}_1,\Gamma_\beta} \|DK_2\|_{\rho,c_2,\Gamma} \max_{k,l} \Gamma_\beta(c_{2,l} - m - c_{1,k}).$$

Since the block is finite dimensional, making the blocks small enough makes its norm small.

Estimates for the lower left block are obtained from the previous ones by switching $1, 2$ and taking m to $-m$.

Now, we turn to estimating the twist condition. Again the strategy is very similar to the one used in checking the non-degeneracy condition. We just check that the matrix we need to invert is arbitrarily close (by taking $|m|$ large enough) to a matrix which is block diagonal and whose blocks are the non-degeneracy conditions of each of the uncoupled solutions.

We proceed to estimate systematically all the ingredients of A defined in (20). We have first

$$\tilde{P}(\theta) = \left(D_{\theta_1} K_1, \quad D_{\theta_2} [\tau^m K_2] \right) \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix} + \Phi(m)$$

$$= (P_1, P_2) + \Phi(m).$$

Since N_1 and N_2 are finite dimensional matrices, P_1 and P_2 are also in $\mathcal{A}_{\rho,\underline{c}_1,\Gamma}$ and $\mathcal{A}_{\rho,\underline{c}_1,\Gamma}$, respectively.

Using that J_∞ is uncoupled and constant for the models (26) we are considering now, we can write:

$$
(125) \quad
\begin{aligned}
(DFJ_\infty^{-1})\,\big(\tilde{K}(\theta)\big)\,&P_1(\theta_1) \\
&= (DFJ_\infty^{-1})\,\big(K_1(\theta_1)\big)\,P_1(\theta_1) \\
&\quad + \int_0^1 (D^2FJ_\infty^{-1})\,\big(K_1(\theta_1) + s\tau^m K_2(\theta_2)\big)\big(\tau^m K_2(\theta_2), P_1(\theta_1)\big)\,ds
\end{aligned}
$$

$$
(126) \quad
\begin{aligned}
(DFJ_\infty^{-1})\,\big(\tilde{K}(\theta)\big)\,&P_2(\theta_2) \\
&= (DFJ_\infty^{-1})\,\big(\tau^m K_2(\theta_2)\big)\,P_2(\theta_2) \\
&\quad + \int_0^1 (D^2FJ_\infty)\,\big(sK_1(\theta_1) + \tau^m K_2(\theta_2)\big)\,\big(K_1(\theta_1),\,P_2(\theta_2)\big)\,ds\,.
\end{aligned}
$$

We denote T_{21} and T_{12} the integral terms in (125) and (126) respectively. We bound from above the i-th component of T_{21} by

$$
\begin{aligned}
\|J^{-1}\|\sum_{j,n}\left|\frac{\partial^2 F_i}{\partial x_j\,\partial x_n}\right|_\rho\,&|(K_2)_{j+m}|_\rho\,|(P_1)_n|_\rho \\
&\leq \|J^{-1}\|\,\|F\|_{C_\Gamma^2}\,\|K_2\|_\rho\,\|P_1\|_{\rho,\underline{c}_1,\Gamma_\beta} \\
&\quad \times \sum_{j,n}\min\big(\Gamma(i-j),\Gamma(i-n)\big)\,\max_{k,l}\Gamma(j+m-\underline{c}_{2,l})\,\Gamma_\beta(n-c_{1,k})\,.
\end{aligned}
$$

The last sum is bounded by

$$
\frac{4}{1-e^{-\beta}}\,\max_{k,l}\Gamma_\beta(i+m-c_{2,l})\cdot\Gamma_\beta(i-c_{1,k})\leq \Phi(m)\,.
$$

Indeed, let $\mathcal{J}_1(i) = \{j,n\in\mathbb{Z}^N\mid |i-j|\leq |i-n|\}$ and $\mathcal{J}_2(i) = \mathbb{Z}^N\setminus\mathcal{J}_1(i)$. Using that $\Gamma(i)\leq\frac{1}{a}\Gamma_\beta(i)\Gamma_\beta(i)$ the sum is bounded by

$$
\begin{aligned}
&\sum_{j,n\in\mathcal{J}_1(i)}\Gamma(i-j)\max_l\Gamma(j+m-c_{2,l})\,\max_k\Gamma_\beta(n-c_{1,k}) \\
&\quad + \sum_{j,n\in\mathcal{J}_2(i)}\Gamma(i-n)\max_l\Gamma(j+m-c_{2,l})\,\max_k\Gamma_\beta(n-c_{1,k}) \\
&\leq \frac{2}{a}\sum_{j\in\mathbb{Z}^N}\Gamma_\beta(i-j)\max_l\Gamma(j+m-c_{2,l})\sum_{n\in\mathbb{Z}^N}\Gamma_\beta(i-n)\max_k\Gamma_\beta(n-c_{1,k})\,.
\end{aligned}
$$

Analogously T_{12} is bounded by

$$
\|J^{-1}\|\,\|F\|_{C_\Gamma^2}\,\|K_1\|_{\rho,\underline{c}_1,\Gamma_\beta}\,\|P_2\|_\rho\,\frac{4}{1-e^{-\beta}}\,\max_{k,l}\Gamma_\beta(i+m-c_{2,l})\,\Gamma_\beta(i-c_{1,k})\,.
$$

Note that $P_1^\top(\theta+\omega)\,T_{21}(\theta)$ and $P_2^\top(\theta+\omega)\,T_{12}(\theta)$ are bounded by $C\,\max_{k,l}\Gamma_\beta(c_{1,k}+m-c_{2,l})$, where C depends on $\|J^{-1}\|$, $\|F\|_{C_\Gamma^2}$, $\|K_1\|_{\rho,\underline{c}_1,\Gamma_\beta}$, $\|K_2\|_{\rho,\underline{c}_2,\Gamma}$, $\|P_1\|_{\rho,\underline{c}_1,\Gamma_\beta}$ and $\|P_2\|_{\rho,\underline{c}_2,\Gamma}$.

Also note that

$$\left|\left[P_1(\theta+\omega)^\top \ \tilde{J} \ P_2(\theta)\right]_i\right| \le \|J\| \ \|P_1\|_{\rho,\underline{c}_1,\Gamma_\beta} \ \|P_2\|_{\rho,\underline{c}_2,\Gamma}$$

$$\times \sum_{i\in\mathbb{Z}^N} \max_{k,l} \ \Gamma_\beta(i-c_{1,k}) \ \Gamma(i+m-c_{2,l})$$

$$\le C \ \max_{k,l} \ \Gamma_\beta \ (c_{2,l}-m-c_{1,k}).$$

Now we consider the terms $P_1^\top(\theta+\omega) \left(DF\tilde{J}^{-1}\right)\left(\tilde{K}(\theta)\right) P_2(\theta)$ and $P_2^\top(\theta+\omega) \left(DF\tilde{J}^{-1}\right)\left(\tilde{K}(\theta)\right) P_1(\theta)$. We evaluate the first one, the other being analogous. Given $n\in\{1,\dots,r\}$

$$\left|\left[P_1^\top \ (\theta+\omega) \left(DF\tilde{J}\right)\left(\tilde{K}(\theta)\right) \ P_2(\theta)\right]_n\right|$$

$$\le \sum_{i,j} \ \left|(P_1)_{i,n}\right|_\rho \|J^{-1}\| \left\|\frac{\partial F_i}{\partial x_j}\right\|_\rho \ |(P_2)_j|_\rho$$

$$\le \|F\|_{C_\Gamma^1} \ \|J^{-1}\| \ \|P_1\|_{\rho,\underline{c},\Gamma_\beta} \ \|P_2\|_{\rho,c_2,\Gamma} \ \sum_{i,j} \max_{k,l} \ \Gamma_\beta(i-c_{1,k}) \ \Gamma(i-j) \ \Gamma(j+m-c_{2,l})$$

$$\le C \ \max_k \ \Gamma_\beta \ (c_{1,k}+m-c_{2,l}).$$

With all these previous estimates we can write:

$$A = \left[(P_1,P_2)\circ T_\omega + \Phi(m)\right]^\top (DFJ_\infty)(\tilde{K}) \ \left[(P_1,P_2)+\Phi(m)\right] - J_\infty P \circ T_\omega$$

$$= \begin{pmatrix} P_1^\top\circ T_{\underline{\omega}_1}\left[(DFJ_\infty^{-1})(K_1) \ P_1\right]-J_\infty^{-1} \ P_1\circ T_{\underline{\omega}_1} & 0 \\ 0 & P_2^\top\circ T_{\omega_2}\left[(DFJ_\infty)(K_2) \ P_2\right]-J_\infty^{-1} \ P_2\circ T_{\omega_2} \end{pmatrix} +$$

$$\Phi(m) = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} + \Phi(m)$$

which shows that it is invertible if $|m|$ is big enough and that the norm of the inverse of A can be bounded from above by $\max(|A_1^{-1}|,|A_2^{-1}|) + \Phi_m$.

$\square$

The estimates about the non-degeneracy with respect to parameters in the construction are automatic since, in the construction in Section 12, which is the one we use here, the matrix Q is the identity, whose norm is bounded by 1 independently of the number of sites considered and independently of the K considered.

9.4. **Adding oscillating sites inductively.**

Recall that we are assuming that $\varepsilon \le \varepsilon^*$ and that we have a set $\Xi_1(\varepsilon^*) \subset D(\nu_0,\kappa_0) \subset \mathbb{R}^l$ of positive measure such that, for all $\omega \in \Xi_1(\varepsilon^*)$, the system (26) has a breather of frequency ω in $\mathcal{A}_{\rho,\{0\},\Gamma_\beta}$. The non-degeneracy and hyperbolicity constants of all these terms are uniformly bounded.

The remaining part to be shown is that given a sequence $\underline{\omega} \in \mathcal{D} \cup \Xi_1^*(\varepsilon^*)^\infty$, we can find a sequence of tori parameterized by $K_{\underline{\omega}^{(n)}} \in \mathcal{A}_{\rho_n,\underline{c}^{(n)},\Gamma_{\beta_n}}$. Here we have that

$$\omega^{(n)} = (\omega_1,\dots\omega_n)$$

is the sequence of truncations of $\underline{\omega}$ and ρ_n, β_n are decreasing sequences so that $\rho_n \to \rho_\infty > 0$, $\beta_n \to \beta_\infty > 0$.

Our unknowns are ρ_n, β_n, the infinite sequence of centers $\underline{c}^{(n)}$ and the embeddings $K_{\underline{\omega}^{(n)}}$.

