Dynamics of PDE, Vol.16, No.4, 325-357, 2019

Long time behavior of the NLS-Szegő equation

Ruoci Sun

Communicated by Thomas Kappeler, received April 25, 2019.

Abstract. We are interested in the influence of filtering the positive Fourier modes to the integrable non linear Schrödinger equation. Equivalently, we want to study the effect of dispersion added to the cubic Szegő equation, leading to the NLS-Szegő equation on the circle $\mathbb{S}^1$

$$i\partial_t u + \epsilon^\alpha \partial_x^2 u = \Pi(|u|^2 u), \qquad 0 < \epsilon < 1, \qquad \alpha \geq 0.$$

There are two sets of results in this paper. The first result concerns the long time Sobolev estimates for small data. The second set of results concerns the orbital stability of plane wave solutions. Some instability results are also obtained, leading to the wave turbulence phenomenon.

Contents

1. Introduction 325
2. The cubic Szegő equation 330
3. Long time behavior for small data 331
4. Orbital stability of the traveling plane wave $\mathbf{e}_m$ 340
5. Comparison to the NLS equation 352
6. Appendix 352
References 356

1. Introduction

We consider the NLS-Szegő equation defined on the circle $\mathbb{S}^1$

$$i\partial_t u + \partial_x^2 u = \Pi(|u|^2 u), \qquad u(0,\cdot) = u_0. \tag{1.1}$$

1991 *Mathematics Subject Classification.* Primary 37, 76, 92, 70; Secondary 34, 35, 82, 80.
Key words and phrases. Cubic Schrödinger equation, Szegő projector, small dispersion, stability, wave turbulence, Birkhoff normal form.
The author would like to express his gratitude towards Patrick Gérard for his deep insight, generous advice and continuous encouragement. He also would like to thank Jean-Marc Delort, Benoit Grébert, Sandrine Grellier and Thomas Kappeler for useful discussions.

Here $\Pi : L^2(\mathbb{S}^1) \to L^2(\mathbb{S}^1)$ denotes the orthogonal projector from $L^2(\mathbb{S}^1)$ onto the space of L^2 boundary values of holomorphic functions on the unit disc,

$$\Pi : \sum_{k\in\mathbb{Z}} u_k e^{ikx} \longmapsto \sum_{k\geq 0} u_k e^{ikx}.$$

We denote by $L^2_+ := \Pi(L^2(\mathbb{S}^1)) \subset L^2(\mathbb{S}^1)$, $H^s_+ := H^s(\mathbb{S}^1) \bigcap L^2_+$, for all $s \geq 0$, and $C^\infty_+ := C^\infty(\mathbb{S}^1) \bigcap L^2_+$.

1.1. Motivation. The NLS-Szegő equation can be seen as the combination of two completely integrable systems: the defocusing cubic Schrödinger equation

$$i\partial_t u + \partial_x^2 u = |u|^2 u, \qquad (t,x) \in \mathbb{R} \times \mathbb{S}^1, \tag{1.2}$$

and the cubic Szegő equation

$$i\partial_t V = \Pi(|V|^2 V), \qquad (t,x) \in \mathbb{R} \times \mathbb{S}^1. \tag{1.3}$$

They have both a Lax Pair structure and the action-angle coordinates, which can be used to obtain their explicit formulas with the inversed spectral method(see Zakharov–Shabat [28], Faddeev–Takhtajan [7], Grébert–Kappeler [19], Gérard [11], for the NLS equation and Gérard–Grellier [12, 14, 16, 17] for the cubic Szegő equation). However, these two Lax pairs cannot be combined in order to give a Lax pair for (1.1). Moreover, the long time behaviors of these two equations are totally different.

The NLS equation (1.2) has a sequence of conservation laws controlling every Sobolev norms(see Faddeev–Takhtajan [7], Grébert–Kappeler [19], Gérard [11]), so all the solutions are uniformly bounded in every H^s space. Moreover, Grébert and Kappeler [19] have proved the existence of the global Birkhoff coordinates for the NLS equation. So the solutions of (1.2) are actually almost periodic on $\mathbb{R}$ valued into $H^s(\mathbb{S}^1)$.

Compared to (1.2), the cubic Szegő equation, which stands for a non-dispersive model, has both the Lax pair structure and the wave turbulence phenomenon. Its long time behavior is extremely sensible according to the different initial data. P.Gérard and S.Grellier have shown that(in [15, 16, 17]) for a G_δ dense subset of initial data in C^∞_+, the solutions may blow up in H^s, for every $s > \frac{1}{2}$ with super–polynomial growth on some sequence of times, while they go back to their initial data on another sequence of times tending to infinity. For another dense subset of initial data in C^∞_+, the solutions are quasi-periodic. (see also Theorem 2.3).

Remark 1.1. Consider the following equation without the Szegő projector Π on $\mathbb{S}^1$:

$$\begin{cases} i\partial_t V = |V|^2 V, \\ V(0,\cdot) = V_0. \end{cases} \tag{1.4}$$

Then $V(t,x) = e^{it|V_0|^2} V_0(x)$ and we have $\|V(t)\|_{H^s} \simeq |t|^s$, for all $s \geq 0$, if $|V_0|$ is not a constant function. Hence, the Szegő projector both accelerates the energy transfer to high frequencies, and facilitates the transition to low frequencies for (1.4).

One wonders about whether filtering the positive Fourier modes can change the long time Sobolev estimates of the cubic defocusing Schrödinger equation. So we introduce equation (1.1). On the other hand, it can also be obtained from the cubic Szegő equation by adding the dispersive term ∂_x^2 to its linear part. In order to see the gradual change of the dispersion, we add the parameter ϵ^α in front of the Laplacian ∂_x^2 to get a more general model, the NLS-Szegő equation (with small dispersion):

$$(1.5) \qquad i\partial_t u + \epsilon^\alpha \partial_x^2 u = \Pi(|u|^2 u), \qquad u(0,\cdot) = u_0, \qquad 0 < \epsilon < 1, \qquad \alpha \geq 0.$$

Equation (1.1) is the special case $\alpha = 0$ for (1.5).

We endow L^2_+ with the canonical symplectic form $\omega(u,v) = \mathrm{Im} \int_{\mathbb{S}^1} \frac{u\bar{v}}{2\pi}$. Equation (1.5) has the Hamiltonian formalism with the energy functional

$$(1.6) \qquad E^{\alpha,\epsilon}(u) = \frac{\epsilon^\alpha}{2} \|\partial_x u\|^2_{L^2} + \frac{1}{4}\|u\|^4_{L^4}, \qquad u \in H^1_+.$$

Besides $E^{\alpha,\epsilon}$, equation (1.5) has two other conservation laws,

$$\begin{cases} Q(u) = \|u\|^2_{L^2}, \\ I(u) = \mathrm{Im} \int_{\mathbb{S}^1} \overline{u}\partial_x u = \|u\|^2_{\dot{H}^{\frac{1}{2}}}, \end{cases}$$

which give the estimate of the solution for low frequencies:

$$\sup_{t\in\mathbb{R}} \|u(t)\|_{H^s} \leq \|u_0\|^{1-2s}_{L^2} \|u_0\|^{2s}_{H^{\frac{1}{2}}}, \qquad \forall s \in [0, \frac{1}{2}].$$

Proceeding as in the case of equation (1.2), one can prove the global existence and uniqueness of the solution of the NLS-Szegő equation in high frequency Sobolev spaces, by using the Brezis–Gallouët type estimate [4], the Aubin–Lions–Simon theorem (see Theorem II.5.16 in Boyer–Fabrie [3]) and the Trudinger type inequality (see Yudovich [27], Vladimirov [26], Ogawa [24] and Gérard–Grellier [12]). Its well-posedness problem in low frequency Sobolev spaces can be dealt with Strichartz's inequality introduced in Bourgain [2]. Only the high frequency Sobolev estimates are considered in this paper.

Proposition 1.2. For every $s \geq \frac{1}{2}$, given $u_0 \in H^s_+$, there exists a unique solution $u \in C(\mathbb{R}, H^s_+)$ of (1.5) such that $u(0) = u_0$. For every $T > 0$, the mapping $u_0 \in H^s_+ \mapsto u \in C([-T,T], H^s_+)$ is continuous.

1.2. Main results. The first result concerns the long time stability around the null solution of the NLS-Szegő equation (1.5). If the initial data u_0 is bounded by ϵ, we look for a time interval I^α_ϵ, in which the solution $u(t)$ is still bounded by $\mathcal{O}(\epsilon)$. Now we state the first result of this paper.

Theorem 1.3. *For every $s > \frac{1}{2}$, there exist two constants $a_s \in (0,1)$ and $K_s > 0$ such that for all $0 < \epsilon \ll 1$ and $u_0 \in H^s_+$, if $\|u_0\|_{H^s} = \epsilon$ and u denotes the solution of* (1.5) *with $u(0) = u_0$, then*

$$(1.7) \qquad \begin{cases} \sup_{|t| \leq \frac{a_s}{\epsilon^{4-\alpha}}} \|u(t)\|_{H^s} \leq K_s \epsilon, & \text{if} \quad \alpha \in [0,2]; \\ \sup_{|t| \leq \frac{a_s}{\epsilon^2}} \|u(t)\|_{H^s} \leq K_s \epsilon, & \text{if} \quad \alpha > 2. \end{cases}$$

Moreover, the time interval $I^\alpha_\epsilon = [-\frac{a_s}{\epsilon^2}, \frac{a_s}{\epsilon^2}]$ is maximal for the case $\alpha > 2$ and $s \geq 1$ in the following sense: for every $0 < \epsilon \ll 1$, there exists $u_0^\epsilon \in C^\infty_+$ such that

$\|u_0^\epsilon\|_{H^s} \simeq \epsilon$ *and for every* $\beta > 0$*, we have*

$$\sup_{|t|\leq \frac{1}{\epsilon^{2+\beta}}} \|u(t)\|_{H^s} \gtrsim \epsilon |\ln \epsilon|^{\frac{1}{2}} \gg \epsilon, \qquad u(0) = u_0^\epsilon.$$

REMARK 1.4. In the case $\alpha \in [0, 2)$, the proof is based on the Birkhoff normal form method, similarly to Bambusi [1], Grébert [18], Gérard–Grellier [13] and Faou–Gauckler–Lubich [8] for instance. However, the time interval $[-\frac{a_s}{\epsilon^{4-\alpha}}, \frac{a_s}{\epsilon^{4-\alpha}}]$ may not be optimal. The resonant term of 6 indices in the homological equation can not be cancelled by the Birkhoff normal form transform.(see subsubsection 3.2.4)

The second set of results concerns the long time H^s-estimates for the solutions of (1.5), if its initial datum is a perturbation of the plane wave $\mathbf{e}_m : x \mapsto e^{imx}$, for some $m \in \mathbb{N}$ and $s \geq 1$. Let $u = u(t, x)$ be the solution of equation (1.5) such that $\|u(0) - \mathbf{e}_m\|_{H^s} = \epsilon$. Its energy functional (1.6) gives the following estimate:

$$\sup_{t\in\mathbb{R}} \|u(t)\|_{H^1} \lesssim_{\|u_0\|_{H^1}} \epsilon^{-\frac{\alpha}{2}}, \qquad \forall 0 < \epsilon < 1, \quad \alpha \geq 0. \tag{1.8}$$

However, no information on the stability of the plane waves $\mathbf{e}_m$ is obtained from (1.8) during the process $\epsilon \to 0^+$. Consider the super-polynomial growth of Sobolev norms in the cubic Szegő equation case (see Gérard–Grellier [15, 17] and Proposition 2.4 in this paper), the occurence of wave turbulence phenomenon for (1.5) depends on the level of its dispersion. We begin with three long time stability results for the polynomial dispersion $\epsilon^\alpha \partial_x^2$ case with $0 \leq \alpha \leq 2$. The following theorem indicates H^1-orbital stability of the traveling waves $\mathbf{e}_m$ for equation (1.5).

