KAM THEORY FOR QUASI-PERIODIC EQUILIBRIA IN 1D QUASIPERIODIC MEDIA–II: LONG-RANGE INTERACTIONS

XIFENG SU AND RAFAEL DE LA LLAVE

Abstract. We consider Frenkel-Kontorova models corresponding to 1 dimensional quasi-crystal with non-nearest neighbor interactions. We formulate and prove a KAM type theorem which establishes the existence of quasi-periodic solutions.

The interactions we consider do not need to be of finite range but do have to decay sufficiently fast with respect to the distance of the position of the atoms. The KAM theorem we present has an a-posteriori format. We do not need to assume that the system is close to integrable. We just assume that there is an approximate solution for the functional equation which satisfies some non-degeneracy conditions.

Quasi-periodic solutions, long-range interactions, quasicrystals, hull functions, KAM theory

[2000] 70K43, i 52C23, 37A60, hanics 37J40, ivisors, KAM theory, Arnol??d diffusion 82B20 tc.) and systems on graphs

1. Introduction

We will consider one dimensional chains of interacting particles and assume that the state of each site is described by a real variable. We also assume that the interaction is translation invariant (later we will also include some regularity and non-degeneracy assumptions of a more technical nature). We do not assume that the interaction is a pair interaction (i.e. we allow many body potentials) nor that is localized in space.

We will be interested in the situation, when the dependence on the variables of the configuration is quasi-periodic. This is a natural assumption in quasi-periodic media (e.g. the standard models of quasi-crystals).

More precisely the configuration of a system is described by $u = \{u_n\}_{n\in\mathbb{Z}}$ with $u_n \in \mathbb{R}$. The models we consider are a particular case of the models in Statistical Mechanics [Rue69]. The formal energy is obtained by assigning a energy to every finite subset of $\mathbb{Z}$. That is, we consider the formal energy functional:

$$\mathscr{S}(u) = \sum_{i\in\mathbb{Z}} \sum_{L\in\mathbb{N}} \widehat{H}_L(u_i\alpha, u_{i+1}\alpha, \dots, u_{i+L}\alpha) \tag{1}$$

where $\widehat{H} : (\mathbb{T}^d)^{L+1} \to \mathbb{R}$, and $\alpha \in \mathbb{R}^d$ which is non-resonant. See (19) for precise conditions on the non-resonance. Note that the form of H_L encodes the quasi-periodic properties of the media.

An interesting particular case of (1) that we will use for illustrations (a complete treatment in this case appears in Appendix A

$$
\begin{aligned}
&\widehat{H}_0(\sigma) = -\widehat{V}(\sigma) \\
&\widehat{H}_L(\sigma_0, \sigma_1, \ldots, \sigma_L) = \frac{1}{2|\alpha|_2^2} A_L |\sigma_0 - \sigma_L|_2^2
\end{aligned}
\tag{2}
$$

with A_L decaying sufficiently fast. Here $\widehat{V} : \mathbb{T}^d \to \mathbb{R}$ and $|\cdot|_2$ is the norm generated by the inner product in $\mathbb{R}^d$. We will deal with case (2) in Appendix A.

Models (2) admit a clear physical interpretation. The quasi-periodic function $\widehat{V}(\theta\alpha)$ may describe the energy of a particle deposited at position θ on a quasi-crystal. The terms $A_L|\sigma_0 - \sigma_L|^2$ describe the energy of interactions between two particles at positions u_0 and u_L (setting $u_i\alpha = \sigma_i$). Hence the model (2) describes deposition of particles which interact with each other on a quasi-crystal. Other physical interpretations are also possible. In the original papers [FK39], the meaning of u was the position of planar dislocations in crystals. We also refer to [Mat09] for several physical applications of 1-dimensional models.

Our main result Theorem 1 for case (1) is a KAM type theorem, which, following [Mos66b, Mos66a, Zeh75, SdlL11, dlL08, CdlL10b] and many others, will be presented in an *a-posteriori* format. That is, we show that, given an approximate solution of the equilibrium equations, which satisfies some appropriate non-degeneracy conditions, then, there exists a true solution nearby. This will be proved by a quasi-Newton method.

The models (1) considered in this paper have two difficulties:

a) the fact that the interaction is quasi-periodic;
b) the fact that the interaction is long-range.

As for the quasi-periodicity we note that , as pointed out in [Mos66b] adding an extra frequency without adding extra parameters makes the KAM theory problematic. This difficulty was overcome in [SdlL11] for nearest neighbor interactions showing by an explicit calculation that the perturbations did not include the terms dangerous to KAM theory.

We also note that the addition of an extra frequency causes difficulties for the Aubry-Mather theory and that in this case there are counterexamples [LS03, DS09] even in the case that the functional is convex. For our results we will not assume convexity (e.g. the A_L in (2) could have either sign provided they decay fast enough). The physical literature [vE99, vEFRJ99, vEFJ01, vEF02] also contains indications that the resonant case of quasi-periodic case is very different from the periodic one. (e.g. there could exist phonon gap–positive Lyapunov exponents even if there is a sliding mode–an invariant circle, a phenomenon which is impossible in the periodic case.) We note that, as pointed out in [DS09], the case when the spatial variables are of dimension higher than one is quite open in the mathematical literature (see [Bur88] for some non-rigorous studies).

Furthermore, the quasi-periodic nature of the interactions also makes it impossible to give a straightforward dynamical meaning to the equilibrium equations.

The long range interactions also make it impossible to obtain a dynamical system interpretation.

The long range difficulty was overcome in the periodic case in [dlL08, CdlL10b]. The quasi-periodic difficulty was overcome in [SdlL11]. In this paper, we do both at the same time. Moreover, we improve some estimates and rearrange some arguments so that the results in this paper require less conditions than that of the previous papers (e.g. we require less decay on the interactions). We also take the opportunity to fix several typos in [dlL08].

Even if some of the calculations are almost verbatim from the previous papers, we have kept them to enhance the readability of the paper and to improve the results through more refined estimates.

The strategy of proof will be a quasi-Newton method that overcomes the small divisors generated by the frequencies. Since there is no dynamical interpretation, we cannot use some of the customary methods of KAM theory such as transformation theory. Nevertheless, we will be able to obtain identities (see Section 5.2 that allow us to reduce to the constant coefficients equation customary in KAM theory. Related ideas have appeared in [Koz83, Mos88, LM01, Ran87, dlL08, CdlL10b, SdlL11].

In Appendix A we present an alternative proof of Theorem 1 for the models in (2). The simplicity of the models allows us to make more detailed estimates as well as more explicit algorithms. This has several advantages:

- Using the structure of the terms we can establish the results with slower decay than what follows for general terms.
- We can present very detailed algorithm which are somewhat similar to those in [CdlL10b] for the periodic long-range case.

From the mathematical point of view, the proof in Appendix A is sighltly different than the proof presented in the main text. We add an external parameter, which simplifies the iterative step and allows the consideration of forces that do not derivar form a potential, and then, at the end, show that, for models with a variational structure, this extra parameter has to vanish. This is similar to the proof of twist mappings theorems going through translated curve theorems. We hope that having both style of proofs in the same paper could have some pedagogical value.

2. General long-range model and its equilibrium configurations

We will seek configurations $\{u_n\}_{n\in\mathbb{Z}}$ which are solutions of the Euler-Lagrange equations of the formal functional $\mathscr{S}$.

We take formal derivative of $\mathscr{S}(u)$ with respect to u_j and set it to zero:

$$\frac{\partial}{\partial u_j}\mathscr{S}(u) = \sum_{L\in\mathbb{N}}\sum_{i=j-L}^{j} \alpha\cdot\partial_{j-i}\widehat{H}_L(u_i\alpha,\ldots,u_{i+L}\alpha) = 0 \qquad \forall j\in\mathbb{Z} \tag{3}$$

where $\partial_j = \frac{\partial}{\partial\sigma_j}$ for $j = 0,1,\ldots,L$. For simplicity, we write $\partial_\alpha^j = \alpha\cdot\partial_j$ in the following. In fact, such operators are commutative, i.e. $\partial_\alpha^j\partial_\alpha^k = \partial_\alpha^k\partial_\alpha^j$ for any $j,\ k\in\mathbb{Z}$.

Note that (3) will make sense (if the interactions H_L decrease fast) even if (1) is only a formal sum. The analysis of this paper is based just on (3).

Note that in the particular case, (2), the equilibrium equations are just:

$$u_{n+1} + u_{n-1} - 2 \cdot u_n + \sum_{j=2}^{\infty} A_j \cdot (u_{n+j} + u_{n-j} - 2 \cdot u_n) + V'(u_n) = 0 \qquad \forall\, n \in \mathbb{Z}. \tag{4}$$

2.1. **Equilibrium equations for hull functions.** We are interested in finding what are called plane-like configurations in homogenization theory which are configurations of the form:

$$u_n = h(n \cdot \omega) = n \cdot \omega + \hat{h}(n \cdot \omega \cdot \alpha) \tag{5}$$

where $\hat{h}$ is a function on $\mathbb{T}^d$ and $n \in \mathbb{Z}, \omega \in \mathbb{R}, \alpha \in \mathbb{R}^d$. In solid state physics, the function h is often referred as "hull" function of the configuration.

Later on, we always denote $\theta = n \cdot \omega$ for variables in $\mathbb{R}$ and

$$\sigma = \theta\alpha$$

for variables in $\mathbb{R}^d$. We often work with the function $\hat{h} : \mathbb{T}^d \to \mathbb{R}$ which is quite convenient.

Therefore, we will write the equilibrium equation (3) in terms of $\hat{h}$. For simplicity, we first introduce the following notations similar to [dlL08], but note that we need to include the higher dimensional phases:

$$\begin{aligned} h^{(j)}(\theta) &= \theta + j\omega + \hat{h}((\theta + j\omega)\alpha) \equiv \theta + j\omega + \hat{h}(\sigma + j\omega\alpha) \\ \gamma_L^{(j)}(\sigma) &= (h^{(j)}(\theta)\alpha, h^{(j+1)}(\theta)\alpha, \ldots, h^{(j+L)}(\theta)\alpha). \end{aligned} \tag{6}$$

In particular, we denote $h(\theta) = h^{(0)}(\theta),\ \gamma_L(\sigma) = \gamma_L^{(0)}(\sigma)$.

