THE STABILITY OF NONLINEAR SCHRÖDINGER EQUATIONS WITH A POTENTIAL IN HIGH SOBOLEV NORMS REVISITED

MYEONGJU CHAE
Department of Applied Mathematics, Hankyong National University,
327 Joongang-ro, Ansong, 17579, Republic of Korea

SOONSIK KWON *
Department of Mathematical Sciences,
Korea Advanced Institute of Science and Technology,
291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea

(Communicated by Feng Zhou)

ABSTRACT. We consider the nonlinear Schrödinger equations with a potential on $\mathbb{T}^d$. For almost all potentials, we show the almost global stability in very high Sobolev norms. We apply an iteration of the Birkhoff normal form, as in the formulation introduced by Bourgain [4]. This result reproves a dynamical consequence of the infinite dimensional Birkhoff normal form theorem by Bambusi and Grebert [2].

1. Introduction. We consider the nonlinear Schrödinger equation with a potential $V$

$$
\begin{cases}
    iu_t = -\Delta u + V * u + 2|u|^2u & x \in \mathbb{T}^d, t \in \mathbb{R}, \\
    u(0, x) = u_0(x) & \in H^s(\mathbb{T}^d)
\end{cases}
$$

where $V$ is a smooth convolution potential. The system (1.1) is an infinite dimensional Hamiltonian system associated with the Hamiltonian functional $H(u, \bar{u}) = \int_{\mathbb{T}^d} |\nabla u|^2 + (V * u)\bar{u} + |u|^4 \, dx$.

The aim of this work is to use a Hamiltonian dynamical method for studying the long time stability of small solutions in $H^s(\mathbb{T}^d)$ for a sufficiently large $s$. In this paper we consider $V$ a random potential $V(x) = \sum_{n \in \mathbb{Z}^d} v_n e^{inx}, \quad v_n = R(1 + |n|)^{-m}\sigma_n,$

$\{\sigma_n\}_{n \in \mathbb{Z}^d}$ being a sequence of i.i.d. random variables uniformly distributed over $[-\frac{1}{2}, \frac{1}{2}]$. In usual Schrödinger equations, the potential $V(x)$ is multiplicative type, but here we choose the convolution type potential so that it makes the formulation

2000 Mathematics Subject Classification. Primary: 35Q55, 37K55.
Key words and phrases. Birkhoff normal form, nonlinear Schrödinger equations, almost global stability.

The first author is supported by NRF(Korea) grant 2011-0028951. The second author is supported by NRF(Korea) grant 2015R1D1A1A01058832.

* Corresponding author.
simpler in Fourier variables. We assume that the potential $V$ is an even function to ensure its Fourier coefficient $v_n \in \mathbb{R}$. We define the measure space for potential:

$$W = \{ V = \sum_{n \in \mathbb{Z}^d} v_n e^{i n \cdot x} | \sigma_n := R^{-1} v_n (1 + |n|)^m \in [-\frac{1}{2}, \frac{1}{2}] \},$$

(1.2)

for some $m \in \mathbb{Z}_+$. We endow $W$ with the product probability measure, that is, if $N$ is a finite subset of $\mathbb{Z}^d$,

$$\mathbb{P}(\{ W = \sum_{n \in \mathbb{Z}^d} v_n e^{i n \cdot x} \in W : \sigma_n \in (a_n, b_n) \subset [-1/2, 1/2] \text{ for } n \in N, \text{ otherwise } \sigma_n = 0 \}) = \prod_{n \in N} (b_n - a_n).$$

As we handle small solutions, we rescale $u(t, x)$ by

$$u(t, x) = \varepsilon q(\varepsilon^2 t, x) = \varepsilon \sum q_n(\varepsilon^2 t) e^{i n \cdot x}.$$  

(1.3)

We consider for initial data $q_0(x) = q(0, x)$ of $O(\varepsilon)$ in $H^s(\mathbb{T}^d)$, and so $\|u_0\|_{H^s} \leq O(\varepsilon^2)$. Then, the equation (1.1) becomes

$$i q_t = -\varepsilon^{-2} \Delta q + \varepsilon^{-2} V * q + |q|^2 q$$

with Hamiltonian

$$H(q, \bar{q}) = \int_{\mathbb{T}^d} \varepsilon^{-2} |\nabla q|^2 + \varepsilon^{-2} (V * q)\bar{q} + |q|^4 dx.$$ 

One can rewrite the infinite dimensional Hamiltonian system in Fourier variables $\{ q_n \}$:

$$i q_n = \frac{\partial H}{\partial \bar{q}_n},$$

(1.4)

where

$$H(q, \bar{q}) = \sum \left( \frac{|n|^2}{\varepsilon^2} + \frac{v_n}{\varepsilon^2} \right) |q_n|^2 + \sum_{n_1-n_2+n_3-n_4=0} q_{n_1} \bar{q}_{n_2} q_{n_3} \bar{q}_{n_4}. $$

(1.5)

We state our main theorem.

**Theorem 1.1.** There exists a subset $\mathcal{V} \subset W$ of full measure, such that for a given $V \in \mathcal{V}$ the following holds: for a given $B > 0$ there exist $C$, $s$, and $\varepsilon(B) > 0$ such that if $\|u(0)\|_{H^s} \leq \varepsilon^2$, the solution $u$ to the Cauchy problem

$$i u_t = -\Delta u + V * u + |u|^2 u, \quad x \in \mathbb{T}^d, t \in \mathbb{R}$$

will satisfy

$$\sup_{|t| < T} \|u(t)\|_{H^s} < 2\varepsilon^2$$

where $T$ can be as large as $\varepsilon^{-B}$.

We remark that the theorem holds true for $|u|^{2p} u$ for any $p \geq 1$. The analysis is similar, thus for simplicity we present $p = 1$ case in this paper.

Theorem 1.1 is a version of **Birkhoff normal form theorems** for the infinite dimensional Hamiltonian system: the Hamiltonian flow associated to a Hamiltonian remains close to the initial state during an arbitrarily long polynomial time ($\varepsilon^{-B}$), if the initial state has a sufficiently small by $\varepsilon$ in $H^s(\mathbb{T}^d)$. Some of instructive expositions for the **Birkhoff normal form** theory can be found in [1, 11, 17]. There are other stability results on (1.1). Eliasson-Kuksin [9] have shown the KAM theorem...
with analytic potential \( V \). Recently, Faou-Grebert [10] showed a Nekhoroshev type theorem (exponential stability) in analytic function spaces. This exponential stability is obtained when one consider analytic data and so the result can be regarded as a limiting case of long time stability in Sobolev norms. It would be also interesting whether it is possible to obtain an exponential stability if one use our argument in the analytic function space setting.

Theorem 1.1 reproves a dynamical consequence of infinite dimensional Birkhoff normal form theorem by Bambusi-Grebert [2]. In [2], the authors construct an abstract Birkhoff normal theorem to infinite dimensional Hamiltonian systems and apply it to PDEs with tame modulus. [2] is systematic and applicable to a wide range of PDE examples.

In this work, we revisit the problem with a more direct approach to the equation. In performing the Birkhoff normal form, we would like to track how the Hamiltonians are changed. Using a sequence of frequency cut-off we obtain a concrete information on the final Hamiltonian to exhibit the stability result. In fact, we are inspired by Bourgain [4], and this work follows a similar line of [4].

In [4] Bourgain consider one dimensional Schrödinger equations with random initial data,

\[
iu_t = -u_{xx} + |u|^2u + \lambda |u|^p u, \quad x \in \mathbb{T}, \quad p > 1,
\]

When \( \lambda = 0 \) (1.6) has been known to be integrable. In [12] the authors proves that the global Birkhoff coordinate exists in this regard. In [4] Bourgain proves that for given \( B > 0 \), the solutions are stable up to time \( \varepsilon^{-B} \) for almost all initial data smaller than \( \varepsilon \) in a high sobolev space (with probability one). The work of [4] does not rely on integrability, however the presence of cubic term \( u^3 \) is essential. As like [2], [4] uses Birkhoff normal form, a nonlinear change of coordinate of symplectic transform to reduce the non resonant terms from the given Hamiltonian. We use the formulation of the sequence of Birkhoff normal form of Bourgain’s to obtain a similar result for (1.1). In use of normal forms, the nonresonancy is inherited from the randomness of initial data for (1.6), whereas the nonresonancy condition for (1.1) is from random potentials. Indeed, the randomness is explicit in the Hamiltonian (1.5) for the (1.1) due to the random potential \( V \). If we define \( \omega_n = |n|^2/\varepsilon^2 + \nu_n/\varepsilon^2 \), the denominator arising in normal form transform is of the form

\[
\Omega(n) = \omega_{n_1} - \omega_{n_2} + \cdots \pm \omega_{n_r}, \quad \mathbf{n} = (n_1, n_2, \ldots, n_r).
\]

As a similar argument to [2] (or [10]), we obtain the lower bound estimate

\[
|\Omega(n)| \geq \gamma R \varepsilon^{-2} C^{-r/m} \mu(n)^{-4r^2} n_m
\]

for most of potentials \( V \).\(^1\) The lower bound here depends on \( \mu(n) \): the third biggest entry among \( |n_j| \)'s where \( n = (n_1, n_2, \ldots, n_r) \), and \( n_m \) denotes the least entry. For (1.6) the randomness is given to the initial data by \( q_n(0) = \varepsilon (1 + |n|)^{-(1 + s)} \sigma_n \). Indeed, rescaling (1.6) by (1.3), one can write the associated Hamiltonian as

\[
H = \sum_{n_1 + n_2 + n_3 = 0} \sum_{n_4} q_{n_1} \bar{q}_{n_2} \bar{q}_{n_3} \bar{q}_{n_4} q_{n_1}^2 + \sum_{n_1 + n_2 + n_3 - n_4 = 0} q_{n_1} \bar{q}_{n_2} \bar{q}_{n_3} \bar{q}_{n_4} q_{n_1}^2 + \sum_{n_1 + n_2 + n_3 - n_4 = 0} q_{n_1} \bar{q}_{n_2} \bar{q}_{n_3} \bar{q}_{n_4} q_{n_1}^2 + 2 \left( \sum |q_{n_1}|^2 \right)^2 - \sum |q_{n_1}|^4 + \sum q_{n_1} \bar{q}_{n_2} \bar{q}_{n_3} \bar{q}_{n_4}.
\]

\(^1\) See a precise probabilistic statement in Lemm 2.1.
The latter equality follows from that all the resonant terms of $q_{n_1}q_{n_2}q_{n_3}q_{n_4}$ are fully resonant on $T$.\(^2\) Note that due to mass conservation law, i.e. \(\{H; \sum |q_n|^2\} = 0\), \(2 \left(\sum |q_n|^2\right)^2\) is merely a constant. If we replace \(|q_n|^2\) by \(J_n = |q_n|^2 - |q_n(0)|^2\), the randomness comes into play in the Hamiltonian:

\[
H = \sum (\frac{n^2}{\varepsilon^2} - 2|q_n(0)|^2) J_n - J_n^2 \sum_{n_1-n_2+n_3-n_4=0}^{n_1^2-n_2^2+n_3^2-n_4^2\neq 0} q_{n_1}q_{n_2}q_{n_3}q_{n_4} + C
\]

with the constant \(C\) depending only on \(q(0)\).\(^3\) In [4] the lower bound estimate of \(\Omega(n)\),

\[
\Omega(n) \geq \varepsilon^2 n_+^{-5r^2} n^{-2s}
\]

holds with large probability, where we denote \(n_+ [n_-]\) is the largest [smallest] entry of \(n = (n_1, n_2, \ldots, n_r)\). Note that the right hand side has also the factor \((n_-)^{-2s}\). Because \(s\) is chosen to be large for the perturbation terms of Hamiltonian to be small enough, the lower bound of (1.9) becomes smaller as increasing \(s\).\(^4\) This small denominator issue can be overcome if coefficients of perturbation terms are appropriately small. In [4] the author performed the normal form transformation to (1.1) inductively to have the series of Hamiltonians and to reach the final one, for which coefficients are small as desired. Once the right induction hypothesis are assumed on the size of coefficients of polynomials in Hamiltonian, the consequential analysis goes straightforward in [4].