THEOREM 1.5. *For all* $\epsilon \in (0, 1)$*,* $\alpha \in [0, 2]$ *and* $m \in \mathbb{N}$*, there exists* $\mathcal{C}_m > 0$ *such that if* $\|u(0) - \mathbf{e}_m\|_{H^1} = \epsilon$*, then we have*

$$\sup_{t\in\mathbb{R}} \inf_{\theta\in\mathbb{R}} \|u(t) - e^{i\theta}\mathbf{e}_m\|_{H^1} \leq \mathcal{C}_m \epsilon^{1-\frac{\alpha}{2}}.$$

For each $t \in \mathbb{R}$, the infimum can be attained when $\theta = \arg u_m(t)$. A similar result is established by Zhidkov [29, Sect. 3.3] and Gallay–Haragus [9, 10] for the 1D cubic Schrödinger equation. In small dispersion case, Theorem 1.5 gives a significant improvement of estimate (1.8). We denote by $S_{\alpha,\epsilon}$ the non linear evolution group defined by (1.5) on $H^{\frac{1}{2}}_+$. In other words, for every $\phi \in H^{\frac{1}{2}}_+$, $t \mapsto S_{\alpha,\epsilon}(t)\phi$ is the solution $u \in C(\mathbb{R}, H^{\frac{1}{2}}_+)$ of equation (1.5) such that $u(0) = \phi$.

COROLLARY 1.6. For every $m \in \mathbb{N}$, we have

$$\sup_{\substack{0<\epsilon<1\\ 0\leq\alpha\leq 2}} \sup_{\|\phi-\mathbf{e}_m\|_{H^1}\leq\epsilon} \sup_{t\in\mathbb{R}} \|S_{\alpha,\epsilon}(t)\phi\|_{H^1} < \infty.$$

Compared to Proposition 2.4 (see Gérard–Grellier [12, 13, 16]), the dispersive term $\epsilon^\alpha \partial_x^2$ counteracts the wave turbulence phenomenon in H^1 norm for equation (1.5), if $0 \leq \alpha \leq 2$. After the change of variable $u(t) = e^{i \arg u_m(t)}(\mathbf{e}_m + \epsilon^{1-\frac{\alpha}{2}} v(t))$, we use a bootstrap argument to get long time orbital stability of the traveling waves $\mathbf{e}_m$ with respect to higher Sobolev norms.

PROPOSITION 1.7. For all $s \geq 1$ and $m \in \mathbb{N}$, there exist two constants $b_{m,s} \in (0, 1)$ and $L_{m,s} > 0$ such that if $0 \leq \alpha < 2$ and $\|u(0) - \mathbf{e}_m\|_{H^s} = \epsilon \in (0, 1)$, then we have

$$\sup_{|t|\leq \frac{b_{m,s}}{\epsilon^{1-\frac{\alpha}{2}}}} \inf_{\theta\in\mathbb{R}} \|u(t) - e^{i\theta}\mathbf{e}_m\|_{H^s} \leq L_{m,s}\epsilon^{1-\frac{\alpha}{2}}. \tag{1.9}$$

We also look for a larger time interval in which the estimate (1.9) holds, by using the Birkhoff normal form transformation. But the coefficients in front of the high frequency Fourier modes in the homological equation may be arbitrarily large, if $\alpha \in (0,2)$. For this reason, we return to the case $\alpha = 0$ and consider equation (1.1).

$$i\partial_t u + \partial_x^2 u = \Pi(|u|^2 u).$$

Then the time interval can be enlarged as $[-\frac{d_{m,s}}{\epsilon^2}, \frac{d_{m,s}}{\epsilon^2}]$ in this case.

THEOREM 1.8. *In the case $\alpha = 0$, for all $s \geq 1$ and $m \in \mathbb{N}$, there exist three constants $d_{m,s}, \epsilon_{m,s} \in (0,1)$ and $K_{m,s} > 0$ such that if $\|u(0) - \mathbf{e}_m\|_{H^s} = \epsilon \in (0, \epsilon_{m,s})$, then we have*

$$\sup_{|t| \leq \frac{d_{m,s}}{\epsilon^2}} \inf_{\theta \in \mathbb{R}} \|u(t) - e^{i\theta}\mathbf{e}_m\|_{H^s} \leq K_{m,s}\epsilon.$$

A similar result is obtained in Faou–Gauckler–Lubich [8] for the focusing or defocusing cubic Schrödinger equation on the arbitrarily dimensional torus. (see Section 5 for the comparison between (1.1) and (1.2))

After stating the stability results, we turn to construct some large solutions for (1.5) with respect to their initial data, if the level of dispersion is exponentially small with respect to the level of perturbation of the plane wave $\mathbf{e}_1 : x \mapsto e^{ix}$. We state the last result of this paper.

THEOREM 1.9. *There exists a constant $K > 0$ such that for all $0 < \delta \ll 1$, we denote by U the solution of the following NLS-Szegő equation with small dispersion*

$$i\partial_t U + \nu^2 \partial_x^2 U = \Pi(|U|^2 U), \qquad U(0,x) = e^{ix} + \delta, \tag{1.10}$$

where $\nu = e^{-\frac{\pi K}{2\delta^2}}$, then we have $\|U(t^\delta)\|_{H^1} \simeq \frac{1}{\delta}$ with $t^\delta := \frac{\pi}{\delta\sqrt{4+\delta^2}}$.

This H^1-instability result indicates that the support of the energy functional of equation (1.10) is transferred to higher Fourier modes. This phenomenon is similar to the cubic Szegő equation case (see Gérard–Grellier [15, 16, 17]) and the 2D cubic NLS equation case (see Colliander–Keel–Staffilani–Takaoka–Tao [5]). Compared to Theorem 1.5, adding the low-level dispersion $e^{-\frac{\pi K}{\delta^2}}\partial_x^2$ fails to change the quality of wave turbulence phenomenon (Proposition 2.4) for the cubic Szegő equation.

The second part of Theorem 1.3 is a consequence of Theorem 1.9. Indeed, if $\alpha > 2$ is fixed, we rescale $u(t,x) = \epsilon U(\epsilon^2 t, x)$ with $e^{-\frac{\pi K}{2\delta^2}} = \nu = \epsilon^{\frac{\alpha-2}{2}}$. Then u solves (1.5) with $u(0,x) = \epsilon(e^{ix} + \delta)$ and

$$\|u(\frac{t^\delta}{\epsilon^2})\|_{H^1} = \epsilon\|U(t^\delta)\|_{H^1} \simeq \frac{\epsilon}{\delta} \simeq \epsilon\sqrt{(\alpha-2)|\ln\epsilon|} \gg \epsilon,$$

while $\frac{t^\delta}{\epsilon^2} \simeq \frac{\sqrt{(\alpha-2)|\ln\epsilon|}}{\epsilon^2} \ll \frac{1}{\epsilon^{2+\beta}}$, for all $\beta > 0$. However, this method does not work in the critical case $\alpha = 2$. If u solves

$$i\partial_t u + \epsilon^2\partial_x^2 u = \Pi(|u|^2 u), \qquad u(0,x) = \epsilon(e^{ix} + \delta),$$

after rescaling $U(t,x) = \epsilon^{-1}u(\epsilon^{-2}t, x)$, we get equation (1.10) with $\nu = 1$, leading to (1.1) with initial datum $U(0,x) = e^{ix} + \delta$. Theorem 1.5 and Theorem 1.8 yield the following two estimates

$$\sup_{t\in\mathbb{R}} \|u(t)\|_{H^1} = \mathcal{O}(\epsilon), \qquad \sup_{|t|\leq \frac{d_{1,s}}{\epsilon^2\delta^2}} \|u(t)\|_{H^s} = \mathcal{O}(\epsilon), \qquad \forall 0<\delta \ll 1, \quad \forall 0<\epsilon<1,$$

for every $s > \frac{1}{2}$. The problem of the optimal time interval in the case $\alpha = 2$ of Theorem 1.3 remains open.

This paper is organized as follows. In Section 2, we recall some basic facts of the cubic Szegő equation and its consequences. In Section 3, we study long time behavior for (1.5) with small data and prove Theorem 1.9 and Theorem 1.3. In Section 4, we study the orbital stability of the plane waves $\mathbf{e}_m$ for (1.5) for every $m \in \mathbb{N}$ and give the proof of Theorem 1.5, Proposition 1.7 and Theorem 1.8. We compare the NLS-Szegő equation with the NLS equation in Section 5.

2. The cubic Szegő equation

In this section, we recall some results of the cubic Szegő equation

$$i\partial_t V = \Pi(|V|^2 V), \qquad V(0,\cdot) = V_0. \tag{2.1}$$

2.1. The Lax pair structure and L^∞-estimate. Given $V \in H^{\frac{1}{2}}_+$, the Hankel operator $H_V : L^2_+ \to L^2_+$ is defined by

$$H_V(h) = \Pi(V\bar{h}).$$

Given $b \in L^\infty(\mathbb{S}^1)$, the Toeplitz operator $T_b : L^2_+ \to L^2_+$ is defined by

$$T_b(h) = \Pi(bh).$$

THEOREM 2.1. *(Gérard–Grellier [12]) Set $V \in C(\mathbb{R}; H^s_+)$ for some $s > \frac{1}{2}$. Then V solves the cubic Szegő equation if and only if H_V satisfies the following evolutive equation*

$$\partial_t H_V = [B_V, H_V]. \tag{2.2}$$

where $B_V := \frac{i}{2}H_V^2 - iT_{|V|^2}$. In other words, (L_V, B_V) is a Lax pair for the cubic Szegő equation.

The equation (2.2) yields that the spectrum of the Hankel operator H_V is invariant under the flow of the cubic Szegő equation. Thus the quantity $\mathrm{Tr}|H_V|$ is conserved. A theorem of Peller ([25] Theorem 2 p. 454) states that

$$\|V\|_{B^1_{1,1}} \simeq \mathrm{Tr}|H_V|.$$

Using the embedding theorem $H^s \hookrightarrow B^1_{1,1} \hookrightarrow L^\infty$, for any $s > 1$, we have the following L^∞ estimate of the Szegő flow.

COROLLARY 2.2. (Gérard–Grellier [12]) Assume $V_0 \in H^s_+$ for some $s > 1$, then we have

$$\sup_{t\in\mathbb{R}} \|V(t)\|_{L^\infty} \lesssim_s \|V_0\|_{H^s}$$

2.2. Wave turbulence. The following theorem indicates its chaotic long time behavior with turbulence phenomenon for general initial data.

THEOREM 2.3. *(Gérard–Grellier* [15, 16, 17]*) 1.There exists a G_δ−dense set $U \subset C_+^\infty$ such that if $V_0 \in U$, then there exist two sequences $(t_n)_{n\in\mathbb{N}}$ and $(t'_n)_{n\in\mathbb{N}}$ tending to infinity such that*

$$\begin{cases} \lim_{n\to+\infty} \frac{\|V(t_n)\|_{H^s}}{|t_n|^p} = +\infty, & \forall s > \frac{1}{2}, \quad \forall p \geq 1, \\ \lim_{n\to+\infty} V(t'_n) = V_0. \end{cases}$$

2. If V_0 is rational, then $V(t)$ is also rational for every $t \in \mathbb{R}$ and the mapping $t \in \mathbb{R} \mapsto V(t) \in C_+^\infty$ is quasi-periodic.