Using the notations (5) and (6) above, we can write (3) more concisely as:

$$\sum_{L \in \mathbb{N}} \sum_{i=-L+j}^{j} \partial_\alpha^{j-i} \widehat{H}_L(\gamma_L^{(i)}(\theta)) = 0 \qquad \forall\, j \in \mathbb{Z}. \tag{7}$$

That is,

$$\sum_{L \in \mathbb{N}} \sum_{k=0}^{L} \partial_\alpha^{k} \widehat{H}_L(\gamma_L^{(j-k)}(\theta)) = 0 \qquad \forall\, j \in \mathbb{Z}. \tag{8}$$

If ω satisfies some Diophantine property defined in (20), (8) holds if and only if $\mathscr{E}[\hat{h}](\theta)$ defined below vanishes identically:

$$\begin{aligned} \mathscr{E}[\hat{h}](\theta) &\equiv \sum_{L \in \mathbb{N}} \sum_{k=0}^{L} \partial_\alpha^{(k)} \widehat{H}_L(\gamma_L^{(-k)}) \\ &\equiv \sum_{L \in \mathbb{N}} \sum_{k=0}^{L} \partial_\alpha^{(k)} \widehat{H}_L(h(\theta - k\omega)\alpha, \ldots, h(\theta - k\omega + L\omega)\alpha) \equiv 0 \quad \forall \theta \in \mathbb{R}. \end{aligned} \tag{9}$$

Note that if (8) holds for some θ it will also holds for $\theta + \omega$. If ω is irrational and h is continuous, we see that (8) holds for a point if and only if $\mathscr{E}[\hat{h}](\theta) = 0$. (Of course, it is true that if (9) holds so does (8).)

2.2. **Non-uniqueness of the equilibria.** We find that the solutions of (9) are not unique. In fact, it is easy to check by substituting directly that if $\hat{h}(\sigma)$ is a solution, then $\hat{h}(\sigma + \beta\alpha) + \beta$ is also a solution for any $\beta \in \mathbb{R}$.

In the following, we consider without loss of generality normalized solution of (9), i.e. we will choose β such that

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} [\hat{h}(\theta\alpha + \beta\alpha) + \beta] d\theta = 0. \tag{10}$$

Indeed, we will establish that the solution of (9) and (10) is locally unique.

Note that given a function $\hat{h}$, there is one and only one β satisfying (10).

3. Preliminaries

3.1. **Spaces of functions we will use.** The quantitative estimates in KAM theory, will require making precise definitions of norms. In this section, we will study the spaces of analytic functions following [dlL08, CdlL10b, SdlL11].

We denote by

$$D_\rho \equiv \{\, \eta \in \mathbb{C}^d/\mathbb{Z}^d \mid \quad |Im(\eta_j)| < \rho \,\}. \tag{11}$$

We denote the Fourier expansion of a periodic mapping $\hat{h}$ on D_ρ by

$$\hat{h}(\eta) = \sum_{k\in\mathbb{Z}^d} \hat{h}_k e^{2\pi i k\cdot\eta}, \tag{12}$$

where $\cdot$ is the Euclidean scalar product in $\mathbb{C}^d$ and $\hat{h}_k$ are the Fourier coefficients. The average of $\hat{h}$ is the 0-Fourier coefficient $\hat{h}_0$.

We denote by $\mathscr{A}_\rho$ the Banach space of such analytic functions on D_ρ which are real for real argument and extend continuously to $\overline{D_\rho}$. We make $\mathscr{A}_\rho$ a Banach space by endowing it with the supremum norm:

$$||\hat{h}||_\rho = \sup_{\eta\in\overline{D_\rho}} |\hat{h}(\eta)|. \tag{13}$$

3.2. **Analysis of the interaction properties.** The goal of this section is to give conditions on the functions H_L which allow the evaluation on the interactions we are considering.

We will assume H_L are analytic in a complex domain and define norms that measure their sizes to state precisely the results. In fact, for a given approximate solution, it suffices to consider domains for the interactions which are defined in a neighborhood of the range of the approximate solution.

We denote

$$\begin{aligned} \mathcal{D}_{L,\hat{h},\delta} = \Big\{ & (\sigma_0, \dots, \sigma_L) \in (\mathbb{C}^d)^{L+1} \mid \\ & \exists\, \sigma \in D_\rho \text{ such that } |\sigma_j - \alpha\hat{h}(\sigma + j\omega\alpha)| \le \delta \quad \forall j = 0, \dots, L \Big\}. \end{aligned} \tag{14}$$

Here, we use the norm in $\mathbb{C}^d$ or $\mathbb{R}^d$ is the supremum of the coordinates.

Since we will deal with mappings from real values into real values in the applications to physical problems, it suffices to consider the domains:

$$\tilde{\mathcal{D}}_{L,\delta} = \{(\sigma_0, \dots, \sigma_L) \in (\mathbb{C}^d)^{L+1} \,\big|\, \quad |Im(\sigma_j)| \le \delta\}. \tag{15}$$

Corresponding to the domains $\mathcal{D}_{L,\hat{h},\delta}, \tilde{\mathcal{D}}_{L,\delta}$, we consider the spaces $\mathcal{H}_{L,\hat{h},\delta}, \mathcal{H}_{L,\delta}$ consisting of functions analytic in the interior and continuous in the whole domain. We endow these spaces with the supremum norm, which makes them Banach spaces.

$$\begin{aligned} \|H_L\|_{L,\hat{h},\delta} &= \sup_{z\in\mathcal{D}_{L,\hat{h},\delta}} |H_L(z)|, \\ \|H_L\|_{L,\delta} &= \sup_{z\in\tilde{\mathcal{D}}_{L,\delta}} |H_L(z)|. \end{aligned} \tag{16}$$

Clearly,

$$\begin{aligned} \mathcal{D}_{L,\hat{h},\delta} &\subseteq \tilde{\mathcal{D}}_{L,\|\hat{h}\|_\rho+\delta}, \\ \mathcal{H}_{L,\hat{h},\delta} &\subseteq \tilde{\mathcal{H}}_{L,\|\hat{h}\|_\rho+\delta}, \\ \|H_L\|_{L,\hat{h},\delta} &\le \|H_L\|_{L,\|\hat{h}\|_\rho+\delta}. \end{aligned} \tag{17}$$

Since L will be unbounded, we will need to estimate the dependence in L for several standard results such as Cauchy estimates and the like.

With the choice of supremum norm in $(\mathbb{C}^d)^{L+1}$, we have

$$\begin{aligned} \sup_{\sigma\in D_\rho} |(\hat{h}(\sigma), \hat{h}(\sigma+\omega\alpha), \dots, \hat{h}(\sigma+L\omega\alpha)) \\ - (\hat{g}(\sigma), \hat{g}(\sigma+\omega\alpha), \dots, \hat{g}(\sigma+L\omega\alpha)| \\ \le \|\hat{h}-\hat{g}\|_\rho. \end{aligned}$$

Therefore, we have

$$\begin{aligned} |H_L(\hat{h}(\sigma), \hat{h}(\sigma+\omega\alpha), \dots, \hat{h}(\sigma+L\omega\alpha)) \\ - H_L(\hat{g}(\sigma), \hat{g}(\sigma+\omega\alpha), \dots, \hat{g}(\sigma+L\omega\alpha))| \\ \le \|DH_L\|_{L^\infty}\|\hat{h}-\hat{g}\|_\rho. \end{aligned}$$

The Cauchy bounds we will use in the following lemma may also have a dependence on L.

Lemma 1. *If* $\Omega \subseteq \tilde{\Omega} \subseteq (\mathbb{C}^d)^{L+1}$ *and* $dist(\Omega, (\mathbb{C}^d)^{L+1} \setminus \tilde{\Omega}) \ge \delta$ *we have:*

$$\|DH_L\|_\Omega \le Cd(L+1)\delta^{-1}\|H_L\|_{\tilde{\Omega}}. \tag{18}$$

3.3. **Diophantine condition.** We will assume that $\alpha \in \mathbb{R}^d$ is non-resonant, i.e.

$$|\alpha \cdot k| \ne 0 \qquad \forall\, k \in \mathbb{Z}^d - \{0\}. \tag{19}$$

We are interested in the rotation number $\omega \in \mathbb{R}$ such that $\omega\alpha$ is a Diophantine vector in the standard sense:

$$|\omega\alpha \cdot k - n| \ge \nu|k|^{-\tau} \qquad \forall\, k \in \mathbb{Z}^d - \{0\},\ n \in \mathbb{Z}. \tag{20}$$

Here ν, τ are positive numbers and we denote the set of ω satisfying (20) by $\mathscr{D}(\nu, \tau; \alpha)$.

It is easy to check that when α satisfies some Diophantine condition, the set $\cup_{\nu>0}\mathscr{D}(\nu,\tau;\alpha)$ has full Lebesgue measure ([SdlL11, Lemma 9]). More refined and quantitative results appear in the recent number theory literature [Kle01, Kle08].

3.4. **Cohomology equations.** It is standard in KAM theory to solve $\hat\phi$ for given $\hat\eta$ with zero average in the equation:

$$\hat\phi(\sigma+\omega\alpha)-\hat\phi(\sigma)=\hat\eta(\sigma) \tag{21}$$

where $\omega\in\mathscr{D}(\nu,\tau)$.

Estimates for (21) are given by the following lemma.

Lemma 2. *Let $\hat\eta\in\mathscr{A}_\rho$ be such that*

$$\int_{\mathbb{T}^d}\hat\eta(\sigma)d\sigma=0. \tag{22}$$

Then, there exists a unique solution $\hat\phi$ of (21) *which satisfies*

$$\int_{\mathbb{T}^d}\hat\phi(\sigma)d\sigma=0. \tag{23}$$

This solution of (21) *satisfies for any $\rho'<\rho$*

$$\|\hat\phi\|_{\rho'}\le C(d,\tau)\cdot\nu^{-1}\cdot(\rho-\rho')^{-\tau}\|\hat\eta\|_\rho \tag{24}$$

Furthermore, any distribution solution of (21) *differs from the solution claimed before by a constant.*

We denote the set of functions $\hat\phi\in\mathscr{A}_\rho$ with zero average by $\mathring{\mathscr{A}}_\rho$.

4. Statement of the main result

4.1. The main result of this paper.

Theorem 1. *Let $h(\theta)=\theta+\hat h(\sigma)$ where $\hat h(\sigma)=\sum_{k\in\mathbb{Z}^d}\hat h_k\cdot e^{2\pi i\cdot k\cdot\sigma}$ with $\hat h_0=0$, $\hat h\in\mathscr{A}^1_\rho$ and $\alpha\in\mathbb{R}^d$ with $|\alpha|=1$ is non-resonant. Denote $\hat l=1+\partial_\alpha\hat h$ and $T_{-\omega\alpha}(\sigma)=\sigma-\omega\alpha$. We assume*

- (H1) *Diophantine properties* (20)*: $|\omega\alpha\cdot k-n|\ge\nu|k|^{-\tau}$, $\quad\forall k\in\mathbb{Z}^d-\{0\},\ n\in\mathbb{Z}$.*
- (H2) *Non-degeneracy condition:*

$$\|\hat l(\sigma)\|_\rho\le N^+,\ \|(\hat l(\sigma))^{-1}\|_\rho\le N^-\ \text{and}\ |\langle\frac{1}{\hat l\cdot\hat l\circ T_{-\omega\alpha}}\rangle|\ge c$$

  for some positive constant c where $\langle f\rangle$ denotes the average of the periodic function f.
- (H3) *The interactions $H_L\in\mathcal{H}_{L,\hat h,\delta}$ for some $\delta>0$. Denote*

$$M_L=\max_{i=0,1,2,3}(\|D^iH_L\|_{L,\hat h,\delta})$$
$$\beta=C\sum_{L\ge2}M_LL^3$$

  where C is a combinatorial constant that will be made explicit during the proof.

