Quasi-periodic breathers in Hamiltonian networks of long-range coupling

Jiansheng Geng [a], Jorge Viveros [b], Yingfei Yi [b,c,*]

[a] *Department of Mathematics and Institute of Mathematical Science, Nanjing University, Nanjing 210093, PR China*
[b] *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA*
[c] *College of Mathematics, Jilin University, Changchun, 130012, PR China*

Abstract

This work is concerned with Hamiltonian networks of weakly and long-range coupled oscillators with either variable or constant on-site frequencies. We derive an infinite dimensional KAM-like theorem by which we establish that, given any N-sites of the lattice, there is a positive measure set of small amplitude, quasi-periodic breathers (solutions of the Hamiltonian network that are quasi-periodic in time and exponentially localized in space) having N-frequencies which are only slightly deformed from the on-site frequencies.
© 2008 Elsevier B.V. All rights reserved.

1. Introduction

During the past two decades or so, there has been an increasing interest in the study of coherent structures in spatially-extended Hamiltonian systems. By studying coherent structures, one is led to the understanding of solutions that are localized in space but with relatively long lifetimes. Perhaps the simplest and most studied examples of structures of this sort are *solitons* which one typically identifies as localized traveling-wave solutions to solitary equations which arise in many problems of physical relevance such as in the formation of surface water waves or in long-distance signal transmission over optical fibers. Another important examples of coherent structures are those which, in an appropriate moving reference frame, are time-periodic and spatially localized. These types of solution, named *breathers*, have been known for over thirty years in the context of the sine-Gordon equation which, among other cases of physical interest, has applications in the study of transmission lines made out of Josephson junctions. In fact, the modified Korteweg–de Vries (KdV) and sine-Gordon equations are well-known to be able to support both types of solutions, soliton-like and breather-like (cf. [1,2]).

The first numerical account we know of regarding the existence of coherent structures of the type we are concerned with in this paper can be found in [46] within the context of pure anharmonic crystal lattices with two distinct types of quartic anharmonicity (fourth-order self-interaction). In this early work it was established (via rotating-wave approximation and lattice Green-function techniques) that stationary localized vibrational modes can occur at any point along the lattice. Prior to this work, the existence of localized vibrational modes had already been established for lattices with impurities and the vibrations had been shown to occur around the impurity. Therefore, this new type of vibrational modes were indeed a feature of the (pure) lattice itself and for this reason they were called intrinsic localized modes (ILMs). The name *discrete breathers* was later coined in [14] for the type of ILMs just described. However, as we pointed out above, the term "breather" had already been in use for some time [1] in the context of the sine-Gordon equation.

---

* Corresponding author at: School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. Tel.: +1 404 894 4568.
E-mail addresses: jgeng@nju.edu.cn (J. Geng), viveros@math.gatech.edu (J. Viveros), yi@math.gatech.edu (Y. Yi).

The importance of breathers or discrete breathers resides in that it is generally believed that their understanding will shed light into the mechanisms of formation of more general coherent structures as well as of localization and transportation of energy in infinite dimensional Hamiltonian systems.

Breather-like coherent structures with two or more incommensurable frequencies are referred to as *quasi-periodic breathers*. More precisely, breathers or quasi-periodic breathers are time-periodic or quasi-periodic, spatially localized solutions of infinite-dimensional Hamiltonian systems. In the case of Hamiltonian networks (or Hamiltonian lattice systems), by localization (or self-localization) we will understand that the solutions, $(q_n)_{n\in\mathbb{Z}}$, oscillate with amplitudes that decay at least exponentially as $|n| \to \infty$. Besides being called ILMs, breathers have often been called discrete solitons, discrete breathers, or dynamical solitons, particularly so in the physics literature. We prefer to use the term breather since, as opposed to solitons, breathers may exchange energy during their interaction and ILMs is too generic a term for our purposes.

To put ideas into context, consider an 1-dimensional infinite lattice of oscillators with unit mass, embedded in a medium with which the oscillators interact. In appropriate dimensionless variables, this corresponds to a Hamiltonian system associated with the standard symplectic structure $\sum_{n\in\mathbb{Z}} dp_n \wedge dq_n$ and real analytic Hamiltonian of the form

$$H = \sum_{n\in\mathbb{Z}} \left(\frac{p_n^2}{2} + V_n(q_n) \right) + W(\{q_n\}). \tag{1.1}$$

In the above, W is the *coupling potential* that models the interaction among the lattice points and V_n is the *on-site potential* that models the interaction between the medium and the lattice point with coordinate q_n. It is usually assumed that, for all $n \in \mathbb{Z}$, $V_n(0) = V'_n(0) = 0$ and $V''_n(0) = \beta_n^2/2 > 0$, as it is the case in most physical systems. The Hamiltonian (1.1) or its associated equations of motion is referred to as a *Hamiltonian network* or a *Hamiltonian lattice system* in which β_n's are called *on-site frequencies*.

Among the Hamiltonian networks of the type (1.1) is the celebrated Frenkel–Kontorova (FK) model which describes an 1-dimensional chain of identical atoms as a dynamical system of lattice points which, in equilibrium, lie equidistant to their nearest neighbors. Furthermore, every atom is assumed to interact with the substrate or medium in the same way as every other atom in the lattice and thus the on-site potential is assumed periodic. As for the interaction among lattice points, only interactions among nearest neighbors is considered. More precisely, the on-site and coupling potentials of the FK model are

$$V_n(x) = V_{FK}(x) \equiv 1 - \cos x, \quad n \in \mathbb{Z}$$

and

$$W = \sum_{n\in\mathbb{Z}} W_{FK}(q_{n+1} - q_n)$$

respectively, where $W_{FK}(x) = gx^2/2$ and g is the elastic constant of the system. The lattice spacing in equilibrium and the period of the substrate potential are taken to be the same (equal to 2π); moreover, there is one and only one atom per energy well of the potential. Observe that the FK model can be obtained from (1.1) under the following simplifying conditions: uniformity in the atom-substrate interaction, i.e., $V_n(x) \equiv V(x)$ for all $n \in \mathbb{Z}$ and $V(x) = V(x + 2\pi)$; expanding V in its Fourier series and considering only the first harmonic (V_{FK}); restricting inter-atomic interactions to nearest neighbors and expanding W in its Taylor series only to consider second-order (harmonic) terms (W_{FK}). The result is the simplest nonlinear type of interaction one can consider. Nevertheless, since it first appeared, nearly seventy years ago, the FK model has proven to be a paradigm for a broad class of Hamiltonian systems capable of sustaining oscillations such as those arising in the adsorption of mono-atomic layers by crystal surfaces, in the insertion of a single atom ("kink") in a perfect crystal lattice, or in metals with dislocations (cf. [13] and references therein).

However, there is still a large class of discrete physical and biological systems that cannot be described by the FK model. To model the general structures and atom dynamics of a crystal lattice, extended FK models of the type (1.1) have been frequently considered with respect to both short and long range couplings of not necessarily identical oscillators. As an illustration of systems that fall out the FK model, we mention the case of vibrational modes in micromechanical cantilever arrays which require the incorporation of damping and driving terms (in case one wants to consider their manipulation) as well as an interaction potential that takes into account influence among lattice neighbors that reaches farther than the simple nearest neighbor (cf. [35,42,43]). Another relevant application of Hamiltonian networks of the type described by (1.1) can be found in modeling of the dynamics of the structure of DNA, in particular, DNA denaturation, i.e., the fluctuations in the opening of base-pairs of double stranded DNA (cf. [13,18,19,40,51]). For example, [18] introduces the idea that concentrated vibrational modes in DNA are responsible for the melting of the bonds within the DNA structure and presents a model for the dynamics of the displacements of nucleotide pairs with a quadratic nearest-neighbor interaction and an attractive-repulsive on-site potential. Hamiltonians that fall within the category described by (1.1) can also be obtained from the discretization of Hamiltonian PDEs such as Klein–Gordon, KdV and nonlinear Schrödinger (NLS) equations. But the physical interest in Hamiltonian networks really comes from dynamics which are far away from those of Hamiltonian PDEs. For instance, condensed matter physics is an area in particular where discreteness plays an important role.

Although breathers are well-known in Hamiltonian PDEs like the sine-Gordon equation and the cubic NLS equation, they appear to be rare and non-robust objects in Hamiltonian PDEs. On the contrary, numerical studies of many physical models find that the existence of breathers and quasi-periodic breathers is a general phenomenon in Hamiltonian networks, suggesting that the localization property is due to the discreteness and the nature of nonlinearities rather than disorder (see the recent review [21] and references therein).

Indeed, there is substantial experimental and numerical evidence (obtained from diverse mathematical models) of the existence of breathers (or ILMs) in many physical lattice models, for instance in superconducting arrays such as Josephson junction (JJ) arrays and Josephson ladders (cf. [36,47] and references therein). Another example is that of antiferromagnetic spin lattice systems (cf. [44] and references therein). Spectroscopy methods have also been used in the detection of breathers within (quasi-)1D crystals (e.g., [45] introduced a mathematical model with a four-site interaction term). Other works describing vibrational states in crystals include [15,29]. A common numerical iterative scheme to determine the existence of breathers is the so-called Rotating Wave Approximation (RWA). RWA was used in [41] to determine the existence of breathers in 1D monoatomic and diatomic lattices of N particles in the absence of a substrate potential ($V_n \equiv 0$) and under a two-body central (anharmonic) interaction potential of physical significance such as the Lennard-Jones or the Born–Mayer plus Coulomb, and under the action of an external driving force provided by a harmonic (sinusoidal) electric field. The results of [41] are in contrast with previous work in the undriven case which showed that the development of breathers is impossible when considering higher than fourth order terms in the Taylor expansion of the interaction potential. The RWA method however, becomes inapplicable when the range of the interaction is longer than nearest-neighbor and even in the nearest-neighbor case, the complexity of the RWA scheme increases rapidly if one wishes to consider higher than second-order Taylor approximations of the interaction term.

The need for mathematical models that can incorporate various systems of the types described above became an important research subject over the past decade, and rigorous methods for the study of breathers in Hamiltonian networks have been developed recently. Below we will mention only those works which are of relevance to this paper.

The existence of breathers in Hamiltonian networks of the type (1.1) was obtained in [34] for the nearest-neighbor interaction potential via an anti-integrability (or anti-continuum) method (cf. [3,4]). More precisely, the authors considered the Hamiltonian (1.1) associated with the coupling potential

$$W_{MA} = \frac{\varepsilon}{2} \sum_{n \in \mathbb{Z}} (q_{n+1} - q_n)^2, \quad n \in \mathbb{Z},$$

where ε is the (sufficiently small) coupling constant. Even though the proof presented in [34] involves identical oscillators (i.e., $V_n = V$ for all n), the authors point out that this condition is not necessary, provided that the excited sites are uniformly anharmonic and the non-excited sites are uniformly nonresonant. Furthermore, the authors state that their method is also applicable for interactions that reach farther than the nearest-neighbor interaction. Existence of breathers for the interaction potential W_{MA} is established near a fixed periodic orbit of the uncoupled Hamiltonian via an implicit-function argument, provided that one either restricts oneself to work in a space of time-reversible solutions, or the phase is "quenched", in addition to that, one must impose conditions of non-resonance (the frequency of the anharmonic mode to be continued as well as the frequencies of all of its harmonics must lie outside the phonon band) and of anharmonicity (the frequency of the mode must be a non-constant function of its action (cf. [5])).

Motivated by proving the exponential stability of breathers conjectured in [34,6] also presents a slightly different proof of their existence by combining Nekhoroshev normal-form and Poincaré continuation-theorem ideas to develop a general theorem which gives a local normal form without requiring the use of action-angle variables. The idea presented in [6] is strong enough to prove the existence of breathers in the case of identical oscillators ($V_n \equiv V$) with a long-range coupling potential; i.e., a system with Hamiltonian (1.1) and interaction potential

$$W_B = \frac{\varepsilon}{4} \sum_{m \neq n} \frac{1}{|m - n|^\alpha} (q_m - q_n)^2,$$

where $\alpha > 1$ and ε is sufficiently small (note that when $\alpha \to \infty$ the interaction becomes nearest-neighbor). The existence of quasi-periodic breathers for the nearest-neighbor interaction potential is established later on in [7] for certain systems with symmetries using the same type of Poincaré–Lyapunov continuation techniques.

In light of the previous results, work started to emerge on interaction potentials of order higher than two. A KAM (Kolmogorov–Arnold–Moser) method for establishing the existence of quasi-periodic breathers with respect to any finite number of incommensurate frequencies was developed in [50] for the system (1.1) of identical oscillators with third-order nearest-neighbor interaction, i.e.,

$$W_Y = \frac{\varepsilon}{3} \sum_{n \in \mathbb{Z}} (q_{n+1} - q_n)^3,$$

where ε is sufficiently small. This work overcame difficulties arising from the infinite multiplicities of normal frequencies, contrasting with previous work on the existence of quasi-periodic solutions in infinite dimensional Hamiltonian systems which required

finite multiplicities (cf. [31,32,37,38,48]). We mention that the proof of the main result in [50] relies heavily on the assumption that the interaction potential is third-order, which is important in avoiding the appearance of the continuous spectrum associated with linearization of the system. The anti-continuity method can also be applied to show the existence of quasi-periodic solutions, but for systems with no more than two incommensurate frequencies; e.g., for the discrete nonlinear Schrödinger equation (cf. [28]).

For the long-range coupling potential

$$W(\{q_n\}) = \frac{1}{3} \sum_{m \neq n} e^{-|n-m|^\alpha} |q_m - q_n|^3$$

the existence of N-quasi-periodic breathers was shown in recent work of the first and the third author [27] with respect to variable on-site frequencies (i.e., treating N of the β_n's as parameters) under certain gap conditions among the frequencies.

Almost-periodic breathers, i.e., spatially localized solutions to (1.1) with infinitely many incommensurate frequencies, have also been investigated. Associated with the third-order, nearest-neighbor interaction potential W_Y, [22] contains a study of the case when frequencies are non-negative random variables with smooth distribution of fast decay at infinity and showed that there is a set $\Omega \subset \mathbb{R}_+^\infty$ of positive probability measure, such that each $\omega \in \Omega$ corresponds to an almost-periodic breather (see also [39] for spatial structures of a more general type).

The present work is concerned with proving the existence of quasi-periodic breathers in an 1-dimensional lattice of oscillators with weak coupling of a long-range type; i.e., in a system with Hamiltonian

$$H = \sum_{n \in \mathbb{Z}} \left(\frac{p_n^2}{2} + V_n(q_n) \right) + \varepsilon W(\{q_n\}), \tag{1.2}$$

where V_n's are on-site potentials satisfying $V_n(0) = V_n'(0) = 0$ and $V_n''(0) = \frac{\beta_n^2}{2} > 0$, ε is sufficiently small, and $W(\{q_n\})$ is the interaction potential of the form

$$W(\{q_n\}) = \frac{1}{p} \sum_{m \neq n} C(m, n)(q_m - q_n)^p, \quad p \geq 3, \tag{1.3}$$

in which $C(m, n)$'s are either *exponential coupling coefficients*, i.e., $C(m, n) = O(e^{-\beta|n-m|})$ as $|n - m| \to \infty$ with respect to a fixed constant $\beta > 0$, or *power-law coupling coefficients*, i.e., $C(m, n) = O(\frac{1}{|n-m|^\alpha})$ as $|n - m| \to \infty$ with respect to a fixed constant $\alpha > 1$.

Our choice of Hamiltonian above is motivated by the need to model physical systems, such as the adsorption of atoms (adatoms) on crystal surfaces, which can exhibit different types of excitation depending on the charge of the adatoms, the nature of the substrate, etc. In such a system, atoms at the equilibrium states are assumed to be a distance a_s apart from their nearest neighbors just as in the case of the FK model. More precisely, the medium is described in terms of a single smooth, periodic-like potential V satisfying $V(na_s) = V'(na_s) = 0$, $V''(na_s) = \frac{\beta_n^2}{2} > 0$, $n \in Z$ (e.g., $V(x) = 1 - \cos \frac{2\pi}{a_s} x$), which has energy wells each occupied by one single atom (thus the wells are also a_s units of distance apart). In [13] the distinction is made among four types of interaction potential: exponential potential, power-law potential, Morse potential, and double-well potential, that model a large class of physically relevant systems involving long-range interactions of identical neutral atoms that interact with one another via their electron clouds. The first two potentials are convex and repulsive, whereas the other two are nonconvex. Here we will focus on the first two potentials, i.e., the exponential potential

$$W_{\exp}(d) = W_0 e^{-\beta(d-a_s)}, \quad \beta > 0 \tag{1.4}$$

and the power-law potential

$$W_{pl}(d) = \frac{W_0}{(d - a_s)^\alpha}, \quad \alpha > 1, \tag{1.5}$$

where W_0 is the strength of the energy of interaction between nearest-neighbor adatoms and $d \geq 0$ is the distance between any two adatoms (not necessarily adjacent) in the chain at a particular (fixed) time. Observe that if $d < a_s$, which can be the case when two adatoms are vibrating in adjacent energy wells, then the system is at a higher energy configuration in regards to its equilibrium configuration. Thus, in order to lower their energy, it is necessary for adatoms to move farther apart, at which point the interaction with the other adatoms in the chain becomes relevant.