2.3. A special case. Set $V_0(x) = V_0^\delta(x) := \delta + e^{ix}$, we denote V^δ the solution of (2.1). Refering to Gérard–Grellier [12 Sect. 6.1, 6.2; 13 Sect. 3; 16 Sect. 4], we have the following explicit formula

$$V^\delta(t,x) = \frac{a^\delta(t)e^{ix} + b^\delta(t)}{1 - p^\delta(t)e^{ix}}, \tag{2.3}$$

where

$$a^\delta(t) = e^{-it(1+\delta^2)}, \quad b^\delta(t) = e^{-it(1+\frac{\delta^2}{2})}(\delta\cos(\omega t) - i\frac{2+\delta^2}{\sqrt{4+\delta^2}}\sin(\omega t)),$$
$$p^\delta(t) = -\frac{2i}{\sqrt{4+\delta^2}}\sin(\omega t)e^{\frac{-it\delta^2}{2}}, \quad \omega = \delta\sqrt{1+\frac{\delta^2}{4}}.$$

PROPOSITION 2.4. (Gérard–Grellier [12, 13, 16]) For $0 < \delta \ll 1$, set $t^\delta := \frac{\pi}{2\omega} = \frac{\pi}{\delta\sqrt{4+\delta^2}} \sim \frac{\pi}{2\delta}$. Let V^δ be the solution of (2.1) with $V^\delta(0,x) = e^{ix} + \delta$, then we have the following estimate

$$\|V^\delta(t)\|_{H^s} \lesssim_s \|V^\delta(t^\delta)\|_{H^s} \simeq_s \frac{1}{\delta^{2s-1}}, \qquad \forall t \in [0, t^\delta].$$

for every $s > \frac{1}{2}$.

PROOF. Expanding formula (2.3) as Fourier series, we have

$$\|V^\delta(t)\|_{H^s}^2 \simeq_s \frac{|a^\delta(t) + b^\delta(t)p^\delta(t)|^2}{(1-|p^\delta(t)|^2)^{2s+1}} = \frac{\|V^\delta(t)\|_{\dot{H}^{\frac{1}{2}}}^2}{(1-|p^\delta(t)|^2)^{2s-1}}$$

with $\|V^\delta(t)\|_{\dot{H}^{\frac{1}{2}}}^2 = \|V^\delta(0)\|_{\dot{H}^{\frac{1}{2}}}^2 = 1$. By the explicit formula of p^δ, we have

$$\|V^\delta(t)\|_{H^s}^2 \simeq_s \left(\frac{4+\delta^2}{4\cos^2(\omega t)+\delta^2}\right)^{2s-1} \leq \|V^\delta(t^\delta)\|_{H^s}^2 \simeq \frac{C_s}{\delta^{4s-2}}$$

with $t^\delta := \frac{\pi}{2\omega} = \frac{\pi}{\delta\sqrt{4+\delta^2}}$.

□

3. Long time behavior for small data

3.1. The case $\alpha \geq 2$. For all $s > \frac{1}{2}$, consider the NLS-Szegő equation with small dispersion and small data.

$$i\partial_t u + \epsilon^\alpha \partial_x^2 u = \Pi(|u|^2 u), \qquad \|u(0)\|_{H^s} = \epsilon, \qquad 0 < \epsilon < 1, \qquad \alpha \geq 0. \tag{3.1}$$

At first, we show how to find the time interval $I_\epsilon^\alpha = [-\frac{a_s}{\epsilon^2}, \frac{a_s}{\epsilon^2}]$, in which the solutions of (3.1) are bounded by $\mathcal{O}(\epsilon)$, for all $\alpha \geq 0$. Then we prove the maximality of I_ϵ^α in the case $\alpha > 2$.

3.1.1. *The bootstrap argument.* The time interval $I_\epsilon^\alpha = [-\frac{a_s}{\epsilon^2}, \frac{a_s}{\epsilon^2}]$ is given by a bootstrap argument.

LEMMA 3.1. Let $a, b, T > 0$, $q > 1$ and $M : [0,T] \longrightarrow \mathbb{R}_+$ be a continuous function satisfying

$$M(\tau) \leq a + bM(\tau)^q, \qquad \text{for} \qquad \text{all} \qquad \tau \in [0,T]$$

Assume that $(qb)^{\frac{1}{q-1}} M(0) \leq 1$ and $(qb)^{\frac{1}{q-1}} a \leq \frac{q-1}{q}$. Then

$$M(\tau) \leq \frac{q}{q-1} a$$

for all $\tau \in [0,T]$.

PROOF. The function $f_q : z \in \mathbb{R}_+ \longmapsto z - bz^q$ attains its maximum at the critical point $z_c = (qb)^{-\frac{1}{q-1}}$. $f_q(z_c) = \frac{q-1}{q}(qb)^{-\frac{1}{q-1}}$. Since $a \leq \max_{z\geq 0} f_q(z) = f_q(z_c)$, there exists $z_- \leq z_c \leq z_+$ such that

$$\{z \geq 0 : f_q(z) \leq a\} = [0, z_-] \cup [z_+, +\infty[$$

and $f_q(z_\pm) = a$. Since $f_q(M(\tau)) \leq a$, $\forall 0 \leq \tau \leq T$ and $M(0) \leq z_c$, we have $M([0,T]) \subset [0, z_-]$. By the concavity of f_q on $[0, +\infty[$, we have $f_q(z) \geq \frac{f_q(z_c)}{z_c} z$ for all $z \in [0, z_c]$. Consequently, $M(\tau) \leq z_- \leq \frac{q}{q-1} a$, for all $0 \leq \tau \leq T$. □

PROOF OF OF ESTIMATE (1.7). For all $\alpha \geq 0$ and $\epsilon \in (0,1)$ fixed, we rescale u as $u = \epsilon\mu$, equation (3.1) becomes

$$\begin{cases} i\partial_t \mu + \epsilon^\alpha \partial_x^2 \mu = \epsilon^2 \Pi(|\mu|^2\mu), \\ \|\mu(0)\|_{H^s} = 1. \end{cases} \tag{3.2}$$

Duhamel's formula of equation (3.2) gives the following estimate:

$$\sup_{0\leq\tau\leq t} \|\mu(\tau)\|_{H^s} \leq \|\mu(0)\|_{H^s} + C_s\epsilon^2 t \sup_{0\leq\tau\leq t} \|\mu(\tau)\|_{H^s}^3 \tag{3.3}$$

Here C_s denotes the Sobolev constant in the inequality $\||\mu|^2\mu\|_{H^s} \leq C_s\|\mu\|_{H^s}^3$. We choose $a_s = \frac{4}{27C_s}$ and the following estimate holds

$$\sup_{|t|\leq \frac{a_s}{\epsilon^2}} \|u(t)\|_{H^s} \leq \frac{3}{2}\epsilon, \tag{3.4}$$

by using Lemma 3.1 with $q = 3$, $T = \frac{a_s}{\epsilon^2}$, $a = M(0) = 1$, $b = C_s\epsilon^2 T$ and $M(t) = \sup_{0\leq\tau\leq t} \|\mu(\tau)\|_{H^s}$. The case $t < 0$ is similar.

□

3.1.2. *Optimality of the time interval if* $\alpha > 2$. In order to prove the optimality of I_ϵ^α in which estimate (3.4) holds, we set $u(0,x) = \epsilon(e^{ix}+\delta)$ and rescale $u(t,x) = \epsilon U(\epsilon^2 t, x)$. Then, we have

$$i\partial_t U + \nu^2 \partial_x^2 U = \Pi(|U|^2 U), \qquad U(0,x) = e^{ix} + \delta,$$

where $\nu := \epsilon^{\frac{\alpha-2}{2}}$. Since the optimality is a consequence of Theorem 1.9, we prove at first Theorem 1.9 by comparing U to the solution of the cubic Szegő equation with the same initial datum,

$$i\partial_t V = \Pi(|V|^2 V), \qquad V(0,x) = e^{ix} + \delta.$$

Proof of Theorem 1.9. We shall estimate their difference $r(t,x) := U(t,x) - V(t,x)$, which satisfies the following equation

$$i\partial_t r + \nu^2 \partial_x^2 r = -\nu^2 \partial_x^2 V + \Pi(V^2 \overline{r} + 2|V|^2 r) + Q(r), \qquad r(0) = 0, \tag{3.5}$$

with $Q(r) := \Pi(\overline{V} r^2 + 2V|r|^2 + |r|^2 r)$. Thus, we can calculate the derivative of $\|r(t)\|_{H^1}^2$,

$$\begin{aligned}
&\partial_t \|r(t)\|_{H^1}^2 \\
=&\partial_t \|r(t)\|_{L^2}^2 + \partial_t \|\partial_x r(t)\|_{L^2}^2 \\
=&2\mathrm{Im}\langle i\partial_t r(t), r(t)\rangle_{L^2} + 2\mathrm{Im}\langle \partial_x(i\partial_t r(t)), \partial_x r(t)\rangle_{L^2} \\
=&2\mathrm{Im}\int_{\mathbb{S}^1} \nu^2 \partial_x V \partial_x \overline{r} + V^2 \overline{r}^2 - \overline{V}|r|^2 r \\
&\quad + 2\mathrm{Im}\int_{\mathbb{S}^1} -\nu^2 \partial_x^3 V \partial_x \overline{r} + V^2 (\partial_x \overline{r})^2 + 2V\partial_x V \overline{r} \partial_x \overline{r} + 4\mathrm{Re}(\overline{V}\partial_x V) r \partial_x \overline{r} \\
&\quad + 2\mathrm{Im}\int_{\mathbb{S}^1} \partial_x \overline{V} r^2 \partial_x \overline{r} + 2\overline{V} r |\partial_x r|^2 + 2\partial_x V |r|^2 \partial_x \overline{r} + 4V\mathrm{Re}(\overline{r}\partial_x r)\partial_x \overline{r} + r^2 (\partial_x \overline{r})^2.
\end{aligned}$$

Then, we have the following estimate

$$\begin{aligned}
&\left|\partial_t \|r(t)\|_{H^1}^2\right| \\
\le& 2\nu^2 \|\partial_x r\|_{L^2}(\|\partial_x V\|_{L^2} + \|\partial_x^3 V\|_{L^2}) + 2\|V\|_{L^\infty}^2 \|r\|_{H^1}^2 + 2\|V\|_{L^\infty}\|r\|_{L^\infty}\|r\|_{L^2}^2 \\
&\quad + 12\|V\|_{L^\infty}\|r\|_{L^\infty}\|\partial_x V\|_{L^2}\|\partial_x r\|_{L^2} + 6\|r\|_{L^\infty}^2 \|\partial_x V\|_{L^2}\|\partial_x r\|_{L^2} \\
&\quad + 12\|V\|_{L^\infty}\|r\|_{L^\infty}\|\partial_x r\|_{L^2}^2 + 2\|r\|_{L^\infty}^2 \|\partial_x r\|_{L^2}^2 \\
\le& \nu^2(2\|\partial_x r\|_{L^2}^2 + \|\partial_x V\|_{L^2}^2 + \|\partial_x^3 V\|_{L^2}^2) + 2\|V\|_{L^\infty}^2 \|r\|_{H^1}^2 \\
&\quad + 12\|V\|_{L^\infty}\|\partial_x V\|_{L^2}\|r\|_{L^\infty}\|\partial_x r\|_{L^2} + \mathcal{O}(\|r\|_{H^1}^3).
\end{aligned}$$