(H4) *Assume that the inverses indicated below exist and have the indicated bounds:*

$$\|(\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1)^{-1}(\gamma_1(\sigma))\|_\rho \le T,$$
$$\left|\left(\int_{\mathbb{T}^d} C_{0,1,1}^{-1}\right)^{-1}\right| \le U$$

where $C_{0,1,1}$ *is defined in* (9).

(H5) $(N^-)^2 T\beta < \frac{1}{2},\ (N^-)^2 UT < \frac{1}{2}$.

(H6) $\|\mathscr{E}[\hat{h}]\|_\rho \le \epsilon$ *where* $\mathscr{E}$ *is defined in* (9).

(H7) *Composition condition: Take* $\iota = dist\Big((\mathbb{C}^d)^{L+1} - \Omega, \gamma_L(\overline{D_\rho})\Big)$ *where* Ω *is the domain of* $\widehat{H}_L$ *and* γ_L *is defined in* (6). *Let* $\delta = \iota$. *We assume that*

$$\|\hat{h}\|_\rho + \rho \le \frac{1}{2}\delta.$$

Assume furthermore that $\epsilon \le \epsilon^*(N^-, N^+, d, \tau, c, \delta, \beta) \cdot \nu^4 \cdot \rho^{4\tau+A}$ *where* $\epsilon^* > 0$ *is a function and* $A \in \mathbb{R}^+$ *(we will make explicit* ϵ^* *and A along the proof).*

Then, there exists a periodic function $\hat{h}^*$ *such that*

$$\mathscr{E}[\hat{h}^*] = 0.$$

Moreover,

$$\|\hat{h} - \hat{h}^*\|_{\frac{\rho}{2}} \le C \cdot \nu^{-2}\rho^{-2\tau-A} \cdot \|\mathscr{E}[\hat{h}]\|_\rho. \tag{25}$$

The solution $\hat{h}^*$ *is the only solution of* $\mathscr{E}[\hat{h}^*] = 0$ *with zero average for* $\hat{h}^*$ *in a ball centered at* $\hat{h}$ *in* $\mathscr{A}_{\frac{3\rho}{8}}$, *i.e.* $\hat{h}^*$ *is the unique solution in the set*

$$\left\{\hat{g} \in \mathscr{A}_{\frac{3\rho}{8}} \mid \langle \hat{g} \rangle = 0,\ \|\hat{g} - \hat{h}\|_{\frac{3\rho}{8}} \le \frac{\nu^2 \cdot \rho^{2\tau}}{2\tilde{C}(N^-, N^+, d, \tau, c, C)}\right\}$$

where $\tilde{C}$ *will be made explicit along the proof.*

5. Description of the proof of main theorem

In this section, we will outline the proofs which is based on an iterative procedure. Several identities are developed to obtain a factorization. Once this factorization is obtained, we can treat the non-nearest neighbor interactions perturbatively. We will give the estimates for the iterative step in the next section.

5.1. **Motivation for the iterative step.** We will use the iterative procedure by modifying the standard Newton method that given an approximate solution $\hat{h}$ of (9), a step of the Newton method consists in finding a solution of

$$D\mathscr{E}[\hat{h}]\widehat{\Delta} = -\mathscr{E}[\hat{h}] \tag{26}$$

where D denotes the derivative of the functional $\mathscr{E}$ with respect to its argument. Then $\hat{h} + \widehat{\Delta}$ will be a better approximate solution of (9).

In fact, we compute that:

(27)
$$(D\mathscr{E}[\hat{h}]\cdot\widehat{\Delta})(\theta) = \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\sum_{j=0}^{L}\partial_\alpha^{(k)}\partial_\alpha^{(j)}\widehat{H}_L(h(\theta-k\omega)\alpha,\dots,h(\theta-k\omega+L\omega)\alpha)\cdot\widehat{\Delta}(\sigma-k\omega\alpha+j\omega\alpha)$$
$$= \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\sum_{j=0}^{L}\partial_\alpha^{(k)}\partial_\alpha^{(j)}\widehat{H}_L(\gamma_L^{(-k)}(\theta))\cdot\widehat{\Delta}(\sigma-k\omega\alpha+j\omega\alpha).$$

5.2. **Useful identities.** Let us follow the idea of [dlL08], but note that in the present case, the functions have more variables so that several of the quantities that in [dlL08] were numbers, now are vectors or matrices and order matters.

A direct calculation implies:

$$\begin{aligned}\frac{d}{d\theta}\mathscr{E}[\hat{h}](\theta) &= \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\frac{d}{d\theta}\partial_\alpha^{(k)}\widehat{H}_L(h(\theta-k\omega)\alpha,\dots,h(\theta-k\omega+L\omega)\alpha)\\ &= \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\sum_{j=0}^{L}\partial_\alpha^{(j)}\partial_\alpha^{(k)}\widehat{H}_L(\gamma_L^{(-k)}(\theta))\cdot(1+\partial_\alpha\hat{h}(\sigma-k\omega\alpha+j\omega\alpha))\end{aligned} \tag{28}$$

Let $\hat{l}(\sigma) = 1+\partial_\alpha\hat{h}(\sigma)$ and then we obtain the important identity for $D\mathscr{E}[\hat{h}]$:

$$\frac{d}{d\theta}\mathscr{E}[\hat{h}](\theta) = (\,D\mathscr{E}[\hat{h}]\cdot\hat{l}\,)(\theta). \tag{29}$$

More conceptually, if we denote $\hat{h}_\beta = \hat{h}(\sigma+\beta\alpha)+\beta$, we have

$$\mathscr{E}[\hat{h}_\beta](\theta) = \mathscr{E}[\hat{h}](\theta+\beta). \tag{30}$$

We take derivative with respect to β and evaluate at $\beta = 0$ and obtain:

$$\frac{d}{d\beta}\Big|_{\beta=0}\hat{h}_\beta(\sigma) = \hat{l}(\sigma). \tag{31}$$

Therefore, we note that (29) is just the derivative with respect to β of (30) evaluated at $\beta = 0$.

5.3. **The quasi-Newton method.** Unfortunately, the equation (26) is hard to solve since it involves difference equations with non-constant coefficients. The trick that works in our case is the one that was used in [SdlL11, dlL08] (see also [LM01, Mos88, Koz83, SZ89]). Namely, we consists in solving the following equation, which is a modification of (26):

$$\hat{l}(\,D\mathscr{E}[\hat{h}]\widehat{\Delta}\,)-\widehat{\Delta}(\,D\mathscr{E}[\hat{h}]\hat{l}\,) = -\hat{l}\mathscr{E}[\hat{h}]. \tag{32}$$

The equation (32) is just the equation (26) multiplied by $\hat{l}$ and added the extra term in $\widehat{\Delta}(D\mathscr{E}[\hat{h}]\hat{l})$ in the left-hand-side. The role of the added extra term will make the left-hand-side of (32) be factorizable.

Due to (29) one can write:

$$\widehat{\Delta}\cdot(\,D\mathscr{E}[\hat{h}]\hat{l}\,) = \widehat{\Delta}\cdot\frac{d}{d\theta}\mathscr{E}[\hat{h}] \tag{33}$$

The reason why this term is small and it does not affect the quadratic character of the procedure will be discussed in Section 6. Another way of dealing with this will be introduced in Appendix A.

5.4. **Solution of the equation of the quasi-Newton method.** The goal of this section is to specify the steps of an algorithm to solve (32). Once we have specified how to break down (32) into several auxiliary problems, we will present estimates for them in Section 6.

We note that several of these steps are very similar to those in [dlL08, SdlL11]. Nevertheless we carry them out in detail because now the variables are higher dimensional and it is not clear a priori that the algebraic operations are still valid. Of course, we also want the paper to be self-contained.

Let

$$\widehat{\Delta} = \hat{l} \cdot \hat{\eta}. \tag{34}$$

The unknowns $\widehat{\Delta}$ and $\hat{\eta}$ are equivalent due to the non-degeneracy assumptions in Theorem 1.

Substituting (34) into (32), we obtain that the equation to be solved for the step of the modified Newton method is:

$$\begin{aligned} &\sum_{L\in\mathbb{N}} \sum_{k=0}^{L} \sum_{j=0}^{L} \partial_\alpha^{(k)} \partial_\alpha^{(j)} \widehat{H}_L(\gamma_L^{(-k)}(\sigma)) \hat{l}(\sigma) \hat{l}^{(j-k)}(\sigma) \hat{\eta}^{(j-k)}(\sigma) \\ &\qquad - \sum_{L\in\mathbb{N}} \sum_{k=0}^{L} \sum_{j=0}^{L} \partial_\alpha^{(k)} \partial_\alpha^{(j)} \widehat{H}_L(\gamma_L^{(-k)}(\sigma)) \hat{l}(\sigma) \hat{l}^{(j-k)}(\sigma) \hat{\eta}(\sigma) \\ &= -\hat{l}(\sigma) \mathscr{E}[\hat{h}](\theta) \end{aligned} \tag{35}$$

where $\hat{l}^{(j)}(\sigma) = \hat{l}(\sigma + j\omega\alpha)$ and $\hat{\eta}^{(j)}(\sigma) = \hat{l}(\sigma + j\omega\alpha)$.

For fixed $L \in \mathbb{N}$, we will analyze the terms that appear in the left-hand-side of (35). We note that, when $j = k = 0, \dots, L$ the term in the first sum of the left-hand-side of (35) cancels the one in the second sum.