To derive our Hamiltonian, we let $V_n(q) = V(na_s + q)$. Then $V_n(0) = V_n'(0) = 0$ and $V_n''(0) = \frac{\beta_n^2}{2}$ for every $n \in \mathbb{Z}$. Let x_m be the absolute position of the mth adatom with respect to a fixed origin and let q_m be its relative position with respect to equilibrium; i.e., $x_m = ma_s + q_m$. Then $d_{m,n} = |x_m - x_n| = |m - n|a_s + \sigma(q_m - q_n)$ where $\sigma = \text{sign}(m - n)$, and, with respect to the exponential
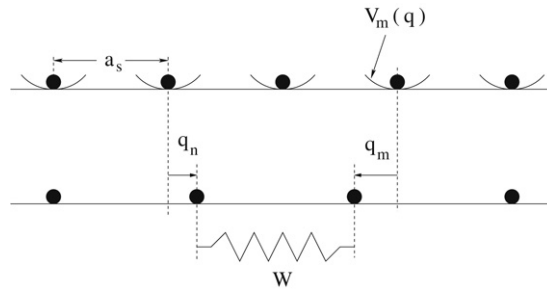
Fig. 1. Mono-atomic lattice with interaction potential of exponential type.

potential, we have

$$
W = \frac{1}{2} \sum_{m \neq n} W_{\exp}(|x_m - x_n|)
$$

$$
= \frac{1}{2} \sum_{m \neq n} W_0 e^{-\beta(|m-n|a_s + \sigma(q_m - q_n) - a_s)} = \frac{1}{2} \sum_{m \neq n} W_0 e^{-\beta(|m-n|-1)a_s} e^{-\beta\sigma(q_m - q_n)}
$$

$$
= \frac{W_0}{2} \sum_{m \neq n} e^{-\beta(|m-n|-1)a_s} \left(1 - (\beta\sigma)(q_m - q_n) + \frac{1}{2!}(\beta\sigma)^2 (q_m - q_n)^2 - \frac{1}{3!}(\beta\sigma)^3 (q_m - q_n)^3 + \cdots \right).
$$

Thus our coupling potential (1.3) with exponential coupling coefficients corresponds to considering only the pth-order interaction terms, where $p \geq 3$, in the Taylor expansion of $W_{\exp}$ in relative coordinates. The interaction of adatoms in relative coordinates are demonstrated in Fig. 1. This idea of identifying various types of dominating factors out of a fundamental interaction potential via Taylor expansions has been frequently used in physics. For instance, if in the expansion above one considers only the nearest-neighbor cubic terms together with all ($m \neq n$) quadratic terms, a potential of the so-called Kac-Baker form is obtained (cf. [13] p. 54).

Based on a similar Taylor expansion of the power-law potential W_{pl}, we can also obtain the interaction potential (1.3) with respect to the power-law coupling coefficients.

Our main results will lie in two categories: one for a family of long range Hamiltonian networks and the other one for a fixed Hamiltonian network. The former case typically occurs when long range coupling of harmonic oscillators is considered, for which, due to the lack of parameters, a finite set of on-site frequencies $\beta'_n s$ need to be treated as parameters in order for quasi-periodic breathers to exist. The later case concerns the coupling of a fixed family of anharmonic oscillators. For each $n \in \mathbb{Z}$, because the on-site potential V_n is assumed to be locally convex, for any $0 < h \ll 1$, the equation

$$
\frac{p^2}{2} + V_n(p) = h
$$

defines a simple closed curve, $\Gamma_n(h)$, on the pq-plane, that encloses the origin. Let

$$
\rho_n(h) = \oint_{\Gamma_n} p \, \mathrm{d}q \tag{1.6}
$$

be the area enclosed by $\Gamma_n(h)$. Then $\rho_n(h) = O(|h|)$ is real analytic and $\rho'_n(h) \neq 0$ for all $0 < h \ll 1$. The potential V_n or the corresponding oscillator is said to be *anharmonic* if $\rho''_n(h) \neq 0$ for all $0 < h \ll 1$. An anharmonic oscillator admits a set of frequencies near 0 which gives parameters needed for the problem. A typical example of anharmomic potential is the sinusoidal one.

This paper is organized as follows. In Section 2 we state our main results into two theorems and also introduce an abstract infinite-dimensional KAM theorem which implies our main results. Sections 3 and 4 are devoted to the proof of the abstract KAM theorem. More precisely, in Section 3, we give a detailed construction of the KAM iteration for one KAM step. The proof of the abstract infinite dimensional KAM theorem is completed in Section 4 where we provide an iteration lemma and show the convergence and measure estimates. Some technical lemmas are included in the Appendix.

2. Main results

2.1. Quasi-periodic breathers for Hamiltonian networks

Consider the Hamiltonian network (1.2) with the long-range interaction potential (1.3). Let $N > 1$ be a prescribed integer and $\mathcal{J} = \{n_1, \ldots, n_N\}$ be a given set of N-sites. In the case of power-law interaction potential, we further assume that the power α in the power-law coupling coefficients is greater than $N + 80$.

We first consider the case with variable on-site frequencies that $\omega = (\beta_{n_1}, \ldots, \beta_{n_N})$ is treated as parameters in a bounded closed region $\mathcal{O}$ in R_+^N, while the remaining β_n's are fixed.

Theorem A. *Consider the Hamiltonian networks* (1.2) *with the interaction potential* (1.3). *Let $N, \mathcal{J}$ be as the above and assume that $|\beta_n| \geq \frac{1}{|n|^2}$ as $|n| \gg 1$. Then there exist a family of Cantor sets $\mathcal{O}_\varepsilon \subset \mathcal{O}$ for $|\varepsilon| \ll 1$ with meas$(\mathcal{O} \setminus \mathcal{O}_\varepsilon) \to 0$, and Whitney smooth maps $\Omega_\varepsilon : \mathcal{O}_\varepsilon \to \mathbb{R}^N$, such that for every $\omega \in \mathcal{O}_\varepsilon$, the Hamiltonian network $\{(1.2), (1.3)\}$ associated with ω admits a small amplitude, linearly stable, quasi-periodic breather $q(t) = (\{q_n(t)\})$ of $N-$frequency $\omega_\varepsilon = \Omega_\varepsilon(\omega)$ that is slightly deformed from ω. Moreover, as $|n| \gg 1$, $|q_n| \sim e^{-\beta|n|}$ in the case with exponential coupling coefficients and $|q_n| \sim |n|^{-\alpha}$ in the case with power-law coupling coefficients.*

For the case with constant on-site frequencies (i.e., all β_n's are fixed), we have the following result.

Theorem B. *Consider the Hamiltonian network* (1.2) *with the interaction potential* (1.3), *in which all on-site oscillators are anharmonic. Let $N, \mathcal{J}$ be as the above and assume that $|\beta_n| \geq \frac{1}{|n|^2}$ as $|n| \gg 1$. Then for any given $0 < r_0 \ll 1$, there exist $r_1 > r_0$ depending on $\mathcal{J}$, a family of Cantor sets $\mathcal{O}_\varepsilon \subset \mathcal{O} =: \{\xi \in R_+^N : r_0 \leq |\xi| \leq r_1\}$ for $|\varepsilon| \ll 1$ with meas$(\mathcal{O} \setminus \mathcal{O}_\varepsilon) \to 0$, and Whitney smooth maps $\Omega_\epsilon : \mathcal{O}_\varepsilon \to \mathbb{R}^N$, such that every $\xi \in \mathcal{O}_\varepsilon$ corresponds to a linearly stable, quasi-periodic breather $q(t) = (\{q_n(t)\})$ of $\{(1.2), (1.3)\}$ of $N-$frequency $\Omega_\epsilon(\xi)$. Moreover, $|q| = O(\sqrt{|\xi|})$, and as $|n| \gg 1$, $|q_n| \sim e^{-\beta|n|}$ in the case with exponential coupling coefficients and $|q_n| \sim |n|^{-\alpha}$ in the case with power-law coupling coefficients.*

Remark. (1) Theorem B clearly holds when $\beta_n \equiv \beta_*$ for some constant $\beta_* > 0$, which is particularly the case for the coupling of identical anharmonic oscillators.

(2) Using essentially the same proof, Theorems A and B can be shown to hold for Hamiltonian $\{(1.2), (1.3)\}$ defined on a higher dimensional lattice (i.e., $n \in \mathbb{Z}^d$, $d > 1$). But at this time, the power α in the case of power-law interaction potential need to be bigger than $N + 40(d + 1)$ for the existence of N-quasi-periodic breathers.

(3) In both cases considered in Theorems A and B, the interaction potential (1.3) starts from third order terms in $q_n - q_m$. In fact, the quadratic interaction potential may cause the appearance of a continuous spectrum — an obstacle for the existence of quasi-periodic breathers. Thus in this sense, Theorems A and B seem to be sharp.

(4) The normal variables $\{w_n\}$ in the nearest-neighbor coupling case considered in [50] grow at most linearly in n, which in turn produces breathers that are super-exponentially localized in space. To the contrary, for the long-range couplings we consider, normal variables grow at most exponentially in n in the case with exponential coupling coefficients and grow super-exponentially in n in the case with power-law coupling coefficients, which in turn only produces breathers that are exponentially localized in space in the exponential case and algebraically localized in space in the power-law case.

(5) Theorem A generalizes the main results in [27] because the on-site frequencies β_n in the present case are allowed to accumulate at any non-negative point as $n \to \infty$.

2.2. An abstract KAM theorem

The abstract KAM theorem that we will formulate can be applied to any type of Hamiltonian network of weakly coupled oscillators with a long-range interaction potential with variable or constant on-site frequencies.

We start by introducing some necessary notation. Given an integer $N > 1$ and real numbers $r, s > 0$, we let $D(r, s)$ be the complex neighborhood of $\mathbb{T}^N \times \{0\} \times \{0\} \times \{0\} \subset \mathbb{T}^N \times R^N \times \ell^1 \times \ell^1$; i.e.,

$$D(r, s) = \{(\theta, I, w, \bar{w}) : |\text{Im}\,\theta| < r, |I| < s^2, \|w\| < s, \|\bar{w}\| < s\},$$

where $|\cdot|$ is the sup-norm of complex vectors and $\|\cdot\|$ stands for the the ℓ^1-norm. Let $\mathcal{O} \subset \mathbb{R}^N$ be such that $|\mathcal{O}| > 0$. Hereafter, for simplicity, we use the same symbol $|\cdot|$ to denote the Lebesgue measure of a set.

Let $F(\theta, I, w, \bar{w})$ be a real analytic function on $D(r, s)$ which depends C^1-Whitney smoothly (i.e., C^1 in the sense of Whitney) on a parameter $\xi \in \mathcal{O}$. In the rest of the paper, all dependencies on ξ are assumed of class C^1-Whitney, thus all derivatives with respect to ξ will be interpreted in this sense. The Taylor–Fourier series expansion of F in $\theta, I, w, \bar{w}$ is given by

$$F(\theta, I, w, \bar{w}) = \sum_{\alpha, \beta} F_{\alpha\beta} w^\alpha \bar{w}^\beta,$$

where $\alpha \equiv (\ldots, \alpha_n, \ldots)$, $\beta \equiv (\ldots, \beta_n, \ldots)$, $\alpha_n, \beta_n \in \mathbb{N}$, are multi-indices with only finitely many non-vanishing components, and

$$F_{\alpha\beta} = \sum_{k \in \mathbb{Z}^N, l \in \mathbb{N}^N} F_{kl\alpha\beta}(\xi) I^l e^{i\langle k, \theta \rangle}.$$

We define the weighted norm of F as

$$\|F\|_{D(r,s),\mathcal{O}} \equiv \sup_{\substack{\|w\|<s \\ \|\bar{w}\|<s}} \sum_{\alpha, \beta} \|F_{\alpha\beta}\| \, |w^\alpha| \, |\bar{w}^\beta|, \tag{2.1}$$

where

$$\|F_{\alpha\beta}\| \equiv \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}} s^{2|l|} e^{|k|r}, \qquad |F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi \in \mathcal{O}} \left\{ |F_{kl\alpha\beta}| + \left| \frac{\partial F_{kl\alpha\beta}}{\partial \xi} \right| \right\}.$$

In the case of a vector-valued function, $G : D(r,s) \times \mathcal{O} \to \mathbb{C}^n$, $n < \infty$, we define its weighted norm simply as

$$\|G\|_{D(r,s),\mathcal{O}} \equiv \sum_{i=1}^{n} \|G_i\|_{D(r,s),\mathcal{O}}.$$

The weighted norm of the Hamiltonian vector field

$$X_F = (F_I, -F_\theta, \{iF_{w_n}\}, \{-iF_{\bar{w}_n}\})$$

associated with a Hamiltonian function F, on $D(r,s) \times \mathcal{O}$, is defined by

$$\|X_F\|_{D(r,s),\mathcal{O}} \equiv \|F_I\|_{D(r,s),\mathcal{O}} + \frac{1}{s^2} \|F_\theta\|_{D(r,s),\mathcal{O}} + \frac{1}{s} \left(\sum_n \|F_{w_n}\|_{D(r,s),\mathcal{O}} + \sum_n \|F_{\bar{w}_n}\|_{D(r,s),\mathcal{O}} \right).$$

Associated with the symplectic structure $dI \wedge d\theta + i \sum_{n \in \mathbb{Z}} dw_n \wedge d\bar{w}_n$, we consider the following family of real-analytic, parametrized Hamiltonians

$$H = \mathcal{N} + \mathcal{P},$$
$$\mathcal{N} = \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}} \Omega_n w_n \bar{w}_n, \tag{2.2}$$
$$\mathcal{P} = \mathcal{P}(\theta, I, w, \bar{w}, \xi),$$

where $(I, \theta, w, \bar{w}) \in D(r,s)$, $\xi \in \mathcal{O} \subset R^N$, $\omega : \mathcal{O} \to \mathbb{R}^N$ is C^1-Whitney smooth, Ω_n, $n \in \mathbb{Z}$, are positive constants independent of ξ, and $\mathcal{P}$ is real-analytic with respect to phase variables and C^1-Whitney smooth in parameter ξ.

Clearly, when $\mathcal{P} = 0$, the Hamiltonian reduces to $\mathcal{N}$ which is completely integrable and admits a family of quasi-periodic solutions $(\theta + \omega(\xi)t, 0, 0, 0)$ corresponding to invariant N-tori in phase space. To show the persistence of some of these N-tori, we need the following assumptions on $\omega(\xi)$, Ω_n and the perturbation $\mathcal{P}$:

(A1) *Non-degeneracy of tangential frequencies*: $\omega : \mathcal{O} \to \mathbb{R}^N$ is non-degenerate in the sense that

$$\det \left(\frac{\partial \omega}{\partial \xi} \right) \neq 0, \quad \xi \in \mathcal{O}.$$

(A2) *Asymptotic condition of normal frequencies*: $\{\Omega_n\}$ are positive and satisfy the asymptotic condition that

$$\Omega_n \geq \frac{1}{|n|^2}, \quad |n| \gg 1.$$

(A3) *Decay properties of the perturbation*: $\mathcal{P} = \bar{\mathcal{P}}$, and if we write $\mathcal{P} = \check{P} + \acute{P} + \grave{P}$, where

$$\check{P} = \check{P}(\theta, I, w, \bar{w}, \xi) = \check{P}(\theta, I, 0, 0, \xi) + \sum_{\substack{n \in \mathbb{Z} \\ \alpha_n + \beta_n \geq 1}} \check{P}_n(\theta, I, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n},$$

$$\acute{P} = \acute{P}(w, \bar{w}, \xi) = \sum_{\substack{n,m \in \mathbb{Z}, n \neq m \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m},$$

$$\grave{P} = \grave{P}(w, \bar{w}, \xi) = \sum_{n \in \mathbb{Z}} O(|w_n|^3), \tag{2.3}$$

then for each n, m lying in the range above, either there is a constant $\beta > 0$ such that

$$\|\check{P}_n(\theta, I, \xi)\| \leq e^{-\beta|n|}, \tag{2.4}$$
$$\|\acute{P}_{nm}(\xi)\| \leq e^{-\beta|n-m|}; \tag{2.5}$$

or there is a constant $\alpha > N + 80$ such that

$$\|\check{P}_n(\theta, I, \xi)\| \leq \frac{1}{|n|^\alpha}, \tag{2.6}$$

$$\|\acute{P}_{nm}(\xi)\| \leq \frac{1}{|n-m|^\alpha}. \tag{2.7}$$

Our abstract KAM theorem states as follows.