The choices of ρ_n, β_n are almost irrelevant for our purposes, so we choose them right away. For example $\rho_n = \rho_\infty + 2^{-n}(\rho_0 - \rho_\infty)$, $\beta_n = \beta_\infty + 2^{-n}(\beta_0 - \beta_\infty)$,

So that now, our only task is to choose a sequence of sites $\underline{c}$, (without loss of generality, we will assume $c_1 = 0$), such that, recursively, we have that $K_{\underline{\omega}^{(n)}} + \tau^{-c_{n+1}} K_{\omega_{n+1}}$ is a very approximate solution of the invariance equation which, furthermore, satisfies uniform hyperbolicity and non-degeneracy conditions. Then, an application of Theorem 3.6 will produce a true solution $K_{\underline{\omega}^{(n+1)}}$.

Of course, we will have to recover the inductive hypothesis we have made to construct this sequence. We will show that, we can ensure that $\|K_{\underline{\omega}^{(n)} - K^*}\|_{\rho_n, \underline{c}^{(n)}, \Gamma_{\beta_n}} \leq \eta/2$ where $\eta > 0$ is the constant introduced in Lemma 9.13 and $K_n^* = K_{\omega_1} + \tau^{-c_2} K_{\omega_2} + \cdots \tau^{-c_n} K_{\omega_n}$.

After this sequence of tori with increasing number of frequencies is produced, we will have to study the limit of the sequence and show that it solves the invariance equation (This will be accomplished in Section 9.5). Note that, since each step changes the number of centers, the convergence of the embeddings cannot be uniform (even in a space of decay functions). Nevertheless, we will show that there is coordinatewise convergence and that this is enough to show that the limit satisfies the invariance equation.

We note that the existence of the sequence and the study of the limit will be accomplished because if we place the centers very far apart from the previously placed ones, we can obtain that the error is small enough to beat the smallness requirements of Theorem 3.6, to ensure that the non-degeneracy and hyperbolicity constants deteriorate an arbitrarily small amount and to ensure the passage to the limit, so that, by recursively assuming that the new center is far away from all the previously placed ones, we can ensure any smallness conditions we wish on the error, on the increment of the distance from the uncoupled solution and on the deterioration of the non-degeneracy and hyperbolicity constants.

We start with K_{ω_1} and K_{ω_2} localized at the node $c_1 = 0$ and c_2 respectively and we take $|c_2|$ big enough so that

$$\tilde{K} = K_{\omega_1} + \tau^{-c_2} \, K_{\omega_2}$$

is a sufficiently approximate solution of $F \circ K - K \circ T_{\underline{\omega}^{(2)}} = 0$ and satisfies both the spectral and the twist non-degeneracy conditions. Then Theorem 3.6 provides the existence of a true invariant torus $K_{\underline{\omega}^{(2)}} \in \mathcal{A}_{\rho_2, \underline{c}^{(2)}, \Gamma_{\beta_2}}$ such that it is non-degenerate and

$$e = \|K_{\underline{\omega}^{(2)}} - \tilde{K}\|_{\rho_2, \underline{c}^{(2)}, \Gamma_{\beta_2}}$$

is small. Actually it can be made as small as we want by taking $|c_2|$ sufficiently big. Remembering that (ω_1, ω_2) is Diophantine (and chosen from the start of the procedure), we see that Theorem 3.6 guarantees that, if we make the initial error small enough, we can produce a solution $K_{\underline{\omega}^{(2)}}$ of the invariance equation with frequency $\underline{\omega}^{(2)}$.

In the $n+1$ step of the process we assume we have the torus $K_{\underline{\omega}^{(n)}} \in \mathcal{A}_{\rho_n, \underline{c}^{(n)}, \Gamma_{\beta_n}}$ localized around the nodes $\underline{c}^{(n)} = (c_1, \ldots, c_n)$, which is non-resonant, that is $K_{\underline{\omega}^{(n)}} \in ND_{\mathrm{loc}}(\rho_n, \Gamma_{\beta_n})$

We consider the parameterization

$$\tilde{K}(\theta) - K_{\underline{\omega}^{(n)}}(\theta_1) + \tau^{m_{n+1}} \, K_{\omega_{n+1}}(\theta_2), \qquad \theta = (\theta_1, \theta_2) \in \mathbb{T}^{nl} \times \mathbb{T}^l,$$

as an approximation for the new torus, which we will denote $K_{\underline{\omega}^{(n+1)}}$, with some $m_{n+1} \in \mathbb{Z}^N$.

By the coupling lemma (Lemma 9.8) if we take a suitable m_{n+1} big enough we obtain

$$E_{n+1} = F \circ \tilde{K} - \tilde{K} \circ T_{\underline{\omega}^{(n+1)}}$$

as small as we want. In particular we take $0 < \delta_{n+1} < \min\left(1, \rho_n/12, (\rho_{n+1}-\rho_n)/6\right)$ and we require

$$C\kappa_{n+1}^4 \ \delta_{n+1}^{-4\nu_{n+1}} \ \|E_{n+1}\|_{\rho_{n+1},\underline{c}^{(n+1)},\Gamma_{\beta_{n+1}}} \leq 1 \,,$$

and

$$C\kappa_{n+1}^2 \ \delta_{n+1}^{-2\nu_{n+1}} \ \|E_{n+1}\|_{\rho_{n+1},\underline{c}^{(n+1)},\Gamma_{\beta_{n+1}}} \leq \frac{e}{2^{n-1}} \,.$$

We denote $c_{n+1} = -m_{n+1}$. Then Theorem 3.6 provides a true invariant torus $K_{\underline{\omega}^{(n+1)}} \in \mathcal{A}_{\rho_{n+1},\underline{c}^{(n+1)},\Gamma_{\beta_{n+1}}}$, non-degenerate and satisfying the estimate

$$\|K_{\underline{\omega}^{(n+1)}} - \tilde{K}\|_{\rho_{n+1},\underline{c}^{(n+1)},\Gamma_{\beta_{n+1}}} \leq \frac{e}{2^{n-1}} \,.$$

9.5. Passage to the limit (Part B of Theorem 3.11).

The issue now is to study the limit $n \to \infty$. Thanks to our weighted spaces and the fact that the solutions we construct have bumps whose distance from each other tends to infinity fast enough, we can prove the following Lemma 9.15 which establishes that for any bounded sets in the lattice, the trajectories of the particles in this set converge uniformly.

Lemma 9.15. *The sequence $\left\{K_{\underline{\omega}^{(n)}}\right\}_{n\geq 1}$ converges component-wise and uniformly on every compact set of $(\mathbb{T}^l)^{\mathbb{N}}$. We denote $K_{\underline{\omega}_\infty}$ the limit obtained in this sense. Furthermore, each component of $K_{\underline{\omega}_\infty}$ is analytic from $(\mathbb{T}^l)^{\mathbb{N}}$ into M.*

Remark 9.16. Here, by analytic on the infinite-dimensional torus $(\mathbb{T}^l)^{\mathbb{N}}$, we mean $K_{\underline{\omega}_\infty}$ writes component-wise

$$(K_{\underline{\omega}_\infty})_i(\underline{\theta}) = \sum_{n\geq 0}(H^{(n)})_i(\theta_1,\ldots,\theta_n),$$

where $(H^{(n)})_i(\theta_1,\ldots,\theta_n)$ are analytic in the usual sense on $(\mathbb{T}^l)^n$ and moreover we have

$$\sum_{n\geq 0}\|(H^{(n)})_i\|_{\rho_n} < \infty,$$

where $D_{\rho_n} \supset (\mathbb{T}^l)^n$.

Proof. We represent $K_{\underline{\omega}_\infty}(\underline{\theta})$ as

$$\lim_{n\to\infty} K_{\underline{\omega}^{(n)}}(\underline{\theta})$$

$$(127) \qquad = K_{\omega_1}(\theta_1) + \sum_{n=1}^{\infty}\left[K_{\underline{\omega}^{(n+1)}}(\theta_1,\ldots,\theta_{n+1}) - K_{\underline{\omega}^{(n)}}(\theta_1,\ldots,\theta_n)\right].$$

We fix $i \in \mathbb{Z}^N$ and we estimate the i-th component of $K_{\underline{\omega}^{(n+1)}} - K_{\underline{\omega}^{(n)}}$. By the triangle inequality

$$\left|\left[K_{\underline{\omega}^{(n+1)}} - K_{\underline{\omega}^{(n)}}\right]_i\right|_{\rho_{n+1}} \leq \left|\left[K_{\underline{\omega}^{(n+1)}} - K_{\underline{\omega}^{(n)}} - \tau^{m_{n+1}} K_{\omega_{n+1}}\right]_i\right|_{\rho_{n+1}}$$

$$(128) \qquad\qquad\qquad + \left|\tau^{m_{n+1}} K_{\omega_{n+1}}\right]_i\right|_{\rho_{n+1}} \,.$$

The first term in the right-hand side of (128) is bounded by

$$\frac{e}{2^{n-1}} \max_{1 \leq k \leq n+1} \Gamma_{\beta_{n+1}} (i - c_k)$$

and the second one is bounded by

$$\left| \left[\tau^{m_{n+1}} K_{\omega_{n+1}} \right]_i \right|_{\rho_1} = \left| \left[K_{\omega_{n+1}} \right]_{i+m_{n+1}} \right|_{\rho_1} \leq \| K_{\omega_{n+1}} \|_{\rho_1, 0, \Gamma} \, \Gamma(i + m_{n+1}).$$

This implies that the i-th component of the sum in (127) is bounded by

$$(129) \qquad \sum_{n=1}^{\infty} \frac{e}{2^{n-1}} + \sum_{n=1}^{\infty} \| K_{\omega_{n+1}} \|_{\rho_1, 0, \Gamma} \, \Gamma(i + m_{n+1}) \,.$$

Therefore the i-th component of (127) converges uniformly on compact sets of $(\mathbb{T}^l)^{\mathbb{N}}$ and $K_{\underline{\omega}_\infty}$ is analytic in the sense of Remark 9.16. $\square$

The following result proves that $K_{\underline{\omega}_\infty}$ is a solution of the invariance equation and therefore is an almost-periodic function of the initial system.

Lemma 9.17. *The limit function $K_{\underline{\omega}_\infty}$ satisfies*

$$F \circ K_{\underline{\omega}_\infty} = K_{\underline{\omega}_\infty} \circ T_{\underline{\omega}_\infty}$$

component-wise.