In this paper, we apply the technique in [4] to higher dimensional case \(T^d\). As opposed to one-dimensional case, the lower bound of the small denominator (1.7) is involved in the third largest frequency \(\mu(n)\). For 1 dimensional (1.1) as well as (1.6) there is no difference in analysis if \(n_+\) and \(\mu(n)\) are replaced with each other in the lower bound estimates (1.7) or (1.9). It is due to that \(n_1 - n_2 + \cdots = 0\) implies \(\mu(n) \geq (n_+)^{1/2}\) on \(T^1\). But this is no longer true for \(T^d\), \(d > 1\). (See the estimates around (4.58) ). Another aspect of our approach is that we are able to see the regularity of the potential \(V\) with respect of \(B\) and \(s\). Indeed, \(m\) is less than \(O(\frac{r}{s})\) and \(s\) is bigger than \(B^3\).

In past decades, the growth of Sobolev norm in Hamiltonian PDE has been intensively studied. On one direction, people show the upper bound of the possible growth of Sobolev norms. (See [3, 5, 21, 7, 19, 20] etc.) Such results use versions of the energy method, exploiting \(X^{s,b}\) analysis, \(I\)-method, or Birkhoff normal forms. Unlike our result, they can handle large data. On the other direction, authors pursue to show the existence of solutions that undergo growth in Sobolev norms (though in most results the growth is arbitrarily large but finite) in [18, 3, 8, 15, 14]. They carefully investigate the resonance nonlinear interactions that arise spread of energy from low frequency modes to high frequency modes. Among those, in [14] Guardia handles the same model as ours, 2D cubic NLS with a convolution potential. Regardless of the long time stability for most of potentials, he has shown

\(^2\)If \(n = (n_1, n_2, n_3, n_4)\) \(n_1 \in \mathbb{Z}\) satisfies \(n_1 - n_2 + n_3 - n_4 = n_1^2 - n_2^2 + n_3^2 - n_4^2 = 0\), it implies \(\{n_1, n_3\} = \{n_2, n_4\}\).

\(^3\)The precise value of \(C\) is \(\sum (n^2/\varepsilon^2|q_n(0)|^2 + (|q_n(0)|^4)^2 + 2|q(0)|^4)\). We use \(\sum J_n = 0\).

\(^4\)The discrepancy \(\varepsilon^2\) and \(\varepsilon^{-2}\) in the lower bounds is not as critical as at first seen. In fact [4] performs the preliminary normal form transform to (1.8) and reduces it to the working hamiltonian

\[H = \sum \varepsilon^2|q_n|^2 + \sum |q_n|^4 + \varepsilon^2 \sum q_n q_m + O(\varepsilon^3)\]

where \(q_m\) denote monomials of degree \(\geq 6\). Roughly speaking, \(\varepsilon^2\) factor in the above makes the discrepancy balance.
that for each convolution potential there exists initial data for which higher Sobolev norms grow to an arbitrarily large number (still finite) within times slightly longer than polynomials. This is an extension of the seminal work of Colliander, Keel, Staffilani, Takaoka, and Tao [8] and [15] for the same equation without potentials. It is noteworthy to mention that even though the potential destroys the exact resonances, the aforementioned energy transfer phenomenon still holds. It no longer holds when $T^d$ is replaced by $\mathbb{R} \times T^{d-1}$. In this less turbulent case, the exact resonances are decisive to the limit dynamics [13]; if the potential satisfies a non-resonance condition of type (1.7), then the weak turbulence phenomena disappear, while the case $V = 0$ exhibits the weak turbulence [16].

We mention that the abstract theorem in [2] is applied to several other equations than (1.6) to obtain Birkhoff normal form theorems. It might be possible that the inductive use of normal form transform in [4] can be applied to reprove the known results on Birkhoff normal form theorems. We have not pushed in this direction, however the method seems quite robust. To our knowledge the similar use of iterative Birkhoff normal form transforms is found in [22]. In [22] Wang proved a long time Anderson localization for the 1-d lattice nonlinear Schrödinger equation. We also remark that in [6] Cohen, Hairer, and Lubich proved a long time stability result for 1-d nonlinear Klein-Gordon equation via modulated Fourier expansion method without using Birkhoff normal form.

The paper is organised as follows: In Section 2, we state preliminary setting of Hamiltonian systems and Birkhoff normal form as well as the estimate of the denominator. Section 3 includes the main analysis of the Birkhoff normal form. We present the reduction of Hamiltonian and the estimate of their coefficients. In Section 4, we provide the proof of main theorem.

**Notations.** We abuse multi indices notation in bold.

\[ n = (n^1, \ldots, n^d) \in \mathbb{Z}^d, \quad |n|^2 = |n^1|^2 + \cdots + |n^d|^2 \]

\[ m = (m^1, \ldots, m^l) \in \mathbb{Z}^d \times \cdots \times \mathbb{Z}^d \text{ \ l \ times} \]

\[ n = (k^1, \ldots, k_L, p^1, \ldots, p_L) = (k, p) \in \mathbb{Z}^d \times \cdots \times \mathbb{Z}^d \text{ \ 2L \ times} \]

\[ l(n) = k^1 - p^1 + \cdots + k_L - p_L \in \mathbb{Z}^d \]

\[ |m| = l, \quad |n| = 2L \]

\[ I_{m_i} = q_{m_i} \]

\[ I_m = I_{m_1} \cdots I_{m_l}, \quad q_n = q_{k^1} q_{k^2} \cdots q_{k^l} q_{p^1} \cdots q_{p_l} \]

\[ \omega_n = \frac{|n|^2}{\varepsilon^2} + \frac{v_n}{\varepsilon^2} \]

We say $m \in m$ if $m = m_j$ for some $j$. Similarly $n \in m \cap n$ if $n \in m$ and $n \in k$ or $n \in p$. In the above notations, $|n|$ denotes the degree of $q_n$, not the length of the vector-valued index $m$. The juxtaposition of two multi-index is written as $(m, m')$ i.e. $(m, m') = (m_1, m_2, \ldots, m_l, m'_1, \ldots, m'_k)$ when $m = (m_1, \ldots, m_l)$ and
\( m' = (m'_1, \ldots, m'_k) \). Also we denote
\[
\begin{align*}
  n_+ & = \text{the largest entry among } |n_j|'s \\
  n_- & = \text{the least entry among } |n_j|'s \\
  \mu(n) & = \text{the third largest entry among } \{|n_j| : j = 1, \ldots, 2L\}, \\
  \Omega(n) & = \sum_{n \in k} \omega_n - \sum_{n \in p} \omega_n.
\end{align*}
\]

On account of \( \sigma_n = R^{-1}v_n(1 + |n|)^m \), we write
\[
\Omega(n) = \varepsilon^{-2} \left[ \sum_{i=1}^{L} (k_i^2 + R^2(1 + |k_i|)^{-m} \sigma_{k_i}) - \sum_{i=1}^{L} (p_i^2 + R^2(1 + |p_i|^2)^{-m} \sigma_{p_i}) \right]. \tag{1.10}
\]

\( n \) is called **non-resonant** if \((k_1, \ldots, k_L)\) is not equal to \((p_1, \ldots, p_L)\) by a permutation. Sometimes we denote the largest entry of \( n \) by \( n^*_1 \), and the next largest entries by \( |n^*_2| \geq |n^*_3| \ldots \) etc.

2. Preliminaries.

2.1. Symplectic transformations. We briefly review basic definitions and set up of the infinite dimensional Hamiltonian system. Starting from the Hamiltonian formulation of the equations, (1.4) and (1.5), we recall elements of Hamiltonian systems such as, phase spaces, symplectic form, Poisson bracket, and symplectic transformation. For more details we refer to [11], [17].

The phase space \( P \) is defined by
\[
P = l^2_s(\mathbb{C}) \times l^2_s(\mathbb{C}), \quad \text{where } l^2_s(\mathbb{C}) := \{ (q_n) \in \mathbb{C}^{\mathbb{Z}^d} | \sum |n|^{2s}|q_n|^2 < \infty \}.
\]

We identify \( q \in H^s(T^d) \) with \((q_n) \in l^2_s(\mathbb{C})\) by \( q = \sum q_n e^{in \cdot x} \) and call \((q, \bar{q})\) a canonical coordinate of \( P \) and Poisson bracket \( \{ \cdot, \cdot \} \) as follows,
\[
\{ F, G \} = i \sum_n \frac{\partial F}{\partial q_n} \frac{\partial G}{\partial \bar{q}_n} - \frac{\partial F}{\partial \bar{q}_n} \frac{\partial G}{\partial q_n}.
\]

A smooth function \( F \in C(P, \mathbb{C}) \) is called a Hamiltonian. The Hamiltonian vector field associated to \( F \) is defined by
\[
X_F = J \nabla F = \left( -i \frac{\partial F}{\partial \bar{q}} \right)^T,
\]

and the Hamiltonian flow associated to \( F \) by the integral curve \((q(t), \bar{q}(t))\) along \( J \nabla F \) such that \((q(t), \bar{q}(t))\) satisfies the ODE
\[
\frac{d}{dt} (q, \bar{q})^T = X_F(q, \bar{q}).
\]

In terms of coordinate \((q_n, \bar{q}_n)\) it is written
\[
i \bar{q}_n = \frac{\partial F}{\partial \bar{q}_n}, \quad n \in \mathbb{Z}^d. \tag{2.11}
\]
A canonical transformation \( \varphi : \mathcal{P}_s \to \mathcal{P}_s \) that preserves the symplectic structure is called a \textit{symplectic transformation}. That is, \( \varphi \) preserves the Poisson bracket
\[
\{ F \circ \varphi, G \circ \varphi \} = \{ F, G \} \circ \varphi.
\] (2.12)
Moreover, it is deduced that if \((q'_n, \bar{q}'_n)\) is given by \((q_n, \bar{q}_n) = \varphi(q'_n, \bar{q}'_n)\), then we have
\[
\frac{d}{dt} q'_n = \frac{\partial H'}{\partial \bar{q}'_n}, \quad H' = H \circ \varphi.
\] For the Birkhoff normal form reduction, we take the Lie transform generated by a Hamiltonian \( F \). This is a \textit{time 1-shift} of a Hamiltonian flow. For a given Hamiltonian \( F \) let us consider the Hamiltonian flow generated by \( F \), and denote the solution at time 1 by
\[
q_n(0) = q'_n, \quad q_n(1) := q_n.
\] (2.13)
Note that the flow exists up to time 1 in a neighborhood of the origin. The map \( q'_n \to q_n \) is called \textit{time 1-shift} by \( F \) which is denote by \( \Phi^F \). The map \( \Phi^F_t : q_n(0) \to q_n(t) \) is defined in a similar way. It is easy to check that \( \Phi^F_t \) is symplectic where the flow exists.

The associated Hamiltonian to \((q'_n, \bar{q}'_n)\) is \( H \circ \Phi^F \). Then, using (2.11), (2.12) and the chain rule, we have
\[
\frac{d}{dt} (H \circ \Phi^F_t) = \{ H, F \} \circ \Phi^F_t.
\]
Roughly speaking, the transform is equivalent to exponentiating by \( \{ \cdot, F \} \).