Theorem C. *Consider the Hamiltonian* (2.2) *and assume* (A1)–(A3). *Then for each* $\gamma > 0$ *sufficiently small, there exists a positive constant* $\varepsilon = \varepsilon(\mathcal{O}, \alpha, \mathcal{N}, \gamma, r, s) \sim \gamma^{\frac{32}{5}}$ *such that if* $\|X_{\mathcal{P}}\|_{D(r,s),\mathcal{O}} < \varepsilon$, *then the following holds. There exist Cantor sets* $\mathcal{O}_\gamma \subset \mathcal{O}$ *with* $|\mathcal{O} \setminus \mathcal{O}_\gamma| = O(\gamma)$ *and maps*

$$\Psi : \mathbb{T}^N \times \mathcal{O}_\gamma \to D(r,s), \qquad \tilde{\omega} : \mathcal{O}_\gamma \to \mathbb{R}^N, \tag{2.8}$$

which are real-analytic in θ *and* C^1-*Whitney smooth in* ξ *with* $\|\Psi - \Psi_0\|_{D(\frac{\varsigma}{2},0),\mathcal{O}_\gamma} \to 0$ *and* $|\tilde{\omega} - \omega|_{\mathcal{O}_\gamma} \to 0$ *as* $\gamma \to 0$, *where* Ψ_0 *is the trivial embedding:* $\mathbb{T}^N \times \mathcal{O} \to \mathbb{T}^N \times \{0,0,0\}$, *such that each* $\xi \in \mathcal{O}_\gamma$ *and* $\theta \in \mathbb{T}^N$ *correspond to a linearly stable, N-frequency quasi-periodic solution* $\Psi(\theta + \tilde{\omega}(\xi)t, \xi) = (\theta + \tilde{\omega}(\xi)t, I(t), \{w_n(t)\})$ *of the equations of motion associated with the Hamiltonian* (2.2). *Moreover,* $|w_n| \sim e^{-\beta|n|}$ *under the conditions* (2.4) *and* (2.5) *and* $|w_n| \sim |n|^{-\alpha}$ *under the conditions* (2.6) *and* (2.7).

Remark. (1) Theorem C is rather different from those for Hamiltonian PDEs. Under the perturbation decay condition (A3) which is natural for Hamiltonian networks, Theorem C allows the normal frequencies to have infinite multiplicity and polynomial decay. On the contrary, Hamiltonian PDEs like NLS, wave, beam and KdV equations always have normal frequencies with finite multiplicity and linear or super-linear growth, and do not satisfy perturbation decay conditions like (A3) in general. The lacking of such a perturbation decay condition in Hamiltonian PDEs makes the respective infinite-dimensional KAM or CWB theory much harder especially in higher space dimensions because the usual failure of the second Melnikov conditions during KAM or Newton iterations (see [8–12,16,17,20,23–26,30–32,37,38,48] and references therein).

(2) The assumption (A3) clearly holds in the case with short-range interaction potential considered in [50]. Thus, Theorem C generalizes the main result of [50] to the case with long-range interaction potential. Due to the iteration scheme we adopt in this paper, the proof of Theorem C is even simpler than that of the analogous theorem in [50].

2.3. Proof of Theorems A and B

We will use the standard procedure to derive action-angle-normal variables for the Hamiltonian (1.2) and show that, in these variables, H is of the form described in Theorem C, from which Theorems A and B will consequently follow.

The desired normal form in the case with variable on-site frequencies (i.e, the case considered in Theorem A) was derived in [27].

To derive the normal form in the case with constant on-site frequencies (i.e., the case considered in Theorem B), we fix a $0 < r_0 \ll 1$ and let $\mathcal{J} = \{n_1, \dots, n_N\}$ be a set of prescribed N-sites. Then a second-order Taylor expansion of the terms of the on-site potential associated with the set $\mathbb{Z}_1 = \mathbb{Z} \setminus \mathcal{J}$ yields

$$H = \sum_{n \in \mathcal{J}} \left(\frac{p_n^2}{2} + V_n(q_n) \right) + \sum_{n \in \mathbb{Z}_1} \left(\frac{p_n^2}{2} + \frac{\beta_n^2 q_n^2}{2} + O(|q_n|^3) \right) + \frac{\varepsilon}{p} \sum_{m \neq n} C(m,n)(q_m - q_n)^p. \tag{2.9}$$

Let $\rho_j(h)$, $j = 1, \dots, N$, be as in (1.6). Then it is clear that there exists a $r_1 > r_0$ such that the inverse $H_{0,j} : [r_0, r_1] \to \mathbb{R}_+$: $\rho_j \mapsto h_j$ of $\rho_j(h)$ is well-defined for each $j = 1, \dots, N$. Also let $(q_n, p_n) = (\frac{1}{\sqrt{\beta_n}} v_n, \sqrt{\beta_n} \bar{v}_n)$ for $n \in \mathbb{Z}_1$ be the normal-coordinate change. A standard action-angle-normal coordinate-reduction procedure leaves (2.9) in the form

$$H = \sum_{j=1}^N H_{0,j}(\rho_j) + \sum_{n \in \mathbb{Z}_1} \frac{\beta_n}{2}(v_n^2 + \bar{v}_n^2 + O(|v_n|^3)) + \frac{\varepsilon}{p} \sum_{m \neq n} C(m,n)(q_m - q_n)^p,$$

where q_m, q_n's are understood as functions of the new variables. Now let $v_n = (w_n + \bar{w}_n)/\sqrt{2}$ and $\bar{v}_n = (w_n - \bar{w}_n)/i\sqrt{2}$ for every $n \in \mathbb{Z}_1$. Then

$$H = \sum_{j=1}^N H_{0,j}(\rho_j) + \sum_{n \in \mathbb{Z}_1} (\beta_n w_n \bar{w}_n + O(|w_n + \bar{w}_n|^3)) + \frac{\varepsilon}{p} \sum_{m \neq n} C(m,n)(q_m - q_n)^p. \tag{2.10}$$

Again, in the above, q_m, q_n's are understood as functions of the new variables. In order to write (2.10) in normal form as described in Theorem C, we first introduce parameters ξ_j by setting $\rho_j = \xi_j + I_j$, $j = 1, \dots, N$, then expand $H_{0,j}(\xi_j + I_j)$ into Taylor series about $I_j = 0$ and disregard constant terms that do not affect the dynamics. Then (2.10) becomes

$$H = \langle \omega(\xi), I \rangle + \sum_{j=1}^N O(|I_j|^2) + \sum_{n \in \mathbb{Z}_1} (\beta_n w_n \bar{w}_n + O(|w_n + \bar{w}_n|^3)) + \frac{\varepsilon}{p} \sum_{m \neq n} C(m,n)(q_m - q_n)^p,$$

where $\omega(\xi) = (H'_{0,1}(\xi_1), \ldots, H'_{0,N}(\xi_N))$. Since each V_{n_j} is anharmonic, we can assume without loss of generality that ω is a diffeomorphism from $\mathcal{O} \equiv \{\xi \in \mathbb{R}^N_+ : r_0 \leq |\xi| \leq r_1\}$ to $\omega(\mathcal{O})$, i.e., the condition (A1) is satisfied. Now introduce the rescaling $I_j \to \varepsilon^{1/2} I_j$, $(w_n, \bar{w}_n) \to (\varepsilon^{1/4} w_n, \varepsilon^{1/4} \bar{w}_n)$, $H \to \varepsilon^{-1/2} H$, and $\varepsilon^{1/4} \to \varepsilon$. The rescaled Hamiltonian reads

$$
\begin{aligned}
H &= \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1} \beta_n w_n \bar{w}_n + \varepsilon \left(\frac{\varepsilon^{p-3}}{p} \sum_{m \neq n} C(m,n)(q_m - q_n)^p + \sum_{n \in \mathbb{Z}_1} O(|w_n + \bar{w}_n|^3) + \varepsilon \sum_{j=1}^N O(|I_j|^2) \right) \\
&=: \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1} \beta_n w_n \bar{w}_n + \varepsilon \mathcal{P}(I, \theta, w, \bar{w}, \xi).
\end{aligned}
$$

With $\Omega_n = \beta_n$, $\forall n \in \mathbb{Z}_1$, the condition (A2) is clearly satisfied. It is also easy to see that $\bar{\mathcal{P}} = \mathcal{P}$, and $\mathcal{P}$ can be decomposed into a sum of three terms $\check{P}$, $\acute{P}$, $\grave{P}$ of the form

$$
\check{P} = \check{P}(I, \theta, w, \bar{w}, \xi) = \check{P}(\theta, I, 0, 0, \xi) + \sum_{\substack{n \in \mathbb{Z}_1 \\ \alpha_n + \beta_n \geq 1}} \check{P}_n(\theta, I, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n},
$$

$$
\acute{P} = \acute{P}(w, \bar{w}, \xi) = \sum_{\substack{n, m \in \mathbb{Z}_1, n \neq m \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m},
$$

$$
\grave{P} = \grave{P}(w, \bar{w}, \xi) = \sum_{n \in \mathbb{Z}_1} O(|w_n|^3),
$$

in which $\check{P}_n(\theta, I, \xi)$, $\acute{P}_{nm}(\xi)$, respectively, satisfy (2.4) and (2.5) (or (2.6) and (2.7)), respectively, in the case that $C(m,n)$ are exponential coupling coefficients (or power-law coupling coefficients). Hence the condition (A3) is also satisfied.

Thus, Theorems A and B follow from Theorem C.

3. KAM step

The remaining sections are devoted to the proof of Theorem C. For simplicity, we only treat the case when $\mathcal{P}$ satisfies (2.4) and (2.5) with $\beta = 1$.

In this section we present the KAM iteration scheme applied to (2.2). This is a succession of infinitely many steps (KAM steps or iterations) whose purpose is to eliminate lower-order θ-dependent terms in $\mathcal{P}$. At each KAM step the perturbation is made smaller at the cost of excluding a small-measure set of parameters. It will be shown that the KAM iterations converge and that, in the end, the total measure of the set of parameters that has been excluded is small.

To begin the KAM iteration, we set $r_0 = r$ and $\gamma_0 = \gamma$.

3.1. Normal form

In order to perform the KAM iteration scheme, we will first write the Hamiltonian (2.2) into a normal form that is more convenient for this purpose. For simplicity, we only outline the derivation of the normal form. Detailed construction and estimation of terms in the normal form is similar to those for the general KAM step which we will show later.

Choose $\varepsilon_* \sim \varepsilon^{\frac{5}{4}}$ and let $K_* = |\ln \varepsilon_*|$, $K_0 = 5K_*$, $\varepsilon_0 = \varepsilon_*^{\frac{5}{4}}$. Also let s_0 be such that $0 < s_0 < \min\{\varepsilon_0, s\}$.

Observe that, using (2.3)–(2.5) in assumption (A3) and Cauchy inequality (Lemma A.2 in the Appendix), one can make s_0 smaller if necessary such that

$$
\|X_{\acute{P} + \grave{P}}\|_{D(r_0, s_0) \times \mathcal{O}} \leq \varepsilon_*.
$$

We now consider the term $\check{P}$. According to (2.4) and the definition of norm (2.1) we have

$$
\begin{aligned}
\check{P} &= \check{P}(\theta, I, 0, 0, \xi) + \sum_{\substack{n \in \mathbb{Z} \\ \alpha_n + \beta_n \geq 1}} \check{P}_n(\theta, I, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} \\
&= \sum_{k,l} \check{P}_{kl} I^l e^{i\langle k, \theta \rangle} + \sum_{\substack{n, k, l \\ \alpha_n + \beta_n \geq 1}} \check{P}_n^{kl\alpha_n\beta_n} I^l e^{i\langle k, \theta \rangle} w_n^{\alpha_n} \bar{w}_n^{\beta_n},
\end{aligned}
$$

where

$$
\|\check{P}_{kl}\| \leq e^{-|k|r_0}, \qquad \|\check{P}_n^{kl\alpha_n\beta_n}\| \leq e^{-|k|r_0} e^{-|n|}. \tag{3.1}
$$

Let

$$R = \sum_{k,|l|\leq 1} \check{P}_{kl} I^l e^{i\langle k,\theta\rangle} + \sum_{\substack{|n|\leq K_*, k \\ 1\leq \alpha_n+\beta_n\leq 2}} \check{P}_n^{k\alpha_n\beta_n} e^{i\langle k,\theta\rangle} w_n^{\alpha_n} \bar{w}_n^{\beta_n}.$$

It is clear that

$$\check{P} - R = \sum_{\substack{|n|>K_* \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^{k\alpha_n\beta_n} e^{i\langle k,\theta\rangle} w_n^{\alpha_n} \bar{w}_n^{\beta_n} + O(|I|^2) + O(|w|^3).$$

It thus follows from (3.1) and Cauchy inequality (Lemma A.2 in the Appendix) that one can make s_0 small enough so that

$$\|X_{\check{P}-R}\| \leq \varepsilon_*.$$

To handle the term R, we will first construct a symplectic transformation $\Phi_* = \Phi_{F_*}^1$ defined as the time-1 map of the Hamiltonian flow associated with a Hamiltonian F_* of the form

$$\begin{aligned}
F_* = \bar{F}_* &= \sum_{k\neq 0,|l|\leq 1} F_{kl} I^l e^{i\langle k,\theta\rangle} + \sum_{k,|n|\leq K_*} (F_n^{k10} w_n + F_n^{k01}\bar{w}_n) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{k,|n|\leq K_*} (F_{nn}^{k20} w_n w_n + F_{nn}^{k02}\bar{w}_n\bar{w}_n) e^{i\langle k,\theta\rangle} + \sum_{k\neq 0,|n|\leq K_*} F_{nn}^{k11} w_n\bar{w}_n e^{i\langle k,\theta\rangle},
\end{aligned}$$

such that all resonant terms $\check{P}_{kl} I^l e^{i\langle k,\theta\rangle}$, $k\neq 0$, $|l|\leq 1$; $\check{P}_n^{k\alpha_n\beta_n} e^{i\langle k,\theta\rangle} w_n^{\alpha_n}\bar{w}_n^{\beta_n}$, $k\neq 0$, $|n|\leq K_*$, $\alpha_n+\beta_n\leq 2$ will be eliminated, and terms $\check{P}_{0l} I^l$, $|l|\leq 1$; $\check{P}_{nn}^{011} w_n\bar{w}_n$, $|n|\leq K_*$ will be added to the normal form part of the new Hamiltonian. More precisely, let F_* satisfy the homological equation

$$\{\mathcal{N}, F_*\} + R = \sum_{|l|\leq 1} \check{P}_{0l} I^l + \sum_{|n|\leq K_*} \check{P}_{nn}^{011} w_n\bar{w}_n.$$

One can show that this homological equation is solvable on the parameter set

$$\mathcal{O}_* = \left\{ \xi \in \mathcal{O} : \begin{array}{l} |\langle k,\omega\rangle| \geq \dfrac{\gamma}{|k|^\tau}, k\neq 0 \\[2mm] |\langle k,\omega\rangle + \Omega_n| \geq \dfrac{\gamma}{|k|^\tau |n|^2}, \\[2mm] |\langle k,\omega\rangle + 2\Omega_n| \geq \dfrac{\gamma}{|k|^\tau |n|^2} \end{array} \right\}.$$

In this way we obtain the transformation Φ_* which transforms the Hamiltonian (2.2) to

$$\begin{aligned}
H_* &= H \circ \Phi_* = \mathcal{N}_* + \check{P}_* + \acute{P} + \grave{P} = \bar{\mathcal{N}}_* + \bar{\check{P}}_* + \bar{\acute{P}} + \bar{\grave{P}}, \\
\mathcal{N}_* &= e_* + \langle \omega_*(\xi), I\rangle + \sum_{|n|\leq K_*} \Omega_n^* w_n\bar{w}_n + \sum_{|n|>K_*} \Omega_n w_n\bar{w}_n,
\end{aligned}$$

where $\omega_* = \omega + \check{P}_{0l}(|l|=1)$, $\Omega_n^* = \Omega_n + \check{P}_{nn}^{011}$, and

$$\check{P}_* = \check{P}^*(\theta, I, w_{n(|n|\leq K_*)}, \bar{w}_{n(|n|\leq K_*)}, \xi) + \sum_{\substack{|n|>K_* \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^*(\theta, I, w_{m(|m|\leq K_*)}, \bar{w}_{m(|m|\leq K_*)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}$$

satisfies

$$\|\check{P}_n^*(\theta, I, w_{m(|m|\leq K_*)}, \bar{w}_{m(|m|\leq K_*)}, \xi)\| \leq e^{-(|n|-K_*)}.$$

The first and second terms in the above expression for $\check{P}_*$ come from $\mathcal{P} \circ \Phi_*$ and $\grave{P} \circ \Phi_* + \acute{P} \circ \Phi_*$, respectively. Moreover, the decay property of $\check{P}_n^*$ follows from the fact that Φ_* depends only on I, θ and w_m, $\bar{w}_m$ for $|m|\leq K_*$.