L^∞-estimate of V is given by Corollary 2.2 and H^s-growth of V is given by Proposition 2.4, for all $s > \frac{1}{2}$. Thus, we have

$$M_\infty := \sup_{0<\delta<1} \sup_{t\in\mathbb{R}} \|V(t)\|_{L^\infty} < +\infty,$$

and there exist $C_1, C_3 > 0$ such that

$$\|\partial_x V(t)\|_{L^2} \le \frac{C_1}{\delta}, \qquad \|\partial_x^3 V(t)\|_{L^2} \le \frac{C_3}{\delta^5},$$

for all $0 \le t \le t^\delta = \frac{\pi}{\delta\sqrt{4+\delta^2}}$. We use a bootstrap argument to estimate the term $\mathcal{O}(\|r\|_{H^1}^3)$. Set

$$T := \sup\{t > 0 : \sup_{0\le\tau\le t} \|r(\tau)\|_{H^1} \le 1\},$$

then we have

$$\sup_{0\le t\le T}\|r(t)\|_{L^\infty}\le C\sup_{0\le t\le T}\|r(t)\|_{H^1}\le C,$$

where C denotes the Sobolev constant $H^1(\mathbb{S}^1)\hookrightarrow L^\infty$. Consequently, for all $0\le t\le\min(T,t^\delta)$, we have

$$\begin{aligned}&\left|\partial_t\|r(t)\|_{H^1}^2\right|\\ \le&\nu^2(\|\partial_x V\|_{L^2}^2+\|\partial_x^3 V\|_{L^2}^2)\\ &+\|r\|_{H^1}^2(2\nu^2+2\|V\|_{L^\infty}^2+12C\|V\|_{L^\infty}\|\partial_x V\|_{L^2}\\ &+6C\|r\|_{L^\infty}\|\partial_x V\|_{L^2}+12\|V\|_{L^\infty}\|r\|_{L^\infty}+2\|r\|_{L^\infty}^2)\\ \le&\nu^2(\frac{C_1^2}{\delta^2}+\frac{C_3^2}{\delta^{10}})+(2+2M_\infty^2+12CM_\infty+2C^2+(12CM_\infty+6C^2)\frac{C_1}{\delta})\|r\|_{H^1}^2\\ \le&K(\frac{\nu^2}{\delta^{10}}+\frac{\|r(t)\|_{H^1}^2}{\delta}),\qquad\forall 0<\delta,\nu<1,\end{aligned}$$

with $K:=\max(C_1^2+C_3^2,2+2M_\infty^2+12CM_\infty+2C^2+(12CM_\infty+6C^2)C_1)$. We set

$$\nu=\epsilon^{\frac{\alpha-2}{2}}=e^{-\frac{\pi K}{2\delta^2}}\Longleftrightarrow\delta=\sqrt{\tfrac{\pi K}{(\alpha-2)|\ln\epsilon|}}.$$

Using Grönwall's inequality, we deduce that

$$\|r(t)\|_{H^1}^2\le\frac{\nu^2}{\delta^9}e^{\frac{\pi K}{2\delta^2}}=\delta^{-9}e^{-\frac{\pi K}{2\delta^2}}\ll 1\ll\delta^{-2},\qquad\forall 0\le t\le t^\delta,\quad\forall 0<\delta\ll 1.$$

Since $\|V(t^\delta)\|_{H^1}\simeq\frac{1}{\delta}$ by Theorem 2.4, we have $\|U(t^\delta)\|_{H^1}=\|V(t^\delta)+r(t^\delta)\|_{H^1}\simeq\frac{1}{\delta}$. □

Fix $\alpha>2$, for every $0<\epsilon\ll 1$, we set

$$\delta=\delta^{\alpha,\epsilon}:=\sqrt{\tfrac{\pi K}{(\alpha-2)|\ln\epsilon|}}\ll 1,\qquad T^{\alpha,\epsilon}:=\tfrac{t^\delta}{\epsilon^2}=\frac{\pi}{2\epsilon^2\delta\sqrt{1+\frac{\delta^2}{4}}}\simeq\frac{\sqrt{(\alpha-2)|\ln\epsilon|}}{\epsilon^2}.$$

Then we have $\|u(T^{\alpha,\epsilon})\|_{H^1}\simeq\epsilon\sqrt{(\alpha-2)|\ln\epsilon|}\gg\epsilon$, while $u(0,x)=\epsilon(e^{ix}+\delta)$. Then the optimality of $I_\epsilon^\alpha=[-\frac{a_s}{\epsilon^2},\frac{a_s}{\epsilon^2}]$ is obtained.

3.2. The case $0\le\alpha<2$. We assume at first that $u(0)\in C_+^\infty$ so that the energy functional of (3.1)

$$E^{\alpha,\epsilon}(u)=\frac{\epsilon^\alpha}{2}\|\partial_x u\|_{L^2}^2+\frac{1}{4}\|u\|_{L^4}^4$$

is well defined. For general initial data $u(0)\in H_+^s$, if $s\in(\frac{1}{2},1)$, we use the density argument $\overline{C_+^\infty}=H_+^s$ and the continuity of the mapping $u(0)\in H_+^s\mapsto C([-\frac{a_s}{\epsilon^{4-\alpha}},\frac{a_s}{\epsilon^{4-\alpha}}];H_+^s)$.

We rescale $u(t,x)\mapsto\epsilon^{-\frac{\alpha}{2}}u(\epsilon^{-\alpha}t,x)$, then (3.1) is reduced to the case $\alpha=0$. It suffices to prove the following estimate

$$\sup_{|t|\le\frac{a_s}{\epsilon^4}}\|u(t)\|_{H^s}=\mathcal{O}(\epsilon)$$

for the NLS-Szegő equation $i\partial_t u+\partial_x^2 u=\Pi(|u|^2u)$ with $\|u(0)\|_{H^s}=\epsilon$.

3.2.1. *Identifying the resonances.* The study of the resonant set of the NLS-Szegő equation is necessary before the Birkhoff normal form transformation. We refer to Eliasson–Kuksin [6] to see the analysis of resonances for a more general non linear term and the KAM theorem for the NLS equation.

We use again the change of variable $u = \epsilon\mu$ and the Duhamel's formula of μ with $\eta(t) = \sum_{k\geq 0} \eta_k(t)e^{ikx} := e^{-it\partial_x^2}\mu(t)$. Then we have

$$\eta_k(t) = \mu_0(t) - i\epsilon^2 \sum_{k_1-k_2+k_3-k=0} \int_0^t e^{-i\tau(k_1^2-k_2^2+k_3^2-k^2)}\eta_{k_1}(\tau)\overline{\eta}_{k_2}(\tau)\eta_{k_3}(\tau)\mathrm{d}\tau,$$

for all $k \geq 0$. Recall the classical identification of the resonant set

$$\begin{cases} k_1 - k_2 + k_3 - k_4 = 0 \\ k_1^2 - k_2^2 + k_3^2 - k_4^2 = 0. \end{cases} \iff \begin{cases} k_1 = k_2 \\ k_3 = k_4 \end{cases} \text{or} \quad \begin{cases} k_1 = k_4 \\ k_2 = k_3 \end{cases}.$$

In order to cancel all the resonances, we apply the transformation

$$v(t) := e^{2it\epsilon^2\|\mu(0)\|_{L^2}^2}\mu(t).$$

As $\|\mu\|_{L^2}$ is a conservation law, we have the following nonlinear Schrödinger equation without resonances

$$i\partial_t v(t) + \partial_x^2 v(t) = \epsilon^2 \left[\Pi(|v(t)|^2 v(t)) - 2\|v(t)\|_{L^2}^2 v(t)\right], \qquad \forall t \in \mathbb{R}. \tag{3.6}$$

Its energy functional is

$$H^\epsilon(v) = \frac{1}{2}\|\partial_x v\|_{L^2}^2 + \frac{\epsilon^2}{4}\left(\|v\|_{L^4}^4 - 2\|v\|_{L^2}^4\right) =: H_0(v) + \epsilon^2 R(v). \tag{3.7}$$

Consider the Fourier modes of $v = \sum_{n\geq 0} v_n e^{inx}$, then we have

$$4R(v) = \sum_{\substack{k_1-k_2+k_3-k_4=0 \\ k_1^2-k_2^2+k_3^2-k_4^2\neq 0}} v_{k_1}\overline{v}_{k_2}v_{k_3}\overline{v}_{k_4} - \sum_{k\geq 0}|v_k|^4.$$

3.2.2. *The Birkhoff normal form.* Equation (3.6) is transferred to another Hamiltonian equation which is closer to the linear Schrödinger equation by the Birkhoff normal form transformation. We look for a symplectomorphism Ψ_ϵ such that the energy functional H^ϵ is reduced to the following Hamiltonian

$$H^\epsilon \circ \Psi_\epsilon(v) = H_0(v) + \epsilon^2\tilde{R}(v) + \mathcal{O}(\epsilon^{4-\alpha}),$$

where $\tilde{R}(v) = -\frac{1}{4}\sum_{k\geq 0}|v_k|^4$. Ψ_ϵ is chosen as the value at time 1 of the Hamiltonian flow of some energy $\epsilon^2 F$.

We fix the value $s > \frac{1}{2}$. Recall that, given a smooth real valued function H, we denote X_H the Hamiltonian vector field, i.e,

$$\mathrm{d}H(v)(h) = \omega(h, X_H(v)).$$

Given two smooth real-valued functions F and G on H_+^s, their Poisson bracket $\{F, G\}$ is defined by

$$\{F, G\}(v) := \omega(X_F(v), X_G(v)) = \frac{2}{i}\sum_{k\geq 0}\left(\partial_{\overline{v}_k}F\partial_{v_k}G - \partial_{\overline{v}_k}G\partial_{v_k}F\right)(v), \tag{3.8}$$

for all $v = \sum_{k\geq 0} v_k e^{ikx} \in H^s_+$. In particular, if F and G are respectively homogeneous of order p and q, then their Poisson bracket is homogeneous of order $p+q-2$.

LEMMA 3.2. Set $F(v) := \sum_{k_1-k_2+k_3-k_4=0} f_{k_1,k_2,k_3,k_4} v_{k_1}\overline{v}_{k_2}v_{k_3}\overline{v}_{k_4}$, with the coefficients

$$f_{k_1,k_2,k_3,k_4} = \begin{cases} \frac{i}{4\left(k_1^2-k_2^2+k_3^2-k_4^2\right)}, & \text{if} \quad k_1^2-k_2^2+k_3^2-k_4^2 \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, F is real-valued and its Hamiltonian field X_F is smooth on H^s_+ such that $\{F, H_0\} + R = \tilde{R}$ and the following estimates hold for all $v \in H^s_+$.

$$\begin{cases} \|X_F(v)\|_{H^s} \lesssim_s \|v\|^3_{H^s}, \\ \|\mathrm{d}X_F(v)\|_{B(H^s)} \lesssim_s \|v\|^2_{H^s}. \end{cases}$$

PROOF. F is well defined because $\sup_{(k_1,k_2,k_3,k_4)\in\mathbb{Z}^4} |f_{k_1,k_2,k_3,k_4}| \leq \frac{1}{4}$, the Sobolev embedding yields that $|F(v)| \leq \frac{1}{4}\left(\sum_{k\geq 0}|v_k|\right)^4 \lesssim_s \|v\|^4_{H^s}$. The Sobolev estimates of $\|X_F(v)\|_{H^s}$ and $\|\mathrm{d}X_F(v)\|_{B(H^s)}$ are given by the Young's convolution inequality $l^1 * l^1 * l^2 \hookrightarrow l^2$. Using (3.8) and the definition of f_{k_1,k_2,k_3,k_4}, we have

$$\begin{aligned} \{F, H_0\}(v) =& i \sum_{k_1-k_2+k_3-k_4=0} (k_1^2-k_2^2+k_3^2-k_4^2) f_{k_1,k_2,k_3,k_4} v_{k_1}\overline{v}_{k_2}v_{k_3}\overline{v}_{k_4} \\ =& -\frac{1}{4} \sum_{\substack{k_1-k_2+k_3-k_4=0 \\ k_1^2-k_2^2+k_3^2-k_4^2\neq 0}} v_{k_1}\overline{v}_{k_2}v_{k_3}\overline{v}_{k_4} \\ =& -R(v) + \tilde{R}(v). \end{aligned}$$

□

Set $\chi_\sigma := \exp(\epsilon^2 \sigma X_F)$ the Hamiltonian flow of $\epsilon^2 F$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\chi_\sigma(v) = \epsilon^2 X_F(\chi_\sigma(v)), \qquad \chi_0(v) = v.$$

We perform the canonical transformation $\Psi_\epsilon := \chi_1 = \exp(\epsilon^2 X_F)$. The next lemma will prove the local existence of χ_σ, for $|\sigma| \leq 1$ and give the estimate of the difference between v and $\Psi_\epsilon^{-1}(v)$