Proof. For every $n \in \mathbb{N}$ one has

$$(130) \qquad\qquad F \circ K_{\underline{\omega}^{(n)}} = K_{\underline{\omega}^{(n)}} \circ T_{\underline{\omega}^{(n)}} \,.$$

We fix a component $i \in \mathbb{Z}^N$. The passage to the limit in the right-hand side of (130) is immediate. For the left-hand side we take n_0 such that $|c_n| > |i|$ for $n > n_0$. Then for $n > n_0$ we have

$$(131) \qquad \left| F_i \circ K_{\underline{\omega}^{(n)}} - F_i \circ K_{\underline{\omega}_\infty} \right|_{\rho_\infty} \leq \sum_j \left| \frac{\partial F_i}{\partial x_j} \right| \, \left| \left[K_{\underline{\omega}^{(n)}} - K_{\underline{\omega}_\infty} \right]_j \right|_{\rho_\infty} \,.$$

We estimate

$$\left| \left[K_{\underline{\omega}^{(n)}} - K_{\underline{\omega}_\infty} \right]_j \right|_{\rho_\infty} \leq \sum_{p=n}^{\infty} \left| \left[K_{\underline{\omega}^{(p)}} - K_{\underline{\omega}^{(p+1)}} \right]_j \right|_{\rho_{p+1}}$$

$$\leq \sum_{p=n}^{\infty} \left[\frac{e}{2^{p-1}} \max_{1 \leq k \leq p+1} \Gamma_{\beta_{p+1}} (j - c_k) + \left| \left[\tau^{-c_{p+1}} K_{\omega_{p+1}} \right]_j \right|_{\rho_{p+1}} \right]$$

$$\leq \sum_{p=n}^{\infty} \left[\frac{e}{2^{p-1}} \max_{1 \leq k \leq p+1} \Gamma_{\beta_\infty} (j - c_k) + \| K_{\omega_{p+1}} \|_{\rho_1, 0, \Gamma} \Gamma(j - c_{p+1}) \right] \,.$$

Then (131) is bounded by

$$\sum_{p=n}^{\infty} \|F\|_{C_{\Gamma}^1} \Big[\frac{e}{2^{p-1}} \sum_j \max_{1 \leq k \leq p+1} \Gamma_{\beta_\infty}(j - c_k)\Gamma(i - j)$$

$$+ \|K_{\omega_{p+1}}\|_{\rho_1,0,\Gamma} \sum_j \Gamma(i - j)\Gamma(j - c_{p+1}) \Big]$$

$$\leq \|F\|_{C_{\Gamma}^1} \Big(\sum_{p=n}^{\infty} \frac{e}{2^{p-1}} \frac{2}{1 - e^{-\beta_\infty}} + \sum_{p=n}^{\infty} \|K_{\omega_{p+1}}\|_{\rho_1,0,\Gamma} \Gamma(i - c_{p+1}) \Big)$$

$$\|F\|_{C_{\Gamma}^1} \Big(\frac{2}{1 - e^{-\beta_\infty}} \frac{e}{2^{n-2}} + C_K \frac{2}{1 - e^{-\beta}} \max_{p \geq n} \Gamma(i - c_{p+1}) \Big),$$

which leads to the desired result. $\qquad\square$

ACKNOWLEDGEMENTS

10. APPENDIX: DECAY FUNCTIONS

This appendix is devoted to the properties of spaces of decay functions.

10.1. Linear and k-linear maps over $\ell^\infty(\mathbb{Z}^N)$.
We are going to consider linear maps from $\ell^\infty(\mathbb{Z}^N)$ into itself such that

$$(132) \qquad \lim_{m \to \infty} \sup_{\substack{u \in \ell^\infty, \, |u| \leq 1 \\ u_j = 0, \, |j - i| \leq m}} (Au)_i = 0, \qquad \forall i \in \mathbb{Z}^M.$$

The condition (132) is equivalent to the fact that A can be written in the form

$$(133) \qquad (Au)_i = \sum_{j \in \mathbb{Z}^N} A_{ij} u_j, \qquad i \in \mathbb{Z}^N,$$

where A_{ij} are linear maps, $u \in \ell^\infty(\mathbb{Z}^N)$ and the series are convergent.

We denote by $\mathcal{L}(\ell^\infty(\mathbb{Z}^N))$ the space of linear maps that satisfy (133). This is a non-trivial assumption, since $\ell^\infty(\mathbb{Z}^N)$ is not reflexive. Furthermore, the space $\mathcal{L}(\ell^\infty(\mathbb{Z}^N))$ is a strict subspace of bounded linear operators on $\ell^\infty(\mathbb{Z}^N)$.

The second assumption we will make is that there exists $C > 0$ and a decay function Γ such that

$$|A_{ij}| \leq C\Gamma(i - j),$$

for all $(i, j) \in (\mathbb{Z}^N)^2$.

In this case,

$$(134) \qquad \sum_{j \in \mathbb{Z}^N} |A_{ij} u_j| \leq \sum_{j \in \mathbb{Z}^N} C\Gamma(i - j)|u_j| \leq \sum_{j \in \mathbb{Z}^N} C\Gamma(i - j)\|u\|_\infty \leq C\|u\|_\infty.$$

Then we define
$$\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N)) = \left\{ A \in \mathcal{L}(\ell^\infty(\mathbb{Z}^N)) \mid \sup_{i,j \in \mathbb{Z}^N} \Gamma(i-j)^{-1}|A_{ij}| < \infty \right\}$$

and we endow it with the norm

(135)
$$\|A\|_\Gamma = \sup_{i,j \in \mathbb{Z}^N} \Gamma(i-j)^{-1}|A_{ij}|.$$

The following lemma has a simple proof that can be found in [FdlLM07]. The most subtle point is that we need to verify that the linear operators are given by the matrix.

Lemma 10.1. *The space $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$ is a Banach space.*

From the definition of the norm of A and the inequalities (134) we deduce that

$$\|Au\|_\infty \le \|A\|_\Gamma \|u\|_\infty$$

for all $u \in \ell^\infty(\mathbb{Z}^N)$.

Remark 10.2. The previous definition will also be used for matrices. If one considers a *finite* set of indexes $I \times J \subset \mathbb{Z}^N \times \mathbb{Z}^N$, we will use

$$\|A\|_\Gamma = \sup_{i \in I, j \in J} |A_{ij}| \Gamma^{-1}(i-j).$$

This is just a way to say that the set of tensors of order 2 with finite indexes are naturally embedded into the set of tensors of order 2 on $\mathbb{Z}^N$ by just setting all the remaining values to 0.

Similarly to the previous definition, we define $\mathcal{L}^k(\ell^\infty(\mathbb{Z}^N))$ as the space of k-linear maps on $\ell^\infty(\mathbb{Z}^N)$ which are represented by a multilinear matrix.

(136)
$$B(u^1, \ldots, u^k)_i = \sum_{(i_1, \ldots, i_k) \in (\mathbb{Z}^N)^k} B_{i,i_1,\ldots,i_k} u^1_{i_1} \ldots u^k_{i_k},$$

where $i, i_1, \ldots, i_k \in \mathbb{Z}^N$, $(u^1, \ldots, u^k) \in (\ell^\infty(\mathbb{Z}^N))^k$ and $B_{i,i_1,\ldots,i_k} \in \mathcal{L}^k(M,M)$.

Given a decay function Γ, we define $\mathcal{L}^k_\Gamma(\ell^\infty(\mathbb{Z}^N))$ as the space of maps in $\mathcal{L}^k(\ell^\infty(\mathbb{Z}^N))$ such

$$|B_{i,i_1,\ldots,i_k}| \le C \min(\Gamma(i-i_1), \ldots, \Gamma(i-i_k)),$$

for some $C \ge 0$.

We define

$$\|B\|_\Gamma = \sup_{i,i_1,\ldots,i_k \in \mathbb{Z}^N} |B_{i,i_1,\ldots,i_k}| \max(\Gamma^{-1}(i-i_1), \ldots, \Gamma^{-1}(i-i_k)).$$

We have the following lemma (see [FdlLM07]).

Lemma 10.3. *The space $\mathcal{L}^k_\Gamma(\ell^\infty(\mathbb{Z}^N))$ is a Banach space.*

The next result provides the Banach algebra property of $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$. This property will be crucial for our estimates. It makes it possible to work with infinite dimensional systems in a way which is not very different from the finite dimensional case.

Lemma 10.4. *If $A, B \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$ then $AB \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$ and we have the estimate*

$$\|AB\|_\Gamma \le \|A\|_\Gamma \|B\|_\Gamma.$$

Proof. It is easy to verify that if A and B can be represented by matrices, so is the product. Since $A, B \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$, we have

$$|A_{ij}| \leq \Gamma(i-j)\|A\|_\Gamma, \quad |B_{jk}| \leq \Gamma(j-k)\|B\|_\Gamma.$$

Therefore, we have:

$$\begin{aligned}
\|AB\|_\Gamma &= \sup_{n,m\in\mathbb{Z}^N} |(AB)_{nm}|\Gamma^{-1}(n-m) \\
&\leq \sup_{n,m\in\mathbb{Z}^N} \sum_{k\in\mathbb{Z}^N} |A_{nk}||B_{km}|\Gamma^{-1}(n-m) \\
&\leq \|A\|_\Gamma\|B\|_\Gamma \sup_{n,m\in\mathbb{Z}^N} \sum_{k\in\mathbb{Z}^N} \Gamma(n-k)\Gamma(k-m)\Gamma^{-1}(n-m).
\end{aligned}$$

Using property (2) of Definition 2.2 we obtain the desired result. $\qquad\square$

By induction on k the same result holds for $k-$linear maps. See [FdlLM07].