When $j \neq k$, we observe that we have four terms involving the mixed derivatives, that is,

$$\begin{aligned} &\partial_\alpha^{(k)} \partial_\alpha^{(j)} \widehat{H}_L(\gamma_L^{(-k)}(\sigma)) \hat{l}(\sigma) \hat{l}^{(j-k)}(\sigma) \hat{\eta}^{(j-k)}(\sigma) \\ &+\partial_\alpha^{(j)} \partial_\alpha^{(k)} \widehat{H}_L(\gamma_L^{(-j)}(\sigma)) \hat{l}(\sigma) \hat{l}^{(k-j)}(\sigma) \hat{\eta}^{(k-j)}(\sigma) \\ &-\partial_\alpha^{(k)} \partial_\alpha^{(j)} \widehat{H}_L(\gamma_L^{(-k)}(\sigma)) \hat{l}^{(j-k)}(\sigma) \hat{l}(\sigma) \hat{\eta}(\sigma) \\ &-\partial_\alpha^{(j)} \partial_\alpha^{(k)} \widehat{H}_L(\gamma_L^{(-j)}(\sigma)) \hat{l}^{(k-j)}(\sigma) \hat{l}(\sigma) \hat{\eta}(\sigma) \end{aligned} \tag{36}$$

We introduce the notations

$$\begin{aligned} &[\mathcal{S}_n \hat{\eta}](\sigma) \equiv \hat{\eta}(\sigma + n \cdot \omega\alpha) - \hat{\eta}(\sigma) \qquad \forall n \in \mathbb{Z},\ \hat{\eta} \in \mathscr{A}_\rho, \\ &C_{j,k,L}(\sigma) \equiv \partial_\alpha^{(k)} \partial_\alpha^{(j)} \widehat{H}_L(\gamma_L^{(-k)}(\sigma)) \hat{l}(\sigma) \hat{l}^{(j-k)}(\sigma). \end{aligned} \tag{37}$$

With the notations (37) above, we can rearrange (36) as:

$$\begin{aligned} & C_{j,k,L}(\sigma)\cdot[\hat{\eta}^{(j-k)}-\hat{\eta}](\sigma) \\ & \qquad -C_{j,k,L}(\sigma+(k-j)\omega\alpha)\cdot[\hat{\eta}^{(j-k)}-\hat{\eta}](\sigma+(k-j)\omega\alpha) \\ & \quad \equiv -\mathcal{S}_{k-j}[C_{j,k,L}\mathcal{S}_{j-k}\hat{\eta}](\sigma) \end{aligned} \tag{38}$$

Therefore, (35) can be written as:

$$\sum_{L\in\mathbb{N}}\sum_{\substack{k,j=0\\k>j}}^{L}\mathcal{S}_{k-j}[C_{j,k,L}\mathcal{S}_{j-k}\hat{\eta}](\sigma)=\hat{l}(\sigma)\mathscr{E}[\hat{h}](\theta) \tag{39}$$

5.5. **Perturbative treatment.** The basic idea we will use is that, under the assumptions of Theorem 1, (39) can be treated as a perturbation of the nearest neighbor interactions.

To accomplish this, it will be crucial to study conditions for the invertibility of the operators $\mathcal{S}_n$. In fact, $\mathcal{S}_n:\mathscr{A}_\rho\to\mathring{\mathscr{A}}_\rho$ is diagonal on Fourier series. Due to the Diophantine property (20), for any given $\hat{\eta}\in\mathring{\mathscr{A}}_\rho$, we can find the solution of

$$\mathcal{S}_n\hat{\eta}=\hat{\eta}. \tag{40}$$

These solutions $\hat{\eta}$ are unique up to additive constants. We will denote by $\mathcal{S}_n^{-1}$ the operator that given $\hat{\eta}$ produces the $\hat{\eta}$ with zero average. This makes it into a linear operator.

Hence, we can define the operators

$$\mathcal{L}_n^{\pm}=\mathcal{S}_{\pm1}^{-1}\mathcal{S}_n$$

acting on $\mathscr{A}_\rho$ and the operators

$$\mathcal{R}_n^{\pm}=\mathcal{S}_n\mathcal{S}_{\pm1}^{-1}$$

defined for $\mathring{\mathscr{A}}_\rho$.

The key observation is that we have

$$\begin{aligned} \|\mathcal{L}_n^{\pm}\|_{\mathscr{A}_\rho} &\le |n| \\ \|\mathcal{R}_n^{\pm}\|_{\mathring{\mathscr{A}}_\rho} &\le |n| \end{aligned} \tag{41}$$

in spite of the fact that $\mathcal{S}_{\pm}^{-1}$ are unbounded operators (see the elementary proof in [dlL08, Lemma 6]).

Therefore, (39) can be written as

$$\begin{aligned} \hat{l}(\sigma)\mathscr{E}[\hat{h}](\theta) &= \mathcal{S}_1[C_{0,1,1}\mathcal{S}_{-1}\hat{\eta}](\sigma)+\sum_{L\ge2}\sum_{k>j}\mathcal{S}_{k-j}[C_{j,k,L}\mathcal{S}_{j-k}\hat{\eta}](\sigma) \\ &= \mathcal{S}_1[C_{0,1,1}+\sum_{L\ge2}\sum_{k>j}\mathcal{S}_1^{-1}\mathcal{S}_{k-j}C_{j,k,L}\mathcal{S}_{j-k}\mathcal{S}_{-1}^{-1}]\mathcal{S}_{-1}\hat{\eta}(\sigma) \\ &\equiv \mathcal{S}_1[[C_{0,1,1}+\mathscr{G}]\mathcal{S}_1\hat{\eta}(\sigma) \end{aligned} \tag{42}$$

We denote

$$\mathscr{G}\equiv\sum_{L\ge2}\sum_{k>j}\mathcal{S}_1^{-1}\mathcal{S}_{k-j}C_{j,k,L}\mathcal{S}_{j-k}\mathcal{S}_{-1}^{-1}\equiv\sum_{L\ge2}\sum_{k>j}\mathcal{L}_{k-j}^{+}C_{j,k,L}\mathcal{R}_{j-k}^{-}. \tag{43}$$

5.6. **Algorithm.** The procedure to solve (42) is to follow the following steps.

(1) It is easy to check that

$$\begin{aligned}\int_{\mathbb{T}^d} \hat{l}(\sigma)\cdot \mathscr{E}[\hat{h}]\, d\sigma &= \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\int_{\mathbb{T}^d} \hat{l}(\sigma)\cdot \partial_\alpha^{(k)}\widehat{H}_L(\gamma_L^{(-k)})\, d\sigma\\ &= \sum_{L\in\mathbb{N}}\sum_{k=0}^{L}\int_{\mathbb{T}^d} \partial_\alpha^{(k)}\widehat{H}_L(\gamma_L(\sigma))\cdot \hat{l}^{(k)}(\sigma)\, d\sigma\\ &= \sum_{L\in\mathbb{N}}\int_{\mathbb{T}^d} d\widehat{H}_L(\gamma_L(\sigma)) = 0.\end{aligned}$$

(2) Solve

$$\mathcal{S}_1\widehat{W}(\sigma) = \hat{l}(\sigma)\cdot\mathscr{E}[\hat{h}]$$

where $\widehat{W} = \widehat{W}^0 + \overline{\widehat{W}}$. More explicitly, $\widehat{W}^0$ with zero average and $\overline{\widehat{W}}$ is some constant such that

$$\int_{\mathbb{T}^d}(C_{0,1,1} + \mathscr{G})^{-1}[\widehat{W}]d\sigma = 0.$$

(3) Solve

$$\mathcal{S}_{-1}[\hat{\eta}](\sigma) = (C_{0,1,1} + \mathscr{G})^{-1}[\widehat{W}](\sigma)$$

(4) Finally, we obtain the improved solution:

$$\hat{h} + \widehat{\Delta} = \hat{h} + \hat{l}\cdot\hat{\eta}.$$

6. Estimates for the iterative step

The goal of this section is to provide precise estimates for the iterative step. Following standard practice in KAM theory, we will denote by C numbers that depend only on combinatorial factors but are independent of the size of the domains considered, the Diophantine constants ν or the size of the error assumed. In our case, we will also require that they are independent of L, the range of the interactions. The meaning of these constants can change from one formula to the other.

6.1. **Estimates for $\overline{\widehat{W}}$.** Due to the assumption (H3) in Theorem 1, we obtain the following estimate:

$$\begin{aligned}\|\mathscr{G}\|_{\mathring{\mathscr{A}}_\rho} &\le \sum_{L\ge 2} M_L (N^+)^2 \sum_{0\le j<k\le L} |j-k|^2\\ &\le C\sum_{L\ge 2} M_L L^3 = \beta.\end{aligned} \tag{44}$$

Hence, by (H4) and (H5) the usual Neumann series shows that the operator $C_{0,1,1} + \mathscr{G}$ is boundedly invertible from $\mathring{\mathscr{A}}_\rho$ to $\mathring{\mathscr{A}}_\rho$. Moreover, we have

$$
\begin{aligned}
\|(C_{0,1,1} + \mathscr{G})^{-1} - C_{0,1,1}^{-1}\|_{\mathring{\mathscr{A}}_\rho} &\leq \|[C_{0,1,1}(Id + C_{0,1,1}^{-1}\mathscr{G})]^{-1} - C_{0,1,1}^{-1}\|_{\mathring{\mathscr{A}}_\rho} \\
&\leq \| \sum_{j=0}^{\infty} (-C_{0,1,1}^{-1}\mathscr{G})^j C_{0,1,1}^{-1} - C_{0,1,1}^{-1}\|_{\mathring{\mathscr{A}}_\rho} \\
&\leq \|C_{0,1,1}^{-1}\|_\rho \sum_{j=1}^{\infty} \|C_{0,1,1}^{-1}\mathscr{G}\|_{\mathring{\mathscr{A}}_\rho}^j \\
&\leq (N^-)^2 T \frac{1}{1-(N^-)^2 T\beta} \|C_{0,1,1}^{-1}\mathscr{G}\|_{\mathring{\mathscr{A}}_\rho} \\
&\leq (N^-)^2 T \frac{(N^-)^2 T\beta}{1-(N^-)^2 T\beta} \\
&\leq (N^-)^2 T.
\end{aligned} \tag{45}
$$

The last inequality follows from the assumption (H5).

The equation for $\overline{\widehat{W}} \in \mathbb{R}$ can be written as:

$$
\int_{\mathbb{T}^d} C_{0,1,1}^{-1} \overline{\widehat{W}} + \int_{\mathbb{T}^d} [(C_{0,1,1} + \mathscr{G})^{-1} - C_{0,1,1}^{-1}][\widehat{W}] d\sigma = - \int_{\mathbb{T}^d} C_{0,1,1}^{-1} \widehat{W}^0.
$$

Therefore, we have

$$
|\overline{\widehat{W}}| \leq U \|(C_{0,1,1} + \mathscr{G})^{-1} - C_{0,1,1}^{-1}\|_{\mathring{\mathscr{A}}_\rho} \cdot (\|\widehat{W}^0\|_\rho + |\overline{\widehat{W}}|) + U(N^-)^2 T \|\widehat{W}^0\|_\rho
$$

By the assumption (H5) of Theorem 1 and the estimates (45), we obtain $|\overline{\widehat{W}}| \leq \|\widehat{W}^0\|_\rho$. Hence

$$
\|\widehat{W}\|_\rho \leq 2\|\widehat{W}^0\|_\rho. \tag{46}
$$

6.2. **Estimates for equations involving small divisors for one iterative step.** We will follow the steps of Algorithm 5.6 but we take care of ensuring that all the steps are well defined and give estimates step by step.