LEMMA 3.3. For $s > \frac{1}{2}$, there exist two constants $\rho_s, C_s > 0$ such that for all $v \in H^s_+$, if $\epsilon\|v\|_{H^s} \leq \rho_s$, then $\chi_\sigma(v)$ is well defined on the interval $[-1,1]$ and the following estimates hold:

$$\begin{aligned} &\sup_{\sigma\in[-1,1]} \|\chi_\sigma(v)\|_{H^s} \leq \frac{3}{2}\|v\|_{H^s}, \\ &\sup_{\sigma\in[-1,1]} \|\chi_\sigma(v) - v\|_{H^s} \leq C_s\epsilon^2\|v\|^3_{H^s}, \\ &\|\mathrm{d}\chi_\sigma(v)\|_{B(H^s)} \leq \exp(C_s\epsilon^2\|v\|^2_{H^s}|\sigma|), \qquad \forall\sigma\in[-1,1]. \end{aligned}$$

Proof. The Lipschitz coefficient of the mapping $v \longmapsto \epsilon^2 X_F(v)$ is bounded by $C_s\epsilon^2\|v\|_{H^s}^2 \le C_s\rho_s^2$, by Lemma 3.2. If ρ_s is sufficiently small, the Hamiltionian flow $(\sigma, v) \mapsto \chi_\sigma(v)$ exists on the maximal interval $(-\sigma^*, \sigma^*)$, by the Picard-Lindelöf theorem. Assume that $\sigma^* < 1$, then Lemma 3.2 and the following integral formula

$$\chi_\sigma(v) = v + \epsilon^2 \int_0^\sigma X_F(\chi_\tau(v))\mathrm{d}\tau, \qquad \forall 0 \le \sigma < \sigma^*. \tag{3.9}$$

yield that

$$\sup_{0\le\tau\le\sigma} \|\chi_\tau(v)\|_{H^s} \le \|v\|_{H^s} + C_s\sigma\epsilon^2 \sup_{0\le\tau\le\sigma} \|\chi_\tau(v)\|_{H^s}^3, \qquad \forall 0 \le \sigma < \sigma^* < 1.$$

By Lemma 3.1 with $M(t) = \sup_{0\le\tau\le t} \|\chi_\tau(v)\|_{H^s}$, $q = 3$, $a = M(0) = \|v\|_{H^s}$ and $b = C_s\epsilon^2$, we have

$$\sup_{|\sigma|\le\sigma^*} \|\chi_\sigma(v)\|_{H^s} \le \frac{3}{2}\|v\|_{H^s},$$

if $\epsilon\|v\|_{H^s} \le \frac{2}{3\sqrt{3C_s}}$. This is a contradiction to the blow-up criterion. Hence $\sigma^* \ge 1$, and we have $\sup_{|\sigma|\le 1} \|\chi_\sigma(v)\|_{H^s} \le \frac{3}{2}\|v\|_{H^s}$, if $\epsilon\|v\|_{H^s} \le \rho_s := \frac{2}{3\sqrt{3C_s}}$. For all $\sigma \in [-1, 1]$, by using Lemma 3.2, we have

$$\|\chi_\sigma(v) - v\|_{H^s} \le |\sigma|\epsilon^2 \sup_{0\le t\le|\sigma|} \|X_F(\chi_t(v))\|_{H^s} \le C_s\epsilon^2 \sup_{0\le t\le|\sigma|} \|\chi_t(v)\|_{H^s}^3 \le C_s\epsilon^2\|v\|_{H^s}^3.$$

if $\epsilon\|v\|_{H^s} \le \rho_s$. We differentiate equation (3.9) and use again Lemma 3.2 to obtain

$$\begin{aligned}\|\mathrm{d}\chi_\sigma(u)\|_{B(H^s)} =& \|\mathrm{Id}_{H^s} + \epsilon^2 \int_0^\sigma \mathrm{d}X_F(\chi_t(u))\mathrm{d}\chi_t(u)\mathrm{d}t\|_{B(H^s)} \\ \le& 1 + \epsilon^2 \Big| \int_0^\sigma \|\mathrm{d}X_F(\chi_t(v))\|_{B(H^s)} \|\mathrm{d}\chi_t(v)\|_{B(H^s)}\mathrm{d}t \Big| \\ \le& 1 + C_s\epsilon^2\|v\|_{H^s}^2 \Big| \int_0^\sigma \|\mathrm{d}\chi_t(v)\|_{B(H^s)}\mathrm{d}t \Big| \\ \le& e^{C_s\epsilon^2|\sigma|\|v\|_{H^s}^2}, \qquad \forall \sigma \in [-1, 1].\end{aligned}$$

Grönwall's inequality is used in the last step.

□

Composed with the symplectomorphism $\Psi_\epsilon = \chi_1$, the energy functional H^ϵ can be reduced to the normal form.

Lemma 3.4. For $s > \frac{1}{2}$, there exists a smooth mapping $Y : H_+^s \longrightarrow H_+^s$ and a constant $C_s' > 0$ such that

$$\begin{cases} X_{H^\epsilon \circ \Psi_\epsilon} = X_{H_0} + \epsilon^2 X_{\tilde{R}} + \epsilon^4 Y, \\ \|Y(v)\|_{H^s} \le C_s'\|v\|_{H^s}^5, \end{cases}$$

for all $v \in H_+^s$ such that $\epsilon\|v\|_{H^s} \le \rho_s$. Set $w(t) := \Psi_\epsilon^{-1}(v(t))$, then we have

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|_{H^s}^2\Big| \le C_s'\epsilon^4\|w(t)\|_{H^s}^6, \tag{3.10}$$

if $\epsilon\|w(t)\|_{H^s} \le \rho_s$.

Proof. We expand the energy $H^\epsilon \circ \chi_1 = H_0 \circ \chi_1 + \epsilon^2 R \circ \chi_1$ with Taylor's formula at time $\sigma = 1$ around 0. Since $\chi_0 = \mathrm{Id}_{H^s_+}$, we have

$$\begin{aligned}
H^\epsilon \circ \chi_1 = & \left(H_0 + \frac{\mathrm{d}}{\mathrm{d}\sigma}[H_0 \circ \chi_\sigma]|_{\sigma=0} + \int_0^1 (1-\sigma) \frac{\mathrm{d}^2}{\mathrm{d}\sigma^2}[H_0 \circ \chi_\sigma] \mathrm{d}\sigma \right) \\
& + \epsilon^2 R + \epsilon^2 \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\sigma}[R \circ \chi_\sigma] \mathrm{d}\sigma \\
= & H_0 + \epsilon^2 \left[\{F, H_0\} + R \right] + \epsilon^4 \int_0^1 (1-\sigma)\{F, \{F, H_0\}\} \circ \chi_\sigma + \{F, R\} \circ \chi_\sigma \mathrm{d}\sigma \\
= & H_0 + \epsilon^2 \tilde{R} + \epsilon^4 \int_0^1 \left[(1-\sigma)\{F, \tilde{R}\} + \sigma\{F, R\} \right] \circ \chi_\sigma \mathrm{d}\sigma
\end{aligned}$$

We set $G(\sigma) := (1-\sigma)\{F, \tilde{R}\} + \sigma\{F, R\}$, $\forall \sigma \in [0,1]$. Since $X_{\{F,R\}}$ and $X_{\{F,\tilde{R}\}}(u)$ are homogeneous of degree 5 with uniformly bounded coefficients, we have

$$\|X_{G(\sigma)}(v)\|_{H^s} \leq (1-\sigma)\|X_{\{F,\tilde{R}\}}(v)\|_{H^s} + \sigma\|X_{\{F,R\}}(v)\|_{H^s} \lesssim_s \|v\|^5_{H^s}, \quad \forall v \in H^s_+,$$

By the chain rule of Hamiltonian vector fields:

$$X_{G(\sigma)\circ\chi_\sigma}(v) = \mathrm{d}\chi_{-\sigma}(\chi_\sigma(v)) \circ X_{G(\sigma)}(\chi_\sigma(v)), \quad \forall v \in H^s_+, \quad \forall \sigma \in [0,1], \tag{3.11}$$

and Lemma 3.3, we have

$$\begin{aligned}
\|X_{G(\sigma)\circ\chi_\sigma}(v)\|_{H^s} & \leq \|\mathrm{d}\chi_{-\sigma}(\chi_\sigma(v))\|_{B(H^s)} \|X_{G(\sigma)}(\chi_\sigma(v))\|_{H^s} \\
& \lesssim_s e^{C_s \epsilon^2 \|v\|^2_{H^s}} \|\chi_\sigma(v)\|^5_{H^s} \lesssim_s \|v\|^5_{H^s},
\end{aligned}$$

for all $v \in H^s_+$ such that $\epsilon\|v\|_{H^s} \leq \rho_s$. Thus we define $Y := \int_0^1 X_{G(\sigma)\circ\chi_\sigma} \mathrm{d}\sigma$ and we have

$$X_{H^\epsilon \circ \Psi_\epsilon} = X_{H^\epsilon \circ \chi_1} = X_{H_0} + \epsilon^2 X_{\tilde{R}} + \epsilon^4 Y.$$

If $\epsilon\|v\|_{H^s} \leq \rho_s$, then $\|Y(v)\|_{H^s} \lesssim_s \|v\|^5_{H^s}$.

Since $\widehat{X_{\tilde{R}}(w)}(k) = -i|w_k|^2 w_k$, $\forall k \geq 0$ and $w(t) = \Psi_\epsilon^{-1}(v(t))$, $w = \sum_{k\geq 0} w_k e^{ikx}$ solves the following infinite dimensional Hamiltonian system of Fourier modes:

$$i\partial_t w_k(t) - k^2 w_k(t) - \epsilon^2 |w_k(t)|^2 w_k(t) = i\epsilon^4 \widehat{Y(w(t))}(k), \quad \forall k \geq 0. \tag{3.12}$$

If $\epsilon\|w(t)\|_{H^s} \leq \rho_s$, then we have

$$\left| \partial_t \|w(t)\|^2_{H^s} \right| \leq 2\epsilon^4 \|Y(w(t))\|_{H^s} \|w(t)\|_{H^s} \lesssim_s \epsilon^4 \|w(t)\|^6_{H^s}.$$

□

3.2.3. *End of the proof of the case $0 \leq \alpha < 2$.*

Proof. Recall that $w(t) = \chi_{-1}(v(t))$ and $\|v(0)\|_{H^s} = 1$. Lemma 3.3 yields that if $\epsilon \leq \rho_s$, then we have

$$\|v(0) - w(0)\|_{H^s} = \|v(0) - \chi_{-1}(v(0))\|_{H^s} \leq C_s \epsilon^2 \|v(0)\|^3_{H^s} \leq C_s.$$

Set $K_s := 3C_s + 1$. Then $\|w(0)\|_{H^s} \leq \frac{K_s}{3}$. We define

$$\epsilon_s := \min \left(\frac{\rho_s}{3K_s}, \quad \frac{1}{\sqrt{8C_s K_s^2}} \right)$$

and

$$T := \sup\{t \geq 0 : \sup_{0\leq\tau\leq t} \|v(\tau)\|_{H^s} \leq 2K_s\}.$$

For all $\epsilon \in (0, \epsilon_s)$ and $t \in [0, T]$, we have $\epsilon\|v(t)\|_{H^s} \leq \rho_s$. Hence Lemma 3.3 gives the following estimate

$$\begin{aligned}\|w(t)\|_{H^s} \leq &\|v(t)\|_{H^s} + \|\chi_{-1}(v(t)) - v(t)\|_{H^s} \\ \leq &\|v(t)\|_{H^s} + C_s\epsilon^2\|v(t)\|^3_{H^s} \\ \leq &2K_s + 8C_sK_s^3\epsilon^2 \\ \leq &3K_s.\end{aligned}$$