Lemma 10.5. *Let $A \in \mathcal{L}_\Gamma^k(\ell^\infty(\mathbb{Z}^N))$ and $B_j \in \mathcal{L}_\Gamma^{n_j}(\ell^\infty(\mathbb{Z}^N))$ for $1 \leq j \leq k$. Then the composition $AB_1 \ldots B_k \in \mathcal{L}_\Gamma^{n_1+\cdots+n_k}(\ell^\infty(\mathbb{Z}^N))$ and*

$$\|AB_1 \ldots B_k\|_\Gamma \leq \|A\|_\Gamma\|B_1\|_\Gamma \ldots \|B_k\|_\Gamma.$$

10.2. **Spaces of differentiable and analytic functions on lattices.** We now define the space of C^r functions, based on the previous weighted norms. Given an open set $\mathcal{B} \subset \mathcal{M}$ we define

$$C_\Gamma^1(\mathcal{B}) = \left\{ \begin{array}{c} F : \mathcal{B} \to \mathcal{M} \mid F \in C^1(\mathcal{B}), \ DF(x) \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N)), \\ \sup_{x\in\mathcal{B}} \|F(x)\| < \infty, \ \sup_{x\in\mathcal{B}} \|DF(x)\|_\Gamma < \infty \end{array} \right\}.$$

We endow $C_\Gamma^1(\mathcal{B})$ with the norm $\|F\|_{C_\Gamma^1} = \max\big(\sup_{x\in\mathcal{B}}\|F(x)\|,$ $\sup_{x\in\mathcal{B}}\|DF(x)\|_\Gamma\big)$. In the definition of $C_\Gamma^1(\mathcal{B})$, when $\mathcal{M}$ is complex, the derivative has to be understood as complex derivative.

We emphasize that the definition of C_Γ^1 includes that $DF(x) \in \mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$ and, in particular, that the derivative of the function is given by the matrix of its partial derivatives. Concretely if $F \in C_\Gamma^1$ then we have the following formula:

$$DF_i(x)v = \sum_{j\in\mathbb{Z}^N} \frac{\partial F_i}{\partial x_j}(x)v_j,$$

where x_j is the variable in M, the j-th component of $\mathcal{M}$.

Now we proceed to define the space of finite differentiable maps. In [JdlL00, FdlLM07], one can find definitions for Hölder spaces, which is useful for other applications (such as thermodynamic formalism).

Definition 10.6. *Given $r \in \mathbb{N}$*

$$C_\Gamma^r(\mathcal{B}) = \left\{ \begin{array}{c} F : \mathcal{B} \to \mathcal{M} \mid F \in C^r(\mathcal{B}), \ D^j F(x) \in C_\Gamma^1(\mathcal{B}), \\ 0 \leq j \leq r-1 \end{array} \right\}.$$

10.3. **Spaces of embeddings from $\mathbb{C}^l$ to $\mathcal{M}$ with decay properties.** In this section, we will consider embeddings from finite dimensional tori into the phase space. We should think of these embeddings as describing some oscillations centered around some sites.

We define the complex strip

$$D_\rho = \big\{z \in \mathbb{C}^l/\mathbb{Z}^l \mid |\mathrm{Im}\, z_i| < \rho, \ i = 1,\ldots,l\big\}.$$

Let $R \geq 1$ be an integer and consider $\underline{c} \in (\mathbb{Z}^N)^R$, i.e.

$$\underline{c} = (c_1, \ldots, c_R).$$

We introduce the following quantity

$$\|f\|_{\rho,\underline{c},\Gamma} = \sup_{i \in \mathbb{Z}^N} \min_{j=1,\ldots,R} \Gamma^{-1}(i - c_j)\|f_i\|_\rho,$$

where

$$\|f_i\|_\rho = \sup_{\theta \in D_\rho} |f_i(\theta)|.$$

Definition 10.7. *We denote*

$$(137) \qquad \mathcal{A}_{\rho,\underline{c},\Gamma} = \left\{ \begin{array}{c} f : D_\rho \to \mathcal{M} \mid f \in C^0(\overline{D}_\rho),\ f\ analytic\ in\ D_\rho, \\ \|f\|_{\rho,\underline{c},\Gamma} < \infty \end{array} \right\}.$$

This space, with the norm $\|\cdot\|_{\rho,\underline{c},\Gamma}$, is a Banach space.

The parameter $\underline{c}$ is the location of the centers of the oscillations of the map $f : \mathbb{T}^l \to \mathcal{M}$. As the argument of f changes, the range of the embedding, will oscillate mainly on the sites in neighborhoods of $c_1, \ldots, c_R$.

The next result is a version of the Cauchy estimates in our context.

Lemma 10.8. *Let $f : D_\rho \to \mathcal{M}$ be an analytic function. Then for all $\delta \in (0, \rho)$, the following holds*

$$\|D_\theta f\|_{\rho-\delta,\underline{c},\Gamma} \leq l\delta^{-1}\|f\|_{\rho,\underline{c},\Gamma}.$$

Proof. Consider the components f_j of f, for $j \in \mathbb{Z}^N$. Each f_j maps D_ρ into M and we have the standard Cauchy estimates for $k = 1, \ldots, l$

$$\|\partial_{\theta_k} f_j\|_{\rho-\delta} \leq \delta^{-1}\|f_j\|_\rho.$$

The result then follows by just multiplying this inequality by $\Gamma^{-1}(i - c_j)$ which is positive, summing with respect to k and taking the supremum for i and the minimum for j. □

Remark 10.9. Note that in the previous Cauchy estimate the bound depends *linearly* on the dimension of the torus l. This will be important when we consider the limit of many dimensions.

On the other hand, taking the supremum over components, makes it clear that the bounds are independent of the dimension of the range. In particular, we can discuss mappings into infinite dimensions. Note that, if we had taken an other norm, the constants would have depended on the dimension of the range.

If we consider a map A from D_ρ into the set of linear maps $\mathcal{L}_\Gamma(\ell^\infty(\mathbb{Z}^N))$, the associated norm is

$$\|A\|_{\rho,\Gamma} = \sup_{i,j \in \mathbb{Z}^N} \sup_{\theta \in D_\rho} \Gamma^{-1}(i - j)|A_{ij}(\theta)| = \sup_{\theta \in D_\rho} \|A(\theta)\|_\Gamma$$

Remark 10.10. The previous definition of the space of analytic maps in the strip can be generalized to any open subset $\mathcal{B}$ of the complex extended manifold $\mathcal{M}^{\mathbb{C}}$. We define

$$\mathcal{A}_{\mathcal{B},\underline{c},\Gamma} = \left\{ \begin{array}{c} F : \mathcal{B} \to \mathcal{M} \mid F \in C^0(\overline{\mathcal{B}}),\ F\ analytic\ in\ \mathcal{B}, \\ \|F\|_{\mathcal{B},\underline{c},\Gamma} < \infty \end{array} \right\}$$

where

$$\|F\|_{\mathcal{B},\underline{c},\Gamma} = \sup_{i\in\mathbb{Z}^N} \min_{j=1,\dots,R} \Gamma^{-1}(i-c_j)\sup_{z\in\mathcal{B}}|F_i(z)|.$$

Note that the space $\mathcal{A}_{\mathcal{B},\underline{c},\Gamma}$ is a closed subspace of $C^1_\Gamma(\mathcal{B})$ for the C^1_Γ topology. Here, the derivatives are understood as complex derivatives.

10.4. **Spaces of localized vectors.** The space of localized vectors $\ell^\infty_{\underline{c},\Gamma}$, plays a role as the space of infinitesimal deformation of the space of localized embeddings defined above. We also isolate a class of linear operators $\mathcal{L}_{\underline{c},\Gamma}$ which send ℓ^∞ into $\ell^\infty_{\underline{c},\Gamma}$.

The key property (Proposition 10.12) is that $\mathcal{L}_{\underline{c},\Gamma}$ is an ideal of the Banach algebra $\mathcal{L}_\Gamma$. It will be important for future developments that the bounds obtained are independent of the parameter $\underline{c}$. This will be an easy consequence of the Banach algebra properties of the decay functions.

Definition 10.11. *Given a decay function Γ and a finite (or infinite) collection of sites $\underline{c} = \{c_k\}_{k\in\mathcal{K}} \subset \mathbb{Z}^N$ we define*

(138)
$$\|v\|_{\underline{c},\Gamma} = \sup_{i\in\mathbb{Z}^N} \inf_{k\in\mathcal{K}} |v_i|\Gamma(i-c_k)^{-1}.$$

We denote

$$\ell^\infty_{\underline{c},\Gamma} = \{v \in (\mathbb{R}^l)^{\mathbb{Z}^N} \mid \|v\|_{\underline{c},\Gamma} < \infty\}.$$

We denote by $\mathcal{L}_{\underline{c},\Gamma}$ the space of linear operators on ℓ^∞ such that

$$(Av)_i = \sum_{j\in\mathbb{Z}^N} A_{ij}v_j,$$

$$|A_{ij}| \le C\min(\sup_{k\in\mathcal{K}}\Gamma(i-c_k),\Gamma(i-j)).$$

We denote by $\|A\|_{\underline{c},\Gamma}$ the best constant C above, i.e.

$$\|A\|_{\underline{c},\Gamma} = \max(\sup_{i,j\in\mathbb{Z}^N}|A_{ij}|\Gamma^{-1}(i-j), \sup_{i,j\in\mathbb{Z}^N}|A_{ij}|\Gamma^{-1}(i-j), \sup_{i,j\in\mathbb{Z}^N}\inf_{k\in\mathcal{K}}|A_{ij}|\Gamma^{-1}(i-c_k))$$

Note that we use $\|\cdot\|_{\underline{c},\Gamma}$ both for the norm in a linear space and in the space of operators. This will not cause any confusion since in this space we will not use the norm of operators from $\ell^\infty_{\underline{c},\Gamma}$ to itself.

The following is an easy exercise.

Proposition 10.12. *We have the following results:*
- *a) The space $\ell^\infty_{\underline{c},\Gamma}$ endowed with $\|\cdot\|_{\underline{c},\Gamma}$ is a Banach space.*
  The embedding $\ell^\infty_{\underline{c},\Gamma} \hookrightarrow \ell^\infty$ is continuous.
  $\ell^\infty_{\underline{c},\Gamma}$ is a closed subspace of ℓ^∞.
- *b) The space $\mathcal{L}_{\underline{c},\Gamma}$ endowed with $\|\cdot\|_{\underline{c},\Gamma}$ is a Banach space.*
  The embedding $\mathcal{L}_{\underline{c},\Gamma} \to \mathcal{L}_\Gamma$ is continuous.
  $\mathcal{L}_{\underline{c},\Gamma}$ is a closed subspace of $\mathcal{L}_\Gamma$.
- *c) Ideal character: If $A \in \mathcal{L}_{\underline{c},\Gamma}$, $B \in \mathcal{L}_\Gamma$, we have*

(139)
$$AB \in \mathcal{L}_{\underline{c},\Gamma} \quad \|AB\|_{\underline{c},\Gamma} \le \|A\|_{\underline{c},\Gamma}\|B\|_\Gamma,$$
$$BA \in \mathcal{L}_{\underline{c},\Gamma} \quad \|BA\|_{\underline{c},\Gamma} \le \|A\|_{\underline{c},\Gamma}\|B\|_\Gamma$$

As a consequence of the above, if $A \in \mathcal{L}_{\underline{c},\Gamma}$, $B \in \mathcal{L}_{\underline{c},\Gamma}$, we have

(140)
$$\|AB\|_{\underline{c},\Gamma},\|BA\|_{\underline{c},\Gamma} \le \|A\|_{\underline{c},\Gamma}\|B\|_{\underline{c},\Gamma}.$$

10.5. **Regularity of the composition operators.** The following propositions (see [JdlL00]) establish the regularity of composition operators and provide estimates for the composition.