We take a formal Taylor series expansion at \( t = 1 \), then we have
\[
H \circ \Phi^F = \sum_{k=0}^{\infty} \frac{1}{k!} \underbrace{\{ \cdots \{ H, F \}, \cdots, F \}}_{k \text{ times}} = H + \{ H, F \} + \frac{1}{2!} \{ \{ H, F \}, F \} + h.o.t.
\] (2.14)
Now we demonstrate how to reduce a lower order polynomials of given Hamiltonian using a time 1-shift. Back to the Hamiltonian
\[
H = \sum_{n|q_n|} \sum_{l(n)=0}^{\infty} \sum_{l(n)=0}^{\infty} a_n q_{n_1} \bar{q}_{n_2} q_{n_3} \bar{q}_{n_4},
\]
we have a shifted Hamiltonian
\[
H \circ \Phi^F = (H_0 + H_1) \circ \Phi^F = H_0 + H_1 + \{ H_0, F \} + \{ H_1, F \} + h.o.t.
\]
by a time 1-shift by \( F \). If we choose
\[
F = \sum_{l(n)=0, \Omega(n)\neq 0} \frac{a_n}{\Omega(n)} q_{n_1} \bar{q}_{n_2} q_{n_3} \bar{q}_{n_4}, \quad \text{where } \Omega(n) = \omega_1 - \omega_2 + \omega_3 - \omega_4,
\] (2.15)
it is straightforward to compute
\[
\{ H_0, F \} = -(H_1 - \sum_{\Omega(n)\neq 0} a_n q_{n_1} \bar{q}_{n_2} q_{n_3} \bar{q}_{n_4}).
\]
We can only reduce ‘non resonant’ monomials with \( \Omega(n) \neq 0 \). Meanwhile, there are abundant resonant monomials in \( H \). This is where the randomness comes to play.
by modulating frequency $\Omega(n)$ so that the denominator is away from zero at a large probability.

2.2. The lower bound of the denominator.

In performing the Birkhoff normal form, we should know that the Hamiltonian has good behaved to the Poisson bracketing. As one see from (2.15), we require a lower bound of the denominators, $\Omega(n)$, to satisfy so called the strongly non resonant condition. The following lemma guarantees the strongly non resonant condition is generically satisfied.

**Lemma 2.1.** Fix $0 < \gamma < 1$ small enough, and $\Omega(n)$, $V$, and $W$ given in (1.2) and (1.10). There exists a set $F_\gamma \subset W$ whose measure is larger than $1 - \gamma$ such that if $V \in F_\gamma$, then

$$|\Omega(n)| \geq \gamma R \frac{1}{m} \langle \mu(n) \rangle^{4r^2} \langle n^- \rangle^{-m}$$

(2.16)

for all non resonant $n = (k_1, \ldots, k_r, p_1, \ldots, p_r)$ and a constant $C = (40/\gamma)^4$.

Our strongly nonresonant condition is controlled by the third largest frequency and the lowest frequency, as well as the regularity parameter $m$ of potential space. In one dimensional NLS (1.6), we have $\mu(n) \lesssim (n_+)^{1/2}$. Thus, the lower bound may be involved in $n_+$. However, in higher dimensional case, this is no longer true. The proof of (2.16) is similar to that in Faou and Grebert [10]. For the convenience of readers, we place it in the Appendix B.

Due to Lemma 2.1, if we set $V = \cup_{\gamma > 0} F_\gamma$, then $P(V) = 1$ and any $V \in V$ satisfies the nonresonant condition (2.16).

2.3. The Poisson brackets.

We use the multi index notation as follows:

$$I_m = I_{m_1} \cdots I_{m_l}, \quad q_n = q_{k_1}q_{k_2} \cdots q_{k_r}q_{p_1} \cdots q_{p_l}, \quad I_1 = I_{i_1} \cdots I_{i_j}.$$  

A straightforward computation shows that

$$\{I_m, I_1\} = 0, \quad \{I_mq_n, I_1\} = I_m\{q_n, I_1\},$$

$$\{c_{mn}I_mq_n, \sum \omega \bar{I}_1\} = \left(\sum_{i=1}^{l} \omega_{k_i} - \sum_{i=1}^{l} \omega_{p_i}\right)c_{mn}I_mq_n,$$  

(2.17)

$$\{c_{mn}I_mq_n, \sum \bar{I}_1^2\} = \left(\sum_{i=1}^{l} I_{k_i} - \sum_{i=1}^{l} I_{p_i}\right)c_{mn}I_mq_n.$$  

(2.18)

We denote the contraction by $I_{m}^\sim = I_m/I_m$ for $m \in m$ and $q_n^\sim = q_n/q_n$ for $n \in k$, or $q_n/q_n$ if $n \in p$. Moreover, when $m = (m_1, \ldots, m_r, m_l)$ and $m_i$ is contracted, we denote the multi index after the contraction by $m^\sim := (\ldots, m_{i-1}, m_{i+1}, \ldots)$. So $I_m^\sim = I_{m^\sim} \text{ etc.}$ We denote the number of $n$ appearing in $m$ by $\sharp n(m)$, then compute that

$$\{q_n, I_1\} = \sum_{n_1 \in k^\sim} \sharp n(1)\sharp n(k_I)^n q_n - \sum_{n_1 \in p^\sim} \sharp n(1)\sharp n(p_I)^n q_n.$$  

For simplicity, we slightly abuse notations, writing

$$\{c_{mn}I_mq_n, a_I I_1^\sim\} = c_{mn}I_mq_n \sum_{n \in n^\sim} a_I I^\sim.$$  

(2.19)

\footnote{We introduce $\sharp n(m)$ for the sake of concreteness only. Mostly we use the upper bound $\sharp n(m) \leq |m|$.}
In the sequel, we will estimate the coefficients $c_{mn}$ for each case. Thus, the equality means that $c_{mn}$ of LHS is replaced by new coefficient $c_{mn}$, (still denoted by $c_{mn}$), with the same upper bound.

\{I_m q_n, I_m' q_n' \} give rise to two types, which are occurred from loss of $I_m$ or a pair of $(q_n, q_n')$. As above, we write

\[
\{I_m q_n, I_m' q_n' \} = \begin{cases}
|n||n'|(|\sum_{n \in m \cap m'} I_{m,n}^2 I_{m'} q_n q_n' \\
|n||n'|(|\sum_{n \in m \cap m'} I_{m,n}^2 I_{m'} q_n q_n' \\
|n||n'|(|\sum_{n \in m \cap m'} q_{n,n'}^2 q_{n,n'}^2 I_{m} I_{m'})
\end{cases}
\]

(2.20)

3. The reduction of Hamiltonians. In this section, we discuss how to iterate the symplectic transformations and show that starting from the Hamiltonian (1.5), we reduce to the final Hamiltonian $H_b$. Then, we show the new Hamiltonian flow associated to $H_b$, still denoted by $\{q_n(t)\}$, remains $\varepsilon$-neighborhood of zero for a long time $T$ with $T \sim \varepsilon^{-B}$ for any given $B$. Define the actions of the phase variables.

\[
I_n^0 = |q_n(0)|^2, \quad I_n(t) = |q_n(t)|^2 = |q_n|^2.
\]

Let $N_a$ and $N_\infty$ be cut-off parameters and we set a large parameter $A(>200B)$. In the middle of the reduction procedure, we have the Hamiltonians of the following form:

\[
H = \sum \omega_n I_n - \sum I_n^2 =: \Sigma_0 + \Sigma_1 + \sum c_m I_m + \sum c_m I_m =: \Sigma_{21} + \Sigma_{22} + \sum_{N_a \leq |n| - \mu(n) < N_\infty} c_{mn} I_m q_n =: \Sigma_3 + \sum_{\mu(n) \geq N_\infty} c_{mn} I_m q_n =: \Sigma_4 + \sum_{A < \deg \leq 2A} c_{mn} I_m q_n =: \Sigma_5 + \sum_{\deg > 2A} c_{mn} I_m q_n =: \Sigma_6 (3.21) + \varepsilon^A \sum c_{mn} I_m q_n =: \Sigma_7.
\]

Here monomials of $\Sigma_2, \Sigma_3, \Sigma_4, \Sigma_5, \Sigma_6, \Sigma_7$ are of degree $\leq A$. Moreover, the degree of $q_n$ in $\Sigma_3$ has at least 4. In $\Sigma_3 \sim \Sigma_7$, $q_n$ is fully nonresonant in the sense that

\[
\{k_1, k_2, \cdots \} \cap \{p_1, p_2, \cdots \} = \emptyset.
\]

In other words, if it were $k_i = p_j$, then the term $q_{k_i}q_{p_j} = |q_{k_i}|^2$ already makes $I_{k_i}$ and is set aside from $q_n$. The decomposition is not unique. For example, $I_m$ can be counted either in $\Sigma_2$ or in one of $\Sigma_5, \Sigma_6, \Sigma_7$.

We set more parameters

\[
s = s_1 + 5\tau, \quad \tau = \frac{10s}{A}, \quad (3.22)
\]

where $5\tau$ is a parameter such that $\sum_{n \in \mathbb{Z}^d} 1/|n|^{5\tau} < \infty$, hence $5\tau > d$. (3.22) is used in the proof of Proposition 3.
We use an induction argument to prove an iteration of the Birkhoff normal forms changes the initial Hamiltonian (1.5) to a final Hamiltonian $H_b$. For this purpose, we impose induction hypotheses to coefficients of Hamiltonians and then we check that the hypotheses are still satisfied after a Birkhoff normal form reduction. We propose induction hypotheses as follows:

$$|c_m| \leq 1 + N_a^{-2s_1} \Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau}$$  \hspace{1cm} (3.23)

$$= 1 + N_a^{-2s_1} I_{(m,N_a)}$$  \hspace{1cm} for $\Sigma_2$

$$|c_{mn}| \leq N_a^{-4s_1} \Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau} \Pi_j(|n_j| \wedge N_a)^{s_1} N_a^{s_\tau}$$  \hspace{1cm} (3.24)

$$= N_a^{-4s_1} I_{(m,N_a)} Q_{(n,N_a)}$$  \hspace{1cm} for $\Sigma_3, \Sigma_5$

$$|c_{m'n}| \leq \Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau} \Pi_j(|n_j| \wedge N_a)^{s_1} N_a^{s_\tau}$$  \hspace{1cm} (3.25)

$$= I_{(m',N_a)} Q_{(n,N_a)}$$  \hspace{1cm} for $\Sigma_4$

$$|c_{mn}| \leq \Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau} \Pi_j(|n_j| \wedge N_a)^{s_1} N_a^{s_\tau}$$  \hspace{1cm} (3.26)

$$= I_{(m,N_a)} Q_{(n,N_a)}$$  \hspace{1cm} for $\Sigma_6$

$$|c_{mn}| \leq \Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau} \Pi_j(|n_j| \wedge N_a)^{s_1} N_a^{s_\tau}$$  \hspace{1cm} (3.27)

$$= I_{(m,N_a)} Q_{(n,N_a)}$$  \hspace{1cm} for $\Sigma_7$.