So we have $\epsilon \sup_{0\leq t\leq T} \|w(t)\|_{H^s} \leq \rho_s$ and $\left|\frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|^2_{H^s}\right| \leq C'_s\epsilon^4\|w(t)\|^6_{H^s}$, by Lemma 3.4. Set $a_s := \frac{1}{3^7K_s^4C'_s}$. We can precise the estimate of $\|w(t)\|_{H^s}$ by limiting $0 \leq t \leq a_s\epsilon^{-4}$:

$$\|w(t)\|^2_{H^s} \leq \|w(0)\|^2_{H^s} + C'_st\epsilon^4\|w(t)\|^6_{H^s} \leq \frac{K_s^2}{9} + 3^6C'_sK_s^6t\epsilon^4 \leq \frac{4K_s^2}{9},$$

for all $0 \leq t \leq \min(T, \frac{a_s}{\epsilon^4})$. Then we have

$$\|v(t)\|_{H^s} \leq \|w(t)\|_{H^s} + \|\chi_1(w(t)) - w(t)\|_{H^s} \leq \|w(t)\|_{H^s} + C_s\epsilon^2\|w(t)\|^3_{H^s} \leq K_s,$$

for all $t \in [0, \frac{a_s}{\epsilon^4}]$. Consequently, we have

$$\sup_{0\leq t\leq \frac{a_s}{\epsilon^4}} \|u(t)\|_{H^s} = \epsilon \sup_{0\leq t\leq \frac{a_s}{\epsilon^4}} \|v(t)\|_{H^s} \leq K_s\epsilon.$$

In the case $t < 0$, we use the same procedure and we replace t by $-t$. □

3.2.4. *The open problem of optimality.* Recall that $H^\epsilon = H_0 + \epsilon^2 R$ is the energy functional of the equation

$$i\partial_t u + \partial_x^2 u = \Pi(|u|^2u), \qquad \|u(0,\cdot)\|_{H^s} = \epsilon. \tag{3.13}$$

with $H_0(v) = \frac{1}{2}\|\partial_x v\|^2_{L^2}$ and $R(v) = \frac{1}{4}(\|v\|^4_{L^4} - 2\|v\|^4_{L^2})$. In order to get a longer time interval in which the solution is uniformly bounded by $\mathcal{O}(\epsilon)$, we expand the Hamiltonian $H^\epsilon \circ \chi_1$ by using the Taylor expansion of higher order to see whether the resonances can be cancelled by the Birkhoff normal form method.

$$\begin{aligned}&H^\epsilon \circ \chi_1 \\ =&H_0 \circ \chi_1 + \epsilon^2 R \circ \chi_1 \\ =&H_0 + \partial_\sigma(H_0 \circ \chi_\sigma)\big|_{\sigma=0} + \frac{1}{2}\partial_\sigma^2(H_0 \circ \chi_\sigma)\big|_{\sigma=0} + \frac{1}{2}\int_0^1 (1-\sigma)^2\partial_\sigma^3(H_0 \circ \chi_\sigma)\mathrm{d}\sigma \\ &+ \epsilon^2\left(R + \partial_\sigma(R \circ \chi_\sigma)\big|_{\sigma=0} + \int_0^1 (1-\sigma)\partial_\sigma^2(H_0 \circ \chi_\sigma)\mathrm{d}\sigma\right) \\ =&H_0 + \epsilon^2\tilde{R} + \frac{\epsilon^4}{2}\{F, R + \tilde{R}\} + \frac{\epsilon^6}{2}\int_0^1 (1-\sigma)\{F, \{F, (1-\sigma)\tilde{R} + (1+\sigma)R\}\} \circ \chi_\sigma \mathrm{d}\sigma\end{aligned}$$

We try to cancel the term $\frac{\epsilon^4}{2}\{F, R + \tilde{R}\}$ by using a canonical transform to $H_1^\epsilon = H^\epsilon \circ \chi_1$ with the following functional

$$G(v) = \sum_{k_1-k_2+k_3-k_4+k_5-k_6=0} g_{k_1,k_2,k_3,k_4,k_5,k_6} v_{k_1}\overline{v}_{k_2}v_{k_3}\overline{v}_{k_4}v_{k_5}\overline{v}_{k_6}.$$

Then we should solve the homological equation $\{G, H_0\} + \frac{1}{2}\{F, R + \tilde{R}\} = 0$.

$$\begin{aligned}&\{G, H_0\}(v)\\ =&i \sum_{k_1-k_2+k_3-k_4+k_5-k_6=0} (k_1^2 - k_2^2 + k_3^2 - k_4^2 + k_5^2 - k_6^2)\\ &g_{k_1,k_2,k_3,k_4,k_5,k_6} v_{k_1} \overline{v}_{k_2} v_{k_3} \overline{v}_{k_4} v_{k_5} \overline{v}_{k_6}.\end{aligned}$$

We can calculate that

$$\begin{aligned}&\frac{1}{2}\{F, R + \tilde{R}\}(v)\\ =&2\mathrm{Im} \sum_{\substack{k_1-k_2+k_3-k_4+k_5-k_6=0\\ k_i, k:=k_1-k_2+k_3\geq 0}} f_{k_1,k_2,k_3,k} v_{k_1} \overline{v}_{k_2} v_{k_3} \overline{v}_{k_4} v_{k_5} \overline{v}_{k_6}\\ &- 4\mathrm{Im} \sum_{\substack{k_1-k_2+k_3-k_4+k_5-k_6=0\\ k_i, k:=k_1-k_2+k_3\geq 0, k_5=k_6}} f_{k_1,k_2,k_3,k} v_{k_1} \overline{v}_{k_2} v_{k_3} \overline{v}_{k_4} v_{k_5} \overline{v}_{k_6}\\ &- 2\mathrm{Im} \sum_{\substack{k_1-k_2+k_3-k_4+k_5-k_6=0\\ k_i, k:=k_1-k_2+k_3\geq 0, k_4=k_5=k_6}} f_{k_1,k_2,k_3,k} v_{k_1} \overline{v}_{k_2} v_{k_3} \overline{v}_{k_4} v_{k_5} \overline{v}_{k_6}\end{aligned}$$

In the first term of the preceding formula, there is a resonant set $k_1^2 - k_2^2 + k_3^2 - k_4^2 + k_5^2 - k_6^2 = 0$ that cannot be cancelled by the other two terms. Thus the resonant subset

$$\begin{cases} k_1 - k_2 + k_3 - k_4 + k_5 - k_6 = 0 \\ k_1^2 - k_2^2 + k_3^2 - k_4^2 + k_5^2 - k_6^2 = 0 \end{cases}$$

should be cancelled before the Birkhoff normal form transform, just like the step $\mu \mapsto v = e^{2it\epsilon^2 \|\mu(0)\|_{L^2}^2} \mu(t)$, which can cancel all the resonances

$$\begin{cases} k_1 - k_2 + k_3 - k_4 = 0 \\ k_1^2 - k_2^2 + k_3^2 - k_4^2 = 0 \end{cases}$$

before we do the canonical transformation $H^\epsilon \mapsto H^\epsilon \circ \chi_1$. We only know that $f_{k_1,k_2,k_3,k_1-k_2+k_3} = f_{k_4,k_5,k_6,k_4-k_5+k_6}$ if

$$\begin{cases} k_1 - k_2 + k_3 - k_4 + k_5 - k_6 = 0 \\ k_1^2 - k_2^2 + k_3^2 - k_4^2 + k_5^2 - k_6^2 = 0. \end{cases}$$

This resonant subset contains the case $k_5 \neq k_6$. The optimality of the time interval for the case $0 \leq \alpha < 2$ remains open. We can see Grébert–Thomann [20] and Haus–Procesi [21] for instance to analyse the resonant set for 6 indices for the quintic NLS equation.

4. Orbital stability of the traveling plane wave $\mathbf{e}_m$

Consider the following NLS-Szegő equation

$$i\partial_t u + \epsilon^\alpha \partial_x^2 u = \Pi(|u|^2 u), \qquad 0 < \epsilon < 1, \qquad 0 \leq \alpha \leq 2. \tag{4.1}$$

We shall prove at first H^1-orbital stability of the traveling waves $\mathbf{e}_m$, for all $m \in \mathbb{N}$. Then, we study their long time H^s-stability, for all $s \geq 1$.

4.1. The proof of Theorem 1.5. We follow the idea of using conserved quantities mentioned in Gallay–Haragus [9] for equation (4.1).

PROOF. For all $m \in \mathbb{N}$, $0 < \epsilon < 1$ and $0 \leq \alpha \leq 2$, we denote $u(0,x) = e^{imx} + \epsilon f(x)$ with $\|f\|_{H^1} \leq 1$. The NLS-Szegő equation has three conservation laws:

$$\begin{cases} Q(u(t)) := \|u(t)\|_{L^2}^2 = \|u(0)\|_{L^2}^2; \\ P(u(t)) := (Du(t), u(t)) = P(u(0)); \\ E^{\alpha,\epsilon}(u(t)) := \frac{\epsilon^\alpha}{2}\|\partial_x u(t)\|_{L^2}^2 + \frac{1}{4}\|u(t)\|_{L^4}^4 = E^{\alpha,\epsilon}(u(0)), \end{cases}$$

with $D = -i\partial_x$ and $(u,v) := \mathrm{Re}\int_{\mathbb{S}^1} \bar{u}v$. Thus the following quantity is conserved,

$$\begin{aligned} &\frac{\epsilon^\alpha}{2}\|Du(t) - mu(t)\|_{L^2}^2 + \frac{1}{4}\||u(t)|^2 - 1\|_{L^2}^2 \\ =&E^{\alpha,\epsilon}(u(t)) - \epsilon^\alpha m P(u(t)) + \frac{|m|^2\epsilon^\alpha - 1}{2}Q(u(t)) + \frac{1}{4} \\ =&\epsilon^2\int_{\mathbb{S}^1} |\mathrm{Re} f(x)e^{-imx}|^2 \mathrm{d}x + \frac{\epsilon^{2+\alpha}}{2}\|Df - mf\|_{L^2}^2 \\ &+ \epsilon^3\int_{\mathbb{S}^1} |f(x)|^2 \mathrm{Re}(f(x)e^{-imx})\mathrm{d}x + \epsilon^4\|f\|_{L^4}^4 \\ \lesssim_m& \epsilon^2. \end{aligned}$$

Then, we have $\sup_{t\in\mathbb{R}} \|Du(t) - mu(t)\|_{L^2} \lesssim_m \epsilon^{1-\frac{\alpha}{2}}$. Recall that $\mathbf{e}_m(x) = e^{imx}$, then the following estimate holds,

$$\|u(t) - u_m(t)\mathbf{e}_m\|_{H^1}^2 = \sum_{n\neq m}(1+n^2)|u_n(t)|^2 \lesssim_m \|Du(t) - mu(t)\|_{L^2}^2 \lesssim_m \epsilon^{2-\alpha}.$$

We have

$$\begin{aligned} &\inf_{\theta\in\mathbb{R}} \|u(t) - u_m(0)e^{i\theta}\mathbf{e}_m\|_{H^1}^2 \\ =&\|u(t) - u_m(0)e^{i(\arg u_m(t) - \arg u_m(0))}\mathbf{e}_m\|_{H^1}^2 \\ =&(1+m^2)\Big||u_m(t)| - |u_m(0)|\Big|^2 + \|u(t) - u_m(t)\mathbf{e}_m\|_{H^1}^2 \end{aligned}$$

and by the conservation of $\|u(t)\|_{L^2}$, we have

$$\begin{aligned} \Big||u_m(t)| - |u_m(0)|\Big|^2 \leq& \Big||u_m(t)|^2 - |u_m(0)|^2\Big| \\ =&\Big|\sum_{n\neq m}|u_n(0)|^2 - \sum_{n\neq m}|u_n(t)|^2\Big| \\ =&\max(\|u(0) - u_m(0)\mathbf{e}_m\|_{L^2}^2, \quad \|u(t) - u_m(t)\mathbf{e}_m\|_{L^2}^2) \\ \lesssim_m& \epsilon^{2-\alpha}. \end{aligned}$$

Thus $\sup_{t\in\mathbb{R}} \|u(t) - u_m(0)e^{i(\arg u_m(t) - \arg u_m(0))}\mathbf{e}_m\|_{H^1} \lesssim_m \epsilon^{1-\frac{\alpha}{2}}$. The proof can be finished by $u_m(0) = 1 + \epsilon f_m = 1 + \mathcal{O}(\epsilon)$. □

The preceding theorem also holds for the defocusing NLS equation on $\mathbb{T}^d$, for $d = 1,2,3$ (in the energy sub-critical case) with $\mathbb{T} = \mathbb{R}/\mathbb{Z} \simeq \mathbb{S}^1$.(see Gallay–Haragus [9, 10]) We refer to Zhidkov [29 Sect. 3.3] for a detailed analysis of the stability of plane waves.