Proposition 10.13. *The mapping defined by*

$$\mathcal{C}(G, h) = G \circ h$$

is locally Lipschitz when considered as

$$\mathcal{C} : C^2_\Gamma \times C^1_\Gamma \mapsto C^1_\Gamma,$$

and we have the estimate

$$\|\mathcal{C}(G, h + \bar{h}) - \mathcal{C}(G, h)\|_{C^1_\Gamma} \leq \|G\|_{C^2_\Gamma}\|\bar{h}\|_{C^1_\Gamma}(1 + \|h\|_{C^1_\Gamma}),$$

Furthermore, when considered as a mapping from $C^3_\Gamma \times C^1_\Gamma$ into C^1_Γ, we have the formula

$$(141) \qquad\qquad D_2\mathcal{C}(G, h)\Delta = (DG \circ h)\,\Delta.$$

We will also need the following estimate on the composition operator.

Lemma 10.14. *Consider two functions $G, h \in C^1_\Gamma$. Then we have*

$$\|G \circ h\|_{C^1_\Gamma} \leq C \max(\|G\|_{C^0}, \|G\|_{C^1_\Gamma}\|h\|_{C^1_\Gamma}).$$

Proof. Clearly, we have

$$\|G \circ h\|_{C^0} \leq \|G\|_{C^0}.$$

We now estimate the norm of the derivatives:

$$D_j(G \circ h)_i = \sum_{k \in \mathbb{Z}^N} D_k G_i \circ h D_j h_k.$$

This leads

$$|D_j(G \circ h)_i| \leq \sum_{k \in \mathbb{Z}^N} \Gamma(i - k)\Gamma(k - j)\|G\|_{C^1_\Gamma}\|h\|_{C^1_\Gamma}$$

$$\leq \Gamma(i - j)\|G\|_{C^1_\Gamma}\|h\|_{C^1_\Gamma}.$$

This ends the proof.   □

The next lemma gives an estimate on the composition of a mapping defined on the manifold and an embedding.

Lemma 10.15. *Let $\mathcal{B} \subset \mathcal{M}$ be a star-like from the origin open set such that $0 \in \mathcal{B}$. Assume that $F \in C^1_\Gamma(\mathcal{B})$ and is analytic. Let $K : D_\rho \to \mathcal{M}$ belong to $\mathcal{A}_{\rho,\underline{c},\Gamma}$, with $\underline{c} \in (\mathbb{Z}^N)^R$ and such that $K(\overline{D}_\rho) \subset \mathcal{B}$.*

(1) Assume that $F(0) = 0$. Then $F \circ K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, and

$$(142) \qquad\qquad \|F \circ K\|_{\rho,\underline{c},\Gamma} \leq R\|F\|_{C^1_\Gamma}\,\|K\|_{\rho,\underline{c},\Gamma}.$$

(2) Let $\mathcal{J} \subset \mathbb{Z}^N$ be a finite set of indexes and assume that $F_j(0) = 0$ for $j \in \mathbb{Z}^N - \mathcal{J}$. Then $F \circ K \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ and

$$(143) \qquad\qquad \|F \circ K\|_{\rho,\underline{c},\Gamma} \leq \|F\|_{C^1_\Gamma}\left(C + R\|K\|_{\rho,\underline{c},\Gamma}\right),$$

where C depends on $\underline{c}$ and $\mathcal{J}$.

In both cases

$$(144) \qquad\qquad \|D(F \circ K)\|_{\rho,\underline{c},\Gamma} \leq R\|F\|_{C^1_\Gamma}\,\|DK\|_{\rho,\underline{c},\Gamma}.$$

Proof. By Definition 2.6 of analytic functions, we have that $F \circ K$ is analytic. To estimate $\|F \circ K\|_{\rho,\underline{c},\Gamma}$ take $i \in \mathbb{Z}^N$ and $j \in \{1,\ldots,R\}$. If $F_i(0) = 0$ we can write

$$F_i\big(K(\theta)\big) = \int_0^1 DF_i\big(sK(\theta)\big)\,K(\theta)ds = \int_0^1 \sum_{p \in \mathbb{Z}^N} \frac{\partial F_i}{\partial x_p}\big(sK(\theta)\big)\,K_p(\theta)ds\,.$$

Taking norms

$$|F_i\big(K(\theta)\big)| \leq \sum_{p \in \mathbb{Z}^N} \|DF\|_\Gamma\,\Gamma(i-p)\,\|K\|_{\rho,\underline{c},\Gamma}\,\max_{1 \leq j \leq R}\,\Gamma(p-c_j)$$

$$\leq R\|F\|_{C_\Gamma^1}\,\|K\|_{\rho,\underline{c},\Gamma}\,\max_{1 \leq j \leq R}\,\Gamma(i-c_j)$$

and then, if $F(0) = 0$, (142) follows.

In the second case $F_m(0) \neq 0$ for $m \in \mathcal{J}$, we have

$$|F_m(K(\theta))| \leq |F_m(0)| + \|F\|_{C_\Gamma^1}\,\|K\|_{\rho,\underline{c},\Gamma}\,\max_{1 \leq j \leq R}\,\Gamma(i-c_j)$$

and then we obtain (143) with

$$C = \max_{m \in \mathcal{J}}\,\min_{1 \leq j \leq R}\,\Gamma^{-1}(m-c_j)\,.$$

The estimate (144) follows from the chain rule and the definitions of the norms. □

11. Appendix: Symplectic geometry on lattices

A symplectic structure on an infinite dimensional manifold is not easy to define (see [CM74], [Bam99]). Fortunately, the KAM theory presented in here uses only very few properties of symplectic geometry.

The aim of the next sections is to develop such ideas and give precise definitions of the theory of symplectic forms we will need. Note that we do not need to develop a systematic geometry. We just need to deal with the standard symplectic form, its primitives, its push-forward and perform just a few operations. This can be readily justified in spite of the difficulties with more sophisticated material.

11.1. **Forms on lattices.** Remember that a form is just an antisymmetric real valued multilinear operator on the tangent space.

We just need to study local forms which are the product of forms in each of the ambient spaces.

We introduce $\pi_i : \mathcal{M} \to M$, the projection $\pi_i(x) = x_i$ for $i \in \mathbb{Z}^N$. Given a collection of smooth k-forms $\gamma_i \in \Lambda^k(M)$, such that $\sup_i \|\gamma_i\| < \infty$, we define a formal form in $\mathcal{M}$ as follows

(145)
$$\gamma = \sum_{i \in \mathbb{Z}^N} \pi_i^* \gamma_i\,,$$

that is

$$\gamma(x)(u_1,\ldots,u_k) = \sum_{i \in \mathbb{Z}^N} \gamma_i(\pi_i(x))(\pi_i u_1,\ldots,\pi_i u_k)$$

for $x \in \mathcal{M}$ and $(u_1,\ldots,u_k) \in (T_x\mathcal{M})^k$. We denote

$$\bar{\Lambda}_\infty^k(\mathcal{M}) = \left\{ \gamma = \sum_i \pi_i^* \gamma_i \right\}$$

the set of such forms.

Of course, this form (145) in general does not define a multilinear function on bounded vector fields, so that it should be understood only formally. Nevertheless, we will show that there are several operations among forms that can be made sense of in the infinite dimensional setting.

Roughly, we will see that these formal forms make sense acting on vectors that decay away from a finite set of centers. We can also push them forward by a decay diffeomorphism and pull them back by a decay embedding. They can also be integrated and, in some weak sense, differentiated. These will be all the operations that we will need. Moreover, we will only need $k = 1, 2$.

When $k = 2$, if each of the γ_i are uniformly non-degenerate, we can define an identification operator defined by

$$(146) \qquad (J_\infty u)_i = J_i \pi_i u \quad i \in \mathbb{Z}^n$$

where J_i is the operator of identification on the i copy of the manifold i.e.

$$\gamma_i(x)(\xi, \eta) = \langle(\xi, J_i(x)\eta) \ \forall \xi, \eta \in T_x M_i.$$

We emphasize that, given the formula (146), it is clear that when the γ_i are uniformly non-degenerate (i.e. $\|J_i(x)\|, \|J_i^{-1}(x)\|$ are bounded uniformly in i, x) we have that the operator J_∞ is bounded and its inverse is also bounded. Note that in the KAM method of [dlLGJV05, FdlLS09a, FdlLS09b], the symplectic properties appear mainly through J, J^{-1} and their invariance properties.

In the main application to the construction of almost periodic solutions, when the system has translation invariance, all the γ_i are identical. Nevertheless we do not assume that the γ_i are given in the standard form. This is useful e.g. in dealing with oscillators, or chemical molecules whose action angle variables are singular.

Let $\gamma \in \bar{\Lambda}_\infty^k(\mathcal{B})$ with $F(\mathcal{B}_1) \subset \mathcal{B}$, we define the pull-back $F^*\gamma$ by

$$(147) \qquad F^*\gamma = \sum_{i \in \mathbb{Z}^N} F^*\gamma_i \,,$$

that is

$$(148) \qquad F^*\gamma(x)(u_1, \ldots, u_k) = \sum_{i \in \mathbb{Z}^N} \gamma_i(F(x))(DF(x)u_1, \ldots, DF(x)u_k).$$

For a general diffeomorphism, the sums in (147), (148) are purely formal. On the other hand, when $F \in C_\Gamma^1$, the sums for $\pi_j F_* \gamma(\pi_{i_1} u_1, \ldots \pi_{i_k} u_k)$ make sense and converge uniformly.

If $\psi : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ is a smooth map we define $\psi^*\gamma$ in the analogous way. It will be important to emphasize for future applications that when γ is a formal form and ψ decay, then $\psi^*\gamma$ is a smooth form in D_ρ.