Since we will lose domain repeatedly we will introduce auxiliary positive numbers $\rho' < \rho'' < \rho$ such that $\rho'' = \rho - \frac{\bar{\sigma}}{2}$ where we denote $\bar{\sigma} = \rho - \rho'$.

In step (2), we estimate using the Banach algebra property

$$
\|\hat{l}(\sigma) \cdot \mathscr{E}[\hat{h}]\|_\rho \leq N^+ \|\mathscr{E}[\hat{h}]\|_\rho.
$$

Then, by Lemma 2, we have

$$
\|\widehat{W}^0\|_{\rho''} \leq C\nu^{-1}\bar{\sigma}^{-\tau} N^+ \|\mathscr{E}[\hat{h}]\|_\rho.
$$

Therefore, due to (46) and Lemma 2, we obtain in step (3)

$$
\|\hat{\eta}\|_{\rho'} \leq C\nu^{-1}\bar{\sigma}^{-\tau} 2\|\widehat{W}^0\|_{\rho''} \leq C\nu^{-2}\bar{\sigma}^{-2\tau} N^+ \|\mathscr{E}[\hat{h}]\|_\rho. \tag{47}
$$

Hence, we will have the estimates for the solution $\widehat{\Delta}$ of (32)

$$
\|\widehat{\Delta}\|_{\rho'} \leq \|\hat{l}(\sigma)\|_{\rho'} \cdot \|\hat{\eta}\|_{\rho'} \leq C\nu^{-2}\bar{\sigma}^{-2\tau} (N^+)^2 \|\mathscr{E}[\hat{h}]\|_\rho. \tag{48}
$$

Now we will use the Taylor estimates to show that the error of the improved approximate solutions is *tame quadratic* in the sense of Nash-Moser theory. The proof consists in showing that $\mathscr{E}[\hat{h}+\widehat{\Delta}]$ is well defined (i.e. we can perform the compositions indicated in the definition) and that we can justify the estimates of the Taylor theorem in $\mathscr{E}$

Formally, we have

$$
\begin{aligned}
\mathscr{E}[\hat{h}+\widehat{\Delta}] &= (\,\mathscr{E}[\hat{h}+\widehat{\Delta}]-\mathscr{E}[\hat{h}]-D\mathscr{E}[\hat{h}]\widehat{\Delta}\,)+\hat{l}^{-1}(\,\hat{l}\mathscr{E}[\hat{h}]+\hat{l}(D\mathscr{E}[\hat{h}]\cdot\widehat{\Delta})\,)\\
&= (\,\mathscr{E}[\hat{h}+\widehat{\Delta}]-\mathscr{E}[\hat{h}]-D\mathscr{E}[\hat{h}]\widehat{\Delta}\,)+\hat{l}^{-1}\widehat{\Delta}(\,D\mathscr{E}[\hat{h}]\cdot\hat{l}\,)\\
&= (\,\mathscr{E}[\hat{h}+\widehat{\Delta}]-\mathscr{E}[\hat{h}]-D\mathscr{E}[\hat{h}]\widehat{\Delta}\,)+\hat{l}^{-1}\widehat{\Delta}\cdot\frac{d}{d\theta}\mathscr{E}[\hat{h}].
\end{aligned}
\tag{49}
$$

The first identity holds just by adding and subtracting appropriate terms. The second equation uses $\widehat{\Delta}$ is the solution of (32) and the third identity is just (33).

If we have the assumptions

$$C\nu^{-2}\bar{\sigma}^{-2\tau}(N^+)^2\|\mathscr{E}[\hat{h}]\|_\rho \le \frac{\delta}{4}, \tag{50}$$

we will have $\|\widehat{\Delta}\|_{\rho'} \le \frac{\delta}{4}$ by (48). Therefore, it is easy to see $\hat{h}+\widehat{\Delta}$ is still in the domain of the error functional $\mathscr{E}$ and we obtain the following estimates (see [dlL08, Lemma 1] or [SdlL11, Lemma 5]).

Using the Cauchy inequality, the Banach algebra property and (48), we have:

$$\left\|\hat{l}^{-1}\widehat{\Delta}\cdot\frac{d}{d\theta}\mathscr{E}[\hat{h}]\right\|_{\rho'} \le C\nu^{-2}\bar{\sigma}^{-2\tau-1}(N^+)^2N^-\|\mathscr{E}[\hat{h}]\|_\rho^2. \tag{51}$$

We also obtain by Taylor's theorem with reminder:

$$
\begin{aligned}
\|\mathscr{E}[\hat{h}+\widehat{\Delta}]-\mathscr{E}[\hat{h}]-D\mathscr{E}[\hat{h}]\widehat{\Delta}\|_{\rho'} &\le \frac{1}{2}\sum_L M_L(L+1)\|\widehat{\Delta}\|_{\rho'}^2\\
&\le C[\sum_L M_L(L+1)]\nu^{-4}\bar{\sigma}^{-4\tau}(N^+)^4\|\mathscr{E}[\hat{h}]\|_\rho^2.
\end{aligned}
\tag{52}
$$

Hence, by (49), (51) and (52) we have

$$\|\mathscr{E}[\hat{h}+\widehat{\Delta}]\|_{\rho'} \le C\left[(N^+)^2N^-+(N^+)^4\sum_L M_L(L+1)\right]\nu^{-4}\bar{\sigma}^{-4\tau-1}\|\mathscr{E}[\hat{h}]\|_\rho^2. \tag{53}$$

6.3. **Estimates for the change of the constants which measure non-degeneracy.** As a consequence of (48), we have the following estimates for the constants that measure the non-degeneracy in Theorem 1.

We use the notation introduced in (6) and denoting by $\tilde{\gamma}_L$ the one corresponding to $\hat{h}+\widehat{\Delta}$ instead of $\hat{h}$. It is easy to see $\|\tilde{\gamma}_L-\gamma_L\|_{\rho'}=\|\widehat{\Delta}\|_{\rho'}$. We first observe that

$$dist\Big(\tilde{\gamma}_L(\Omega),(\mathbb{C}^d)^{L+1}-Domain(H_L)\Big) \ge dist\Big(\gamma_L(\Omega),(\mathbb{C}^d)^{L+1}-Domain(H_L)\Big)-\|\tilde{\gamma}_L-\gamma_L\|_{\rho'}. \tag{54}$$

We see that the new function $\hat{h}+\widehat{\Delta}$ satisfies the assumption (H3) with

$$\tilde{\delta}=\delta-C\nu^{-2}(\rho-\rho')^{-2\tau}(N^+)^2\|\mathscr{E}[\hat{h}]\|_\rho.$$

Note that M_L do not need to be changed because they are the supremum of functions over an smaller set. And by Cauchy estimates and the mean value theorem, we have

$$\begin{aligned}&\|\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\tilde{\gamma}_1(\sigma)) - \partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\gamma_1(\sigma))\|_{\rho'}\\ \le& 2M_1\|\widehat{\Delta}\|_{\rho'} \le 2M_1 C\nu^{-2}(\rho-\rho')^{-2\tau}(N^+)^2\epsilon.\end{aligned}$$

We define

$$\begin{aligned}\chi &= C\nu^{-2}(\rho-\rho')^{-2\tau}(N^+)^2\epsilon,\\ \chi' &= C\nu^{-2}(\rho-\rho')^{-2\tau-1}(N^+)^2\epsilon.\end{aligned}$$

Because $\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\gamma_1(\sigma))$ is invertible for all σ, we obtain by the Neumann series that if χ is small enough , so is $\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1 \circ \tilde{\gamma}_1$.

Adding and subtracting , we also get

$$\begin{aligned}\|\widetilde{C}_{0,1,1} - C_{0,1,1}\|_{\rho'} \le& \|\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\tilde{\gamma}_1(\sigma)) - \partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\gamma_1(\sigma))\|_{\rho'}\|\hat{l}(\sigma)\hat{l}(\sigma+\omega\alpha)\|_{\rho'}\\ &+ \|\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\tilde{\gamma}_1(\sigma))\|_{\rho'}\|\partial_\alpha\widehat{\Delta}(\sigma)\|_{\rho'}\|\hat{l}(\sigma)(\sigma+\omega\alpha)\|_{\rho'}\\ &+ \|\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1(\tilde{\gamma}_1(\sigma))\|_{\rho'}\|\partial_\alpha\widehat{\Delta}(\sigma+\omega\alpha)\|_{\rho'}\|\hat{l}(\sigma)\|_{\rho'}\\ \le& 2M_1\chi(N^+)^2 + 2M_1\chi N^+ = 2M_1((N^+)^2+N^+)\chi.\end{aligned}$$

We use the same notations as in Theorem 1 but use the ˜ to indicate that they are evaluated at the function $\hat{h} + \widehat{\Delta}$. Therefore, it is easy to check that we have

$$\begin{aligned}
\widetilde{T} &\equiv \left\|(\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1)^{-1}(\tilde{\gamma}_1(\sigma))\right\|_{\rho'} \\
&\leq T + \left\|(\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1)^{-1}(\tilde{\gamma}_1(\sigma)) - (\partial_\alpha^{(0)}\partial_\alpha^{(1)}\widehat{H}_1)^{-1}(\gamma_1(\sigma))\right\|_{\rho'} \\
&\leq T + C\delta^{-1}\chi \\
\widetilde{N}^+ &\equiv \left\|1 + \partial_\alpha(\hat{h} + \widehat{\Delta})\right\|_{\rho'} \leq N^+ + \left\|\partial_\alpha\widehat{\Delta}\right\|_{\rho'} \leq N^+ + \chi' \\
\widetilde{N}^- &\equiv \left\|(1 + \partial_\alpha(\hat{h} + \widehat{\Delta}))^{-1}\right\|_{\rho'} \leq N^- + \sum_{j=1}^{\infty}\left\|(-\frac{\partial_\alpha\widehat{\Delta}}{1 + \partial_\alpha\hat{h}})^j(1 + \partial_\alpha\hat{h})^{-1}\right\|_{\rho'} \\
&\leq N^- + \frac{\|\partial_\alpha\widehat{\Delta}\|_{\rho'}(N^-)^2}{1 - \|\partial_\alpha\widehat{\Delta}\|_{\rho'}} \leq N^- + \frac{\chi'(N^-)^2}{1-\chi'} \\
\widetilde{U}^{-1} &\equiv \left|\int_{\mathbb{T}^d}\widetilde{C}_{0,1,1}^{-1}\right| = \left|\int_{\mathbb{T}^d}\left\{C_{0,1,1}[Id + C_{0,1,1}^{-1}(\widetilde{C}_{0,1,1} - C_{0,1,1})]\right\}^{-1}\right| \\
&\geq U^{-1} - \frac{(N^-)^4T^2 2M_1((N^+)^2 + N^+)\chi}{1 - (N^-)^2T2M_1((N^+)^2 + N^+)\chi} \\
|\tilde{c} - c| &\equiv \left|\langle\frac{1}{(\hat{l} + \widehat{\Delta})\cdot(\hat{l} + \widehat{\Delta})\circ T_{-\omega\alpha}}\rangle - \langle\frac{1}{\hat{l}\cdot\hat{l}\circ T_{-\omega\alpha}}\rangle\right| \\
&= \left|\langle\frac{\hat{l}\cdot\widehat{\Delta}\circ T_{-\omega\alpha} + \widehat{\Delta}\cdot(\hat{l} + \widehat{\Delta})\circ T_{-\omega\alpha})}{(\hat{l} + \widehat{\Delta})\cdot(\hat{l} + \widehat{\Delta})\circ T_{-\omega\alpha}\cdot\hat{l}\cdot\hat{l}\circ T_{-\omega\alpha}}\rangle\right| \\
&\leq \frac{2N^+ + \chi}{(N^-)^2(N^- - \chi)^2}\chi.
\end{aligned} \tag{55}$$