Here we denote $a \wedge b = \min(a, b)$. Note that there is no smallness hypothesis on $\Sigma_2$. In fact, eventually, $\Sigma_2$ need not to be small as it is fully resonant term. However, the Poisson bracket with $\Sigma_2$ produces other terms $\Sigma_3, \Sigma_7$. Thus, we require the hypothesis on $\Sigma_2$. In $\Sigma_k$, we put $|n_j|$ into the decreasing order, $|n_1| \geq |n_2| \geq |n_3| \geq \ldots$. We denote

$$\Pi_j(|m_j| \wedge N_a)^{2s_1} N_a^{2\tau} := I_{(m,N_a)}$$

$$\Pi_j(|n_j| \wedge N_a)^{s_1} N_a^{s_\tau} := Q_{(n,N_a)}$$

It follows that

$$I_{(m,N_a)} I_{(m',N_a)} = I_{((m,m'),N_a)}, \quad Q_{(n,N_a)} Q_{(n',N_a)} = Q_{((n,n'),N_a)}.$$  \hspace{1cm} (3.28)

One can verify the initial Hamiltonian (1.5) fits into the above description up to the initially given constant. For (1.5), $N_1 = 1$ and $c_{mn} = 1$. Note that

$$\sum_{n_1 - n_2 + n_3 - n_4 = 0} q_{n_1} q_{n_2} q_{n_3} q_{n_4} = 2 \left( \sum I_m \right)^2 - \sum I_m^2 + \sum_{\{n_1,n_3\} \cap \{n_2,n_4\} = \emptyset} q_{n_1} q_{n_2} q_{n_3} q_{n_4}.$$

Then for the initial Hamiltonian (1.5),

$$c_m = 0 \text{ for } \Sigma_2, \Sigma_4, \quad c_{mn} = 1 \text{ for } \Sigma_3, \Sigma_4, \quad c_{m'n} = 0 \text{ for } \Sigma_5, \Sigma_6, \Sigma_7$$

Now we explain on the form of (3.21) and the coefficient bounds (3.23) – (3.27). First of all, $c_{mn}$ is naturally bounded by product form:

$$c_{mn} \leq C \Pi_j(|m_j| \wedge N_a)^{2s_1} \Pi_j(|n_j| \wedge N_a)^{s_1},$$

then the sum $\sum c_{m'n} I_m q_n$ converges due to $|q_{n_j}| \leq \varepsilon |n_j|^{-\sigma}$ for each $j$. To obtain Theorem 1.1, $\sum c_{m'n} I_m q_n$ does not only converge but also is smaller than $\varepsilon^B$. For this purpose, we choose large parameters $A$, and $N_{\infty}$, such that the sum of monomials in $\Sigma_4, \Sigma_5, \Sigma_6$, and $\Sigma_7$ are small. $\Sigma_3$ may contain harmful terms when the cut-off parameter $N_a$ is small. But by iteration, we push $N_a$ to larger number and to obtain the smallness from the factor $N_a^{-4s_1}$. To be consistent with this, we impose
the condition $N_a \leq |n_-|$. Then we have $N_a^{-4s} \prod_{j=1}^4(|n_j| \wedge N_a)^{s_1} N_a^r \geq 1$. Hence, the induction hypothesis (3.24) holds true for (1.5) for any $N_a$.

For a given parameter $N_a < N_{a+1}$, we want to remove harmful nonresonant terms of $N_a \leq |n_-| \leq N_{a+1}$ in $\Sigma_3$ via the Birkhoff normal form transformation. For this purpose we choose Hamiltonian for time 1-shift

$$F = -i \sum_{N_a \leq |n_-| < N_{a+1}} \frac{c_{mn}}{\Omega(n)} I_m q_n,$$

then by (2.17) we have

$$\{F, \sum \omega_n I_n \} = - \sum_{N_a \leq |n_-| < N_{a+1}} c_{mn} I_m q_n.$$

Now, we explain how we proceed normal forms. We set a increasing sequence of parameters,

$$N_1 = 1 < N_2 < \cdots < N_a < N_{a+1} < \cdots < N_b \leq N_{\infty}.$$ 

For a fixed $N_a$, in the middle of procedure, Hamiltonians are of the form (3.21). Then we take the Poisson bracket with $F$, $\{F, \Sigma_k\}$ for each $k = 1, \cdots, 7$, and check the generated polynomials in $\{F, \Sigma_k\}$ can be put into one of $\Sigma'_k$s by showing the corresponding induction hypothesis still holds (see (3.43)). Moreover we show $H \circ \Phi_F$ allows the decomposition (3.21) satisfying the induction hypothesis with respect to $N_a$ (Proposition 1). In this step, $\Sigma_3$ consists of polynomials with a frequency cut-off $N_{a+1}$ or that with an extra $\varepsilon$ multiplied. We iterate the Birkhoff normal forms until all polynomials in $\Sigma_3$ with $N_a \leq |n_-| \leq N_{a+1}$ are put into $\Sigma_7$. Next, we increase the cut-off parameter $N_a$ to $N_{a+1}$ and rearrange the Hamiltonian as in (3.21) with respect to $N_{a+1}$ (Proposition 3). We iterate this procedure until $N_a$ reaches a sufficiently large $N_b$, for which we will have a desired estimates on coefficients.

In the following we show how to obtain $H_{a+1}$ from $H_a$ with details. It will be summarized in Proposition 3. First, we study the sums that $\{F, H_a\}$ generates. In the sequel, $H$ stands for $H_a$, taken off the subscript for notational simplicity.

$\{F, H\}$ gives rise to $\{F, \Sigma_k\}$ for $k = 1, \ldots, 7$ and each case results in several types of sum.

(i) $\{F, \Sigma I^2\}$

$\{F, \Sigma I^2\}$ is only of $\Sigma_3$ type:

$$\{F, \Sigma I^2\} = \sum_{N_a \leq |n_-| < N_{a+1}} \sum_{n \in \Omega(n)} c_{mn} \pi_1(n) I_n I_m q_n = \sum_{n \in \Omega(n)} |n| c_{mn} \pi_1(n) I_n I_m q_n.$$

We write the sum $\sum_{n \in \Omega(n)} c_{mn} \pi_1(n) I_n - \sum_{n \in \Omega(n)} c_{mn} \pi_1(n) I_n$ as $\sum_{n \in \Omega(n)} \pi_1(n) I_n$, and bound them by $\sum_{n \in \Omega(n)} |n| I_n$. We apply the estimate of $\Omega(n)$ in (2.16) with noting that $r(-\text{degree of } q_n) \leq A$, $n_- \leq N_{a+1}$ for the monomial in $F$, and obtain

$$\Omega(n)^{-1} \leq \varepsilon^2 \gamma^{-1} R^{-1} N^A \tau^{N_{a+1}} C^{A/m}.$$  

(3.29)

By (2.16) and (3.23), we estimate

$$|n| c_{mn} \Omega(n) \leq A^{-1} R^{-1} \varepsilon^2 N_{\tau^{N_{a+1}}} C^{A/m} N_a^{-A} I_{(m,N_a)} Q(n,N_a).$$
By a trivial bound $I(m, N_a) \leq I(m, m^*), N_a$ we have
\[
\left| \frac{c_{mn}}{\Omega(n)} \right| \leq \varepsilon N_a^{-4s_1} I(m, n), N_a) Q(n, N_a)
\] (3.30)
under a condition
\[
\varepsilon^{-1} R^{-1} A N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} \leq 1.
\] (3.31)

(ii) $\{F, \Sigma_2\}$
Note that $\{F, \Sigma_2\} = 0$ due to frequency separations. $\{F, \Sigma_2\}$ becomes of $\Sigma_3$ type or $\Sigma_5$ type. Using (2.19)
\[
\{F, \sum q_l I_l \} \leq A^2 \sum_{n \in n^*} \sum_{n l} \frac{c_{mn} q_l}{\Omega(n)} I_m I_l^{-n} q_n.
\]
If $n \in m \cap 1$, then $|(m, 1) \setminus \{n\}| = |m| + |l| - 1$. To obtain the coefficient for $I_m I_l^{-n}$, we make product for $m_j \in (m, l^{-n})$, and denote
\[
I((m, l^{-n}), N_a) = \Pi_{j=1}^{[m]+|l|-1} (|m_j| \wedge N_a) N_a^{2s_1} N_a^{-2r}
\]
i-1) $\Sigma_3$ type
We estimate separately the cases of $|a| \leq 1$ and $|a| \leq N_a^{-2s_1} I(1, N_a)$. If $|a| \leq 1$, we have
\[
A^2 \left| \frac{c_{mn} q_l}{\Omega(n)} \right| \leq A^2 \varepsilon N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} N_a^{-4s_1} I(m, N_a) Q(n, N_a)
\] (3.32)
\[
\leq A^2 \varepsilon^{-1} R^{-1} N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} N_a^{-4s_1} N_a^{-2s_1} N_a^{2s_1} N_a^{-2r} Q(n, N_a)
\]
\[
\leq \varepsilon N_a^{-4s_1} I((m, 1^{-n}), N_a) Q(n, N_a)
\] (3.33)
under a condition
\[
A^2 \varepsilon^{-1} R^{-1} \varepsilon N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} \leq 1.
\] (3.34)
On the other hand, if $|a| \leq N_a^{-2s_1} I(1, N_a)$, we have a factor $(|a| \wedge N_a) N_a^{2s_1} = N_a^{2s_1}$ due to the loss of $I_n$, and
\[
A^2 \left| \frac{c_{mn} q_l}{\Omega(n)} \right| \leq A^2 \varepsilon^{-1} R^{-1} N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} N_a^{-4s_1} N_a^{-2s_1} I(m, N_a) I(1, N_a) Q(n, N_a)
\]
\[
\leq A^2 \varepsilon^{-1} R^{-1} N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} N_a^{-4s_1} N_a^{-2s_1} N_a^{2s_1} N_a^{2r} \times \Pi_{j=1}^{[m]+|l|-1} (|m_j| \wedge N_a) N_a^{2s_1} N_a^{-2r} Q(n, N_a)
\]
\[
\leq \varepsilon N_a^{-4s_1} I((m, 1^{-n}), N_a) Q(n, N_a)
\] (3.35)
under a condition
\[
A^2 \varepsilon^{-1} R^{-1} \varepsilon N_\infty^{4A^2} N_{a+1}^{m^*} C^{A/m} N_a^{2r} \leq 1.
\] (3.36)
Note that we have an extra $\varepsilon$ in the coefficient $c_{(m, l^{-n})n}$ when $\{F, \Sigma_2\}$ results in $\Sigma_3$.
i-2) $\Sigma_5$ type
The estimate and the condition are similar.

(iii) \{F, \Sigma_3\}
It generates one of types of \(\Sigma_2, \Sigma_3, \Sigma_4, \text{or } \Sigma_5\).

\[
\{F, \Sigma_3\} \leq \sum_{m,n,m',n'} A^2 \left( \sum_{m \in n'/m} \frac{c_{mn,m'n'}}{\Omega(n)} I_m^{n,n'} q_m q_{n'} + \sum_{n \in m'/n} \frac{c_{mn,m'n'}}{\Omega(n)} I_m^{n,n'} q_m q_{n'} \right)
\]

\[
= S_1 + S_2 + S_3,
\]

where \(n, n'\) run through \(N_a \leq n_{-1} < N_{a+1}\), \(N_a \leq |n'| < N_\infty\) respectively in the sum \(\sum_{n,n'}\). Let us treat \(S_1\) first and then \(S_3\), for \(S_2\) is treated similarly.

iii-1) \(\Sigma_3\) type of \(S_1\)

\[
A^2 \frac{c_{mn,m'n'}}{\Omega(n)} \leq A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-4s_1} N_a^{-4s_1} I_{(m',N_a)} Q_{(n,N_a)} Q_{(n',N_a)}
\]

\[
\leq A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-4s_1} N_a^{-4s_1} N_{a-2s_1 + 2\tau}
\]

\[
\Pi_{j=1}^{[m']-1} \left(|m_j| \wedge N_aight)^{2s_1} N_a^{-2s_1 + 2\tau}
\]

\[
\Pi_{j=1}^{[m']-1} \left(|m_j| \wedge N_aight)^{s_1} N_a^{-s_1 + 3\tau}
\]

\[
\varepsilon I_{((m,m^{--}),N_a)} Q_{((n,n'),N_a)}
\]

under a condition

\[
A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-2s_1 + 2\tau} \leq 1. \tag{3.37}
\]

iii-2) \(\Sigma_4\) type of \(S_1\)