An easy computation shows that if $F \in C^\infty(\mathcal{B}_1)$ and $G \in C^\infty(\mathcal{B}_2)$, with $F(\mathcal{B}_1) \subset \mathcal{B}_2$, then

$$(G \circ F)^* = F^* \circ G^*.$$

Also, if $\psi(D_\rho) \subset \mathcal{B}_1$ we have

$$(149) \qquad (\psi \circ F)^* = F^* \circ \psi^*.$$

Again, this is a formal computation for diffeomorphisms, but, when $F, G \in C_\Gamma^1$, then, the calculation can be justified. If $G \in C_\Gamma^1, F \in \mathcal{A}_{\rho,c,\Gamma}$, then $(G \circ F)^*\gamma$ is a well defined form.

Definition 11.1. *Given* $\gamma \in \tilde{\Lambda}_\infty^k$ *we define*

$$d\gamma = \sum_{i \in \mathbb{Z}^N} d\gamma_i.$$

We clearly have that $d^2\gamma = 0$.

Lemma 11.2. *Let* $F \in C_\Gamma^2(\mathcal{B}_1)$, $\psi : \mathcal{A}_{\rho,c.\Gamma} \to \mathcal{M}$ *and* $\gamma \in \tilde{\Lambda}_\infty^k(\mathcal{B})$ *so that the composition makes sense.*

$$F^* d\gamma = d(F^*\gamma),$$
$$\psi^* d\gamma = d(\psi^*\gamma).$$

Proof. It consists mainly in going over the formal computation, but paying attention to the fact that all the steps can be justified by the convergence. We carry out explicitly the first one and we let the other one to the reader.

$$F^* d\gamma = F^* \Big(\sum_i d\gamma_i\Big) = \sum_i F^* d\gamma_i = \sum_i d(F^*\gamma_i)$$
$$= d\sum_i F^*\gamma_i = d(F^*\gamma),$$

where we have used that γ_i are true differential forms. $\square$

In the case that $\gamma = \sum_{i \in \mathbb{Z}^N} \pi_i^*\gamma_i$, with $\gamma_i \in \Lambda^k(M)$,

$$d\gamma = \sum_{i \in \mathbb{Z}^N} d(\pi_i^*\gamma_i) = \sum_{i \in \mathbb{Z}^N} \pi_i^* d\gamma_i,$$

where $d\gamma_i$ is the exterior differential of γ_i in M, and if $\psi : D_\rho \to \mathcal{M}$,

$$\psi^*\gamma = \sum_{i \in \mathbb{Z}^N} \psi^* \pi_i^*\gamma_i = \sum_{i \in \mathbb{Z}^N} \psi_i^*\gamma_i.$$

We can also define the contraction operator. Given a smooth vector field X in $\mathcal{M}$ and a k-form $\gamma \in \tilde{\Lambda}_\infty^k$ we set

$$(i_X\gamma)(x)(u_1,\ldots,u_{k-1}) = \sum_{j \in \mathbb{Z}^N} (i_X\gamma_j)(x)(u_1,\ldots,u_{k-1}).$$

Hence we can also introduce the Lie derivative for formal forms by the usual formula

$$\mathcal{L}_X\gamma = i_X d\gamma + d(i_X\gamma).$$

Lemma 11.3. *Let* $\gamma = \sum_{i \in \mathbb{Z}^N} \pi_i^*\gamma_i \in \bar{\Lambda}_\infty^k$ *be a formal form and* $\psi : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ *a map with decay, i.e.* $\psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ *with* $\underline{c} = (c_1,\ldots,c_R)$. *Then for* $0 < \delta < \rho$, $\psi^*\gamma$ *is a well-defined k-form in* $D_{\rho-\delta}$.

As a consequence if $F \in C_\Gamma^1(\mathcal{B})$ is analytic, $\psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$ and $\psi(D_\rho) \subset \mathcal{B}$, by Lemma 10.15 and (149) we have that $\psi^* F^*\gamma$ is a well-defined k-form in $D_{\rho-\delta}$.

Proof of Lemma 11.3. By definition of γ, we have

$$(\psi^*\gamma)(\theta)(u_1,\ldots,u_k) = \sum_{i \in \mathbb{Z}^N} \gamma_i(\psi_i(\theta))(D\psi_i(\theta)u_1,\ldots,D\psi_i(\theta)u_k)$$

for $\theta \in \mathbb{T}^l$ and $u_1, \ldots, u_k \in T_\theta \mathbb{T}^l$. Then

$$|(\psi^* \gamma)(\theta)(u_1, \ldots, u_k)|$$

$$\leq \sum_{i \in \mathbb{Z}^N} \|\gamma_i\| \sum_{m_1} |D_{m_1} \psi_i(\theta)(u_1)_{m_1}| \cdots \sum_{m_k} |D_{m_1} \psi_i(\theta)(u_k)_{m_k}|$$

$$\leq \sum_{i \in \mathbb{Z}^N} \|\gamma_i\| \max_j \Gamma(i - c_j) |D\psi| \|u_1\| \cdots \max_j \Gamma(i - c_j) |D\psi| \|u_k\|$$

$$\leq R^k \|\gamma\| \|D\psi\|_{\rho - \delta, \underline{c}, \Gamma} \|u_1\| \cdots \|u_k\|.$$

We have used that $\sum_{i \in \mathbb{Z}^N} \max_j \Gamma(i - c_j) \leq R$. This proves that the series is absolutely convergent and so $\psi^* \gamma$ is well-defined on $T^* \mathbb{T}^l$. $\qquad \square$

Therefore, by the previous construction, by pulling back formal forms on the lattice to the torus, one obtains well-defined quantities.

Lemma 11.4. *For every function $\psi \in \mathcal{A}_{\rho, \underline{c}, \Gamma}$, we have*

$$\psi^* d\gamma = d(\psi^* \gamma).$$

Proof. By the definition of $d\gamma$ and the convergence of the series, we have

$$\psi^*(d\gamma) = \psi^* \Big(\sum_i d\pi_i^* \gamma_i \Big) = \sum_i \psi^* d\pi_i^* \gamma_i.$$

But by the definition of the exterior differentiation, we have

$$\sum_i \psi^* d\pi_i^* \gamma_i = \sum_i d\psi^* \pi_i^* \gamma_i = d \sum_i \psi^* \pi_i^* \gamma_i = d\psi^* \gamma.$$

$\qquad \square$

11.2. **Some symplectic geometry on lattices.** In this section we discuss the elements of symplectic geometry that we will need. This will play a role in the vanishing lemma Lemma 6.1 in Section 6.

Consider a finite-dimensional exact symplectic manifold $(M, \Omega = d\alpha)$ and the associated lattice

$$\mathcal{M} = \ell^\infty(\mathbb{Z}^N).$$

Let α_∞ and Ω_∞ be defined by

$$\alpha_\infty = \sum_{j \in \mathbb{Z}^N} \pi_j^* \alpha, \qquad \Omega_\infty = \sum_{j \in \mathbb{Z}^N} \pi_j^* \Omega.$$

Then $\alpha_\infty \in \bar{\Lambda}_\infty^1$ and $\Omega_\infty \in \bar{\Lambda}_\infty^2$. Moreover note that $d\alpha_\infty = \Omega_\infty$. We introduce the following definitions.

Definition 11.5. *We say that a C_Γ^1 function $F : \mathcal{M} \to \mathcal{M}$ is symplectic if the following identity holds for any $z \in \mathcal{M}$*

$$DF^\top(z) J_\infty(F(z)) DF(z) = J_\infty(z)$$

Definition 11.6. *We say that a C_Γ^1 function $F : \mathcal{M} \to \mathcal{M}$ is exact symplectic on $\mathcal{M}$ if there exists a one-form $\tilde{\alpha}$ defined on $T\mathcal{M}$ with matrix $\tilde{A}$ such that*

- *For every $j \in \mathbb{Z}^N$, there exists a smooth function W_j on M such that*

$$\tilde{\alpha}_j = dW_j$$

  where d is the exterior differentiation on M.

- *The following formula holds component-wise on the lattice*

$$DF(z)^\top \hat{A}_\infty(F(z)) = \hat{A}_\infty(z) + \tilde{A}(z).$$

The previous definitions are completely equivalent to the standard definitions of symplectic and exact symplectic maps in the finite-dimensional case, but they are among the mildest ones that we can imagine in infinite dimensions. The following is a straightforward result.

Lemma 11.7. *Let $F \in C^1_\Gamma$ be a map from $\mathcal{M}$ into itself. If F is exact symplectic then it is symplectic.*

Remark 11.8. Through a localized embedding, this is even easier.

Since F is exact symplectic, for every decay function $\psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, there exists a smooth function W_ψ defined on the torus such that

$$\psi^* F^* \alpha_\infty = \psi^* \alpha_\infty + dW_\psi.$$

By the property of the exterior differentiation and the fact that, by the hypotheses, $F \circ \psi \in \mathcal{A}_{\rho,\underline{c},\Gamma}$, we have

$$(F \circ \psi)^* d\alpha_\infty = d((F \circ \psi)^* \alpha_\infty) = d(\psi^* \alpha_\infty + dW_\psi) = d(\psi^* \alpha_\infty) = \psi^* d\alpha_\infty.$$

Since $\Omega_\infty = d\alpha_\infty$ this gives the desired result.

We now turn to the symplectic geometry of vector fields. We will always be considering vector-fields of the form

$$X = J_\infty \nabla H$$

where the operator ∇ has to be understood w.r.t. the inner product on $\ell^2(\mathbb{Z}^N)$. The following result is proved.

Proposition 11.9. *Aasume that the vector-field X previously defined has decay. Then it generates flows consisting of exact symplectic diffeomorphisms.*

Proof. Since X has decay, the operation $i_X \Omega_\infty$ makes sense and one has

$$i_X \Omega_\infty = dH.$$

The proof then follows the standard one by using the fact that decay vector fields generate decay diffeomorphisms. $\square$

12. Appendix : Construction of deformations of symplectic maps which are not exact symplectic

In the construction of the invariant torus, we are going to use a family of maps F_λ such that F_0 is exact symplectic and F_λ is symplectic for all λ but not exact symplectic for $\lambda \neq 0$. Indeed, these maps will be used to kill some averages in the invariance equations, so it will be important that, by choosing λ appropriately we can obtain all the possible cohomology obstructions to exactnesn. The change on cohomology is more or less proportional to the change in the parameter λ.