Remark 4.1. Obtaining the estimate $\sup_{t\in\mathbb{R}} \|u(t)-u_m(t)\mathbf{e}_m\|_{L^\infty} \lesssim_m \epsilon^{1-\frac{\alpha}{2}}$ by the Sobolev embedding $H^1(\mathbb{S}^1)\hookrightarrow L^\infty$, we can also proceed by using the following estimate, which is uniform on x and t,

$$|u(t,x)|^2-1=|u_m(t)|^2-1+\mathcal{O}_m(\epsilon^{1-\frac{\alpha}{2}}).$$

The notation $\mathcal{O}_m$ means that $\sup_{(t,x)\in\mathbb{R}\times\mathbb{S}^1}\Big||u(t,x)|^2-|u_m(t)|^2\Big|\lesssim_m \epsilon^{1-\frac{\alpha}{2}}$. Integrating the preceding term with respect to x, we have

$$||u_m(t)|^2-1|\leq \||u(t)|^2-1\|_{L^2}+\|\mathcal{O}_m(\epsilon^{1-\frac{\alpha}{2}})\|_{L^2}\lesssim_m \epsilon^{1-\frac{\alpha}{2}}$$

Thus $|u_m(t)|=1+\mathcal{O}_m(\epsilon^{1-\frac{\alpha}{2}})$ and $u_m(t)=e^{i\arg u_m(t)}+\mathcal{O}_m(\epsilon^{1-\frac{\alpha}{2}})$. Then we have

$$\sup_{t\in\mathbb{R}}\|u(t)-e^{i\arg u_m(t)}\mathbf{e}_m\|_{H^1}^2\lesssim_m \epsilon^{2-\alpha}. \tag{4.2}$$

Recall that if $z=1+\mathcal{O}(\epsilon)$ then $e^{i\arg z}=1+\mathcal{O}(\epsilon)$.

4.2. Long time H^s-stability. For every $s\geq 1$, we suppose that $\|u(0)-\mathbf{e}_m\|_{H^s}\leq\epsilon$. Thanks to estimate (4.2), we change the variable $u\mapsto v=v^{m,\alpha,\epsilon}(t,x)=\sum_{n\geq 0}v_n(t)e^{inx}\in C^\infty(\mathbb{R}\times\mathbb{S}^1)$ such that

$$u(t,x)=e^{i\arg u_m(t)}(e^{imx}+\epsilon^{1-\frac{\alpha}{2}}v(t,x)) \tag{4.3}$$

to study H^s-stability of plane waves $\mathbf{e}_m$ and we have

$$I_m:=\sup_{0\leq\alpha\leq 2}\sup_{0<\epsilon<1}\sup_{t\in\mathbb{R}}\|v(t)\|_{H^1}<+\infty. \tag{4.4}$$

Proposition 4.2. For every $s\geq 1$, $m\in\mathbb{N}$, $\epsilon\in(0,1)$ and $\alpha\in[0,2)$, if u is smooth and solves (4.1) with $u(0,x)=e^{imx}+\epsilon f(x)$ and $\|f\|_{H^s}\leq 1$, v is defined by formula (4.3), then we have

$$\begin{cases} v_m(t)\in\mathbb{R}, \qquad \forall t\in\mathbb{R},\\ \sup_{t\in\mathbb{R}}|v_m(t)|\lesssim_m \epsilon^{\min(\frac{\alpha}{2},1-\frac{\alpha}{2})},\\ \sup_{t\in\mathbb{R}}|\partial_t v_m(t)|\lesssim_m \epsilon^{1-\frac{\alpha}{2}},\\ \|v(0)\|_{H^s}\lesssim_{m,s}\epsilon^{\frac{\alpha}{2}}.\end{cases}$$

Moreover, there exists a smooth function $\varphi=\varphi_m:\mathbb{R}\to\mathbb{R}/2\pi\mathbb{Z}\simeq\mathbb{S}^1$ and $\epsilon_m^*\in(0,1)$ such that for every $1<2^-<2$, we have

$$\begin{cases}\arg u_m(t)=-(1+m^2\epsilon^\alpha)t+\epsilon^{\min(1,2-\alpha)}\varphi(t),\\ \sup_{0\leq\alpha\leq 2^-}\sup_{0<\epsilon<\epsilon_m^*}\sup_{t\in\mathbb{R}}|\varphi'(t)|<+\infty.\end{cases}$$

The parameter v satisfies the following equation

$$\begin{aligned} &i\partial_t v+\epsilon^\alpha\partial_x^2 v-H_{e^{2imx}}(v)-(1-m^2\epsilon^\alpha+\epsilon^{\min(1,2-\alpha)}\varphi'(t))v\\ =&\epsilon^{\min(\frac{\alpha}{2},1-\frac{\alpha}{2})}\varphi'(t)e^{imx}+\epsilon^{1-\frac{\alpha}{2}}\Pi(e^{-imx}v^2+2e^{imx}|v|^2)+\epsilon^{2-\alpha}\Pi(|v|^2v),\end{aligned} \tag{4.5}$$

where $H_{e^{2imx}}(v):=\Pi[e^{2imx}\bar{v}]$ denotes the Hankel operator of symbol $\mathbf{e}_{2m}$.

Proof. Since $u_m(t)=e^{i\arg u_m(t)}(1+\epsilon^{1-\frac{\alpha}{2}}v_m(t))$, we have $1+\epsilon^{1-\frac{\alpha}{2}}v_m(t)=|u_m(t)|\in\mathbb{R}$. So $v_m(t)\in\mathbb{R}$, for all $t\in\mathbb{R}$. By using the conservation law $\|\cdot\|_{L^2}$ and estimate (4.4), we have

$$1+2\epsilon\mathrm{Re}f_m+\epsilon^2=\|u(0)\|_{L^2}^2=\|u(t)\|_{L^2}^2=1+2\epsilon^{1-\frac{\alpha}{2}}v_m(t)+\epsilon^{2-\alpha}\|v(t)\|_{L^2}^2,$$

which yields that $\sup_{t\in\mathbb{R}} |v_m(t)| \lesssim_m \epsilon^{\min(\frac{\alpha}{2}, 1-\frac{\alpha}{2})}$. Recall that

$$u(0,x) = \sum_{n\geq 0} u_n(0)e^{inx} = e^{imx} + \epsilon f(x).$$

Then we have $u_m(0) = 1 + \epsilon f_m = 1 + \mathcal{O}(\epsilon)$ and $|e^{i\arg u_m(0)} - 1| \lesssim \epsilon$. Thus we have

$$\epsilon^{1-\frac{\alpha}{2}} \|v(0)\|_{H^s} \lesssim (1+m^2)^{\frac{s}{2}} |e^{i\arg u_m(0)} - 1| + \epsilon \|f\|_{H^s} \lesssim_{m,s} \epsilon.$$

We define $\theta(t) := \arg(u_m(t))$. Combing the following two formulas

$$i\partial_t u + \epsilon^\alpha \partial_x^2 u = e^{i\theta(t)} \left[\epsilon^{1-\frac{\alpha}{2}} (i\partial_t v + \epsilon^\alpha \partial_x^2 v - \theta'(t) v) - (m^2\epsilon^\alpha + \theta'(t))e^{imx}\right]$$

$$\Pi[|u|^2 u] = e^{i\theta(t)} \Big[e^{imx} + \epsilon^{1-\frac{\alpha}{2}} (2v + \Pi(e^{2imx}\overline{v})) + \epsilon^{2-\alpha} \Pi(e^{-imx}v^2 + 2e^{imx}|v|^2) + \epsilon^{3(1-\frac{\alpha}{2})} \Pi(|v|^2 v) \Big]$$

we obtain that

$$\begin{aligned} &\epsilon^{1-\frac{\alpha}{2}} [i\partial_t v + \epsilon^\alpha \partial_x^2 v - H_{e^{2imx}}(v) - (2 + \theta'(t))v] \\ =&(1 + m^2\epsilon^\alpha + \theta'(t))e^{imx} + \epsilon^{2-\alpha}\Pi(e^{-imx}v^2 + 2e^{imx}|v|^2) + \epsilon^{3(1-\frac{\alpha}{2})}\Pi(|v|^2 v), \end{aligned} \tag{4.6}$$

where $H_{e^{2imx}}(v) := \Pi[e^{2imx}\bar{v}]$. The Fourier mode $v_m(t)$ satisfies the following equation

$$\begin{aligned} &\epsilon^{1-\frac{\alpha}{2}} \left[i\partial_t v_m(t) - m^2\epsilon^\alpha v_m(t) - \overline{v_m(t)} - (2+\theta'(t))v_m(t) \right] \\ =&1 + m^2\epsilon^\alpha + \theta'(t) + \epsilon^{2-\alpha}\Pi(e^{-imx}v^2 + 2e^{imx}|v|^2)_m(t) + \epsilon^{3(1-\frac{\alpha}{2})}\Pi(|v|^2 v)_m(t). \end{aligned}$$

Estimate (4.4) yields that

$$\sup_{t\in\mathbb{R}} |\epsilon^{2-\alpha}\Pi(e^{-imx}v^2 + 2e^{imx}|v|^2)_m(t) + \epsilon^{3(1-\frac{\alpha}{2})}\Pi(|v|^2 v)_m(t)| \lesssim_m \epsilon^{2-\alpha}.$$

Thus, we have

$$\epsilon^{1-\frac{\alpha}{2}} \left[i\partial_t v_m(t) - (3 + m^2\epsilon^\alpha + \theta'(t))v_m(t) \right] = 1 + m^2\epsilon^\alpha + \theta'(t) + \mathcal{O}_m(\epsilon^{2-\alpha}). \tag{4.7}$$

The imaginary part and the real part of (4.7) give respectively the two following estimates:

$$\sup_{t\in\mathbb{R}} |\partial_t v_m(t)| \lesssim_m \epsilon^{1-\frac{\alpha}{2}};$$

$$\begin{aligned} 1 + m^2\epsilon^\alpha + \theta'(t) &= \frac{-2\epsilon^{1-\frac{\alpha}{2}} v_m(t) + \mathcal{O}_m(\epsilon^{2-\alpha})}{1 + \epsilon^{1-\frac{\alpha}{2}} v_m(t)} \\ &= \frac{\mathcal{O}_m(\epsilon^{\min(1,2-\alpha)})}{1 + \mathcal{O}_m(\epsilon^{\min(1,2-\alpha)})} = \mathcal{O}_m(\epsilon^{\min(1,2-\alpha)}). \end{aligned}$$

for all $0 < \epsilon \ll 1$. Then we define $\varphi(t) := \frac{(1+m^2\epsilon^\alpha)t + \theta(t)}{\epsilon^{\min(1,2-\alpha)}}$. Consequently, there exists $\epsilon_m^* \in (0,1)$ such that

$$\begin{cases} \arg u_m(t) = -(1 + m^2\epsilon^\alpha)t + \epsilon^{\min(1,2-\alpha)}\varphi(t) \\ \sup_{0\leq\alpha\leq 2^-} \sup_{0<\epsilon<\epsilon_m^*} \sup_{t\in\mathbb{R}} |\varphi'(t)| < +\infty. \end{cases}$$

We replace $\theta'(t)$ by $-1 - m^2\epsilon^\alpha + \epsilon^{\min(1,2-\alpha)}\varphi'(t)$ in (4.6) and we obtain (4.5). □

4.2.1. *Proof of Proposition* 1.7. For every $n \in \mathbb{N}$, we define the projector $\mathbb{P}_n : L^2_+ \to L^2_+$ such that

$$\mathbb{P}_n(\sum_{k\geq 0} v_k e^{ikx}) = \sum_{j=0}^{n} v_k e^{ikx}.$$

Now we prove Proposition 1.7 by using a bootstrap argument.