At least three entries of \((n, n')\) are bigger than \(N_\infty\). We have

\[
A^2 \frac{c_{mn,m'n'}}{\Omega(n)} \leq A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-4s_1} N_a^{-2s_1 + 2\tau}
\]

\[
\times \Pi_{j=1}^{[m']-1} \left(|m_j| \wedge N_aight)^{s_1} N_a^{-s_1 + 3\tau}
\]

\[
\Pi_{j=1}^{[m']-1} \left(|m_j| \wedge N_aight)^{2s_1} N_a^{-2s_1 + 2\tau}
\]

\[
\varepsilon I_{((m,m^{--}),N_a)} Q_{((n,n'),N_a)}
\]

under a condition

\[
\varepsilon A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-3s_1 + 3\tau} \leq 1 \tag{3.38}
\]

iii-3) \(\Sigma_3\) case of \(S_3\)
We have

\[
A^2 \frac{c_{mn,m'n'}}{\Omega(n)} \leq A^2 \gamma^{-1} R^{-1} \varepsilon^2 N_\infty^{4A^2} N_{a+1}^{m} C^{A/m} N_a^{-4s_1} N_a^{-4s_1} N_{a+1}^{2s_1 + 2\tau}
\]

\[
I_{((m,m'),N_a)} \Pi_{j=1}^{[m']-1} \left(|m_j| \wedge N_aight)^{s_1} N_a^{-s_1 + 3\tau}
\]

\[
\leq \varepsilon N_a^{-4s_1} I_{((m,m'),N_a)} Q_{((n,n^{--}),N_a)} \tag{3.39}
\]
We have
\[ \varepsilon \gamma^{-1} R^{-1} A^2 N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-2s} \leq 1. \]

iii-4) Σ₄ case of S₃
We have
\[ A^2 \frac{cmm'c'n'}{\Omega(n)} \leq A^2 \gamma^{-1} R^{-1} \varepsilon N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-8s} I_{(m,N_a)} I_{(m',N_a)} Q_{(n,N_a)} Q_{(n',N_a)} \]
\[ \leq A^2 \gamma^{-1} R^{-1} \varepsilon N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-8s} N_a^{5(s+1)} I_{(m,m'),N_a)} \]
\[ \Pi_{j \geq 1}^{|n|+|n'|-2} (|n_j| \wedge N_a)^{s} N_a^r \]
\[ \leq \varepsilon I_{(m,m'),N_a)} Q_{(n'=n,n'=n)} \]

under
\[ \varepsilon \gamma^{-1} R^{-1} A^2 N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-3s} \leq 1. \]

S₁, S₂, S₃ generate Σ₅ terms. Since it is estimated similarly to Σ₃, we omit the detail here.
We postpone the case where \( \{F, \Sigma_3\} \) generates Σ₂ in the end of the part (iv).

(iv) \{F, Σ₄\}
It gives rise to terms of type Σ₃, Σ₄, or Σ₅.

iv-1) Σ₃ type
It is obtained from reduction of a pair of \((q_n, q_n')\), which is the third case in (2.20).
Let us estimate the coefficient bound of
\[ \sum_{n \in n' \cap n'} A^2 \frac{cmm'c'n'}{\Omega(n)} I_m I_{m'} q_n^\sim q_n'^\sim. \]
We have
\[ A^2 \frac{cmm'c'n'}{\Omega(n)} \]
\[ \leq A^2 \gamma^{-1} R^{-1} \varepsilon N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-4s} N_a^{2(s+1)} I_{(m,m'),N_a)} \]
\[ \times \Pi_{j=1}^{|m|+|m'|} (|m_j| \wedge N_a)^{s} N_a^r \]
\[ \leq A^2 \gamma^{-1} R^{-1} \varepsilon N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-4s} I_{(m,m'),N_a)} \]
\[ \times \Pi_{j=1}^{|m|+|m'|} (|m_j| \wedge N_a)^{s} N_a^r \]
\[ \leq A^2 \gamma^{-1} R^{-1} \varepsilon N_{\infty}^4 N_{a+1}^m C^{A/m} N_a^{-4s} I_{(m,m'),N_a)} \]
\[ \times \Pi_{j=1}^{|m|+|m'|} (|m_j| \wedge N_a)^{s} N_a^r \]
\[ \leq \varepsilon N_a^{-4s} I_{(m,m'),N_a)} Q_{(n=n,n=n)} \]
\[ \leq R^{-1} A^2 N_{\infty}^4 N_{a+1}^m C^{A/m} \leq 1. \]

iv-2) Σ₄ type
We treat the first and second reduction cases in (2.20) and the third one separately. Let us estimate the coefficient bound of

\[
\sum_{n \in \mathfrak{m} \cap \mathfrak{n}'} A^2 \frac{c_{mn'} n'}{\Omega(n)} I_{mn} q_n q_{n'}.
\]

We have

\[
A^2 \frac{c_{mn'} n'}{\Omega(n)} \leq A^2 \gamma - R^{-1} \varepsilon N^A N^m_{a+1} C^{A/m} N^{-4s_1} N_a^{2s_2} I_{mn}^{m+n} q_n q_{n'}.
\]

If all the three biggest index among \(\{ n, n' \} \) arise in \( n' \), we bound the right hand side by

\[
A^2 \gamma - R^{-1} \varepsilon N^A N^m_{a+1} C^{A/m} N^{-4s_1} N_a^{2s_2} I_{mn}^{m+n} q_n q_{n'}.
\]

under \( \varepsilon \gamma - R^{-1} A^2 N^A N^m_{a+1} C^{A/m} N^{-4s_1} N_a^{2s_2} \leq 1 \). Next we estimate the coefficient bound of

\[
\sum_{n \in \mathfrak{m} \cap \mathfrak{n}'} A^2 \frac{c_{mn'} n'}{\Omega(n)} I_{mn} q_n q_{n'}.
\]

A 3) The case that \( \{ F, \Sigma_4 \} \) generate \( \Sigma_5 \) term are estimated as same as \( \Sigma_3 \).

\( \{ F, \Sigma_4 \} \) and \( \{ F, \Sigma_4 \} \) can generate \( \Sigma_2 \) terms when \( \mathfrak{m} = q_{n'} \). In the induction hypotheses (3.23)-(3.27) we see that the coefficient’s bounds for \( \Sigma_3 \) and \( \Sigma_4 \) are assumed to be smaller than for \( \Sigma_2 \); For \( \Sigma_3 \) it is obvious, and for \( \Sigma_4 \) it is from \( I_{mn} q_{(n \cap N)} q_{(n \cap N)} \). So by estimates in (iii) and (iv), we have the coefficients of the generated \( \Sigma_2 \) term bounded by

\[
|c_m| \leq \varepsilon N_a^{-2s_1} I_{mn} q_{(n \cap N)}.
\]
(v) Similarly we have \( \{ F, \Sigma_5 \} = \varepsilon \Sigma_5 + \varepsilon \Sigma_6 \), \( \{ F, \Sigma_6 \} = \varepsilon \Sigma_6 \), and \( \{ F, \Sigma_7 \} = \varepsilon \Sigma_7 \). Let us verify \( \{ F, \Sigma_6 \} = \varepsilon \Sigma_6 \). According to (2.20), we have to show

\[
A|\mathbf{m}| \frac{|c_{mn} c_{m'n'}|}{\Omega(n)} \leq \varepsilon (|\mathbf{m}| + |\mathbf{m}'|)(|\mathbf{n}| + |\mathbf{n}'|) I_{(m,m')}(\mathbf{n},N_a) Q((n,n'),N_a)
\]

when a loss of \( I_n \) occurs, and

\[
A|\mathbf{n}| \frac{|c_{mn} c_{m'n'}|}{\Omega(n)} \leq \varepsilon I_{(m,m'),N_a} Q((n,n',n'^{n-n}),N_a)
\]

when a pair of \((q_n, q_n')\) is contracted. For the former, we have

\[
A|\mathbf{m}| \frac{|c_{mn} c_{m'n'}|}{\Omega(n)} \leq A|\mathbf{m}'| \gamma R^{-1} \varepsilon^{2} N_{\infty}^{4A^2} N_{a+1}^{m} C^{A/m} N_{\alpha}^{2s+2t} I_{(m,N_a)} I_{(m,N_a)} Q(n,N_a) Q(n',N_a)
\]

\[
\leq A|\mathbf{m}'| \gamma R^{-1} \varepsilon^{2} N_{\infty}^{4A^2} N_{a+1}^{m} C^{A/m} N_{\alpha}^{2s+2t} I_{(m,m'^{n-n})} Q((n,n'),N_a)
\]

under the condition \( A \gamma R^{-1} N_{\infty}^{4A^2} N_{a+1}^{m} C^{A/m} N_{\alpha}^{2s+2t} \leq 1 \). The other cases are similar. We omit details.

Overall the conditions for \( A, s, \varepsilon, \{ N_a, N_{\infty} \} \) is reduced to (3.22) and

\[
\varepsilon \gamma R^{-1} N_{\infty} A^2 N_{\infty}^{4A^2} N_{\alpha}^{m} C^{A/m} \leq 1.
\]

We assume that \( A, s, \varepsilon, \{ N_a \} \) satisfy

\[
\tau = \frac{10s}{A}, N_{\infty} = \varepsilon^{-\frac{1}{10s}}, s > A^3, \text{ and } \varepsilon \gamma R^{-1} C^{A/m} \leq 1.
\]

So far, we have proven

\[
\{ F, \Sigma_0 \} = -\Sigma_3 \bigg|_{N_a \leq n < N_{a+1}}
\]

\[
\{ F, \Sigma_1 \} = \varepsilon \Sigma_4,
\]

\[
\{ F, \Sigma_2 \} = \varepsilon \Sigma_4,
\]

\[
\{ F, \Sigma_3 \} = \varepsilon (\Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5),
\]

\[
\{ F, \Sigma_4 \} = \varepsilon (\Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5),
\]

\[
\{ F, \Sigma_5 \} = \varepsilon \Sigma_5 + \varepsilon \Sigma_6,
\]

\[
\{ F, \Sigma_6 \} = \varepsilon \Sigma_6,
\]

\[
\{ F, \Sigma_7 \} = \varepsilon \Sigma_7.
\]

The point is that we have the extra \( \varepsilon \) in front of \( \Sigma_3 \) in \( \Sigma_i \) when \( i \geq 1 \) (The corresponding estimates are (3.30), (3.33), (3.35), (3.39), and (3.40)). Let us define

\[
H_F := H \circ \Phi_F.
\]

Recalling (2.14), the Taylor series expansion formula of \( H \circ \Phi_F \) centered at \( t = 0 \), we obtain

\[
H_F = \sum_{k=0}^{l} \frac{1}{k!} \{ F, H \}^{(k)} + \frac{1}{l!} \int_0^1 (1 - t)^l \{ F, H \}^{l+1} \circ \Phi_F dt.
\]
We denote
\[ \{F, H\}^{(k)} := \{F, \cdots [F, H], \cdots \} \]
\( k \) times
and \( \{F, H\}^{(0)} = H \). Under the initial condition \( \|q(0)\|_{H^s} \leq \varepsilon \) and a consistency condition to be proved in Section 4, the remainder converges, so we simply write
\[ H_F = \sum_{k=0}^{\infty} \frac{1}{k!} \{H, F\}^{(k)}. \]

**Proposition 1.** By the induction argument we have
\[ H_F = \Sigma_0 + \Sigma_1 + (1 + 5\varepsilon)(\Sigma_2 + \Sigma_4 + \Sigma_5 + \Sigma_6 + \Sigma_7) + \sum_{n+1 \leq |n_-| < N_{\infty}} +6\varepsilon\Sigma_3. \]