The construction of F_λ will be done in this section by considering flows which are locally but not globally Hamiltonian. We emphasize that the diffeomorphisms introduced will be quite simple. They will just deform a finite number of sites on the lattice. In the case that the phase space is $\mathbb{T}^l \times \mathbb{R}^{2d-l}$ endowed with the standard symplectic form the map F_λ will be given by $A_i \to A_i + \lambda_i$ where A_i are the variables symplectic conjugate to angles. We note that the obstruction to

exactness are the integrals of the forms $A_i d\phi_i$ around a cycle in the torus along ϕ_i. The rest of the section is devoted to make a geometrically natural construction that works in all manifolds.

Consider $(M_i, \Omega_i = d\alpha_i)_{i \in \mathbb{Z}^N}$ a family of finite-dimensional exact symplectic manifolds and denote $\mathcal{M}$ the phase space of the associated lattice map.

Let $\mathcal{J} \subset \mathbb{Z}^N$ be a *finite* set of indexes. We denote by $H^1(M_i)$ the first de Rham cohomology group of the manifold M_i and assume that it is non-trivial. Consider $(\delta_k^i)_{k=1,\ldots,l}$ a basis of $H^1(M_i)$. Since Ω_i are non-degenerate, one can construct a family of vector fields Y_i^λ on M_i with indexes $i \in \mathcal{J}$ such that

$$i_{Y_i^\lambda} \Omega_i = \sum_{k=1}^{l} \lambda_k \delta_k^i.$$

Note that Y_i^λ only depends on $x_i \in M$. Now we introduce the vector-field X_λ on the lattice $\mathcal{M}$ defined by

$$(X_\lambda)_j(x) = \begin{cases} 0 & \text{if } j \notin \mathcal{J}, \\ Y_j^\lambda(\pi_j(x)) & \text{if } j \in \mathcal{J}. \end{cases}$$

By construction, we have $X_0 = 0$. Furthermore, the family of vector-fields X_λ is symplectic for all λ. Indeed, consider a decay function ψ and compute

$$\mathcal{L}_{X_\lambda} \Omega_\infty = d \sum_{j \in \mathbb{Z}^N} i_{(X_\lambda)_j} \Omega_j = d \sum_{j \in \mathcal{J}} \sum_{k=1}^{l} \lambda_k \delta_k^i = 0,$$

where we have used that the last sum is finite by construction of X_λ and $(\delta_k^i)_{k=1,\ldots,l}$ are closed forms. We obtain that X_λ is symplectic. Notice also that all but a finite number of the components of X_λ are zero and so $DX_\lambda(x) \in \mathcal{L}_\Gamma$, i.e. X_λ is a decay vector field. If $x \in \mathcal{M}$, denote $\varphi(s, \lambda, x)$ the flow generated by X_λ. That is:

$$\frac{d}{ds}\varphi(s, \lambda, x) = X_\lambda(\varphi(s, \lambda, x)), \qquad \varphi(0, \lambda, x) = x.$$

The existence and uniqueness of $\varphi(s, \lambda, x)$ is ensured by the theorem of existence and uniqueness of solutions for Lipschitz differential equations in Banach spaces. See [Hal80] for instance. Here the Banach space is $\ell^\infty(\mathbb{Z}^N)$.

Given an exact symplectic map F satisfying $F(0) = 0$ and $F \in C_\Gamma^1$, we define the family of maps we want to construct by

$$F_\lambda = \varphi(\lambda, \lambda, .) \circ F.$$

We have the following easy lemma

Lemma 12.1. *For all $s \in \mathbb{R}$, we have*

(1) $\varphi(s, 0, x) = x$.
(2) *For all $j \in \mathbb{Z}^N$, $\varphi_j(s, \lambda, .)$ only depends on x_j and*
$$\varphi_j(s, \lambda, x) = x_j, \qquad j \notin \mathcal{J}.$$

Proof. (1) follows directly from the fact that $X_0 = 0$. The first part of (2) follows from the fact that $(X_\lambda)_j$ only depends on x_j. Moreover if $j \notin \mathcal{J}$, $(X_\lambda)_j = 0$ and then φ_j is constant in s. Therefore $\varphi_j(s, \lambda, x) = \varphi_j(0, \lambda, x) = x_j$. $\qquad \square$

As a consequence we have that

$$F_0 = F$$

and

$$(F_\lambda)_j = F_j, \qquad \text{for } \lambda \in \mathbb{R}^l, \quad j \notin \mathcal{J}.$$

Since φ_j is not constant for only a finite set of indexes, for λ small $\varphi(1, \lambda, .)$ is well-defined on the range of F. Moreover, since φ is uncoupled, i.e. $\pi_i \varphi(s, \lambda, .)$ depends only on x_i we have that $\varphi(\lambda, \lambda, .) \in C^1_\Gamma$. On the other hand, $(F_\lambda)_j(0) = F_j(0) = 0$ for $\lambda \notin \mathcal{J}$. Therefore F satisfies the assumptions of Lemma 10.15.

Finally, the following lemma ends the details of the construction.

Lemma 12.2. *For all λ, the map $F_\lambda \in C^1_\Gamma$ is symplectic, but it is not exact symplectic for $\lambda \neq 0$.*

Indeed, we have that if Ψ is the embedding given by the coordinates in $\mathcal{J}$, and $[\cdot]$ denotes the cohomology class on the torus expressed in the basis of the forms δ_k, we have

(150)
$$[\Psi^* F_\lambda^* \alpha_\infty] = \lambda$$

Proof. Let $\psi \in \mathcal{A}_{\rho, \underline{c}, \Gamma}$ be a decay function. We want to prove that for any λ

$$\psi^* F_\lambda^* \Omega_\infty = \psi^* \Omega_\infty.$$

By construction of F_λ we have

$$\psi^* F_\lambda^* \alpha_\infty - \psi^* F_0^* \alpha_\infty = \psi^* F_0^* \varphi^*(\lambda, \lambda, .) \alpha_\infty - \psi^* F_0^* \varphi^*(0, \lambda, .) \alpha_\infty$$
$$= \int_0^1 \frac{d}{ds} \left(\psi^* F_0^* \varphi^*(s, \lambda, .) \alpha_\infty \right) ds$$
$$= \int_0^1 \psi^* F_0^* \left(\left(\frac{d}{ds} \varphi^* i \right)(s, \lambda, .) \alpha_\infty \right) \lambda \, ds$$
$$= \int_0^1 \psi^* F_0^* \mathcal{L}_{X_\lambda} \alpha_\infty \, ds$$
$$= \int_0^1 \left[\psi^* F_0^* d(i_{X_\lambda} \alpha_\infty) + i_{X_\lambda} d\alpha_\infty \right] ds$$
$$= d \int_0^1 \psi^* F_0^* (i_{X_\lambda} \alpha_\infty) \, ds + \int_0^1 \psi^* F_0^* \sum_{j \in \mathcal{J}} \sum_{k=1}^l \lambda_k \delta_k^i \, ds \, .$$

Since δ_k^j are closed forms, taking exterior differential at both sides of the previous formula we get $\psi^* F_\lambda^* \Omega_\infty - \psi^* F_0^* \Omega_\infty = 0$. Finally, using that F_0 is symplectic we get that F_λ is symplectic.

Moreover, if $\lambda \neq 0$,

$$\psi^* F_\lambda^* \alpha_\infty - \psi^* F_0^* \alpha_\infty = dW_\infty + E,$$

where $W_\infty = \int_0^\lambda \psi^* F_0^*(i_{X_\lambda}\alpha_\infty)\, ds$ and E is not a differential. The formula (150), follows easily from the expression for E above. We note that that

$$[\Psi^* F_\lambda^* \alpha_\infty] = [\Psi^* F_\lambda^* \alpha_\infty - \Psi^* F_0^* \alpha_\infty]$$

$$= [d\int_0^1 \Psi^* F_0^*(i_{X_\lambda}\alpha_\infty)\, ds + \int_0^1 \Psi^* F_0^* \sum_{j\in\mathcal{J}}\sum_{k=1}^l \lambda_k \delta_k^i\, ds\,.]$$

$$= 0 + \int_0^1 [\Psi^* F_0^* \sum_{j\in\mathcal{J}}\sum_{k=1}^l \lambda_k \delta_k^i]\, ds\,.]$$

$$= \int_0^1 \Psi_\#(F_0)_\# \sum_{j\in\mathcal{J}}\sum_{k=1}^l \lambda_k [\delta_k^i]\, ds\,.]$$

$$= \lambda$$

$\square$

References

[AG96]   G. Arioli and F. Gazzola. Periodic motions of an infinite lattice of particles with nearest neighbor interaction. *Nonlinear Anal.*, 26(6):1103–1114, 1996.

[AGT96]   G. Arioli, F. Gazzola, and S. Terracini. Multibump periodic motions of an infinite lattice of particles. *Math. Z.*, 223(4):627–642, 1996.

[AKK01]   S. Aubry, G. Kopidakis, and V. Kadelburg. Variational proof for hard discrete breathers in some classes of Hamiltonian dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B*, 1(3):271–298, 2001.

[Bam99]   Dario Bambusi. On the Darboux theorem for weak symplectic manifolds. *Proc. Amer. Math. Soc.*, 127(11):3383–3391, 1999.

[BEMW07]   Burak Bilki, Mehmet Erdubak, Muhittin Mungan, and Yves Weisskopf. Structure formation of a later of adatoms of a quasicrystaline substrate: Molecular dynamics study. *Phys. Rev. B*, 75(045437):1–6, 2007.

[BK95]   J. Bricmont and A. Kupiainen. Coupled analytic maps. *Nonlinearity*, 8(3):379–396, 1995.

[BK98]   Oleg M. Braun and Yuri S. Kivshar. Nonlinear dynamics of the Frenkel-Kontorova model. *Phys. Rep.*, 306(1-2):108, 1998.

[BK04]   O. M. Braun and Y. S. Kivshar. *The Frenkel-Kontorova model*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 2004. Concepts, methods, and applications.

[BV02]   D. Bambusi and D. Vella. Quasi periodic breathers in Hamiltonian lattices with symmetries. *Discrete Contin. Dyn. Syst. Ser. B*, 2(3):389–399, 2002.

[CF05]   Jean-René (ed.) Chazottes and Bastien (ed.) Fernandez. *Dynamics of coupled map lattices and of related spatially extended systems. Lectures delivered at the school-forum CML 2004, Paris, France, June 21 – July 2, 2004.* Lecture Notes in Physics 671. Berlin: Springer, 2005.