Now let us write the second term in $\check{P}_*$ as

$$\begin{aligned}
&\sum_{\substack{|n|>K_* \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^*(\theta, I, w_{m(|m|\leq K_*)}, \bar{w}_{m(|m|\leq K_*)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n} \\
&= \sum_{\substack{|n|>K_0 \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^*(\theta, I, w_{m(|m|\leq K_*)}, \bar{w}_{m(|m|\leq K_*)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n} + \sum_{\substack{K_*<|n|\leq K_0 \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^*(\theta, I, w_{m(|m|\leq K_*)}, \bar{w}_{m(|m|\leq K_*)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}.
\end{aligned}$$

It is easy to see that, on $D(r_0, s_0) \times \mathcal{O}_*$, the norm of the vector field associated with the first term above is bounded by ε_*^2. However, due to the condition (A2), terms of the form $\check{P}_{mn}^* w_m\bar{w}_n + \check{P}_{nm}^* w_n\bar{w}_m$ in the second term will not be canceled by solving

a homological equation, hence they need to be included in the normal part of the Hamiltonian. More precisely, let

$$R_* = \sum_{k,|l|\leq 1} \check{P}_{kl}^* I^l e^{i\langle k,\theta\rangle} + \sum_{\substack{|m|\leq K_*,|n|\leq K_0,k \\ 1\leq \alpha_m+\beta_n\leq 2}} \check{P}_{mn}^{*k\alpha_m\beta_n} e^{i\langle k,\theta\rangle} (w_m^{\alpha_m}\bar{w}_n^{\beta_n} + \bar{w}_m^{\alpha_m}w_n^{\beta_n})$$

and

$$\begin{aligned}
F_{**} = \bar{F}_{**} &= \sum_{k\neq 0,|l|\leq 1} f_{kl}e^{i\langle k,\theta\rangle} I^l + \sum_{k,|n|\leq K_0} (f_n^{k10}w_n + f_n^{k01}\bar{w}_n)e^{i\langle k,\theta\rangle} \\
&+ \sum_{k,|m|\leq K_*,|n|\leq K_0} (f_{nm}^{k20}w_nw_m + f_{nm}^{k02}\bar{w}_n\bar{w}_m)e^{i\langle k,\theta\rangle} + \sum_{k\neq 0,|m|\leq K_*,|n|\leq K_0} (f_{nm}^{k11}w_n\bar{w}_m + f_{mn}^{k11}w_m\bar{w}_n)e^{i\langle k,\theta\rangle} \\
&+ \sum_{k,K_*<|n|\leq K_0} (f_{nn}^{k20}w_nw_n + f_{nn}^{k02}\bar{w}_n\bar{w}_n)e^{i\langle k,\theta\rangle} + \sum_{k\neq 0,K_*<|n|\leq K_0} f_{nn}^{k11}w_n\bar{w}_ne^{i\langle k,\theta\rangle}
\end{aligned}$$

satisfy the homological equation

$$\{\mathcal{N}_*, F_{**}\} + R_* = \sum_{|l|\leq 1} \check{P}_{0l}^* I^l + \sum_{|m|\leq K_*,|n|\leq K_0} (\check{P}_{nm}^{*011}w_n\bar{w}_m + \check{P}_{mn}^{*011}w_m\bar{w}_n) + \sum_{K_*<|n|\leq K_0} \check{P}_{nn}^{*011}w_n\bar{w}_n.$$

It is easy to see that the above equation is solvable on the domain

$$\mathcal{O}_0 = \left\{ \xi \in \mathcal{O}_* : \begin{array}{l}
|\langle k,\omega_*\rangle| \geq \dfrac{\gamma}{|k|^\tau}, k\neq 0 \\[2mm]
|\langle k,\omega_*\rangle + \Omega_n^*| \geq \dfrac{\gamma}{|k|^\tau|n|^2}, |n|\leq K_* \\[2mm]
|\langle k,\omega_*\rangle + \Omega_n| \geq \dfrac{\gamma}{|k|^\tau|n|^2}, |n| > K_* \\[2mm]
|\langle k,\omega_*\rangle + \Omega_m^* + \Omega_n^*| \geq \dfrac{\gamma}{|k|^\tau|m|^2|n|^2}, |m|,|n|\leq K_* \\[2mm]
|\langle k,\omega_*\rangle + \Omega_m^* + \Omega_n| \geq \dfrac{\gamma}{|k|^\tau|m|^2|n|^2}, |m|\leq K_*, |n| > K_* \\[2mm]
|\langle k,\omega_*\rangle + \Omega_m^* - \Omega_n^*| \geq \dfrac{\gamma}{|k|^\tau|m|^2|n|^2}, k\neq 0, |m|,|n|\leq K_* \\[2mm]
|\langle k,\omega_*\rangle + \Omega_m^* - \Omega_n| \geq \dfrac{\gamma}{|k|^\tau|m|^2|n|^2}, k\neq 0, |m|\leq K_*, |n| > K_* \\[2mm]
|\langle k,\omega_*\rangle + 2\Omega_n| \geq \dfrac{\gamma}{|k|^\tau|n|^2}, |n| > K_*
\end{array} \right\}.$$

Now under the symplectic transformation $\Phi_{**} = \Phi_{F_{**}}^1$, we have

$$\begin{aligned}
H_0 &= H_* \circ \Phi_{**} = \mathcal{N}_0 + \mathcal{P}_0 = \bar{\mathcal{N}}_0 + \bar{\mathcal{P}}_0, \\
\mathcal{N}_0 &= e_0 + \langle \omega_0(\xi), I\rangle + \langle A^0 z^0, \bar{z}^0\rangle + \sum_{|n|>K_0} \Omega_n w_n\bar{w}_n,
\end{aligned}$$

where

$$\begin{aligned}
e_0 &= e_* + \check{P}_{00}^*, \\
\omega_0 &= \omega_* + \check{P}_{0l}^*(|l|=1), \\
\langle A^0 z^0, \bar{z}^0\rangle &= \sum_{|n|\leq K_0} \Omega_n^{**} w_n\bar{w}_n + \sum_{|m|\leq K_*,|n|\leq K_0,m\neq n} (\check{P}_{mn}^{*011}w_m\bar{w}_n + \check{P}_{nm}^{*011}w_n\bar{w}_m), \\
\Omega_n^{**} &= \Omega_n^* + \check{P}_{nn}^{*011}, \\
\mathcal{P}_0 &= \check{P}_0 + \acute{P}_0 + \grave{P}_0, \\
\check{P}_0 &= \check{P}(\theta, I, z^0, \bar{z}^0, \xi) = \check{P}^0(\theta, I, w_{n(|n|\leq K_0)}, \bar{w}_{n(|n|\leq K_0)}, \xi) + \sum_{\substack{|n|>K_0 \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^0(\theta, I, w_{m(|m|\leq K_0)}, \bar{w}_{m(|m|\leq K_0)}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n} \\
&=: \check{P}^0(\theta, I, z^0, \bar{z}^0, \xi) + \sum_{\substack{|n|>K_0 \\ \alpha_n+\beta_n\geq 1}} \check{P}_n^0(\theta, I, z^0, \bar{z}^0, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}, \\
\|\check{P}_n^0(\theta, I, z^0, \bar{z}^0, \xi)\| &\leq e^{-(|n|-K_0)}, \quad |n| > K_0, \\
\acute{P}_0 &= \acute{P}(w, \bar{w}, \xi) = \sum_{\substack{n\neq m \\ \alpha_n+\beta_n,\alpha_m+\beta_m\geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}_{nm}^0(\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}w_m^{\alpha_m}\bar{w}_m^{\beta_m},
\end{aligned}$$

$$\|\acute{P}^0_{nm}(\xi)\| \le e^{-|n-m|},$$
$$\grave{P}_0 = \grave{P}(w, \bar{w}, \xi) = \sum_{n \in \mathbb{Z}} O(|w_n|^3).$$

In the above,

$$z^0 = (\ldots, w_n, \ldots)_{|n| \le K_0}, \qquad \bar{z}^0 = (\ldots, \bar{w}_n, \ldots)_{|n| \le K_0},$$

and A^0 is a Hermitian matrix with $\dim(A^0) \le K_0$ and $\|X_{\mathcal{P}_0}\|_{D(r_0, s_0), \mathcal{O}_0} \le \varepsilon_*^{\frac{5}{4}} = \varepsilon_0$.

Suppose that after the νth KAM step one arrives at the following Hamiltonian,

$$H \equiv H_\nu = \mathcal{N} + \mathcal{P} = \bar{\mathcal{N}} + \bar{\mathcal{P}} = \mathcal{N} + \check{P} + \acute{P}_0 + \grave{P}_0,$$

where

$$\mathcal{N} = \mathcal{N}_\nu = \langle \omega(\xi), I \rangle + \langle Az, \bar{z} \rangle + \sum_{|n| > K} \Omega_n w_n \bar{w}_n,$$

$$\check{P} = \check{P}_\nu = \check{P}(\theta, I, z, \bar{z}, \xi) + \sum_{\substack{|n| > K \\ \alpha_n + \beta_n \ge 1}} \check{P}_n(\theta, I, z, \bar{z}, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n}$$

$$= \check{P}^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi) + \sum_{\substack{|n| > K_\nu \\ \alpha_n + \beta_n \ge 1}} \check{P}_n^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n}$$

are defined on a domain $D(r, s) \times \mathcal{O} = D(r_\nu, s_\nu) \times \mathcal{O}_\nu$, $K = K_\nu$ is a positive constant,

$$z = z^\nu = (\ldots, w_n, \ldots)_{|n| \le K}, \qquad \bar{z} = \bar{z}^\nu = (\ldots, \bar{w}_n, \ldots)_{|n| \le K},$$

$\mathcal{P} = \mathcal{P}_\nu$ is such that $\|X_{\mathcal{P}}\| < \varepsilon$ for some $\varepsilon = \varepsilon_\nu$ and

$$\|\check{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r,s), \mathcal{O}} \le e^{-(|n|-K)}, \quad |n| > K.$$

In what follows, we will show how to construct a symplectic transformation, $\Phi = \Phi_\nu$, which, in smaller frequency and phase domains, carries the Hamiltonian $H = H_\nu$ into the next KAM cycle. In the remaining part of this section, all constants labeled $c_1 \ldots c_{12}$ are positive and independent of the iteration process. We will denote the tensor (or direct) product between two matrices, $A = (a_{ij})$ and $B = (b_{nl})$, of sizes $m \times n$ and $k \times l$, respectively, as $A \otimes B$. Recall that this product results in a new $mk \times nl$ matrix given by

$$A \otimes B = (a_{ij}B) = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \cdots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}.$$

We will also adopt the following definition for the operator norm of a matrix A: if $A = (a_{ij})$, then $\|A\| = \max\{\sup_i(\sum_j |a_{ij}|), \sup_j(\sum_i |a_{ij}|)\}$.

Let $K_+ = 5K$. In the KAM step detailed below, terms $w_n, \bar{w}_n$ with $K < |n| \le K_+$ will be added to the new normal components $z^+, \bar{z}^+$. To facilitate the calculations when solving a homological equation later on, we will also adopt the following notation when writing $\mathcal{N}$,

$$\mathcal{N} = \bar{\mathcal{N}} = e + \langle \omega(\xi), I \rangle + \langle Az, \bar{z} \rangle + \sum_{K < |n| \le K_+} \Omega_n w_n \bar{w}_n + \sum_{|n| > K_+} \Omega_n w_n \bar{w}_n$$

$$=: e + \langle \omega(\xi), I \rangle + \langle \tilde{A}z^+, \bar{z}^+ \rangle + \sum_{|n| > K_+} \Omega_n w_n \bar{w}_n,$$

where $\tilde{A}$ is a Hermitian matrix with $\dim(\tilde{A}) \le K_+$ given by

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & \Omega_n \end{pmatrix}_{K < |n| \le K_+}$$

and $z^+ = (\ldots, w_n, \ldots)_{|n| \le K_+}, \bar{z}^+ = (\ldots, \bar{w}_n, \ldots)_{|n| \le K_+}$.

3.2. Truncation

Let us start by expanding $\check{P}$, as obtained after the νth KAM iteration, into its Taylor–Fourier series,

$$\check{P} = \sum_{k,l,\alpha,\beta} \check{P}_{kl\alpha\beta} e^{i\langle k,\theta\rangle} I^l z^\alpha \bar{z}^\beta + \sum_{\substack{k,l,n,\alpha,\beta \\ |n|>K, \alpha_n+\beta_n \geq 1}} \check{P}_{kln\alpha\beta} e^{i\langle k,\theta\rangle} I^l z^\alpha \bar{z}^\beta w_n^{\alpha_n} \bar{w}_n^{\beta_n},$$

where $k \in \mathbb{Z}^N$, $l \in \mathbb{N}^N$ and the multi-index α (β) runs over the set $\alpha \equiv (\ldots, \alpha_m, \ldots)_{|m|\leq K}$, $\alpha_m \in \mathbb{N}$ (resp. $\beta \equiv (\ldots, \beta_m, \ldots)_{|m|\leq K}$, $\beta_m \in \mathbb{N}$). Now let R be the following truncation of $\check{P}$:

$$\begin{aligned}
R(\theta, I, z, \bar{z}, w, \bar{w}) &= \sum_{k,|l|\leq 1} \check{P}_{kl} e^{i\langle k,\theta\rangle} I^l + \sum_k (\langle \check{P}^{k10}, z\rangle + \langle \check{P}^{k01}, \bar{z}\rangle) e^{i\langle k,\theta\rangle} + \sum_{k,K<|n|\leq K_+} (\check{P}_n^{k10} w_n + \check{P}_n^{k01} \bar{w}_n) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_k (\langle \check{P}^{k20} z, z\rangle + \langle \check{P}^{k11} z, \bar{z}\rangle + \langle \check{P}^{k02} \bar{z}, \bar{z}\rangle) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{k,K<|n|\leq K_+} (\langle \check{P}_n^{k20} z, w_n\rangle + \langle \check{P}_n^{k11} z, \bar{w}_n\rangle + \langle \overline{\check{P}^{(-k)11}}_n \bar{z}, w_n\rangle + \langle \check{P}_n^{k02} \bar{z}, \bar{w}_n\rangle) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{k,K<|n|\leq K_+} (\check{P}_{nn}^{k20} w_n w_n + \check{P}_{nn}^{k11} w_n \bar{w}_n + \check{P}_{nn}^{k02} \bar{w}_n \bar{w}_n) e^{i\langle k,\theta\rangle},
\end{aligned}$$

i.e., R consists of all terms on the right hand side of $\check{P}$ of the above form.

Remark 3.1. We observe that, due to their decay property, terms in the Taylor–Fourier expansion of $\check{P}$ corresponding to $|n| > K_+$ are small enough to be postponed to the next KAM step. Similarly, due to the decay property of $\check{P}$ and the fact that $\acute{P}_0$ starts from third order terms, there are no coupling terms of the form $\sum_{\substack{n\neq m \\ K<|n|,|m|\leq K_+}} w_n \bar{w}_m$ in R. If $\acute{P}_0$ started from second-order terms, then the couplings between different oscillators would be so strong that there may be the appearance of continuous spectrum.

Associated with terms in the normal form $\mathcal{N}$, we rewrite R as

$$\begin{aligned}
R(\theta, I, z^+, \bar{z}^+) &= R_0 + R_1 + R_2 \\
&= \sum_{k,|l|\leq 1} P_{kl} e^{i\langle k,\theta\rangle} I^l + \sum_k (\langle R^{k10}, z^+\rangle + \langle R^{k01}, \bar{z}^+\rangle) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_k (\langle R^{k20} z^+, z^+\rangle + \langle R^{k11} z^+, \bar{z}^+\rangle + \langle R^{k02} \bar{z}^+, \bar{z}^+\rangle) e^{i\langle k,\theta\rangle},
\end{aligned}$$

where

$$R^{k10} = \begin{pmatrix} \check{P}^{k10} \\ \check{P}_n^{k10} \end{pmatrix}_{K<|n|\leq K_+},$$

$$R^{k01} = \begin{pmatrix} \check{P}^{k01} \\ \check{P}_n^{k01} \end{pmatrix}_{K<|n|\leq K_+},$$

$$R^{k20} = \begin{pmatrix} \check{P}^{k20} & \frac{1}{2}(\check{P}_n^{k20})^\top \\ \frac{1}{2}\check{P}_n^{k20} & \check{P}_{nn}^{k20} \end{pmatrix}_{K<|n|\leq K_+},$$

$$R^{k11} = \begin{pmatrix} \check{P}^{k11} & (\overline{\check{P}_n^{(-k)11}})^\top \\ \check{P}_n^{k11} & \check{P}_{nn}^{k11} \end{pmatrix}_{K<|n|\leq K_+},$$

$$R^{k02} = \begin{pmatrix} \check{P}^{k02} & \frac{1}{2}(\check{P}_n^{k02})^\top \\ \frac{1}{2}\check{P}_n^{k02} & \check{P}_{nn}^{k02} \end{pmatrix}_{K<|n|\leq K_+}.$$

Since $\bar{R} = R$, it is clear that $\overline{P_{(-k)l}} = P_{kl}$, $\overline{R^{(-k)10}} = R^{k01}$, $\overline{R^{(-k)01}} = R^{k10}$, $\overline{R^{(-k)20}} = R^{k02}$, $\overline{R^{(-k)11}}^\top = R^{k11}$, and $\overline{R^{(-k)02}} = R^{k20}$.