**Proof.** By the induction argument we have Proposition 1. That \( H + \{F, H\} = \Sigma_0 + \Sigma_1 + \Sigma_3 \mid_{N_{a+1} \leq n_- < N_{\infty}} + 4\varepsilon\Sigma_3 \)
\[ + (1 + 2\varepsilon)(\Sigma_2 + \Sigma_4 + \Sigma_6) + (1 + 3\varepsilon)\Sigma_5 + (1 + \varepsilon)\Sigma_7. \]

For \( k \geq 2 \) we assume the induction hypothesis:
\[ \frac{\{F, H\}^{(k)}}{4^k} = \varepsilon^{k-1} \sum_{2,3,4,5} + \varepsilon^k \sum_{2,3,4,5,6,7} \]
where we denote \( \sum_i + \sum_j \) by \( \sum_{i,j} \) for simplicity. If \( k = 2 \), it is straightforward that
\[ \frac{\{F, \{F, H\}\}}{4^2} = \varepsilon \sum_{2,3,4,5} + \varepsilon^2 \sum_{2,3,4,5,6,7} \]
due to (3.43). Similarly, the induction hypothesis holds for \( k + 1 \)th step;
\[ \frac{\{F, H\}^{(k+1)}}{4^{k+1}} = \frac{1}{3} \{F, \{F, H\}^{(k)}\} = \varepsilon^k \sum_{2,3,4,5} + \varepsilon^{k+1} \sum_{2,3,4,5,6,7}. \]
It holds that
\[ \sum_{k \geq 2} \frac{\{F, H\}^{(k)}}{k!} = \frac{32}{3} \varepsilon \sum_{2,3,4,5} + \varepsilon^2 \sum_{2,3,4,5,6,7} \]
with \( \frac{4^k}{k!} \leq \frac{32}{3} \). The proposition follows by adding (3.45) and (3.46). \( \Box \)

So far we have removed monomials of \( N_a \leq |n_-| < N_{a+1} \) in \( \Sigma_3 \) and obtain an extra \( \varepsilon \) factor in front of \( \Sigma_3 \). We will go on until the increasing exponent is bigger than \( A \) so that we have \( \varepsilon^A\Sigma_3 \), which joins \( \Sigma_6 \). Let us consider the normal form transform of \( H_F \) by the associated Hamiltonian \( \varepsilon F \), where we use the same notation \( F \) to denote
\[ F = \Sigma_3 \bigg|_{N_a \leq |n_-| < N_{a+1}} \frac{c_{mn}}{\Omega(n)} I_{m} q_{n} \]
with \( c_{mn}I_{m} q_{n} \) the monomial of \( \Sigma_3 \) of \( H_F \). Here, \( \Sigma_3 \bigg|_{N_a \leq |n_-| < N_{a+1}} \) means the summation of term with condition \( N_a \leq |n_-| < N_{a+1} \). Similarly to Proposition 1 we compute \( H_F \circ \Phi_{\varepsilon F} \). For notational simplicity, we use Proposition 1 in the form of
\[ H_F = \Sigma_0 + \Sigma_1 + \Sigma_2 + \Sigma_3 \bigg|_{N_{a+1} \leq |n_-| < N_{\infty}} + \varepsilon\Sigma_3 + \Sigma_4 + \Sigma_5 + \Sigma_6 + \Sigma_7. \]
Proposition 2. By the induction argument we have

\[ H_F \circ \Phi_{\varepsilon F} = \Sigma_0 + \Sigma_1 + \Sigma_2 + \Sigma_3 |_{N_{a+1} \leq |n_-| < N_\infty} + \varepsilon^2 \Sigma_4 + \Sigma_5 + \Sigma_6 + \Sigma_7. \]

Proof. First we note that

\[ \{\varepsilon F, \Sigma_0\} = -\varepsilon \Sigma_3 |_{N_a \leq |n_-| < N_{a+1}}. \]

Hence

\[ H_F + \{\varepsilon F, \Sigma_0\} = \sum_{0,1,2} + (1 + \varepsilon) \Sigma_3 |_{N_{a+1} \leq |n_-| < N_\infty} + \sum_{4,5,6,7}. \]

On the other hand, we have

\[ \{\varepsilon F, H_F - \Sigma_0\} = \{\varepsilon F, \sum_{1,2} + (1 + \varepsilon) \Sigma_4 + \sum_{4,5,6,7} \} \]

\[ = \varepsilon (\varepsilon^2 + 3\varepsilon) \Sigma_4 + \varepsilon (\varepsilon^2 + 2\varepsilon) \Sigma_5 + \varepsilon (\varepsilon^2 + 4\varepsilon) \Sigma_6 + \varepsilon^2 \Sigma_7 \]

\[ = \varepsilon^2 \Sigma_3 + \varepsilon^2 \Sigma_4 + \varepsilon^2 \Sigma_5 + \varepsilon^2 \Sigma_6 \]

according to the policy in front of Proposition 1. We have

\[ H_F + \{\varepsilon F, H_F\} = \sum_{0,1,2} + (1 + \varepsilon) \Sigma_3 |_{N_{a+1} \leq |n_-| < N_\infty} + \varepsilon^2 \Sigma_4 + (1 + \varepsilon^2) \sum_{4,5,6} + \Sigma_7 \]

Note that we have

\[ \{\varepsilon F, H_F\} = (1 + \varepsilon^2) \Sigma_3 + \varepsilon^2 \sum_{4,5,6} + \varepsilon \Sigma_7. \]

Assume the induction hypotheses hold for \( k \geq 2 \)

\[ \frac{1}{k!} \{\varepsilon F, H_F\}^{(k)} = \varepsilon^k \sum_{3,4,5,6} + \varepsilon^{k-1} \Sigma_7. \]

The \( k = 2 \) case is established by the same computations as above. Then it is straightforward that the \( k + 1 \)-step holds:

\[ \frac{1}{(k + 1)!} \{\varepsilon F, H_F\}^{(k+1)} = \frac{1}{k + 1} \{\varepsilon F, \frac{1}{k!} \{\varepsilon F, H_F\}^{(k)}\} \]

\[ = \frac{2 \varepsilon^2 (k+1)}{k + 1} (\Sigma_3 + \Sigma_4 + \Sigma_5 + \Sigma_6) + \frac{\varepsilon^{2k}}{k + 1} \Sigma_7. \]

So we have

\[ H_F \circ \Phi_{\varepsilon F} = \sum_{k=0}^{\infty} \frac{1}{k!} \{\varepsilon F, H_F\} \]

\[ = \sum_{0,1,2} + (1 + \varepsilon) \Sigma_3 |_{N_{a+1} \leq |n_-| < N_\infty} + \varepsilon^2 \Sigma_4 + (1 + \varepsilon^2) \sum_{4,5,6} \]

\[ + (\Sigma_{k \geq 2} \varepsilon^{2k}) \sum_{3,4,5,6} + (\Sigma_{k \geq 2} \varepsilon^{2k-1}) \Sigma_7 \]

\[ = \Sigma_0 + \Sigma_1 + \Sigma_2 + \Sigma_3 |_{N_{a+1} \leq |n_-| < N_\infty} + \varepsilon^2 \Sigma_4 + \Sigma_5 + \Sigma_6 + \Sigma_7 \]

as desired. \( \square \)
We can repeat the above procedure all over again by taking the normal form transformation with $\Phi_{\varepsilon F}$. Denote
\[
H_F \circ \Phi_{\varepsilon F} \circ \Phi_{\varepsilon^2 F} \cdots \circ \Phi_{\varepsilon^k F} := H_{F^{(k)}}.
\]
Inductively, we have the following proposition.

**Proposition 3.** If $k > A$ we have
\[
H_{F^{(k)}} = \Sigma_0 + \Sigma_1 + \Sigma_2 + \Sigma_3\bigg|_{N_{a+1} \leq |n_-| < N_\infty} + \Sigma_4 + \Sigma_5 + \Sigma_6 + \Sigma_7.
\]

For such $k$ we denote $H_{F^{(k)}}$ by $H_{a+1}$. The coefficients for $H_{a+1}$ satisfy the induction hypotheses in (3.23) – (3.26) replacing $N_a$ by $N_{a+1}$.

**Proof.** We show only the second assertion. At the time of reaching $k > A$ the coefficients for $H_{a+1}$ remain to be bounded as (3.23) – (3.26). To upgrade $a$ to $a+1$ we check (3.23) – (3.26) separately as follows.

The monomials of $\Sigma_3$ satisfy $|n_+| \geq N_{a+1}$, while the other monomials goes to $\Sigma_2$, for which we impose no condition. So there is at least one $m_j$ of $|m_j| \geq N_{a+1}$, and for this $m_j$ we have
\[
|m_j| \wedge N_a \leq \frac{N_{a+1}}{N_a}|m_j| \wedge N_{a+1},
\]
\[
N_a^{-2s_1} I_{(m,N_a)} \leq N_{a+1}^{-2s_1} \left( \frac{N_{a+1}}{N_a} \right)^{2s_1} I_{(m,N_a)}
\]
\[
\leq N_{a+1}^{-2s_1} \left( \frac{N_{a+1}}{N_a} \right)^{2s_1} \left( \frac{N_{a+1}}{N_a} \right)^{-2s_1} I_{(m,N_{a+1})}.
\]

The monomials of $\Sigma_3$ satisfy $N_{a+1} \leq |n_-|$ and the degree of $q_a$ is at least 4. We have
\[
N_a^{-4s_1} I_{(m,N_a)} Q_{(n,N_a)} \leq N_{a+1}^{-4s_1} \left( \frac{N_{a+1}}{N_a} \right)^{4s_1} \left( \frac{N_{a+1}}{N_a} \right)^{-4s_1} I_{(m,N_{a+1})} Q_{(n,N_{a+1})}.
\]

The hypothesis (3.24) for $\Sigma_4$ is automatically upgraded. The monomials of $\Sigma_5$ are of degree bigger than $A$. We have
\[
N_a^{-4s_1} I_{(m,N_a)} Q_{(n,N_a)} \leq N_{a+1}^{-4s_1} \left( \frac{N_{a+1}}{N_a} \right)^{4s_1} \left( \frac{N_{a+1}}{N_a} \right)^{-4s_1} I_{(m,N_{a+1})} Q_{(n,N_{a+1})}.
\]

Since $\tau = \frac{10}{A}$, the extra coefficient is smaller than 1.

Lastly $\Sigma_6$, $\Sigma_7$ are automatically upgraded. \[\square\]

We set $N_1 = 1$ and $N_2 = N_\infty$, and perform Proposition 1 to Proposition 3. Then $\Sigma_3$ is empty. (The emptiness of $\Sigma_3$ is not crucial for the following analysis). The final Hamiltonian is written as
\[
H_b = \Sigma_0 + \Sigma_1 + \Sigma_2 : \text{ fully resonant terms}
\]
\[+
\sum_{A \leq \text{deg} \leq 2A} a_{mn} I_m q_n + \sum_{\mu(n) > N_\infty} b_{mn} I_m q_n + \sum_{\text{deg} > 2A} c_{mn} I_m q_n
\]
\[+
\varepsilon^A \sum_{\text{deg} < A} d_{mn} I_m q_n,
\]

(3.48)
where

\[ |a_{mn}| \leq N_b^{-4s_1} I_{(m, N_a)} Q_{(n, N_b)}, \]  
\[ |b_{mn}| \leq I_{(m, N_b)} Q_{(n, N_b)}, \]  
\[ |c_{mn}| \leq I_{(m, N_b)} Q_{(n, N_b)}, \]  
\[ |d_{mn}| \leq I_{(m, N_b)} Q_{(n, N_b)}. \]  

(3.49)

(3.50)

(3.51)

(3.52)

The bound (3.49) corresponds to the sums, Σ4, Σ5, (3.50) to Σ3, (3.51) to the sum Σ6, and (3.52) to Σ7.