[CM74]   Paul R. Chernoff and Jerrold E. Marsden. *Properties of infinite dimensional Hamiltonian systems*. Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 425.

[CP95]   Luigi Chierchia and Paolo Perfetti. Second order Hamiltonian equations on $\mathbf{T}^\infty$ and almost-periodic solutions. *J. Differential Equations*, 116(1):172–201, 1995.

[CY07a]   Shun Qing Chen and Xiao Ping Yuan. A KAM theorem for reversible systems of infinite dimension. *Acta Math. Sin. (Engl. Ser.)*, 23(10):1777–1796, 2007.

[CY07b]   K. W. Chung and Xiaoping Yuan. Existence and stability of quasi-periodic breathers in networks of Ginzburg-Landau oscillators. *Phys. D*, 227(1):43–50, 2007.

[dlL01]   Rafael de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[dlLW04] R. de la Llave and C. E. Wayne. Whiskered and low dimensional tori in nearly integrable Hamiltonian systems. *Math. Phys. Electron. J.*, 10:Paper 5, 45 pp. (electronic), 2004.

[DPW92] T. Dauxois, M. Peyrard, and C. R. Willis. Localized breather-like solution in a discrete Klein-Gordon model and application to DNA. *Phys. D*, 57(3-4):267–282, 1992.

[DRAW02] T. Dauxois, S. Ruffo, E. Arimondo, and M. Wilkens, editors. *Dynamics and thermodynamics of systems with long-range interactions*, volume 602 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 2002. Lectures from the conference held in Les Houches, February 18–22, 2002.

[Dua94] Pedro Duarte. Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 11(4):359–409, 1994.

[Dua08] P. Duarte. Elliptic isles in families of area-preserving maps. *Ergodic Theory Dynam. Systems*, 28(6):1781–1813, 2008.

[Dur96] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[FBGGn05] L.M. Floria, C. Baesens, and J. Gómez-Gardeñes. The Frenkel-Kontorova model. [CF05].

[FdlLM07] E. Fontich, R. de la Llave, and P. Martin. Invariant manifolds of hyperbolic sets and their decay properties in weakly coupled map lattices. 2007. preprint.

[FdlLM11a] Ernest Fontich, Rafael de la Llave, and Pau Martín. Dynamical systems on lattices with decaying interaction I: a functional analysis framework. *J. Differential Equations*, 250(6):2838–2886, 2011.

[FdlLM11b] Ernest Fontich, Rafael de la Llave, and Pau Martín. Dynamical systems on lattices with decaying interaction II: hyperbolic sets and their invariant manifolds. *J. Differential Equations*, 250(6):2887–2926, 2011.

[FdlLS09a] Ernest Fontich, Rafael de la Llave, and Yannick Sire. Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions. *J. Differential Equations*, 246(8):3136–3213, 2009.

[FdlLS09b] Ernest Fontich, Rafael de la Llave, and Yannick Sire. A method for the study of whiskered quasi-periodic and almost-periodic solutions in finite and infinite dimensional Hamiltonian systems. *Electron. Res. Announc. Math. Sci.*, 16:9–22, 2009.

[FP99] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices. I. Qualitative properties, renormalization and continuum limit. *Nonlinearity*, 12(6):1601–1627, 1999.

[FP02] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices. II. Linear implies nonlinear stability. *Nonlinearity*, 15(4):1343–1359, 2002.

[FP04a] G. Friesecke and R. L. Pego. Solitary waves on Fermi-Pasta-Ulam lattices. III. Howland-type Floquet theory. *Nonlinearity*, 17(1):207–227, 2004.

[FP04b] G. Friesecke and R. L. Pego. Solitary waves on Fermi-Pasta-Ulam lattices. IV. Proof of stability at low energy. *Nonlinearity*, 17(1):229–251, 2004.

[FSW86] J. Fröhlich, T. Spencer, and C. E. Wayne. Localization in disordered, nonlinear dynamical systems. *J. Statist. Phys.*, 42(3-4):247–274, 1986.

[FW94] Gero Friesecke and Jonathan A. D. Wattis. Existence theorem for solitary waves on lattices. *Comm. Math. Phys.*, 161(2):391–418, 1994.

[FW98] S. Flach and C. R. Willis. Discrete breathers. *Phys. Rep.*, 295(5):181–264, 1998.

[Gal08] G. Gallavotti, editor. *The Fermi-Pasta-Ulam problem*, volume 728 of *Lecture Notes in Physics*. Springer, Berlin, 2008. A status report.

[Gra74] Samuel M. Graff. On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, 15:1–69, 1974.

[GVY08] Jiansheng Geng, Jorge Viveros, and Yingfei Yi. Quasi-periodic breathers in Hamiltonian systems with long-range coupling. *Physica D*, 237(22):2866–2922, 2008.

[GY07] Jiansheng Geng and Yingfei Yi. A KAM theorem for Hamiltonian networks with long ranged couplings. *Nonlinearity*, 20(6):1313–1342, 2007.

[Hal80] Jack K. Hale. *Ordinary differential equations*. Robert E. Krieger Publishing Co. Inc., Huntington, N.Y., second edition, 1980.

[HP70]   M.W. Hirsch and C.C. Pugh. Stable manifolds and hyperbolic sets. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 133–163. Amer. Math. Soc., Providence, R.I., 1970.

[HPPS70]   M. Hirsch, J. Palis, C. Pugh, and M. Shub. Neighborhoods of hyperbolic sets. *Invent. Math.*, 9:121–134, 1969/1970.

[HPS77]   M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*. Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 583.

[IK00]   Gérard Iooss and Klaus Kirchgässner. Travelling waves in a chain of coupled nonlinear oscillators. *Comm. Math. Phys.*, 211(2):439–464, 2000.

[Ioo00]   Gérard Iooss. Travelling waves in the Fermi-Pasta-Ulam lattice. *Nonlinearity*, 13(3):849–866, 2000.

[Jam01]   Guillaume James. Existence of breathers on FPU lattices. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(6):581–586, 2001.

[JdlL00]   Miaohua Jiang and Rafael de la Llave. Smooth dependence of thermodynamic limits of SRB-measures. *Comm. Math. Phys.*, 211(2):303–333, 2000.

[JS05]   Guillaume James and Yannick Sire. Travelling breathers with exponentially small tails in a chain of nonlinear oscillators. *Comm. Math. Phys.*, 257(1):51–85, 2005.

[KK08]   Alexander I. Komech and Andrew A. Komech. Global attraction to solitary waves in models based on the Klein-Gordon equation. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 4:Paper 010, 23, 2008.

[KK10]   Alexander Komech and Andrew Komech. On global attraction to solitary waves for the Klein-Gordon field coupled to several nonlinear oscillators. *J. Math. Pures Appl. (9)*, 93(1):91–111, 2010.

[MA94]   R. S. MacKay and S. Aubry. Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators. *Nonlinearity*, 7(6):1623–1643, 1994.

[Mos66a]   J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.

[Mos66b]   J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Mos67]   J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[Per03]   Paolo Perfetti. A KAM theorem for infinite-dimensional discrete systems. *Math. Phys. Electron. J.*, 9:Paper 1, 15 pp. (electronic), 2003.

[Pey04]   Michel Peyrard. Nonlinear dynamics and statistical physics of DNA. *Nonlinearity*, 17(2):R1–R40, 2004.

[Pös90]   J. Pöschel. Small divisors with spatial structure in infinite-dimensional Hamiltonian systems. *Comm. Math. Phys.*, 127(2):351–393, 1990.

[PS99]   Victor A. Pliss and George R. Sell. Robustness of exponential dichotomies in infinite-dimensional dynamical systems. *J. Dynam. Differential Equations*, 11(3):471–513, 1999.

[PS04]   M Peyrard and Y. Sire. Breathers in biomolecules ? *Conference "Energy Localisation and transfer in Crystals, Biomolecules and Josephson Arrays". Advances series in nonlinear dynamics*, 22(2):391–418, 2004.

[Pyk96]   Randall Pyke. Virial relations for nonlinear wave equations and nonexistence of almost periodic solutions. *Rev. Math. Phys.*, 8(7):1001–1039, 1996.

[Rug02]   Hans Henrik Rugh. Coupled maps and analytic function spaces. *Ann. Sci. École Norm. Sup. (4)*, 35(4):489–535, 2002.

[Rüs75]   H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical Systems, Theory and Applications (Battelle Rencontres, Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38, Berlin, 1975. Springer.

[Rüs76a]   H. Rüssmann. On a new proof of Moser's twist mapping theorem. In *Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I*, volume 14, pages 19–31, 1976.

[Rüs76b]   H. Rüssmann. On optimal estimates for the solutions of linear difference equations on the circle. *Celestial Mech.*, 14(1):33–37, 1976.

[Sev99]   M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, pages 568–572. Kluwer Acad. Publ., Dordrecht, 1999.

[Sir05]   Yannick Sire. Travelling breathers in Klein-Gordon lattices as homoclinic orbits to *p*-tori. *J. Dynam. Differential Equations*, 17(4):779–823, 2005.

[SW99]   A. Soffer and M. I. Weinstein. Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations. *Invent. Math.*, 136(1):9–74, 1999.

[Thi97]   W. Thirring. *Classical Mathematical Physics*. Springer-Verlag, New York, third edition, 1997.

[Tre94]   D.V. Treschev. Hyperbolic tori and asymptotic surfaces in Hamiltonian systems. *Russ. J. Math. Phys.*, 2(1):93–110, 1994.

[You99]   Jiangong You. Perturbations of lower-dimensional tori for Hamiltonian systems. *J. Differential Equations*, 152(1):1–29, 1999.

[Yua02]   Xiaoping Yuan. Construction of quasi-periodic breathers via KAM technique. *Comm. Math. Phys.*, 226(1):61–100, 2002.

[Zeh75]   E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76a] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

[Zeh76b] Eduard Zehnder. Moser's implicit function theorem in the framework of analytic smoothing. *Math. Ann.*, 219(2):105–121, 1976.

DEPARTARMNENT DE MATEMÀTICA APLICADA I ANÀLISI, VIA 585. 08007 BARCELONA. SPAIN
E-mail address: `fontich@maia.ub.es`

SHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY ST. ATLANTA, GA 30332-0160, USA
E-mail address: `rafael.delallave@math.gatech.edu`

UNIVERSITÉ PAUL CÉZANNE, LABORATOIRE LATP UMR 6632, MARSEILLE, FRANCE
E-mail address: `sire@cmi.univ-mrs.fr`