Now, if we write $H = \mathcal{N} + R + (\mathcal{P} - R)$, then, from our definition of norms, it follows that

$$\|X_R\|_{D(r,s),\mathcal{O}} \leq \|X_{\mathcal{P}}\|_{D(r,s),\mathcal{O}} \leq \varepsilon.$$

Let $r_+ = \frac{r}{2} + \frac{r_0}{4}$ and $\eta = \varepsilon^{\frac{1}{4}}$. Since

$$\mathcal{P} - R = \sum_{|n| > K_+, \alpha_n + \beta_n \geq 1} \check{P}_n(\theta, I, z, \bar{z}, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} + O(|I|^2 + |I||w| + |w|^3)$$

and

$$\|\check{P}_n(\theta, I, z, \bar{z}, \xi)\| \leq e^{-(|n| - K)},$$

one has by Cauchy inequality (Lemma A.2 in the Appendix) that

$$\|X_{\mathcal{P} - R}\|_{D(r_+ + \frac{r - r_+}{2}, \eta s), \mathcal{O}} \leq \sum_{|n| > K_+} e^{-(|n| - K)} + O(s) \leq c_1 \varepsilon^{\frac{5}{4}}, \tag{3.2}$$

provided that

(C0) $s \leq \varepsilon$.

3.3. The homological equation

Below we show that one can find a Hamiltonian function F, defined on a domain $D_+ = D(r_+, s_+)$ such that, the time-1 map $\Phi = \Phi_F^1$ associated with the Hamiltonian vector field X_F, is a (symplectic) map from D_+ to D which transforms H into H_+, the Hamiltonian of the next KAM cycle. Let

$$F = F(\theta, I, z^+, \bar{z}^+) = F_0 + F_1 + F_2$$

satisfy the homological equation

$$\{\mathcal{N}, F\} + R = \check{P}_{00} + \langle \omega', I \rangle + \langle R^{011} z^+, \bar{z}^+ \rangle, \tag{3.3}$$

where

$$F_0 = \sum_{\substack{k \neq 0 \\ |l| \leq 1}} F_{kl} e^{i\langle k, \theta \rangle} I^l,$$

$$F_1 = \sum_{\substack{k \\ |n| \leq K_+}} (f_n^{k10} w_n + f_n^{k01} \bar{w}_n) e^{i\langle k, \theta \rangle} = \sum_k (\langle F^{k10}, z^+ \rangle + \langle F^{k01}, \bar{z}^+ \rangle) e^{i\langle k, \theta \rangle},$$

$$\begin{aligned}
F_2 &= \sum_{k, |m| \leq K, |n| \leq K_+} (f_{nm}^{k20} w_n w_m + f_{nm}^{k02} \bar{w}_n \bar{w}_m) e^{i\langle k, \theta \rangle} + \sum_{k \neq 0, |m| \leq K, |n| \leq K_+} (f_{nm}^{k11} w_n \bar{w}_m + f_{mn}^{k11} w_m \bar{w}_n) e^{i\langle k, \theta \rangle} \\
&\quad + \sum_{k, K < |n| \leq K_+} (f_{nn}^{k20} w_n w_n + f_{nn}^{k02} \bar{w}_n \bar{w}_n) e^{i\langle k, \theta \rangle} + \sum_{k \neq 0, K < |n| \leq K_+} f_{nn}^{k11} w_n \bar{w}_n e^{i\langle k, \theta \rangle} \\
&= \sum_k (\langle F^{k20} z^+, z^+ \rangle + \langle F^{k02} \bar{z}^+, \bar{z}^+ \rangle) e^{i\langle k, \theta \rangle} + \sum_{k \neq 0} \langle F^{k11} z^+, \bar{z}^+ \rangle e^{i\langle k, \theta \rangle},
\end{aligned}$$

$$\omega' = \int \frac{\partial \check{P}}{\partial I} d\theta |_{z^+ = \bar{z}^+ = w = \bar{w} = 0, I = 0}.$$

Lemma 3.1. *Eq.* (3.3) *is equivalent to the following system*

$$\begin{aligned}
&\langle k, \omega \rangle F_{kl} = i \check{P}_{kl}, \quad k \neq 0, |l| \leq 1, \\
&(\langle k, \omega \rangle I - \tilde{A}) F^{k10} = i R^{k10}, \\
&(\langle k, \omega \rangle I + \tilde{A}) F^{k01} = i R^{k01}, \\
&(\langle k, \omega \rangle I - \tilde{A}) F^{k20} - F^{k20} \tilde{A} = i R^{k20}, \\
&(\langle k, \omega \rangle I + \tilde{A}) F^{k11} - F^{k11} \tilde{A} = i R^{k11}, \quad k \neq 0, \\
&(\langle k, \omega \rangle I + \tilde{A}) F^{k02} + F^{k02} \tilde{A} = i R^{k02}.
\end{aligned} \tag{3.4}$$

Proof. (3.3) is equivalent to the following system of equations

$$\begin{aligned}
&\{\mathcal{N}, F_0\} + R_0 = \check{P}_{00} + \langle \omega', I \rangle, \\
&\{\mathcal{N}, F_1\} + R_1 = 0, \\
&\{\mathcal{N}, F_2\} + R_2 = \langle R^{011} z^+, \bar{z}^+ \rangle.
\end{aligned} \tag{3.5}$$

By simple comparison of coefficients, we see that the first equation in (3.5) is equivalent to the first equation in (3.4). Also note that

$$\{\mathcal{N}, F_1\} = i\sum_k ((\langle k, \omega \rangle F^{k10}, z^+) - \langle \tilde{A}z^+, F^{k10} \rangle)e^{i\langle k, \theta \rangle} + i\sum_k ((\langle k, \omega \rangle F^{k01}, \bar{z}^+) + \langle \tilde{A}\bar{z}^+, F^{k01} \rangle)e^{i\langle k, \theta \rangle}$$

$$= i\sum_k \langle ((\langle k, \omega \rangle I - \tilde{A})F^{k10}, z^+ \rangle e^{i\langle k, \theta \rangle} + i\sum_k \langle ((\langle k, \omega \rangle I + \tilde{A})F^{k01}, \bar{z}^+ \rangle e^{i\langle k, \theta \rangle},$$

and

$$\{\mathcal{N}, F_2\} = i\sum_k ((\langle k, \omega \rangle F^{k20}z^+, z^+) - \langle F^{k20}z^+, \tilde{A}z^+ \rangle - \langle \tilde{A}z^+, (F^{k20})^\top z^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$+ i\sum_{k \neq 0} ((\langle k, \omega \rangle F^{k11}z^+, \bar{z}^+) + \langle F^{k11}z^+, \tilde{A}\bar{z}^+ \rangle - \langle \tilde{A}z^+, (F^{k11})^\top \bar{z}^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$+ i\sum_k ((\langle k, \omega \rangle F^{k02}\bar{z}^+, \bar{z}^+) + \langle F^{k02}\bar{z}^+, \tilde{A}\bar{z}^+ \rangle + \langle \tilde{A}\bar{z}^+, (F^{k02})^\top \bar{z}^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$= i\sum_k ((\langle k, \omega \rangle F^{k20}z^+, z^+) - \langle (\tilde{A}F^{k20} + F^{k20}\tilde{A})z^+, z^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$+ i\sum_{k \neq 0} ((\langle k, \omega \rangle F^{k11}z^+, \bar{z}^+) + \langle (\tilde{A}F^{k11} - F^{k11}\tilde{A})z^+, \bar{z}^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$+ i\sum_k ((\langle k, \omega \rangle F^{k02}\bar{z}^+, \bar{z}^+) + \langle (\tilde{A}F^{k02} + F^{k02}\tilde{A})\bar{z}^+, \bar{z}^+ \rangle)e^{i\langle k, \theta \rangle}$$

$$= i\sum_k \langle ((\langle k, \omega \rangle F^{k20} - \tilde{A}F^{k20} - F^{k20}\tilde{A})z^+, z^+ \rangle e^{i\langle k, \theta \rangle} + i\sum_{k \neq 0} \langle ((\langle k, \omega \rangle F^{k11} + \tilde{A}F^{k11} - F^{k11}\tilde{A})z^+, \bar{z}^+ \rangle e^{i\langle k, \theta \rangle}$$

$$+ i\sum_k \langle ((\langle k, \omega \rangle F^{k02} + \tilde{A}F^{k02} + F^{k02}\tilde{A})\bar{z}^+, \bar{z}^+ \rangle e^{i\langle k, \theta \rangle}.$$

From the second and the third equations in (3.5), we see that $F^{k10}, F^{k01}, F^{k20}, F^{k11}, F^{k02}$ satisfy the corresponding equations in (3.4). $\quad\square$

Let

$$\mathcal{O}_+ = \left\{ \xi \in \mathcal{O} : \begin{array}{l} |\langle k, \omega \rangle^{-1}| \leq \dfrac{|k|^\tau}{\gamma}, k \neq 0 \\[2mm] \|((\langle k, \omega \rangle I + \tilde{A})^{-1}\| \leq K_+^4 \dfrac{|k|^\tau}{\gamma}, \\[2mm] \|((\langle k, \omega \rangle I + \tilde{A} \otimes I + I \otimes \tilde{A})^{-1}\| \leq K_+^8 \dfrac{|k|^\tau}{\gamma}, \\[2mm] \|((\langle k, \omega \rangle I + \tilde{A} \otimes I - I \otimes \tilde{A})^{-1}\| \leq K_+^8 \dfrac{|k|^\tau}{\gamma}, k \neq 0 \end{array} \right\}.$$

Then the first three equations in (3.4) can be solved in this region. Solvability of the remaining equations in (3.4) is a consequence of the following elementary result from matrix theory.

Lemma 3.2. *Let A, B, C be $n \times n$, $m \times m$ and $n \times m$ matrices, respectively, and let X be a $n \times m$ unknown matrix. Then the matrix equation*

$$AX - XB = C,$$

is solvable if and only if $I_m \otimes A - B \otimes I_n$ is nonsingular. Moreover,

$$\|X\| \leq \|(I_m \otimes A - B \otimes I_n)^{-1}\| \cdot \|C\|.$$

Proof. See [33,49]. $\quad\square$

Note that $(F^{k20})^\top$ satisfies the same equation as F^{k20}. Hence by uniqueness of solutions, $F^{k20} = (F^{k20})^\top$. Similarly, $F^{k02} = (F^{k02})^\top$.

Observe that, taking the conjugate transpose of the fourth equation in (3.4) and replacing k with $-k$, one obtains

$$(\langle k, \omega \rangle I + (\bar{\tilde{A}})^\top)(\overline{F^{(-k)20}})^\top + (\overline{F^{(-k)20}})^\top (\bar{\tilde{A}})^\top = i(\overline{R^{(-k)20}})^\top.$$

Using the fact that $(\bar{\tilde{A}})^\top = \tilde{A}$, $(\overline{F^{(-k)20}})^\top = \overline{F^{(-k)20}}$ and $(\overline{R^{(-k)20}})^\top = (R^{k02})^\top = R^{k02}$, one has that

$$(\langle k, \omega \rangle I + \tilde{A}) \overline{F^{(-k)20}} + \overline{F^{(-k)20}} \tilde{A} = iR^{k02};$$

i.e., $\overline{F^{(-k)20}}$ satisfies the same equation as F^{k02}. Thus $\overline{F^{(-k)20}} = F^{k02}$.

In an entirely analogous way one can also show that $\overline{F_{(-k)l}} = F_{kl}$, $(\overline{F^{(-k)11}})^{\top} = F^{k11}$, $\overline{F^{(-k)10}} = F^{k01}$, $\overline{F^{(-k)01}} = F^{k10}$ and $\overline{F^{(-k)02}} = F^{k20}$. Thus

$$\bar{F} = F.$$

We will now proceed to estimate the norm of X_F and to study properties of Φ_F^1. To this end, we will need the following

Lemma 3.3. *Let $A = (a_{ij}(\xi))_{K \times K}$ be an invertible matrix depending differentiably on a parameter $\xi \in \mathcal{O}$, and $\|A^{-1}\| \leq L$, $\|\partial_\xi A\| \leq M$. Then*

$$\|\partial_\xi A^{-1}\| \leq L^2 M.$$

Proof. Because $AA^{-1} = I$, then $(\partial_\xi A)A^{-1} + A(\partial_\xi A^{-1}) = 0$, hence $\partial_\xi A^{-1} = -A^{-1}(\partial_\xi A)A^{-1}$. Thus

$$\|\partial_\xi A^{-1}\| \leq \|A^{-1}\|^2 \|\partial_\xi A\| \leq L^2 M. \quad \square$$

Lemma 3.4. *Let $D_i = D(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $0 < i \leq 4$. If*

(C1) $K_+^{18} \leq \varepsilon^{-\frac{1}{4}}$,

   then there is a constant $c_2 > 0$ such that

$$\|X_F\|_{D_3, \mathcal{O}_+} \leq c_2 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} \varepsilon^{\frac{3}{4}}.$$

Proof. From the definition of $\mathcal{O}_+$ we see that

$$\sup_{\xi \in \mathcal{O}_+} \|\partial_\xi \langle k, \omega \rangle\| \leq |k|,$$

$$\sup_{\xi \in \mathcal{O}_+} \|\partial_\xi (\langle k, \omega \rangle I + \tilde{A})\| \leq (|k| + K_+),$$

$$\sup_{\xi \in \mathcal{O}_+} \|\partial_\xi (\langle k, \omega \rangle I + \tilde{A} \otimes I + I \otimes \tilde{A})\| \leq (|k| + K_+^2),$$

$$\sup_{\xi \in \mathcal{O}_+} \|\partial_\xi (\langle k, \omega \rangle I + \tilde{A} \otimes I - I \otimes \tilde{A})\| \leq (|k| + K_+^2).$$

Then, according to Lemmas 3.1–3.3, we have

$$|F_{kl}|_{\mathcal{O}_+} \leq |\langle k, \omega \rangle|^{-2} |k| |\check{P}_{kl}|_{\mathcal{O}_+} \leq \gamma^{-2} |k|^{2\tau + 1} |\check{P}_{kl}|_{\mathcal{O}_+}, \quad k \neq 0, \ |l| \leq 1;$$

$$\|F^{kij}\|_{\mathcal{O}_+} \leq \gamma^{-2} K_+^9 |k|^{2\tau + 1} \|R^{kij}\|_{\mathcal{O}_+}, \quad i \neq j, \ 1 \leq i + j \leq 2;$$

$$\|F^{k11}\|_{\mathcal{O}_+} \leq \gamma^{-2} K_+^{18} |k|^{2\tau + 1} \|R^{k11}\|_{\mathcal{O}_+}, \quad k \neq 0.$$

It follows that

$$\frac{1}{s^2} \|F_\theta\|_{D_3, \mathcal{O}_+} \leq \frac{1}{s^2} \left(\sum_{k, |l| \leq 1} |F_{kl}| \cdot s^{2|l|} \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} + \sum_{1 \leq i + j \leq 2} \sum_k \|F^{kij}\| \cdot \|z^+\| \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \right)$$

$$\leq \frac{\gamma^{-2} K_+^{18}}{s^2} \left(\sum_{k, |l| \leq 1} |\check{P}_{kl}| \cdot s^{2|l|} \cdot |k|^{2\tau + 2} \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} + \sum_{1 \leq i + j \leq 2} \sum_k \|R^{kij}\| \cdot \|z^+\| \cdot |k|^{2\tau + 2} \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \right)$$

$$\leq c_3 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} K_+^{18} \|X_R\| \leq c_3 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} \varepsilon^{\frac{3}{4}}.$$

A similar derivation yields

$$\|F_I\|_{D_3, \mathcal{O}_+} = \sum_{|l| = 1} |F_{kl}| e^{|k|(r - \frac{1}{4}(r - r_+))} \leq c_4 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} \varepsilon^{\frac{3}{4}}.$$

$$\|X_{F_1}\|_{D_3, \mathcal{O}_+} \leq \frac{1}{s} \left(\sum_n \|F_{1_{w_n}}\| + \sum_n \|F_{1_{\bar{w}_n}}\| \right) \leq \frac{1}{s} (\|F_{1_{z^+}}\| + \|F_{1_{\bar{z}^+}}\|)$$

$$\leq c_5 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} K_+^{18} \|X_{R_1}\| \leq c_5 \gamma^{-2}(r - r_+)^{-(2\tau + N + 1)} \varepsilon^{\frac{3}{4}}.$$

$$\|X_{F_2}\|_{D_3,\mathcal{O}_+} \le \frac{1}{s}\left(\sum_n \|F_{2_{w_n}}\| + \sum_n \|F_{2_{\bar{w}_n}}\|\right) \le \frac{1}{s}(\|F_{2_{z^+}}\| + \|F_{2_{\bar{z}^+}}\|)$$

$$\le c_6\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}K_+^{18}\|X_{R_2}\| \le c_6\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{3}{4}}.$$

Putting together the estimates above, Lemma 3.4 follows. $\square$

Now let $D_{i\eta} = D(r_+ + \frac{i}{4}(r-r_+), \frac{i}{4}\eta s), 0 < i \le 4$.