6.4. **A direct proof of the convergence of the procedure in the analytic case.** We consider a system which satisfies the hypotheses of Theorem 1. We label with a sub-index n all the elements corresponding to the n iterative step. We start with a function $\hat{h}_0 \in \mathscr{A}^1_{\rho_0}$ which defined in a domain parameterized by ρ_0. We choose a sequence of parameters

$$\rho_n = \rho_{n-1} - \frac{\rho_0}{4}2^{-n} = \rho_0(1 - \frac{1}{4}\sum_{j=0}^{n}2^{-j}). \tag{56}$$

(Note that the factor $\frac{1}{4}$ in (56) was omitted in [dlL08]. As we see now, this only leads to a redefinition of the constant Cin (57).)

We try the iterative step so that the n iterative step starts with a function $\hat{h}_n$ defined in a domain of radius ρ_n and ends up with a function $\hat{h}_{n+1} = \hat{h}_n + \widehat{\Delta}_n$ defined in a domain of radius ρ_{n+1}.

Since the non-degeneracy conditions (55) are bounded uniformly, by (53), we have:

$$
\begin{aligned}
\epsilon_n &\le C\nu^{-4}\rho_0^{-4\tau-1}2^{(n+1)(4\tau-1)}\epsilon_{n-1}^2 \\
&\le (C\nu^{-4}\rho_0^{-4\tau-1})^{1+2}2^{(4\tau-1)(n+1+2n)}(\epsilon_{n-2}^2)^2 \\
&\cdots \\
&\le (C\nu^{-4}\rho_0^{-4\tau-1})^{1+2^1+\ldots+2^n}2^{(4\tau-1)(n+1+2^1n+\ldots+2^n1)}\epsilon_0^{2^n} \\
&\le (C\nu^{-4}\rho_0^{-4\tau-1})^{2^{n+1}-1}2^{(4\tau-1)2^nB}\epsilon_0^{2^n},
\end{aligned}
\tag{57}
$$

where $B = \sum_{j=0}^{\infty}(j+1)2^{-j}$. It is easy to see by making ϵ_0 small enough, the right-hand-side of (57) decreases faster than any exponential.

If we apply n times the inductive step, we see that the distance from the range of $\hat{\eta}_L$ to the complement of the domain of definition of $\widehat{H}_L$ is at least

$$
\begin{aligned}
&\delta - \sum_{j=0}^{n} \|\widehat{\Delta}_j\|_{\rho_j} \\
\ge&\delta - \sum_{j=0}^{n} C\nu^{-2}(\rho_0 2^{-j-1})^{-2\tau}(N^+)^2(C\nu^{-4}\rho_0^{-4\tau-1})^{2^{j+1}-1}2^{(4\tau-1)2^jB}\epsilon_0^{2^j}
\end{aligned}
\tag{58}
$$

Note that if ϵ_0 is small enough, this is bounded from below by $\frac{3}{4}\delta$ independent of n. And we see that (50) is satisfied independently of n when n is large enough.

In summary, under smallness conditions in ϵ_0, we conclude that the iterative step can be carried out infinitely often and the assumptions on the non-degeneracy constants make in the estimates for the step remain valid.

6.5. **Remarks on the finite differentiable case.** The method developed in this paper can be adapted to produce results in the case that the interactions H_L are finitely differentiable. This is well known to experts in KAM theory. In this short subsection, we just summarize, without any proof, some of the considerations involved. Of course, this section does not affect any results in this paper.

We note if H_L are C^{r+A} with A large enough and $r \ge r_0 > d/2$ the procedure presented here, the result Theorem 1 can be adapted to producing results in which the Sobolev norm H^r is used to measure the proximity of functions. This is done in great detail in [CdlL10b] in the periodic case and in [SdlL11].

Note that, irregardless of the space that we are working on, we can use the manipulations presented here which allow to compute the corrections in the quasi-Newton equations by solving homology equations, performing algebraic operations and taking derivatives. To obtain the estimates of the iterative step, we just use the estimates of the cohomology equations as well as estimates on the derivative and estimates on the Taylor expansion. The solution of the cohomology equations is bounded from one Sobolev space to another of lower index. The estimates on products and Taylor estimates remain true in the case of Sobolev spaces for H^r provided that $r > d/2$ and that the functions involved are sufficiently differentiable. The estimates needed differ little from those in [CdlL09].

Hence, applying the method described here, we obtain estimates for the change $||\Delta||_{H^{r-A}} \leq C(1 + ||e||_{H^r})||e||_{H^r}$ and the new error satisfies $||\tilde{e}||_{H^{r-A}} \leq C||h||_{H^r}$. There are abstract implicit function theorems that show that, in these circumstances, if we alternate applying the Quasi-Newton method and some smoothing, we obtain a convergent algorithm (one needs to take care of how do the non-degeneracy condition change). There are several such abstract implicit function theorems that show that indeed this is the case. One particularly well suited for steps such as the step here can be found in Appendix A in [CdlL10b].

Results in C^r regularity can be also obtained aligning the lines outlined above. A general method which yields better regularity was developed [Mos66b, Mos66a] and streamlined in [Zeh75]. The method is based on showing that C^r functions $r \notin \mathbb{N}$ are characterized by quantitative estimates the speed of convergence of analytic approximations in decreasing domains of analyticity. Hence, one can deduce more or less automatically.

The fact that we have an a-priori format of the theorem, including local uniqueness, leads more or less automatically (as shown in [CdlL10b]) to several results, including a bootstrap of regularity (all sufficiently smooth solutions are analytic for analytic problems). More importantly, they lead to a numerically accessible criterion for breakdown that was proved to be very effective (namely that the analytic circle break down if and only if some of the Sobolev norms blow up (provided that we can check that some non-degeneracy conditions remain true). The Sobolev norms are easy to compute numerically and, developing a proof based on Sobolev norms is useful for the numerical implementation.

Appendix A. A specific model

The goal of this appendix is to present in detail the treatment of a concrete model. We hope that this direct treatment can be read directly and serve as a motivation for the general constructions. (Actually, we present a small modification that leads to a slightly more general result). In this model, we can also have very explicit algorithms, similar to the algorithms that were implemented in [CdlL09, CdlL10a]. We also obtain better estimates taking advantage of the fact that the interactions are pair interactions, an assumption that happens very often in Physics.

In this appendix, we use a slight modification of the general method studied in the paper, which is better suited for numerics and leads to slightly more general results. The method consists in adding an extra parameter to the unknowns. This allows to deal with forces that do not derive from a potential. We also establish a "vanishing lemma" which shows that, when the forces derive from a potential, the extra parameter vanishes. This is very reminiscent of the proof of invariant circle theorems using translated curve theorems, standard in KAM theory. The motivation is that, as we have seen, the Newton method requires solving equations of the form (21), which have an obstruction. In the main text, we showed that, for variational problems, this obstruction vanishes at each step In this appendix, we just add an extra parameter that fixes this obstruction without using any variational structure. When the iteration has finished, we use the variational structure to show that the extra parameter is zero (vanishing lemma).

The models studied in this appendix are described by the energy,

$$\mathscr{S}(\{u_n\}_{n\in\mathbb{Z}}) = \sum_{i\in\mathbb{Z}}[\frac{1}{2}(u_n - u_{n+1})^2 + \frac{1}{2}\sum_{j=2}^{\infty} A_j(u_n - u_{n+j})^2 - V(u_n)] \tag{59}$$

where $\sum_{j=2}^{\infty} A_j$ is sufficiently small and $V(\theta) = \widehat{V}(\theta \cdot \alpha)$ for any $\theta \in \mathbb{R}$ and $\alpha \in \mathbb{R}^d$ which is non-resonant satisfying (19), $\widehat{V}$ is an analytic function.

Remark 1. *When $A_L = CL^{-\gamma}$ the models* (59) *are related to the Hierarchical model of* [Dys69, Dys71].

Note that $\gamma = 2$ is the critical case in the hierarchical model. The results in this appendix apply to any $\gamma > 2$. The paper [CdlL10a] *considers numerically the destruction of invariant circles for many of these models in the periodic case and finds evidence that the breakdown is different for $\gamma = 2$, We think it would be interesting to carry out similar investigations in the quasi-periodic case.*

A.1. **Equilibrium equations and hull function approach.** The goal of this section is to formulate the equilibrium equations and reformulate them in terms of the hull function.

A.1.1. *The equilibrium equations.* Heuristically, we take the formal derivative of the formal functional (59) with respect to u_n and obtain the equilibrium equation:

$$u_{n+1} + u_{n-1} - 2 \cdot u_n + \sum_{j=2}^{\infty} A_j \cdot (u_{n+j} + u_{n-j} - 2 \cdot u_n) + V'(u_n) = 0 \qquad \forall\, n \in \mathbb{Z}. \tag{60}$$

In contrast with (59), the left-hand-side of (60) converges uniformly when we make assumptions on the decay of the interactions with the distance. Our study will be based exclusively on the equation (60) here.

A.1.2. *The equilibrium equations in terms of hull functions.* We are interested in finding what are called *plane-like configurations* in homogenization theory. These are configurations of the form:

$$u_n = h(n \cdot \omega) = n \cdot \omega + \hat{h}(n \cdot \omega \cdot \alpha) \tag{61}$$

where $\hat{h}$ is a function on $\mathbb{T}^d$ and $n \in \mathbb{Z}, \omega \in \mathbb{R}, \alpha \in \mathbb{R}^d$. In solid state physics, the function h is often referred as *"hull"* function of the configuration.

Hence, by (5), we can write (60) in terms of the hull function h:

$$h(\theta+\omega)+h(\theta-\omega)-2h(\theta)+\sum_{j=2}^{\infty} A_j\cdot[h(\theta+\omega)+h(\theta-\omega)-2h(\theta)]+V'(h(\theta)) = 0, \tag{62}$$

where $\theta,\ \omega \in \mathbb{R}$.