4. Proof of Theorem 1.1.

4.1. Estimates on the symplectic transformations.

Before proceeding to prove Theorem 1.1, we mention that the stability condition is preserved under the symplectic transforms. Indeed, the new Hamiltonian flow \( \tilde{q}_n \) is obtained from a time-1 shift of the evolution

\[ \dot{q}_n = \frac{\partial F}{\partial \bar{q}_n}, \quad q_n(0) = q_n, \quad q_n(1) = \tilde{q}_n. \]

Using the definition of \( F \) we estimate that \( q_n \sim \tilde{q}_n \) as follows:

\[
\sum |n|^2 |q_n(1) - q_n(0)|^2 \leq \sum |n|^2 \int_0^1 |\text{Im} \tilde{q}_n \frac{\partial F}{\partial \bar{q}_n}| \, dt \\
\leq \varepsilon^2 N_\infty^{A} N_{a+1}^{m} C^{A/m} \sum \sum |c_{mn}| |I_m||q_n| \\
\leq \varepsilon^2 N_\infty^{A} N_{a+1}^{m} C^{A/m} \\
\sum \sum |n|^2 |n_1^s| |n_2^s| |q_n||q_{n_1}| |I_{(m, N_a)}| I_{m_j} |\Pi_{j \geq 3}(|n_j| \wedge N_a)^{s_1} N_\infty^s |q_{n_j}| \\
\leq \varepsilon^2 N_\infty^{A} N_{a+1}^{m} C^{A/m} A^{s_2} \sum r^{s} \sum |n|^2 |\Pi_{j \geq 3}(|n_j|)^{-s} (-1)^{s_1} \varepsilon |n_j| \\
\leq \varepsilon^2 N_\infty^{A} N_{a+1}^{m} C^{A/m} A^{s_2} \sum r^{s} \varepsilon |n_j| \\
\leq \varepsilon^2 N_\infty^{A} N_{a+1}^{m} C^{A/m} A^{s_2} \varepsilon^2 \leq \varepsilon^2
\]

under a condition

\[ A^s \varepsilon \leq 1, \]  
\[ A^s \varepsilon \leq 1, \]  

(4.53)

and (3.41).

4.2. Proof of Theorem 1.1.

Fix \( A = 200B \). Assume that all parameters satisfy size conditions given so far. (See Appendix A for summary.) As we have only \( N_b = N_2 = N_\infty \), we denote it by \( N \).

We prove that the Hamiltonian flow \( \{q_n(t)\} \) generated by \( H_b \) remains

\[ \|q(t)\|_{H^s} \leq 2\varepsilon \quad \text{for} \quad t \leq \varepsilon^{-B}. \]

The flow \( \{q_n(t)\} \) is given by

\[ iq_n = \frac{\partial H_b}{\partial \bar{q}_n}, \]
hence

\[ \left| \frac{d}{dt} |q_n|^2 \right| \leq \sum_n (|a_{mn}| + |b_{mn}| + |c_{mn}| + |d_{mn}|) \left| \operatorname{Im} \left( \frac{\partial (I_m q_n)}{\partial q_n} \right) q_n \right|, \quad (4.54) \]

where \( a_{mn}, b_{mn}, c_{mn} \) are bounded by (3.49) – (3.52), and \( c_{mn} = 0 \) if the degree of \( I_m q_n \) is less than \( 2A \). The consistency assumption for \( \{q_n\} \) is

\[ \|q\|_{H^s} \leq \varepsilon. \quad (4.55) \]

By (4.54), we have

\[ \sum_n |n|^{2s} |q_n(t)|^2 - |q_n(0)|^2 | \leq \int_0^T \sum_{m,n} |n|^{2s} |a_{mn}| + |b_{mn}| + |c_{mn}| + |d_{mn}| |q_n| \left| \frac{\partial (I_m q_n)}{\partial q_n} \right|, \quad (4.56) \]

We want to bound the integrand of (4.56) by \( \varepsilon B \). Note that the sums are taken over \((m, n)\) with \( l(n) = 0 \). If \( n = (n_1, \ldots, n_l) \), we bound \( |n_1|^l \leq l|n_2|^s \) from the relation \( n_1 - n_2 + n_3 \cdots = 0 \). Moreover, we use the estimates for \( |n|^{2s} \),

\[ |n| \leq A : \ |n|^{2s} \leq A^s |n_1|^s |n_2|^s, \]

\[ |n| = l > A : \ |n|^{2s} \leq l^s |n_1|^s |n_2|^s \leq A^s N^{ls/A} |n_1|^s |n_2|^s = A^s N^{ls/10} |n_1|^s |n_2|^s, \]

using \( l/A < N^{l/A} \). Then, we estimate

\[ \sum_n \sum_{m,n} |n|^{2s} |a_{mn}| |I_m| |q_n| \]

\[ \leq N^{-4s} A^s \sum_{m,n} |n_1|^s |n_2|^s I_{(m, n_0)} \Pi_j (|n_j| \land N)^{s_1} N_b^{s_1} N_j^{11/10} |I_m| |q_n| \]

\[ \leq N^{-4s} A^s N^{2(s_1 + 2\tau)} \sum_{m,n} |n_1|^s |n_2|^s |q_{n_1}| |q_{n_2}| |I_m| |N_j^{2\tau} I_m| \]

\[ \times \Pi_{j \geq 3} (|n_j| \land N)^{s_1} N_{2\tau}^r |q_{n_j}| \]

\[ \leq N^{-s_1} A^s \|q\|_{H^s}^2 \left( \sum_{r \geq 2} r \varepsilon^r \right)^2 \]

\[ \leq N^{-s_1} A^s \varepsilon \leq \varepsilon A/100 \]

by (4.53). Here we used \( |q_{n_j}| \leq \varepsilon |n_j|^{-s} \) and \( N^{2\tau} \sim \varepsilon^{-1/4} \) to bound

\[ (|m_j| \land N)^{2s_1} N^{2\tau} I_{m_j} \leq \varepsilon^{1000} |m_j|^{-2(s_1 - s)} = \varepsilon^{1000} |m_j|^{-20s/10}, \quad (4.57) \]

and in turn we estimate

\[ \sum_m \Pi_j (|m_j| \land N)^{2s_1} N^{2\tau} I_{m_j} \leq \sum_{r \geq 2} \varepsilon \sum_{|m| = r} \Pi_j^r |m_j|^{-20s/10} \leq \sum_{r \geq 2} \varepsilon. \]
In the sum $\Sigma_4$, the degree of $q_n$ is less than $A$ and $\mu(n) > N$. We have
\[
\sum_{m,n} |n|^{2s} |l_{mn}| |l_m| |q_n| \\
\leq A^s \sum_{m,n} |n|^{2s} |l_{mn}| |l_m| |q_n| |\Pi_j(|m_j| \wedge N)^{2s_1} N^{2s \tau} I_{m_j} \Pi_j(|n_j| \wedge N)^{s_1} N^\tau |q_{n_j}| \\
\leq A^s N^{-s} \|q\|^2_{H^s} \langle \tau \rangle \leq \varepsilon^{A/200}.
\]
(4.58)

Now we estimate the $c_{mn}$ part. The needed $\varepsilon^{A/100}$ factor is obviously from that $\Sigma_6$ consists of $I_m q_n$ with the degree bigger than $2A$. Let us assume $|m| > A/2$. The other case $|n| > A$ can be treated similarly.
\[
\sum_{n} \sum_{m,n} |n|^{2s} |c_{mn}| |l_m| |q_n| \\
\leq \sum_{n} \sum_{m,n} |n|^{2s} |m| |l_{n,m,n}| \Pi_j(|n_j| \wedge N)^{s_1} N^\tau |l_m| |q_n| \\
\leq \sum_{m,n} |n|^{2s} |m| \Pi_j(|m_j| \wedge N)^{s_1} N^{2s \tau} I_{m_j} \\
\Pi_j \geq 3(|n_j| \wedge N)^{s_1} N^{2s \tau} |q_{n_j}| \\
\leq \|q\|^2_{H^s} \sum_{|m| > A/2} |m| |l_{m,N_0}| \sum_{|n| \geq 4} |n| \Pi_j |l_{n,m}|^{-20s/A} \sum_{r \geq 4} r \Pi_j = 3 \sum_{r \geq 4} r \sum_{n_j} \varepsilon^{\frac{10r}{100}} |n_j|^{-20s/A} \\
\leq \varepsilon^2 \sum_{r \geq A/2} \sum_{r \geq A} \varepsilon^{\frac{10r}{100}} \sum_{r \geq 4} \varepsilon^{\frac{10r}{100}} \leq \varepsilon^{A/100}
\]

where we used the consistent assumption (4.55) and the relation (4.57). The contribution of $|d_{mn}|$ can be similarly bounded due to $\varepsilon^A < N^{-4s_1}$. \hfill \Box

**Appendix A. Summary of parameters.**

We introduce many parameters in the analysis. Here, for reader’s convenience, we summarize the size relation of parameters. In the Theorem 1.1, parameter $B > 0$ is given. Then we consecutively choose $A, s, \tau, N_\infty$ to satisfy the following:

\[
A > 200B, \quad s > A^3 \\
\begin{aligned}
&\ \quad s = s_1 + 5\tau, \quad \tau = \frac{10s}{A}, \\
&N_\infty = \varepsilon^{-\frac{10s}{10\tau}},
\end{aligned}
\]

and further we choose $m, s, A$ are so that

\[
m \leq \frac{s}{A}.
\]

Also we choose $\varepsilon = \varepsilon(C, A, s)$ such that

\[
\varepsilon A^5 \leq 1, \quad \varepsilon^{\frac{1}{2}} \gamma^{-1} R^{-1} C^{A/m} \leq 1,
\]
Choosing \( A = 200B, \ s = A^3, \ m = A^2 \), we have \( A^2 \ll \tau = 50A^2 \ll s_1 \). In fact, the conditions imposed in Section 3 are reduced to

\[ \varepsilon \gamma^{-1} R^{-1} A^2 N_\infty^m C^{A/m} \leq 1, \]

and it can easily be verified from the above choices. Note that in this work, we need only \( N_1 \) and \( N_\infty \).

**Appendix B. Proof of Lemma 2.1.**

Lemma 2.1 is the direct consequence of Proposition 4. We closely follow the argument in [10]. The difference here is that the exponent of \( \langle n_+ \rangle \) involves only \( r \). In the original version in [10] it comes with \( \langle n_+ \rangle^{-4rm} \) in Proposition 4.

**Lemma B.1.** Fix \( \gamma \) and \( m > d/2 \). There exists a set \( F'_\gamma \subset \mathcal{W} \) whose measure is larger than \( 1 - \gamma \) such that if \( V \in F'_\gamma \) then for any \( r \)

\[ \left| \Omega(n) - b \right| \geq \gamma \frac{\varepsilon_1}{m} \langle n_+ \rangle^{-3(r+1)} \langle n_- \rangle^{-m} \]

for any non resonant \( n = (n_1, \ldots, n_{2r}) \) and for any \( b \in \mathbb{Z} \).