Lemma 3.5. *If*

(C2) $c_2\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{1}{2}} < 1,$

   then

$$\Phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \le t \le 1,$$

and moreover,

$$\|D\Phi_F^t - I\|_{D_{1\eta}} < c_7\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{3}{4}}.$$

Proof. Let

$$\|D^m F\|_{D,\mathcal{O}_+} = \max\left\{\left\|\frac{\partial^{|i|+|l|+|\alpha|+|\beta|}F}{\partial\theta^i\,\partial I^l\,\partial(z^+)^\alpha\,\partial(\bar{z}^+)^\beta}\right\|_{D,\mathcal{O}_+}, \ |i|+|l|+|\alpha|+|\beta| = m \ge 2\right\}.$$

We note that F is a polynomial of order 1 in I and of order 2 in z^+, $\bar{z}^+$. It thus follows from Lemma 3.4 and Cauchy inequality (Lemma A.2 in the Appendix) that

$$\|D^m F\|_{D_2,\mathcal{O}_+} < c_8\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{3}{4}},$$

for any $m \ge 2$.

Using the integral equation

$$\Phi_F^t = id + \int_0^t X_F \circ \Phi_F^s \mathrm{d}s$$

and Lemma 3.4, one sees easily that $\Phi_F^t : D_{2\eta} \to D_{3\eta}, -1 \le t \le 1$. Moreover, since

$$D\Phi_F^t = Id + \int_0^t (DX_F)D\Phi_F^s \mathrm{d}s = Id + \int_0^t J(D^2 F)D\Phi_F^s \mathrm{d}s,$$

where J denotes the standard symplectic matrix, it follows that

$$\|D\Phi_F^t - I\| \le 2\|D^2 F\| \le c_7\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{3}{4}},$$

where $c_7 = 2c_8$. $\square$

3.4. The new Hamiltonian

Let $\Phi = \Phi_F^1, s_+ = \frac{1}{8}\eta s, D_+ = D(r_+, s_+)$ and

$$\mathcal{N}_+ = e_+ + \langle\omega_+, I\rangle + \langle A^+ z^+, \bar{z}^+\rangle + \sum_{|n|>K_+}\Omega_n w_n\bar{w}_n,$$

$$\mathcal{P}_+ = \check{P}_+ + \acute{P}_0 + \grave{P}_0,$$

where

$$e_+ = e + \check{P}_{00},$$
$$\omega_+ = \omega + \check{P}_{0l}(|l| = 1),$$
$$A^+ = \tilde{A} + R^{011},$$
$$z^+ = (\dots, w_n, \dots)_{|n|\le K_+}, \qquad \bar{z}^+ = (\dots, \bar{w}_n, \dots)_{|n|\le K_+},$$
$$\check{P}_+ = \int_0^1 \{(1-t)\{\mathcal{N}, F\} + R, F\} \circ \Phi_F^t \mathrm{d}t + (\check{P} - R)\circ\Phi_F^1 + \int_0^1 \{\acute{P}_0 + \grave{P}_0, F\}\circ\Phi_F^t \mathrm{d}t.$$

Then $\Phi : D_+ \times \mathcal{O}_+ \to D$ and, by Taylor's second-order formula,

$$
\begin{aligned}
H_+ &\equiv H \circ \Phi = (\mathcal{N} + R) \circ \Phi + (\mathcal{P} - R) \circ \Phi \\
&= \mathcal{N} + \{\mathcal{N}, F\} + R + \int_0^1 (1-t)\{\{\mathcal{N}, F\}, F\} \circ \Phi_F^t \mathrm{d}t + \int_0^1 \{R, F\} \circ \Phi_F^t \mathrm{d}t + (\check{P} - R) \circ \Phi_F^1 + (\acute{P}_0 + \grave{P}_0) \circ \Phi_F^1 \\
&= \mathcal{N} + \{\mathcal{N}, F\} + R + \check{P}_+ + \acute{P}_0 + \grave{P}_0 \\
&= \mathcal{N}_+ + \mathcal{P}_+ + \{\mathcal{N}, F\} + R - \check{P}_{00} - \langle \omega', I \rangle - \langle R^{011} z^+, \bar{z}^+ \rangle \\
&= \mathcal{N}_+ + \mathcal{P}_+.
\end{aligned}
$$

We now show that H_+ has properties similar to those of H.

Observe that, since $(\overline{\tilde{A}})^\top = \tilde{A}$ and $(\overline{R^{011}})^\top = R^{011}$, we have $(\overline{A^+})^\top = A^+$, i.e., A^+ is a Hermitian matrix. Then, from the assumptions on $\check{P}$, we further have that there is a constant $c_9 > 0$ such that

$$
|\omega_+ - \omega|_{\mathcal{O}_+} \le c_9 \varepsilon, \qquad \|A^+ - \tilde{A}\|_{\mathcal{O}_+} \le c_9 \varepsilon.
$$

It thus follows that, if

(C3) $c_9 K_+^{\tau+9} \varepsilon < \gamma - \gamma_+$,

then whenever $|k| \le K_+$,

$$
|\langle k, \omega + P_{0l} \rangle^{-1}| \le \frac{|\langle k, \omega \rangle^{-1}|}{1 - |\langle k, \omega \rangle^{-1}| c_9 |k| \varepsilon} \le \frac{|k|^\tau}{\gamma_+}, \quad k \ne 0,
$$

$$
\|(\langle k, \omega + P_{0l} \rangle I + A^+)^{-1}\| \le \frac{\|(\langle k, \omega \rangle I + \tilde{A})^{-1}\|}{1 - \|(\langle k, \omega \rangle I + \tilde{A})^{-1}\| c_9 |k| \varepsilon} \le K_+^4 \frac{|k|^\tau}{\gamma_+},
$$

$$
\|(\langle k, \omega + P_{0l} \rangle I + A^+ \otimes I + I \otimes A^+)^{-1}\| \le \frac{\|(\langle k, \omega \rangle I + \tilde{A} \otimes I + I \otimes \tilde{A})^{-1}\|}{1 - \|(\langle k, \omega \rangle I + \tilde{A} \otimes I + I \otimes \tilde{A})^{-1}\| c_9 |k| \varepsilon} \le K_+^8 \frac{|k|^\tau}{\gamma_+},
$$

$$
\begin{aligned}
\|(\langle k, \omega + P_{0l} \rangle I + A^+ \otimes I - I \otimes A^+)^{-1}\| &\le \frac{\|(\langle k, \omega \rangle I + \tilde{A} \otimes I - I \otimes \tilde{A})^{-1}\|}{1 - \|(\langle k, \omega \rangle I + \tilde{A} \otimes I - I \otimes \tilde{A})^{-1}\| c_9 |k| \varepsilon} \\
&\le K_+^8 \frac{|k|^\tau}{\gamma_+}, \quad k \ne 0.
\end{aligned}
$$

The above implies that, in the next KAM step, small denominator conditions are automatically satisfied when $|k| \le K_+$.

Let $R(t) = (1-t)(\mathcal{N}_+ - \mathcal{N}) + tR$. Then $\mathcal{P}_+$ can be rewritten as

$$
\begin{aligned}
\mathcal{P}_+ &= \int_0^1 (1-t)\{\{\mathcal{N}, F\}, F\} \circ \Phi_F^t \mathrm{d}t + \int_0^1 \{R, F\} \circ \Phi_F^t \mathrm{d}t + (\mathcal{P} - R) \circ \Phi_F^1 \\
&= \int_0^1 \{R(t), F\} \circ \Phi_F^t \mathrm{d}t + (\mathcal{P} - R) \circ \Phi_F^1.
\end{aligned}
$$

Hence

$$
X_{\mathcal{P}_+} = \int_0^1 (\Phi_F^t)^* X_{\{R(t), F\}} \mathrm{d}t + (\Phi_F^1)^* X_{(\mathcal{P}-R)}.
$$

By Lemma 3.5, if

(C4) $c_7 \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \varepsilon^{\frac{3}{4}} \le 1$,

then

$$
\|D\Phi_F^t\|_{D_{1\eta}} \le 1 + \|D\Phi_F^t - I\|_{D_{1\eta}} \le 2, \quad -1 \le t \le 1.
$$

Furthermore, by Lemma A.4 (see Appendix) and (3.2), we also have

$$
\|X_{\{R(t), F\}}\|_{D_{2\eta}} \le c_{10} \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \eta^{-2} \varepsilon^{\frac{7}{4}},
$$

$$
\|X_{(\mathcal{P}-R)}\|_{D_{2\eta}} \le c_1 \varepsilon^{\frac{5}{4}}.
$$

Let $c_0 = \max\{c_1, \dots, c_{10}, c_{11}, c_{12}\}$, where c_{11}, c_{12} will be defined later on, and let

$$
\varepsilon_+ = 4c_0 \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \varepsilon^{\frac{5}{4}}.
$$

Then

$$\|X_{\mathcal{P}_+}\|_{D_+,\mathcal{O}_+} \le 2c_1\varepsilon^{\frac{5}{4}} + 2c_{10}\gamma^{-2}(r-r_+)^{-(2\tau+N+1)}\varepsilon^{\frac{5}{4}} \le \varepsilon_+.$$

The reality of $\mathcal{P}_+$ is verified easily because, if F satisfies $\bar{F} = F$ and G satisfies $\bar{G} = G$, then their Poisson bracket $\{F, G\}$ also satisfies $\overline{\{F, G\}} = \{\bar{F}, \bar{G}\} = \{F, G\}$.

We now examine the decay property of $\check{P}_+$. More precisely, write

$$\check{P}_+ = \check{P}^+(\theta, I, z^+, \bar{z}^+, \xi) + \sum_{|n|>K_+, \alpha_n+\beta_n\ge 1} \check{P}_n^+(\theta, I, z^+, \bar{z}^+, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}.$$

We will show that

$$\|\check{P}_n^+(\theta, I, z^+, \bar{z}^+, \xi)\|_{D_+,\mathcal{O}_+} \le e^{-(|n|-K_+)}, \quad |n| > K_+.$$

Since F only involves the normal components $w_n, \bar{w}_n$ for $|n| \le K_+$, so does $\{\mathcal{N}, F\}$. Therefore, $\int_0^1(1-t)\{\{\mathcal{N}, F\}, F\} \circ \Phi_F^t dt$ only involves the normal components $w_n, \bar{w}_n$ for $|n| \le K_+$. Since

$$\grave{P}_0 = \sum_n O(|w_n|^3),$$

we conclude that $\{\grave{P}_0, F\}$ also only involves normal components $w_n, \bar{w}_n$ for $|n| \le K_+$. Consequently, the same is true about $\int_0^1\{\grave{P}_0, F\} \circ \Phi_F^t dt$. Now, since R is a truncation of $\check{P}$, in order to establish the decay property above, it suffices to consider the following two terms: $(\check{P} - R)$ and $\int_0^1\{\check{P} + \grave{P}_0, F\} \circ \Phi_F^t dt$. Let us consider the first term. Recall that

$$\check{P} = \check{P}(\theta, I, z, \bar{z}, \xi) + \sum_{|n|>K, \alpha_n+\beta_n\ge 1} \check{P}_n(\theta, I, z, \bar{z}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n},$$

$$\|\check{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r,s),\mathcal{O}_+} \le e^{-(|n|-K)};$$

$$\grave{P}_0 = \sum_{\substack{n\ne m \\ \alpha_n+\beta_n, \alpha_m+\beta_m \ge 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m \ge 3}} \grave{P}_{nm}^0(\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}w_m^{\alpha_m}\bar{w}_m^{\beta_m},$$

$$\|\grave{P}_{nm}^0(\xi)\|_{D(r,s),\mathcal{O}_+} \le e^{-|n-m|}.$$

Since R only involves normal components $w_n, \bar{w}_n$ for $|n| \le K_+$, terms corresponding to normal components $w_n, \bar{w}_n$ for $|n| > K_+$ in $\check{P} - R$ are those corresponding to $|n| > K_+$ in $\check{P}$, for which we already have the decay property

$$\|\check{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r_+,s_+),\mathcal{O}_+} \le e^{-(|n|-K)} \le e^{-(|n|-K_+)}.$$

It remains to obtain decay estimates for $\int_0^1\{\check{P} + \grave{P}_0, F\} \circ \Phi_F^t dt$. Once again, we only need to consider terms corresponding to normal components $w_n, \bar{w}_n$ for $|n| > K_+$. Note that F is independent of such components and thus so will be

$$\int_0^1\left\{\check{P}(\theta, I, z, \bar{z}, \xi) + \sum_{K<|n|\le K_+, \alpha_n+\beta_n\ge 1} \check{P}_n(\theta, I, z, \bar{z}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}, F\right\} \circ \Phi_F^t dt.$$

By the same token,

$$\int_0^1\left\{\sum_{\substack{n\ne m, |n|, |m|\le K_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m \ge 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m \ge 3}} \grave{P}_{nm}^0(\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}w_m^{\alpha_m}\bar{w}_m^{\beta_m}, F\right\} \circ \Phi_F^t dt$$

is independent of normal components $w_n, \bar{w}_n$ for $|n| > K_+$. It only remains to consider the following terms

$$\int_0^1\left\{\sum_{|n|>K_+, \alpha_n+\beta_n\ge 1} \check{P}_n(\theta, I, z, \bar{z}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}, F\right\} \circ \Phi_F^t dt = \int_0^1 \sum_{|n|>K_+, \alpha_n+\beta_n\ge 1} \{\check{P}_n(\theta, I, z, \bar{z}, \xi), F\} \circ \Phi_F^t w_n^{\alpha_n}\bar{w}_n^{\beta_n} dt$$

$$= \sum_{|n|>K_+, \alpha_n+\beta_n\ge 1} \left(\int_0^1\{\check{P}_n(\theta, I, z, \bar{z}, \xi), F\} \circ \Phi_F^t dt\right) w_n^{\alpha_n}\bar{w}_n^{\beta_n}, \tag{3.6}$$

and

$$\int_0^1 \left\{ \sum_{\substack{n \neq m, |n| > K_+, |m| \leq K_+ \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}^0(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F \right\} \circ \Phi_F^t \, dt$$

$$= \int_0^1 \sum_{|n| > K_+} \left\{ \sum_{\substack{|m| \leq K_+ \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}^0(\xi) w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F \right\} \circ \Phi_F^t w_n^{\alpha_n} \bar{w}_n^{\beta_n} \, dt$$

$$= \sum_{|n| > K_+} \left(\int_0^1 \left\{ \sum_{\substack{|m| \leq K_+ \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}^0(\xi) w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F \right\} \circ \Phi_F^t \, dt \right) w_n^{\alpha_n} \bar{w}_n^{\beta_n}. \tag{3.7}$$

Let

$$\tilde{P}_n = \check{P}_n(\theta, I, z, \bar{z}, \xi) + \sum_{\substack{|m| \leq K_+ \\ \alpha_n + \beta_n, \alpha_m + \beta_m \geq 1 \\ \alpha_n + \beta_n + \alpha_m + \beta_m \geq 3}} \acute{P}_{nm}^0(\xi) w_m^{\alpha_m} \bar{w}_m^{\beta_m}.$$

We will combine (3.6) and (3.7) to obtain the decay property of

$$\sum_{|n| > K_+, \alpha_n + \beta_n \geq 1} \left(\int_0^1 \{ \tilde{P}_n, F \} \circ \Phi_F^t \, dt \right) w_n^{\alpha_n} \bar{w}_n^{\beta_n}.$$

By relaxing decay properties of $e^{-(|n|-K)}$, $e^{-|n-m|}$ to $e^{-(|n|-K_+)}$ we have, by Lemma A.3 in the Appendix that

$$\| \{ \tilde{P}_n, F \} \|_{D(r-\sigma, \frac{1}{2}s)} \leq c_{11} \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \sigma^{-1} s^{-2} \varepsilon^{\frac{3}{4}} e^{-(|n|-K_+)}.$$

It follows by Cauchy inequality (Lemma A.2 in the Appendix) that

$$\| X_{\{\tilde{P}_n, F\}} \|_{D(r-2\sigma, \frac{1}{4}s)} \leq c_{12} \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \sigma^{-2} s^{-4} \varepsilon^{\frac{3}{4}} e^{-(|n|-K_+)}.$$