When $\omega \cdot k \neq 0$ for every $k \in \mathbb{Z}^d - \{0\}$ (62) is equivalent to

$$\begin{aligned}&\hat{h}(\sigma + \omega\alpha) + \hat{h}(\sigma - \omega\alpha) - 2\hat{h}(\sigma) + \\ &\sum_{j=2}^{\infty} A_j \, [\hat{h}(\sigma + j\omega\alpha) + \hat{h}(\sigma - j\omega\alpha) - 2\hat{h}(\sigma)] + \partial_\alpha V(\sigma + \alpha\hat{h}(\sigma)) = 0,\end{aligned} \tag{63}$$

where $\sigma,\ \alpha \in \mathbb{R}^d$ and $\partial_\alpha = \alpha \cdot \nabla$.

Again we note that $\sum_{j=2}^{\infty} A_j < \infty$, the functional equation (63) is well defined.

Later on, we always denote $\theta = n \cdot \omega$ for variables in $\mathbb{R}$ and $\sigma = \theta\alpha$ for variables in $\mathbb{R}^d$.

A.2. **Motivation: Quasi-Newton iteration.** We will introduce and solve more general equilibrium equation in terms of hull function $\hat{h}$ and extra parameter λ:

$$\begin{aligned}\mathscr{E}[\hat{h}, \lambda](\sigma) \equiv\ & \hat{h}(\sigma + \omega\alpha) + \hat{h}(\sigma - \omega\alpha) - 2\hat{h}(\sigma) \\ & + \sum_{j=2}^{\infty} A_j \cdot [\hat{h}(\sigma + j\omega\alpha) + \hat{h}(\sigma - j\omega\alpha) - 2\hat{h}(\sigma)] \\ & + \widehat{U}(\sigma + \alpha\hat{h}(\sigma)) + \lambda = 0\end{aligned} \tag{64}$$

where $\widehat{U}$ is a function on $\mathbb{T}^d$.

Remark 2. *We find it more convenient to deal with* (64) *studying also the case where the forces do not derive from a potential.*

Given the solution of (64)*, it suffices to show that* $\widehat{U} = \partial_\alpha \widehat{V}$ *implies* $\lambda = 0$ *in* [SdlL11]. *(See Lemma 5 and its proof in Section A.5.)*

A.2.1. *The quasi-Newton method.* In the next, we will use quasi-Newton iteration to find the solution of (64).

Suppose we have the approximate solution $[\hat{h}, \lambda]$ of (64). Here, we will heuristically devise a procedure to produce a much more approximate solution $[\hat{h}+\widehat{\Delta}, \lambda+\delta]$ of (64). Following the Newton's method, to improve an approximate solution $\hat{h}, \lambda$ we compute "first order" approximation to $\mathscr{E}[\hat{h} + \hat{\Delta}, \lambda + \delta] - \mathscr{E}[\hat{h}, \lambda]$ and require that $\hat{\Delta}, \delta$ are such that this first order increment is $-\mathscr{E}[\hat{h}, \lambda]$.

We write

$$\mathscr{E}[\hat{h}, \lambda] \equiv e \tag{65}$$

for the initial state of the Newton iteration. We think of e (and its derivatives) heuristically as small.

Taking derivatives with respect to $\hat{h}, \lambda$ in (65), we are lead to the Newton method:

$$\begin{aligned}&\widehat{\Delta}(\sigma + \omega\alpha) + \widehat{\Delta}(\sigma - \omega\alpha) - 2\widehat{\Delta}(\sigma) \\ &+ \sum_{j=2}^{\infty} A_j \cdot [\widehat{\Delta}(\sigma + j\omega\alpha) + \widehat{\Delta}(\sigma - j\omega\alpha) - 2\widehat{\Delta}(\sigma)] \\ &+ \partial_\alpha \widehat{U}(\sigma + \alpha \cdot \hat{h}(\sigma)) \cdot \widehat{\Delta}(\sigma) + \delta = -e(\theta).\end{aligned} \tag{66}$$

Our next goal is to simplify (66) so that it becomes readily solvable. We take the derivative with respect to θ in (64), the initial information for the Newton step, we get:

$$\begin{aligned}&\partial_\alpha\hat{h}(\sigma+\omega\alpha)+\partial_\alpha\hat{h}(\sigma-\omega\alpha)-2\partial_\alpha\hat{h}(\sigma)\\&+\sum_{j=2}^{\infty}A_j\cdot[\partial_\alpha\hat{h}(\sigma+j\omega\alpha)+\partial_\alpha\hat{h}(\sigma-j\omega\alpha)-2\partial_\alpha\hat{h}(\sigma)]\\&+\partial_\alpha\widehat{U}(\sigma+\alpha\cdot\hat{h}(\sigma))(1+\partial_\alpha\hat{h}(\sigma))=e'(\theta)\end{aligned}\tag{67}$$

where $\partial_\alpha = \alpha\cdot\nabla$.

Denote $\hat{l}(\sigma) = 1+\partial_\alpha\hat{h}(\sigma)$, then (67) can be written as:

$$\begin{aligned}&\hat{l}(\sigma+\omega\alpha)+\hat{l}(\sigma-\omega\alpha)-2\hat{l}(\sigma)\\&+\sum_{j=2}^{\infty}A_j\cdot[\hat{l}(\sigma+j\omega\alpha)+\hat{l}(\sigma-j\omega\alpha)-2\hat{l}(\sigma)]\\&+\partial_\alpha\widehat{U}(\sigma+\alpha\cdot\hat{h}(\sigma))\cdot\hat{l}(\sigma)=e'(\theta)\end{aligned}\tag{68}$$

Substituting (68) into (66) and omitting the term $e'(\theta)\cdot\widehat{\Delta}(\sigma)$ which we argue for the moment heuristically is quadratic, we have the Quasi-Newton equation for $[\widehat{\Delta},\delta]$:

$$\begin{aligned}&[\widehat{\Delta}(\sigma+\omega\alpha)+\widehat{\Delta}(\sigma-\omega\alpha)]\cdot\hat{l}-[\hat{l}(\sigma+\omega\alpha)+\hat{l}(\sigma-\omega\alpha)]\cdot\widehat{\Delta}\\&+\sum_{j=2}^{\infty}A_j[\widehat{\Delta}(\sigma+j\omega\alpha)+\widehat{\Delta}(\sigma-j\omega\alpha)]\cdot\hat{l}-\sum_{j=2}^{\infty}A_j[\hat{l}(\sigma+j\omega\alpha)+\hat{l}(\sigma-j\omega\alpha)]\cdot\hat{l}\\&=-(e+\delta)\cdot\hat{l}(\sigma).\end{aligned}\tag{69}$$

It is easy to check (we have given full details in the general case, but this ones are much simpler) that the following identities hold.

Lemma 3.

$$\hat{l}\cdot(D_1\mathscr{E}[\hat{h},\lambda]\,\widehat{\Delta})-\widehat{\Delta}\cdot[D_1\mathscr{E}[\hat{h},\lambda]\,\hat{l}]=-(\mathscr{E}[\hat{h},\lambda]+\delta)\cdot\hat{l}.\tag{70}$$

where D_1 is the functional derivative with respect to $\hat{h}$.

$$\mathscr{E}[\hat{h}+\widehat{\Delta},\lambda+\delta]=e'\frac{\widehat{\Delta}(\sigma)}{\hat{l}(\sigma)}+R\tag{71}$$

where $R=\widehat{U}(\sigma+\alpha\cdot(\hat{h}+\widehat{\Delta})(\sigma))-\widehat{U}(\sigma+\alpha\cdot\hat{h}(\sigma))-\partial_\alpha\widehat{U}(\sigma+\alpha\cdot\hat{h})\cdot\widehat{\Delta}(\sigma)$.

A.3. **Algorithm.** Let $\widehat{\Delta} = \hat{l} \cdot \hat{\eta}$. The left hand side of (69) will be

$$\begin{aligned} &\hat{l}(\sigma) \cdot [\hat{l}(\sigma + \omega\alpha) \cdot (\hat{\eta}(\sigma + \omega\alpha) - \hat{\eta}(\sigma)) \\ &\quad + \hat{l}(\sigma - \omega\alpha) \cdot (\hat{\eta}(\sigma - \omega\alpha) - \hat{\eta}(\sigma))] \\ &+ \sum_{j=2}^{\infty} A_j \hat{l}(\sigma) \cdot [\hat{l}(\sigma + j\omega\alpha) \cdot (\hat{\eta}(\sigma + j\omega\alpha) - \hat{\eta}(\sigma)) \\ &\quad + \hat{l}(\sigma - j\omega\alpha)(\hat{\eta}(\sigma - j\omega\alpha) - \hat{\eta}(\sigma))]. \end{aligned} \tag{72}$$

It is easy to see that

$$[\mathcal{S}_2\hat{\eta}](\sigma) = [\mathcal{S}_1\hat{\eta}](\sigma + \omega) + [\mathcal{S}_1\hat{\eta}](\sigma). \tag{73}$$

Therefore, (69) will be written as:

$$\begin{aligned} &\mathcal{S}_1[\hat{l} \circ T_{-\omega\alpha} \cdot \hat{l}]\mathcal{S}_{-1}[\hat{\eta}](\sigma) \\ &+ \sum_{j=2}^{\infty} \mathcal{S}_j[A_j \cdot \hat{l} \circ T_{-j\omega\alpha} \cdot \hat{l}]\mathcal{S}_{-j}[\hat{\eta}](\sigma) = (e + \delta) \cdot \hat{l}(\sigma). \end{aligned}$$

By (73), we have

$$\mathcal{S}_1[C_{0,1,1} + \mathscr{G}]\mathcal{S}_{-1}[\hat{\eta}](\sigma) = (e + \delta) \cdot \hat{l}(\sigma)$$

where $C_{0,1,1} = \hat{l} \circ T_{-\omega\alpha} \cdot \hat{l}$ and $\mathscr{G} = \sum_{j=2}^{\infty} \mathcal{S}_1^{-1}\mathcal{S}_j[A \cdot \hat{l} \circ T_{-j\omega\alpha} \cdot \hat{l}]\mathcal{S}_{-j}\mathcal{S}_{-1}^{-1}$.