Consider the case in which we have two independent random variables \( x, y \), uniformly distributed over \([-M, M]\). Define a set \( A \subset \mathbb{R}^2 \) by \( I = \{(x, y)||x-y+c| \leq \eta\} \) for \( c \in \mathbb{R} \) then

\[ \mathbb{P}[(x, y) \in A] = |A| = \int_{-M}^{M} \int_{-\min{|y+c|, \eta}}^{\max{|y+c|, \eta}} \mathcal{d}x \mathcal{d}y \leq 4M\eta. \]

For \( A \) not to be empty, \( c \) satisfies \( |c| \leq 2M + \eta \). Similarly, let \( x_1, \ldots, x_n \) be independent random variables, which are uniformly distributed over \([-M, M]\), and

\[ A = \{|a_1x_1 + a_2x_2 + \cdots + a_nx_n + c| \leq \eta\}. \]

Assuming \( |a_n| = \max\{|a_1|, \ldots, |a_n|\} \), we have

\[ \mathbb{P}[(x_1, \ldots, x_n) \in B] \leq \int_{-M}^{M} \int_{-M-M}^{M-M} \ldots \int_{-M-M}^{M-M} \chi_{-M,M} dx_1 \ldots dx_n \]

\[ \leq (2M)^n \frac{2\eta}{|a_n|}. \]

For \( A \) not to be empty, \( d \) satisfies \( |d| \leq M(|a_1| + \cdots |a_n|) + \eta \).

**Proof of Lemma B.1.** Let \( n = (n_1, \ldots, n_{2r}) = (k, p) \) and \( b \in \mathbb{Z} \) be given and \( \eta(n) \) be chosen later. By (2.59) we have

\[ \mathbb{P}_{n,b} \left\{(\sigma_k, \sigma_p) \in [-1/2, 1/2]^{2r} : \left| \sum_{i=1}^{r} |k_i|^2 + \frac{\varepsilon_1 \sigma_k}{\langle k_i \rangle^m} - \sum_{i=1}^{r} |p_i|^2 + \frac{\varepsilon_1 \sigma_p}{\langle p_i \rangle^m} - b \right| \leq \eta \right\} \]

\[ \leq 2\eta \varepsilon_1^{-1} \langle n_- \rangle^{-m}. \]

Since \( \mathbb{P}_{n,b} \left[ \ldots \right] = 0 \) for \( |b| \leq \frac{1}{2} \left( \sum_{i=1}^{2r} |n_i|^2 + \frac{\varepsilon_1}{(\langle n_i \rangle^m)} \right) + \eta \), summing over \( |b| \leq 2r(|n_+|^2 + \varepsilon_1(n_-)^{-m}) \), we have

\[ \mathbb{P}_{n} \left[ \ldots \right] \leq 4\eta(n)(2\varepsilon_1^{-1} |n_+|^2 (n_-)^m + 1). \]

Therefore for any non resonant \( n = (n_1, \ldots, n_{2r}) \) and for any \( b \in \mathbb{Z} \), we have

\[ \mathbb{P} \left[ V = \{w \in \mathcal{W} : |\Omega(n) - b| \leq \eta(n) \} \right] \leq 10\varepsilon_1^{-1} \sum_{n=(n_1, \ldots, n_{2r})} \eta(n)(r|n_+|^2 (n_-)^m). \]
The choice of \( \eta(n) \leq \gamma \frac{\varepsilon_1}{10} (n_+)^{-2(r+1)}(n_-)^{-m} \) gives
\[
P[V] \leq \gamma \sum_{n=(n_1,\ldots,n_{2r})} \langle n_+ \rangle^{-2r} \leq \sum_{n=(n_1,\ldots,n_{2r})} \Pi_{l=1}^{2r} \langle n_l \rangle^{-2} \leq \gamma.
\]
Using \( \langle n_+ \rangle^r \leq r \), we set \( \eta = \gamma \frac{\varepsilon_1}{10} (n_+)^{-3(r+1)}(n_-)^{-m} \) and conclude the lemma. \( \square \)

**Proposition 4.** Fix \( \gamma \) and \( m > d/2 \). There exists a set \( F'_\gamma \subset W \) whose measure is larger than \( 1 - \gamma \) such that if \( V \in F'_\gamma \) then for any \( r \)
\[
|\Omega(n) + \lambda_1 w_{l_1} + \lambda_2 w_{l_2}| \geq \frac{\varepsilon_1}{10} \gamma (\gamma/40)^{\frac{m}{r}} \langle n_+ \rangle^{-4r^2} \langle n_- \rangle^{-m}
\]
for any \( n = (n_1,\ldots,n_{2r}) \), for any \( l_1,l_2 \in \mathbb{Z} \), and for any \( \lambda_1, \lambda_2 \in \{0,1,-1\} \) such that \((n,l_1,l_2)\) is non resonant.

**Proof.** Firstly, if \( \lambda_1 = \lambda_2 = 0 \), it holds trivially due to Lemma B.1.
Secondly, \( \lambda_2 = 0, \lambda_1 = \pm 1 \) : we note that
\[
|\Omega(n)| \leq 3r|n_+|^2 := (*) .
\]
If \( |l_1|^2 \geq 2(*) \), we have
\[
|\Omega(n) + \lambda_1 w_{l_1}| > (*).
\]
If \( |l_1|^2 \leq 2(*) \), we apply Lemma B.1 to \( n' = (n,l_1) \) to have
\[
|\Omega(n) + \lambda_1 w_{l_1}| \geq \gamma \frac{\varepsilon_1}{10} \langle \sqrt{6rn_+} \rangle^{-3(r+2)} \langle n_- \rangle^{-m} \geq \gamma \frac{\varepsilon_1}{10} \langle n_+ \rangle^{-2} \langle n_- \rangle^{-m} ,
\]
(2.60)
by using \( r \leq \langle n_+ \rangle^r \).
The case when \( \lambda_1, \lambda_2 \) has the same sign is treated similarly.
It remains to consider the form \( |\Omega(n) + \lambda_1 w_{l_1} - w_{l_2}| \). We assume \( |l_1| \leq |l_2| \) wlog and further
\[
|l_2|^2 - |l_1|^2 \leq 3(*)
\]
(2.61)
because \( |\Omega(n)| < (*) \) combining \( |l_2|^2 - |l_1|^2 > 3(*) \) leads to \( |\Omega(n) + w_{l_1} - w_{l_2}| < 2(*) \).
By the triangle inequality it holds that
\[
|\Omega(n) + w_{l_1} - w_{l_2}| \geq |\Omega(n) + |l_1|^2 - |l_2|^2| - |w_{l_1} - w_{l_2} - (|l_1|^2 - |l_2|^2)| |
\]
Note that
\[
|w_{l_1} - w_{l_2} - l_1^2 + l_2^2| \leq \frac{\varepsilon_1 \sigma_{l_1}}{|l_1|^m} - \frac{\varepsilon_1 \sigma_{l_2}}{|l_2|^m} \leq \frac{2\varepsilon_1}{|l_1|^m}.
\]
Since \( |l_1|^2 - |l_2|^2 \in \mathbb{Z} \), by Lemma B.1 we have
\[
|\Omega(n) + |l_1|^2 - |l_2|^2| \geq \gamma \frac{\varepsilon_1}{10} \langle n_- \rangle^{-3} \langle n_- \rangle^{-1} := (**).
\]
So if \( \frac{2\varepsilon_1}{|l_1|^m} \leq \frac{(**)}{2} \), it leads to \( |\Omega(n) + w_{l_1} - w_{l_2}| \geq \frac{(**)}{2} \).
Consider the last case \( \frac{2\varepsilon_1}{|l_1|^m} > \frac{(**)}{2} \) under (2.61), that is
\[
|l_1| < (40/\gamma)^{\frac{3}{m}} \langle n_+ \rangle^{3} \langle n_- \rangle^{-m}, \quad |l_2| < |l_1| + 2 \sqrt{(*)}.
\]
Applying Lemma B.1 to \( \max(|n_+||l_2|) \leq \min(|n_+|, |40/\gamma|^{\frac{m}{3}} \langle n_+ \rangle^{\frac{3(m+3)}{m}} \langle n_- \rangle^{3} + 2|n_+|) \) we conclude
\[
|\Omega(n) + w_{l_1} - w_{l_2}| \geq \frac{\varepsilon_1}{10} \gamma (\gamma/40)^{\frac{3(m+3)}{m}} \langle n_+ \rangle^{\frac{3(m+3)}{m}} \langle n_- \rangle^{-3} \langle n_- \rangle^{-(r+1)}.
\]
(2.62)
Comparing the lower bounds (2.60) and (2.62) we have the proposition. \( \square \)
Acknowledgements. A part of this work was carried out while the first author was visiting University of Wisconsin at Madison. We are grateful to Tadahiro Oh for suggesting references.

REFERENCES

[1] D. Bambusi, Birkhoff normal form for some nonlinear PDEs, Comm. Math. Phys., 234 (2003), 253–285.
[2] D. Bambusi and B. Grebert, Birkhoff normal form for PDE’s with tame modulus, Duke Math. J., 135 (2006), 507–567.
[3] J. Bourgain, On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE, Intern. Math. Res. Notices, 6 (1996), 277–304.
[4] J. Bourgain, On diffusion in High-dimensional Hamiltonian systems and PDE, J. d’analyse Math., 80 (2000), 1–34.
[5] J. Bourgain, Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations, Ergodic Theory Dynam. Syst., 24 (2004), 1331–1357.
[6] D. Cohen, E. Hairer and C. Lubich Long-time analysis of nonlinearity perturbed wave equations via modulated Fourier expansions, Arch. Rational Mech. Anal., 187 (2008), 341–368.
[7] J. Colliander, S. Kwon and T. Oh, A remark on normal forms and the “upside-down” 1 - method for periodic NLS: growth of higher Sobolev norms, J. d’analyse Math., 118 (2012), 55–82.
[8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation, Invent. Math., 181 (2010), 39–113.
[9] L. H. Eliasson and S. B. Kuksin, KAM for nonlinear Schrödinger equation, Annals of Math., 172 (2010), 371–435.
[10] E. Faou, and B. Grebert, A Nekhoroshev type theorem for the nonlinear Schrödinger equation on the d-dimensional torus, Analysis and PDE, 6 (2013), 1243–1262.
[11] B. Grebert, Birkhoff normal form and Hamiltonian PDEs, Partial Differential Equations and Applications, 1–46, Sémin. Congr., 15, Soc. Math. France, Paris, 2007.
[12] B. Grebert, T. Kappeler and J. Pöschel, Normal form theory for the NLS equatons, preprint, arXiv:0907.3938.
[13] B. Grebert, E. Paturel and L. Thomann, Modifed scattering for the cubic Schrödinger equation on product spaces: the non resonance case, preprint, arXiv:1502.07699.
[14] M. Guardia, Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential, Comm. Math. Phys., 329 (2014), 405–434.
[15] M. Guardia, and V. Kaloshin, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation, J. Eur. Math. Soc., 17 (2015), 71–149.
[16] Z. Hani, B. Pausader, N. Tzvetkov and N. Visciglia, Modified scattering for the cubic Schrödinger equation on product spaces and applications, Forum of mathematics, Pi, 3 (2015), e4, 63pages.
[17] T. Kappeler and J. Pöschel, KdV and KAM, A Series of Modern Surveys in Mathematics, Springer, 2003.
[18] S. B. Kuksin, Oscillations in space-periodic nonlinear Schrodinger equations, Geom. Funct. Anal., 7 (1997), 338–363.
[19] V. Sohinger, Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on R, Indiana Univ. Math. J., 60 (2011), 1487–1516.
[20] V. Sohinger, Bounds on the growth of high Sobolev norms of solutions to 2D Hartree equations, Discret. Contin. Dyn. Syst., 32 (2012), 3733–3771.
[21] G. Staffilani, On the growth of high sobolev norms of solutions for kdv and Schrödinger equations, Duke Math. J., 86 (1997), 109–142.
[22] W.-M. Wang, Long time anderson localization for the nonlinear random Schrödinger equation, J. Stat. Physics, 134 (2009), 953–968.

Received February 2015; revised November 2015.

E-mail address: mchae@hknu.ac.kr
E-mail address: soonsikk@kaist.edu