Therefore, by Lemma 3.5, if

(C5) $c_{11} \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \eta^{-2} \varepsilon^{\frac{3}{4}} \leq \frac{1}{2}$,

(C6) $c_{12} c_2 (\gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \eta^{-2} \varepsilon^{\frac{3}{4}})^2 \leq \frac{1}{2}$,

then

$$\left\| \int_0^1 \{ \tilde{P}_n, F \} \circ \Phi_F^t \, dt \right\|_{D(r_+, s_+)}$$

$$\leq \| \{ \tilde{P}_n, F \} \circ \Phi_F^t \|_{D(r_+, s_+)}$$

$$\leq \| \{ \tilde{P}_n, F \} \|_{D(r_+, s_+)} + \| \{ \tilde{P}_n, F \} \circ \Phi_F^t - \{ \tilde{P}_n, F \} \|_{D(r_+, s_+)}$$

$$\leq \| \{ \tilde{P}_n, F \} \|_{D(r_+, s_+)} + \| X_{\{\tilde{P}_n, F\}} \|_{D_{2\eta}} \| \Phi_F^t - id \|_{D_{1\eta}}$$

$$\leq c_{11} \gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \eta^{-2} \varepsilon^{\frac{3}{4}} e^{-(|n|-K_+)}$$

$$\quad + c_{12} c_2 (\gamma^{-2} (r - r_+)^{-(2\tau+N+1)} \eta^{-2} \varepsilon^{\frac{3}{4}})^2 e^{-(|n|-K_+)}$$

$$\leq e^{-(|n|-K_+)}.$$

Note in this KAM step, the normal components w_n, $\bar{w}_n$ with $K < |n| \leq K_+$ are involved but, at this time, the perturbation is of order $O(\varepsilon)$, which means

$$|w_n| \sim \varepsilon \sim e^{-|n|}. \tag{3.8}$$

This completes one step of KAM iterations. $\square$

4. Proof of Theorem C

Let $r_0, s_0, \varepsilon_0, \gamma_0, K_0, \mathcal{O}_0, H_0, \mathcal{N}_0, \mathcal{P}_0$ be as given in the beginning of Section 3. For each $\nu = 0, 1, \ldots$, we will label all index-free quantities by ν and all +-indexed quantities by $\nu + 1$, in Section 3. This defines, for all $\nu = 1, 2, \ldots$, the following sequences:

$$r_\nu = r_0 \left(1 - \sum_{i=2}^{\nu+1} 2^{-i} \right),$$

$$\varepsilon_\nu = 4c_0 \gamma_{\nu-1}^{-2} (r_{\nu-1} - r_\nu)^{-(2\tau+N+1)} \varepsilon_{\nu-1}^{\frac{5}{4}},$$

$$\gamma_\nu = \gamma_0 \left(1 - \sum_{i=2}^{\nu+1} 2^{-i} \right),$$

$$s_\nu = \frac{1}{8} \eta_{\nu-1} s_{\nu-1} = 2^{-3\nu} \left(\prod_{i=0}^{\nu-1} \varepsilon_i \right)^{\frac{1}{4}} s_0, \quad \eta_\nu = \varepsilon_\nu^{\frac{1}{4}},$$

$$K_\nu = 5K_{\nu-1},$$

$$D_\nu = D(r_\nu, s_\nu),$$

$$\tilde{D}_\nu = D \left(r_{\nu+1} + \frac{1}{4} (r_\nu - r_{\nu+1}), \frac{1}{4} \eta_\nu s_\nu \right),$$

$$H_\nu = \bar{H}_\nu = \mathcal{N}_\nu + \mathcal{P}_\nu,$$

$$\mathcal{N}_\nu = \bar{\mathcal{N}}_\nu = e_\nu + \langle \omega_\nu(\xi), I \rangle + \langle A^\nu z^\nu, \bar{z}^\nu \rangle + \sum_{|n|>K_\nu} \Omega_n w_n \bar{w}_n,$$

$$\mathcal{O}_\nu = \left\{ \xi \in \mathcal{O}_{\nu-1} : \begin{array}{l} |\langle k, \omega_{\nu-1} \rangle^{-1}| \leq \dfrac{|k|^\tau}{\gamma_{\nu-1}}, k \neq 0 \\[2ex] \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}^{\nu-1})^{-1}\| \leq K_\nu^4 \dfrac{|k|^\tau}{\gamma_{\nu-1}} \\[2ex] \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}^{\nu-1} \otimes I + I \otimes \tilde{A}^{\nu-1})^{-1}\| \leq K_\nu^8 \dfrac{|k|^\tau}{\gamma_{\nu-1}} \\[2ex] \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}^{\nu-1} \otimes I - I \otimes \tilde{A}^{\nu-1})^{-1}\| \leq K_\nu^8 \dfrac{|k|^\tau}{\gamma_{\nu-1}}, k \neq 0 \end{array} \right\},$$

where

$$\tilde{A}^{\nu-1} = \begin{pmatrix} A^{\nu-1} & 0 \\ 0 & \Omega_n \end{pmatrix}_{K_{\nu-1} < |n| \leq K_\nu}.$$

4.1. Iteration lemma

The preceding analysis may be summarized in the following

Lemma 4.1. *Given γ to be sufficiently small, there is an ε sufficiently small such that the following holds for all $\nu = 0, 1, \ldots$.*

(a) *H_ν is real analytic on $D_\nu \times \mathcal{O}_\nu$,*

$$\mathcal{N}_\nu = \bar{\mathcal{N}}_\nu = e_\nu + \langle \omega_\nu(\xi), I \rangle + \langle \tilde{A}^\nu z^{\nu+1}, \bar{z}^{\nu+1} \rangle + \sum_{|n|>K_{\nu+1}} \Omega_n w_n \bar{w}_n,$$

$$\mathcal{P}_\nu = \bar{\mathcal{P}}_\nu = \check{P}_\nu + \acute{P}_0 + \grave{P}_0,$$

and moreover,

$$|\omega_{\nu+1} - \omega_\nu|_{\mathcal{O}_\nu} \leq c_0 \varepsilon_\nu,$$

$$\|\tilde{A}^{\nu+1} - \tilde{A}^\nu\|_{\mathcal{O}_\nu} \leq c_0 \varepsilon_\nu,$$

$$\|X_{\mathcal{P}_\nu}\|_{D_\nu, \mathcal{O}_\nu} \leq \varepsilon_\nu,$$

$$\check{P}_\nu = \check{P}^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi) + \sum_{|n|>K_\nu, \alpha_n+\beta_n \geq 1} \check{P}_n^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n},$$

$$\acute{P}_0 = \sum_{\substack{n \neq m \\ \alpha_n+\beta_n, \alpha_m+\beta_m \geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m \geq 3}} \acute{P}_{nm}^0(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m},$$

$$\dot{P}_0 = \sum_n O(|w_n|^3),$$

with

$$\|\check{P}_n^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi)\|_{D_\nu, \mathcal{O}_\nu} \le e^{-(|n|-K_\nu)},$$
$$\|\acute{P}_{nm}^0(\xi)\|_{D_\nu, \mathcal{O}_\nu} \le e^{-|n-m|}.$$

(b) *There is a symplectic transformation*

$$\Phi_\nu : \tilde{D}_\nu \times \mathcal{O}_{\nu+1} \to D_\nu$$

with

$$\|D\Phi_\nu - I\|_{D_\nu \times \mathcal{O}_{\nu+1}} \le \varepsilon_\nu$$

such that

$$H_{\nu+1} = \bar{H}_{\nu+1} = H_\nu \circ \Phi_\nu.$$

Proof. It suffices to verify conditions (C0)–(C6) for all $\nu = 0, 1, \ldots$.

By the choice of s_0, the condition (C0) clearly holds for $\nu = 0$. By induction, suppose it hold for some $\nu - 1$. Then it is easy to see that one can make $\gamma = \gamma_0$ small enough so that

$$s_\nu = \frac{1}{8}\eta_{\nu-1}s_{\nu-1} < \frac{1}{8}\varepsilon_{\nu-1}^{\frac{5}{4}} < \varepsilon_\nu.$$

Hence (C0) holds for all ν.

To verify the conditions (C1)–(C6), we note that these easily follow from conditions:

(D1) $K_{\nu+1} \le \varepsilon_\nu^{-\frac{1}{80}}$,

(D2) $c_0\gamma_\nu^{-2}(r_\nu - r_{\nu+1})^{-(2\tau+N+1)}\varepsilon_\nu^{\frac{1}{4}} \le \frac{1}{2}$

for all $\nu = 0, 1, \ldots$.

Let us first take ε (hence ε_0) sufficiently small such that

$$\varepsilon_0 < \min\left\{\frac{\gamma_0^{10}r_0^{5(2\tau+N+1)}}{2^{5(2\tau+N+1)}c_0^5}(\Psi(r_0))^{-1}, \frac{\delta}{2}\right\},$$

where

$$\Psi(r_0) = \prod_{i=1}^{\infty}[(r_{i-1} - r_i)^{-5(2\tau+N+1)}]^{(\frac{4}{5})^i}$$

which is easily seen to be well-defined. Then

$$c_0\gamma_0^{-2}(r_0 - r_1)^{-(2\tau+N+1)}\varepsilon_0^{\frac{1}{4}} \le \frac{1}{2},$$

i.e., (D2) holds for $\nu = 0$. Since $K_0 = -\frac{5}{4}\ln\varepsilon$, we see that (D1) also holds for $\nu = 0$.

Using an induction argument, one can show that, in fact, for any $\nu \ge 1$,

$$\begin{aligned}
c_0\gamma_\nu^{-2}(r_\nu - r_{\nu+1})^{-(2\tau+N+1)}\varepsilon_\nu^{\frac{1}{4}} &= c_0\gamma_\nu^{-2}(r_\nu - r_{\nu+1})^{-(2\tau+N+1)}(4c_0\gamma_\nu^{-2}(r_{\nu-1} - r_\nu)^{-(2\tau+N+1)}\varepsilon_{\nu-1}^{\frac{5}{4}})^{\frac{1}{4}} \\
&\le (2^{4(2\tau+N+1)}c_0^5\gamma_\nu^{-10}(r_{\nu-1} - r_\nu)^{-5(2\tau+N+1)}\varepsilon_{\nu-1}^{\frac{5}{4}})^{\frac{1}{4}} \le (2^{4(2\tau+N+1)}c_0^5\gamma_0^{-10}\Psi(r_0)\varepsilon_0)^{\frac{1}{4}(\frac{5}{4})^\nu} \\
&\le \left(\frac{r_0^{5(2\tau+N+1)}}{2^{2\tau+N+1}}\right)^{\frac{1}{4}(\frac{5}{4})^\nu} \le \frac{1}{2},
\end{aligned}$$

and

$$K_{\nu+1} = 5^{\nu+1}K_0 \le \varepsilon_\nu^{-\frac{1}{80}},$$

i.e., (D1) and (D2) hold true. $\square$

4.2. Convergence

Let $\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\nu-1}$, $\nu = 1, 2, \ldots$. An induction argument shows that $\Psi^\nu : \tilde{D}_\nu \times \mathcal{O}_{\nu+1} \to D_0$ and

$$H_0 \circ \Psi^\nu = H_\nu = \mathcal{N}_\nu + \mathcal{P}_\nu$$

for all $\nu = 1, 2, \ldots$.

Let $\tilde{\mathcal{O}} = \cap_{\nu=0}^\infty \mathcal{O}_\nu$. Using Lemma 4.1 and standard arguments (e.g. [31,38]), one can conclude that $H_\nu, e_\nu, \mathcal{N}_\nu, \mathcal{P}_\nu, \Psi^\nu$ and ω_ν converge uniformly on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$ to, say, $H_\infty, e_\infty, \mathcal{N}_\infty, \mathcal{P}_\infty, \Psi^\infty$ and ω_∞, respectively, in which case it is clear that

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

Since

$$\varepsilon_\nu = 4c_0 \gamma_{\nu-1}^{-2} (r_{\nu-1} - r_\nu)^{-(2\tau+N+1)} \varepsilon_{\nu-1}^{\frac{5}{4}} \le (4c_0 \gamma_0^{-2} \Psi(r_0)\varepsilon_0)^{(\frac{5}{4})^\nu},$$

we have, by Lemma 4.1, that

$$X_{\mathcal{P}^\infty}|_{D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}} \equiv 0.$$

Let Φ_H^t denote the flow of any Hamiltonian vector field X_H. Since $H_0 \circ \Psi^\nu = H_\nu$, we have

$$\Phi_{H_0}^t \circ \Psi^\nu = \Psi^\nu \circ \Phi_{H_\nu}^t. \tag{4.1}$$

The uniform convergence of Ψ^ν and X_{H_ν} implies that one can pass the limit in the above and conclude that

$$\Phi_{H_0}^t \circ \Psi^\infty = \Psi^\infty \circ \Phi_{H_\infty}^t,$$

on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$. It thus follows that

$$\Phi_{H_0}^t(\Psi^\infty(\mathbb{T}^N \times \{\xi\})) = \Psi^\infty \Phi_{\mathcal{N}_\infty}^t(\mathbb{T}^N \times \{\xi\}) = \Psi^\infty(\mathbb{T}^N \times \{\xi\}),$$

for all $\xi \in \tilde{\mathcal{O}}$. Hence $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $\xi \in \tilde{\mathcal{O}}$. The frequencies $\omega_\infty(\xi)$ associated with $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ are slightly deformed from the unperturbed ones, $\omega(\xi)$, and moreover, it follows from the constancy of the normal matrix $A^\infty = A^\infty(\xi)$ that the invariant torus $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ is linearly stable.

Let $\theta \in T^N$ and $\Psi(\theta + \tilde{\omega}(\xi)t, \xi) = (\theta + \tilde{\omega}(\xi)t, I(t), \{w_n(t)\})$. Then it follows from (3.8) that $|w_n| \sim e^{-|n|}$. $\square$

4.3. Measure estimates

For each $\nu = 0, 1, 2, \ldots$, recall that, for $|k| \le K_{\nu+1}$, small denominator conditions are automatically satisfied. For $|k| > K_{\nu+1}$, let us consider the most complicated case, $\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu$. Since $\tilde{A}^\nu$ is Hermitian with $\dim(\tilde{A}^\nu) \le K_{\nu+1}$, $\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu$ is also Hermitian with $\dim(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu) \le K_{\nu+1}^2$.

Lemma 4.2. *Let $\mu_1, \ldots, \mu_K$ be the eigenvalues of a Hermitian matrix A and let P with $\bar{P}^\top P = I$ be such that $A = P^\top \Lambda P$, with $\Lambda = (\mu_j)_{1 \le j \le K}$. If*

$$\min\{|\mu_1|, \ldots, |\mu_K|\} \ge l,$$

then

$$\|A^{-1}\| \le \frac{K^2}{l}.$$

Proof. Since $\bar{P}^\top P = I$, we have $\|P\| \le K$, and hence

$$\|A^{-1}\| \le \|P\|^2 \|\Lambda^{-1}\| \le \frac{K^2}{l}. \quad \square$$

Let μ be an eigenvalue of $(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu)$. According to Lemma 4.2, as long as $|\mu| \ge \frac{\gamma_\nu}{|k|^\tau |m|^2 |n|^2}$ for some $|m|, |n| \le K_{\nu+1}$, one will have

$$\|(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu)^{-1}\| \le K_{\nu+1}^8 \frac{|k|^\tau}{\gamma_\nu}.$$

As a consequence of this, we need to exclude the following parameter set

$$\mathcal{R}^{\nu 1}_{kmn} = \left\{ \xi \in \mathcal{O}_{\nu-1} : |\mu| < \frac{\gamma_\nu}{|k|^\tau |m|^2 |n|^2} \right\}.$$

Similarly, we will also need to exclude the following parameter sets

$$\mathcal{R}^{\nu 2}_k = \left\{ \xi \in \mathcal{O}_{\nu-1} : |\langle k, \omega_\nu \rangle| < \frac{\gamma_\nu}{|k|^\tau} \right\},$$

$$\mathcal{R}^{\nu 3}_{kn} = \left\{ \xi \in \mathcal{O}_{\nu-1} : |\zeta| < \frac{\gamma_\nu}{|k|^\tau |n|^2} \right\},$$

$$\mathcal{R}^{\nu 4}_{kmn} = \left\{ \xi \in \mathcal{O}_{\nu-1} : |\lambda| < \frac{\gamma_\nu}{|k|^\tau |m|^2 |n|^2} \right\},$$

where ζ, λ denote eigenvalues of $(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu)$, $(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I + I \otimes \tilde{A}^\nu)$, respectively.