We present the following procedure to improve an approximate solution:

(1) Choose $\delta = -\langle \hat{l} \cdot e \rangle$ such that

$$\int_{\mathbb{T}^d} (e + \delta) \cdot \hat{l}(\sigma) d\sigma = 0.$$

(2) Solve

$$\mathcal{S}_1 \widehat{W}(\sigma) = (e + \delta) \cdot \hat{l}(\sigma)$$

where $\widehat{W} = \widehat{W}^0 + \overline{\overline{W}}$. More explicitly, $\widehat{W}^0$ with zero average and $\overline{\overline{W}}$ is some constant such that

$$\int_{\mathbb{T}^d} (C_{0,1,1} + \mathscr{G})^{-1}[\widehat{W}] d\sigma = 0.$$

(3) Solve

$$\mathcal{S}_{-1}[\hat{\eta}](\sigma) = (C_{0,1,1} + \mathscr{G})^{-1}[\widehat{W}](\sigma)$$

(4) Finally, we obtain the improved solution:

$$\widehat{\Delta} = \hat{l} \cdot \hat{\eta}.$$

Remark 3. *Note that the algorithm indicated in the above steps is very efficient. Each of the steps is linear either on a discretization evaluating the function in a grid of points or discretizing the function in Fourier series, but there are efficient methods (FFT) to change from space discretizations to Fourier discretizations. Hence, if we discretize the functions by evaluating them in N points and keeping N Fourier coefficients, a Newton step requires only $O(N \ln(N))$ operations and only $O(N)$ storage. In the periodic case, these algorithms were implemented in* [CdlL09, CdlL10a]. *Implementations in the quasi-periodic case are in progress.*

A.4. **Estimate for the quasi-Newton method.** The key observation is that

$$\begin{aligned} \|\mathcal{S}_1^{-1}\mathcal{S}_n\|_\rho &\le |n| \\ \|\mathcal{S}_n\mathcal{S}_{-1}^{-1}\|_\rho &\le |n| \end{aligned} \tag{74}$$

in spite of the fact that $\mathcal{S}_\pm^{-1}$ are unbounded operators (see [dlL08]).

It is easy to get the following lemma:

Lemma 4. *If* $\sum |A_L|L^2 \ll 1$*, the operator* $C_{0,1,1} + \mathscr{G}$ *is boundedly invertible from* $\mathscr{A}_\rho$ *to* $\mathscr{A}_\rho$*.*

Note that the condition of Lemma 4 is that the second moment in the coefficients is finite and small. Superficially, the condition **H3** in Theorem 1 would require that the third moment is small.

The key observation is that for the models (2), we have

$$C_{0.L.L} = A_L$$

and $C_{i,j,L} = 0$ for all other values of i, j, $i \le j$. Hence

$$\|\mathscr{G} \le \sum_{i\le j}\sum_L L\|C_{i,j,L}\|L \le \sum_L |A_L|L^2$$

□

A.5. **Proof of vanishing lemma.**

Lemma 5 (Vanishing lemma)**.** *Consider a solution of* (64) *with the stated periodic condition. If* $\widehat{U} = \partial_\alpha\widehat{V}$*, then* $\lambda = 0$*.*

Proof. We multiply (64) by $\hat{l}(\sigma) = 1 + \partial_\alpha\hat{h}(\sigma)$ and compute $\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}$ of all the terms one by one. We observe that

$$\begin{aligned} \widehat{\mathscr{L}_k} &= \hat{h}_k\cdot[e^{2\pi ik\cdot j\omega\alpha} + e^{-2\pi ik\cdot j\omega\alpha} - 2] \\ &= 2\,(\cos j\cdot\omega\alpha\cdot k - 1)\,\hat{h}_k \end{aligned} \tag{75}$$

where $\mathscr{L} = \hat{h}(\sigma + j\omega\alpha) + \hat{h}(\sigma - j\omega\alpha) - 2\hat{h}(\sigma)$ and $j \ge 1$.

By the definition of $\hat{l}$, we have

$$\hat{l}_k = \delta_{k,0} + 2\pi i\, k\cdot\alpha\,\hat{h}_k. \tag{76}$$

Hence, we obtain

$$\begin{aligned} (\mathscr{L}\cdot\hat{l})_0 &= \sum_{k\in\mathbb{Z}^d}\widehat{\mathscr{L}_k}\cdot\hat{l}_{-k} \\ &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}}[2\,(\cos j\omega\alpha\cdot k - 1)\,\hat{h}_k][-2\pi i\,k\cdot\alpha\,\hat{h}_{-k}] \\ &= \sum_{k\in\mathbb{Z}^d\setminus\{0\}} -[2\,(\cos j\omega\alpha\cdot k - 1)\,\hat{h}_{-k}][2\pi i\,k\cdot\alpha\,\hat{h}_k] \\ &= -\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\widehat{\mathscr{L}_{-k}}\cdot\hat{l}_k \\ &= 0 \end{aligned}$$

We note that this produces the formula:

$$\lambda = -\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \widehat{U}(\theta\alpha + \hat{h}(\theta\alpha)) \cdot \hat{l}(\theta\alpha) d\theta$$

We also observe that

$$\int_{-T}^{T} \partial_\alpha \widehat{V}(\theta\alpha + \hat{h}(\theta\alpha)\alpha) \cdot \hat{l}(\theta\alpha) d\theta = \widehat{V}(h(T)\alpha) - \widehat{V}(h(-T)\alpha)$$

when $\widehat{U} = \partial_\alpha \widehat{V}$. So it is bounded independent of T. When we divide the integral by $2T$ and take the limit as $T \to \infty$. We obtain 0. This finishes the proof of the lemma. □

A.6. **Statement of the result for the specific models.** Eventually, it is not difficult to obtain the following theorem.

Theorem 2. *Consider the models* (2) *with* $\widehat{V}$ *an analytic function. Let* $\mathscr{E}[\hat{h}, \lambda] \le \epsilon$ *for some* $\hat{h} \in \mathscr{A}^1_\rho$. *Assume* ω *and* $\hat{l}(\sigma)$ *satisfies the Diophantine properties (H1) and the Non-degeneracy condition (H2) in Theorem 1 respectively. Assume*

(H3') $\beta = C \sum_{L\ge 2} A_L L^2 \le \frac{1}{2}$.

When ϵ *is small enough, we are able to find the equilibrium solution to* (60).

Note that (H3') is a weakening of (H3) in Theorem 1. In our case $C_{0,1,1} = 1$, hence (H4) in Theorem 1 is automatic and the assumptions in the above theorem.

Acknowledgements

The authors would like to thank the Dept. of Mathematics at Nanjing Univ. specially prof. C-Q. Cheng for a very warm hospitality which was crucial to finishing this paper. The work of R.L has been supported by NSF grants.

References

[Bur88] S. E. Burkov. Ground states of two-dimensional quasicrystals. *J. Statist. Phys.*, 52(1-2):453–461, 1988.

[CdlL09] Renato Calleja and Rafael de la Llave. Fast numerical computation of quasi-periodic equilibrium states in 1D statistical mechanics, including twist maps. *Nonlinearity*, 22(6):1311–1336, 2009.

[CdlL10a] Renato Calleja and Rafael de la Llave. Computation of the breakdown of analyticity in statistical mechanics models: numerical results and a renormalization group explanation. *J. Stat. Phys.*, 141(6):940–951, 2010.

[CdlL10b] Renato Calleja and Rafael de la Llave. A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. *Nonlinearity*, 23(9):2029–2058, 2010.

[dlL08] Rafael de la Llave. KAM theory for equilibrium states in 1-D statistical mechanics models. *Ann. Henri Poincaré*, 9(5):835–880, 2008.

[DS09] Andrea Davini and Antonio Siconolfi. Exact and approximate correctors for stochastic Hamiltonians: the 1-dimensional case. *Math. Ann.*, 345(4):749–782, 2009.

[Dys69] Freeman J. Dyson. Existence of a phase-transition in a one-dimensional Ising ferromagnet. *Comm. Math. Phys.*, 12(2):91–107, 1969.

[Dys71] Freeman J. Dyson. An Ising ferromagnet with discontinuous long-range order. *Comm. Math. Phys.*, 21:269–283, 1971.

[FK39] J. Frenkel and T. Kontorova. On the theory of plastic deformation and twinning. *Acad. Sci. URSS, J. Physics*, 1:137–149, 1939.

[Kle01] Dmitry Kleinbock. Some applications of homogeneous dynamics to number theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 639–660. Amer. Math. Soc., Providence, RI, 2001.

[Kle08] Dmitry Kleinbock. An extension of quantitative nondivergence and applications to Diophantine exponents. *Trans. Amer. Math. Soc.*, 360(12):6497–6523, 2008.

[Koz83] S. M. Kozlov. Reducibility of quasiperiodic differential operators and averaging. *Trudy Moskov. Mat. Obshch.*, 46:99–123, 1983.

[LM01] M. Levi and J. Moser. A Lagrangian proof of the invariant curve theorem for twist mappings. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 733–746. Amer. Math. Soc., Providence, RI, 2001.

[LS03] Pierre-Louis Lions and Panagiotis E. Souganidis. Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting. *Comm. Pure Appl. Math.*, 56(10):1501–1524, 2003.

[Mat09] Daniel C. (ed.) Mattis. *The many-body problem. An encyclopedia of exactly solved models in one dimension. 3rd printing with revisions and corrections.* Hackensack, NJ: World Scientific. xxvi, 964 p., 2009.

[Mos66a] Jürgen Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.

[Mos66b] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Mos88] Jürgen Moser. A stability theorem for minimal foliations on a torus. *Ergodic Theory Dynam. Systems*, 8^*(Charles Conley Memorial Issue):251–281, 1988.

[Ran87] David Rana. *PROOF OF ACCURATE UPPER AND LOWER BOUNDS TO STABILITY DOMAINS IN SMALL DENOMINATOR PROBLEMS*. ProQuest LLC, Ann Arbor, MI, 1987. Thesis (Ph.D.)–Princeton University.

[Rue69] David Ruelle. *Statistical mechanics: Rigorous results*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[SdlL11] Xifeng Su and Rafael de la Llave. KAM theory for quasi-periodic equlibria 1-D quasiperiodic media. 2011. MP_ARC #11-56.

[SZ89] Dietmar Salamon and Eduard Zehnder. KAM theory in configuration space. *Comment. Math. Helv.*, 64(1):84–132, 1989.

[vE99] T. S. van Erp. Frenkel Kontorova model on quasiperiodic substrate potentials. 1999. Thesis.

[vEF02] T. S. van Erp and A. Fasolino. Aubry transition studied by direct evaluation of the modulation functions of infinite incommensurate systems. *Europhys. Lett.*, 59(3):330–336, 2002.

[vEFJ01] T. S. van Erp, A. Fasolino, and T. Janssen. Structural transitions and phonon localization in Frenkel Kontorova models with quasi-periodic potentials. *Ferroelectrics*, 250:421–424, 2001.

[vEFRJ99] T. S. van Erp, A. Fasolino, O. Radulescu, and T. Janssen. Pinning and phonon localization in Frenkel-Kontorova models on quasiperiodic substrates. *Physical Review B*, 60(9):6522–6528, 1999.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

Academy of Mathematics and Systems Science,Chinese Academy of Sciences, No.55 Zhongguancun East Road, Beijing, 100190, CHINA
E-mail address: billy3492@gmail.com, xfsu@amss.ac.cn

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332
E-mail address: rafael.delallave@math.gatech.edu