Now consider the resonant sets

$$\mathcal{R}^\nu = \bigcup_{|k| > K_{\nu+1}} \left(\left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 1}_{kmn} \right) \bigcup \mathcal{R}^{\nu 2}_k \bigcup \left(\bigcup_{|n| \leq K_{\nu+1}} \mathcal{R}^{\nu 3}_{kn} \right) \bigcup \left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 4}_{kmn} \right) \bigcup \mathcal{R}^{\nu 5}_k \right),$$

where

$$\mathcal{R}^{\nu 5}_k = \left\{ \xi \in \mathcal{O}_{\nu-1} : \begin{array}{l} |\langle k, \omega_\nu \rangle + \Omega_n^\nu| < \dfrac{\gamma_\nu}{|k|^\tau |n|^2}, \, |n| > K_{\nu+1} \\[2mm] |\langle k, \omega_\nu \rangle + \Omega_m^\nu \pm \Omega_n^\nu| < \dfrac{\gamma_\nu}{|k|^\tau |m|^2 |n|^2}, \, |m| \leq K_{\nu+1}, |n| > K_{\nu+1} \\[2mm] |\langle k, \omega_\nu \rangle + 2\Omega_n| < \dfrac{\gamma_\nu}{|k|^\tau |n|^2}, \, |n| > K_{\nu+1} \end{array} \right\}.$$

It is clear that

$$\mathcal{O} \setminus \tilde{\mathcal{O}} \subseteq \bigcup_{\nu \geq 0} \mathcal{R}^\nu.$$

Lemma 4.3. *For fixed k and ν, there is a constant $C_1 > 0$ such that*

$$\left| \left(\left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 1}_{kmn} \right) \bigcup \mathcal{R}^{\nu 2}_k \bigcup \left(\bigcup_{|n| \leq K_{\nu+1}} \mathcal{R}^{\nu 3}_{kn} \right) \bigcup \left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 4}_{kmn} \right) \bigcup \mathcal{R}^{\nu 5}_k \right) \right| \leq C_1 \frac{\gamma}{|k|^{\tau+1}}.$$

Proof. Let μ is an eigenvalue of the Hermitian $(\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu)$. It is well-known that μ depends on ξ smoothly and there exists a unit eigenvector ψ associated with μ which smoothly depends on ξ (see e.g. [17]). It follows that

$$\mu = \langle (\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu) \psi, \psi \rangle,$$

and hence

$$\partial_\xi \mu = \langle \partial_\xi (\langle k, \omega_\nu \rangle I + \tilde{A}^\nu \otimes I - I \otimes \tilde{A}^\nu) \psi, \psi \rangle.$$

Thus

$$|\partial_\xi \mu| \geq |\langle \partial_\xi (\langle k, \omega_0 \rangle + \Omega_m^0 - \Omega_n^0) \psi, \psi \rangle| - \varepsilon_0 |k| = O(|k|).$$

The cases for $\mathcal{R}^{\nu 2}_k, \mathcal{R}^{\nu 3}_{kn}, \mathcal{R}^{\nu 4}_{kmn}, \mathcal{R}^{\nu 5}_k$ can be handled in an entirely analogous way. Thus

$$\left| \left(\left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 1}_{kmn} \right) \bigcup \mathcal{R}^{\nu 2}_k \bigcup \left(\bigcup_{|n| \leq K_{\nu+1}} \mathcal{R}^{\nu 3}_{kn} \right) \bigcup \left(\bigcup_{|m|, |n| \leq K_{\nu+1}} \mathcal{R}^{\nu 4}_{kmn} \right) \bigcup \mathcal{R}^{\nu 5}_k \right) \right|$$

$$\leq C_1 \sum_{m,n} \frac{\gamma}{|k|^{\tau+1} |m|^2 |n|^2} \leq C_1 \frac{\gamma}{|k|^{\tau+1}}.$$

This proves the lemma. $\quad \square$

Lemma 4.4.

$$\left| \mathcal{O} \setminus \tilde{\mathcal{O}} \right| \leq \left| \bigcup_{v \geq 0} \mathcal{R}^v \right| = O(\gamma).$$

Proof. Let $\tau \geq N$. By Lemma 4.3, we have that

$$\left| \mathcal{O} \setminus \tilde{\mathcal{O}} \right| \leq \left| \bigcup_{v \geq 0} \mathcal{R}^v \right| = O\left(\sum_{v \geq 0} \sum_{|k| > K_{v+1}} \frac{\gamma}{|k|^{\tau+1}} \right) = O\left(\sum_{v \geq 0} \frac{\gamma}{K_{v+1}} \right) = O(\gamma). \quad \square$$

The measure estimate is now complete.

Acknowledgements

Appendix

Lemma A.1.

$$\|FG\|_{D(r,s),\mathcal{O}} \leq \|F\|_{D(r,s),\mathcal{O}} \|G\|_{D(r,s),\mathcal{O}}.$$

Proof. Since

$$(FG)_{kl\alpha\beta} = \sum_{k',l',\alpha',\beta'} F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'},$$

one has that

$$
\begin{aligned}
\|FG\|_{D(r,s),\mathcal{O}} &= \sup_{\substack{\|w\|<s \\ \|\bar{w}\|<s}} \sum_{k,l,\alpha,\beta} |(FG)_{kl\alpha\beta}| s^{2l} |w^\alpha| |\bar{w}^\beta| e^{|k|r} \\
&\leq \sup_{\substack{\|w\|<s \\ \|\bar{w}\|<s}} \sum_{k,l,\alpha,\beta} \sum_{k',l',\alpha',\beta'} |F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}| s^{2l} |w^\alpha| |\bar{w}^\beta| e^{|k|r} \\
&\leq \|F\|_{D(r,s),\mathcal{O}} \|G\|_{D(r,s),\mathcal{O}}. \quad \square
\end{aligned}
$$

Lemma A.2 (*Generalized Cauchy Inequalities*).

$$\|F_\theta\|_{D(r-\sigma,s)} \leq \frac{1}{\sigma} \|F\|_{D(r,s)},$$

$$\|F_I\|_{D(r,\frac{1}{2}s)} \leq \frac{4}{s^2} \|F\|_{D(r,s)},$$

$$\|F_w\|_{D(r,\frac{1}{2}s)} \leq \frac{2}{s} \|F\|_{D(r,s)},$$

$$\|F_{\bar{w}}\|_{D(r,\frac{1}{2}s)} \leq \frac{2}{s} \|F\|_{D(r,s)}.$$

Proof. See [38]. $\quad \square$

Let $\{\cdot, \cdot\}$ denote the Poisson bracket, i.e., for smooth functions F, G,

$$\{F, G\} = \left\langle \frac{\partial F}{\partial I}, \frac{\partial G}{\partial \theta} \right\rangle - \left\langle \frac{\partial F}{\partial \theta}, \frac{\partial G}{\partial I} \right\rangle + i \sum_n \left(\frac{\partial F}{\partial w_n} \frac{\partial G}{\partial \bar{w}_n} - \frac{\partial F}{\partial \bar{w}_n} \frac{\partial G}{\partial w_n} \right).$$

Lemma A.3. *There exists a constant $c > 0$ such that if*

$$\|F_n\|_{D(r,s)} < e^{-|n|}, \qquad \|G\|_{D(r,s)} < \varepsilon,$$

then

$$\|\{F_n, G\}\|_{D(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-2} \|F_n\|_{D(r,s)} \|G\|_{D(r,s)} \leq c\sigma^{-1}s^{-2}\varepsilon e^{-|n|}.$$

Proof. By Lemmas A.1 and A.2,

$$\|\langle F_{n_I}, G_\theta \rangle\|_{D(r-\sigma, \frac{1}{2}s)} < 4\sigma^{-1}s^{-2}\|F_n\| \cdot \|G\|,$$

$$\|\langle F_{n_\theta}, G_I \rangle\|_{D(r-\sigma, \frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|F_n\| \cdot \|G\|,$$

$$\left\|\sum_m F_{n_{w_m}} G_{\bar{w}_m}\right\|_{D(r,\frac{1}{2}s)} \leq \sum_m \|F_{n_{w_m}}\|_{D(r,\frac{1}{2}s)} \|G_{\bar{w}_m}\|_{D(r,\frac{1}{2}s)}$$
$$\leq \|F_{n_w}\|_{D(r,\frac{1}{2}s)} \|G_{\bar{w}}\|_{D(r,\frac{1}{2}s)}$$
$$\leq 4s^{-2}\|F_n\| \cdot \|G\|,$$

$$\left\|\sum_m F_{n_{\bar{w}_m}} G_{w_m}\right\|_{D(r,\frac{1}{2}s)} \leq \sum_m \|F_{n_{\bar{w}_m}}\|_{D(r,\frac{1}{2}s)} \|G_{w_m}\|_{D(r,\frac{1}{2}s)}$$
$$\leq \|F_{n_{\bar{w}}}\|_{D(r,\frac{1}{2}s)} \|G_w\|_{D(r,\frac{1}{2}s)}$$
$$\leq 4s^{-2}\|F_n\| \cdot \|G\|.$$

It follows that

$$\|\{F_n, G\}\|_{D(r-\sigma, \frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|F_n\|_{D(r,s)}\|G\|_{D(r,s)} \leq c\sigma^{-1}s^{-2}\varepsilon e^{-|n|}. \quad \square$$

Lemma A.4. *There exists a constant $c > 0$ such that if*

$$\|X_F\|_{D(r,s)} < \varepsilon', \qquad \|X_G\|_{D(r,s)} < \varepsilon'',$$

for some $\varepsilon', \varepsilon'' > 0$, then

$$\|X_{\{F,G\}}\|_{D(r-\sigma,\eta s)} < c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon'',$$

for any $0 < \sigma < r$ and $0 < \eta \ll 1$. In particular, if $\eta \sim \varepsilon^{\frac{1}{4}}, \varepsilon' \sim \varepsilon, \varepsilon'' \sim \varepsilon^{\frac{3}{4}}$, then

$$\|X_{\{F,G\}}\|_{D(r-\sigma,\eta s)} \sim \varepsilon^{\frac{5}{4}}.$$

Proof. See [23]. $\quad \square$

References

[1] M.J. Ablowitz, D.J. Kaup, A.C. Newell, H. Segur, Method for solving the sine-Gordon equation, Phys. Rev. Lett. 30 (1973) 1262–1264.
[2] M.J. Ablowitz, D.J. Kaup, A.C. Newell, H. Segur, Nonlinear-evolution equations of physical significance, Phys. Rev. Lett. 31 (1973) 125–127.
[3] S. Aubry, Anti-integrability in dynamical and variational problems, Physica D 86 (1995) 284–296.
[4] S. Aubry, Breathers in nonlinear lattices: Existence, linear stability and quantization, Physica D 103 (1997) 201–250.
[5] S. Aubry, Discrete breathers: Localization and transfer of energy in discrete Hamiltonian nonlinear systems, Physica D 216 (2006) 1–30.
[6] D. Bambusi, Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators, Nonlinearity 9 (1996) 433–457.
[7] D. Bambusi, D. Vella, Quasi periodic breathers in Hamiltonian lattices with symmetries, Discrete Cont. Dyn. Sys. Ser. B 2 (2002) 389–399.
[8] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. Math. 148 (1998) 363–439.
[9] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.
[10] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, Int. Math. Res. Notices (1994) 475–497.
[11] J. Bourgain, Green's function estimates for lattice Schrödinger operators and applications, in: Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005.
[12] J. Bourgain, Nonlinear Schrödinger equations, in: Park City Series, vol. 5, American Mathematical Society, Providence, RI, 1999.
[13] O.M. Braun, Y.S. Kivshar, Nonlinear dynamics of the Frenkel–Kontorova model, Phys. Rep. 306 (1998) 1–108.
[14] D.K. Campbell, M. Peyrard, Chaos and order in non-integrable model field theories, in: Chaos (Woods Hole, MA, 1989), Amer. Inst. Phys, New York, 1990, pp. 305–334.
[15] A. Carpio, L.L. Bonilla, Oscillatory wave fronts in chains of coupled nonlinear oscillators, Phys. Rev. E 67 (2003) 621–656.
[16] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Comm. Math. Phys. 211 (2000) 498–525.
[17] W. Craig, C.E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure. Appl. Math. 46 (1993) 1409–1498.
[18] T. Dauxois, M. Peyrard, C.R. Willis, Localized breather-like solution in a discrete Klein–Gordon model and application to DNA, Physica D 57 (1992) 267–282.
[19] T. Dauxois, M. Peyrard, Dynamics and thermodynamics of a nonlinear model for DNA denaturation, Phys. Rev. E 47 (1993) 684–685.
[20] L.H. Eliasson, S.B. Kuksin, KAM for the non-linear Schrödinger equation, Ann. Math. (in press).
[21] S. Flach, C.R. Willis, Discrete breathers, Phys. Rep. 295 (1998) 181–264.
[22] J. Fröhlich, T. Spencer, C.E. Wayne, Localization in disordered nonlinear dynamical systems, J. Stat. Phys. 42 (1986) 247–274.
[23] J. Geng, J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Differentail Equations 209 (2005) 1–56.
[24] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2006) 343–372.
[25] J. Geng, J. You, KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity 19 (2006) 2405–2423.

[26] J. Geng, Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Differential Equations 233 (2007) 512–542.
[27] J. Geng, Y. Yi, A KAM theorem for Hamiltonian networks with long ranged couplings, Nonlinearity 20 (2007) 1313–1342.
[28] M. Johansson, S. Aubry, Existence and stability of quasi-periodic breathers in the discrete nonlinear Schrödinger equation, Nonlinearity 10 (1997) 1151–1178.
[29] O. Kresse, L. Truskinovsky, Mobility of lattice defects: Discrete and continuum approaches, J. Mech. Phys. Solids 51 (2003) 1305–1332.
[30] S.B. Kuksin, Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum, Funct. Anal. Appl. 21 (1987) 192–205.
[31] S.B. Kuksin, Nearly integrable infinite dimensional Hamiltonian systems, in: Lecture Notes in Mathematics, vol. 1556, Springer, Berlin, 1993.
[32] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, Ann. Math. 143 (1996) 149–179.
[33] Y. Li, Y. Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, Trans. Amer. Math. Soc. 357 (4) (2004) 1565–1600.
[34] R.S. Mackay, S. Aubry, Proof of existence of breathers of time-reversible or Hamiltonian networks of weakly coupled oscillators, Nonlinearity 7 (1994) 1623–1643.
[35] P. Maniadis, S. Flach, Nonlinear response and discrete breather excitation in driven micro-mechanical cantilever arrays. arXiv:cond-math/0504298, v1, 12 April 2005.
[36] J.J. Mazo, T.P. Orlando, Discrete breathers in Josephson arrays, Chaos 13 (2) (2003) 733–743.
[37] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comment. Math. Helv. 71 (1996) 269–296.
[38] J. Pöschel, A KAM Theorem for some nonlinear partial differential equations, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 23 (1996) 119–148.
[39] J. Pöschel, Small divisors with spatial structure in infinite dimensional Hamiltonian systems, Comm. Math. Phys. 127 (1990) 351–393.
[40] E.W. Prohofsky, K.C. Lu, L.L. Van Zandt, B.F. Putnam, Breathing models and induced resonant melting of the double helix, Phys. Lett. A 70 (1979) 492–494.
[41] T. Rösler, J.B. Page, Intrinsic localized modes in driven anharmonic lattices with realistic potentials, Phys. Lett. A 204 (1995) 418–426.
[42] M. Sato, B.E. Hubbard, A.J. Sievers, Observation of locked intrinsic localized vibrational modes in a micromechanical oscillator array, Phys. Rev. Lett. 90 (4) (2003).
[43] M. Sato, B.E. Hubbard, A.J. Sievers, B. Ilic, H.G. Craighead, Optical manipulation of intrinsic localized vibrational energy in cantilever arrays, Europhys. Lett. http://arxiv.org/abs/nlin/0403031.
[44] M. Sato, A.J. Sievers, Direct observation of the discrete character of intrinsic localized modes in an antiferromagnet, Lett. Nat. 432 (2004) 486–487.
[45] B.I. Swanson, J.A. Brozik, S.P. Love, G.F. Strouse, A.P. Shreve, Observation of intrinsically localized modes in a discrete low-dimensional material, Phys. Rev. Lett. 82 (16) (1999).
[46] S. Takeno, K. Kisoda, A.J. Sievers, Intrinsic localized vibrations modes in anaharmonic crystals, Prog. Theoret. Phys. Suppl. 94 (1988) 242–269.
[47] E. Trías, J.J. Mazo, A. Brinkman, T.P. Orlando, Discrete breathers in Josephson ladders, Physica D 156 (2001) 98–138.
[48] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Comm. Math. Phys. 127 (1990) 479–528.
[49] J. You, Perturbation of lower dimensional tori for Hamiltonian systems, J. Differential Equations 152 (1999) 1–29.
[50] X. Yuan, Construction of quasi-periodic breathers via KAM technique, Comm. Math. Phys. 226 (2002) 61–100.
[51] C.T. Zhang, Soliton excitations in deoxyribonucleic acid (DNA) double helices, Phys. Rev. A 35 (1987) 886–891.