Proof. At the beginning, we suppose that $u(0) \in C^\infty_+$. In the general case $u(0) \in H^s_+$, the proof can be completed by using the continuity of the mapping $u(0) \mapsto u$ from H^s_+ to $C([-\frac{b_{s,m}}{\epsilon^{1-\frac{\alpha}{2}}}, \frac{b_{s,m}}{\epsilon^{1-\frac{\alpha}{2}}}], H^s_+)$. We use the same transformation $u \mapsto v$ as (4.3). Proposition 4.2 yields that there exists $A_{m,s} \geq 1$ such that $\|v(0)\|_{H^s} \lesssim_{m,s} \epsilon^{\frac{\alpha}{2}} \leq A_{m,s}$. By using estimate (4.4), we have

$$\sup_{\epsilon\in(0,1)} \sup_{t\in\mathbb{R}} \|\mathbb{P}_{2m}(v(t))\|_{H^s} \leq (1+4m^2)^{\frac{s}{2}} I_m$$

We define that $L_{m,s} := \max(2(1+4m^2)^{\frac{s}{2}} I_m, 2A_{m,s}+1)$ and

$$T := \sup\{t > 0 : \sup_{0\leq\tau\leq t} \|v(\tau)\|_{H^s} \leq 2L_{m,s}\}.$$

Rewrite equation (4.5) with Fourier modes and we have

$$i\partial_t v_n - (1 + (n^2 - m^2)\epsilon^\alpha + \epsilon^{\min(1,2-\alpha)}\varphi'(t))v_n = \epsilon^{1-\frac{\alpha}{2}}[Z(v(t))]_n, \qquad \forall n \geq 2m+1,$$

with $Z(v) = \sum_{n\geq 0}[Z(v)]_n e^{inx} = \Pi(e^{-imx}v^2 + 2e^{imx}|v|^2) + \epsilon^{1-\frac{\alpha}{2}}\Pi(|v|^2 v)$. Then we have

$$\|Z(v) - \mathbb{P}_{2m}(Z(v))\|_{H^s} \lesssim_s \|v\|^2_{H^s} + \epsilon^{1-\frac{\alpha}{2}}\|v\|^3_{H^s}. \tag{4.8}$$

Then we have

$$\begin{aligned}
|\partial_t \|v(t) - \mathbb{P}_{2m}(v(t))\|^2_{H^s}| \leq & 2\epsilon^{1-\frac{\alpha}{2}} \sum_{n\geq 2m+1} (1+n^2)^s |v_n(t)||[Z(v(t))]_n| \\
\leq & 2\epsilon^{1-\frac{\alpha}{2}} \|v(t)\|_{H^s} \|Z(v(t)) - \mathbb{P}_{2m}(Z(v(t)))\|_{H^s} \\
\lesssim_s & \epsilon^{1-\frac{\alpha}{2}} \|v(t)\|^3_{H^s} + \epsilon^{2-\alpha}\|v(t)\|^4_{H^s}.
\end{aligned}$$

For all $t \in [0, T]$, we have

$$\begin{aligned}
\|v(t)\|^2_{H^s} = & \|\mathbb{P}_{2m}v(t)\|^2_{H^s} + \|v(t) - \mathbb{P}_{2m}(v(t))\|^2_{H^s} \\
\leq & (1+4m^2)^s I_m^2 + C_s(\epsilon^{1-\frac{\alpha}{2}}\|v(t)\|^3_{H^s} + \epsilon^{2-\alpha}\|v(t)\|^4_{H^s})t + \|v(0)\|^2_{H^s} \\
\leq & \frac{1}{4}L^2_{m,s} + 32C_s L^4_{m,s}\epsilon^{1-\frac{\alpha}{2}} t + A^2_{m,s}.
\end{aligned}$$

Define $b_{m,s} = \frac{1}{64C_s L^2_{m,s}}$ and we have $\|v(t)\|_{H^s} \leq L_{m,s}$, for all $t \in [0, \frac{b_{s,m}}{\epsilon^{1-\frac{\alpha}{2}}}]$. The case $t < 0$ is similar. □

4.2.2. *Homological equation.* We try to improve Proposition 1.7 and get a longer time interval in which the solution v is still bounded by $\mathcal{O}_{m,s}(1)$, by using the Birkhoff normal form transformation. Recall the symplectic form $\omega(u,v) = \operatorname{Im}\int_{\mathbb{S}^1} u\overline{v}\frac{\mathrm{d}\theta}{2\pi}$ on the energy space H^1_+ and the Poisson bracket for two smooth real-valued functionals $F, G : C^\infty_+ \to \mathbb{R}$

$$\{F, G\}(v) = \frac{2}{i}\sum_{k\geq 0}(\partial_{\overline{v}_k}F\partial_{v_k}G - \partial_{\overline{v}_k}G\partial_{v_k}F)(v),$$

for all $v = \sum_{k\geq 0} v_k e^{ikx} \in C_+^\infty$. For all $0 \leq \alpha < 2$ and $0 < \epsilon \ll 1$, equation (4.5) is a non autonomous Hamiltonian equation. Its energy functional is

$$\begin{aligned}&\mathcal{H}^{m,\alpha,\epsilon}(t,v)\\ =&\mathcal{H}_0^{m,\alpha,\epsilon}(v) + \epsilon^{1-\frac{\alpha}{2}}\mathcal{H}_1^m(v) + \frac{\epsilon^{2-\alpha}}{4}\mathcal{N}_4(v) + \epsilon^{\min(\frac{\alpha}{2},1-\frac{\alpha}{2})}\varphi'(t)(\mathcal{L}_m(v) + \frac{\epsilon^{1-\frac{\alpha}{2}}}{2}\mathcal{N}_2(v)),\end{aligned}$$

with

$$\begin{cases}\mathcal{H}_0^{m,\alpha,\epsilon}(v) = \frac{\epsilon^\alpha}{2}\|\partial_x v\|_{L^2}^2 + \frac{1-m^2\epsilon^\alpha}{2}\|v\|_{L^2}^2 + \frac{1}{2}\int_{\mathbb{S}^1}\mathrm{Re}(e^{2imx}\overline{v}^2),\\ \mathcal{H}_1^m(v) = \mathrm{Re}(\int_{\mathbb{S}^1} e^{-imx}|v|^2 v),\\ \mathcal{N}_p(v) = \|v\|_{L^p}^p, \qquad p=2 \quad \text{or} \quad 4,\\ \mathcal{L}_m(v) = \mathrm{Re} v_m.\end{cases}$$

We want to cancel all the high frequencies in the term $\mathcal{H}_1^m(v)$ by composing $\mathcal{H}^{m,\alpha,\epsilon}$ with the Hamiltonian flow of some auxiliary functional $\epsilon^{1-\frac{\alpha}{2}}\mathcal{F}_m$. In order to get the appropriate $\mathcal{F}_m$, we need to solve the following homological system

$$\begin{cases}\{\mathcal{F}_m, \mathcal{H}_0^{m,\alpha,\epsilon}\}(v) + \mathcal{H}_1^m(v) = \mathcal{R}_m(v)\\ \{\mathcal{F}_m, \mathcal{L}_m\}(v) = -\tilde{\mathcal{N}}_2^m(v) := -\sum_{n\geq 2m+1}|v_n|^2\end{cases}$$

such that $\mathcal{R}_m$ depends only on finitely many Fourier modes of v. The remaining coefficient in front of $\epsilon^{1-\frac{\alpha}{2}}$ would be $\mathcal{R}_m + \varphi'(t)\epsilon^{\min(\frac{\alpha}{2},1-\frac{\alpha}{2})}(-\tilde{\mathcal{N}}_2^m + \frac{\mathcal{N}_2}{2})$. One can prove the following proposition.(see also Proposition 4.4 and Appendix for the proof in the special case $\alpha = 0$)

Proposition 4.3. For every $m \in \mathbb{N}$, the homogenous functional $\mathcal{F}_m$ of degree 3 is defined as

$$\mathcal{F}_m(v) = \sum_{\substack{j-l+k=m\\ j,k,l\in\mathbb{N}}} \mathrm{Re}(a_{j,l,k} v_j \overline{v}_l v_k), \qquad \forall v \in \bigcup_{n\geq 0}\mathbb{P}_n(C_+^\infty),$$

for some $a_{k,l,j} = a_{j,l,k} \in \mathbb{C}$. Then we have the following formula

$$\{\mathcal{F}_m, \mathcal{H}_0^{m,\alpha,\epsilon}\}(v) + \mathcal{H}_1^m(v) = \mathrm{Re}\left(\mathrm{Reson}^{low}(v) + \mathrm{Reson}^{\geq 2m+1}(v)\right),$$

where

$$\begin{aligned}&\mathrm{Reson}^{low}(v)\\ =&\sum_{0\leq j,k\leq 2m} c_{j,j+k-m,k} v_j \overline{v}_{j+k-m} v_k\\ &-\sum_{\substack{j-l+k=m\\ j,l,k\in\mathbb{N}, l\leq 2m}} i a_{j,l,k} v_j v_k\left[(1+(l^2-m^2)\epsilon^\alpha)\overline{v}_l + v_{2m-l}\right],\end{aligned}$$

for some $c_{j,j+k-m,k} = c_{k,j+k-m,j} \in \mathbb{C}$ and

$$\begin{aligned}
&\mathrm{Reson}^{\geq 2m+1}(v)\\
&= \sum_{k\geq 2m+1}\sum_{j=0}^{m-1} 2(-i\overline{a}_{2m-j,m+k-j,k}\\
&\quad + (1+2(m-j)(k-j)\epsilon^{\alpha})ia_{j,k,k+m-j}+1)v_j\overline{v}_k v_{k+m-j}\\
&+ \sum_{k\geq 2m+1}\sum_{j=0}^{m-1} 2((1-2(k-m)(m-j)\epsilon^{\alpha})ia_{2m-j,m+k-j,k}\\
&\quad - i\overline{a}_{j,k,k+m-j}+1)v_{2m-j}\overline{v}_{m+k-j}v_k\\
&+ \sum_{k\geq 2m+1} (2(ia_{m,k,k}-i\overline{a}_{m,k,k}+1)v_m|v_k|^2\\
&\quad + ((1-2(k-m)^2\epsilon^{\alpha})ia_{k,2k-m,k}+1)v_k^2\overline{v}_{2k-m})\\
&+ \sum_{k\geq 2m+1}\sum_{j\geq k+1} 2((1-2(j-m)(k-m)\epsilon^{\alpha})ia_{k,j+k-m,j}+1)v_k\overline{v}_{j+k-m}v_j.
\end{aligned}$$

The term Reson^{low} depends only on the small Fourier modes $v_1, v_2, \cdots, v_{3m}$. We try to find a bounded sequence $(a_{j,l,k})_{j-l+k=m}$ such that $\mathrm{Reson}^{\geq 2m+1} = 0$ in order to cancel the term $\mathcal{H}_1^m$. However, the coefficient $(1-2(j-m)(k-m)\epsilon^{\alpha})$ in front of the parameter $a_{k,j+k-m,j}$ may have an arbitrarily small absolute value if $\alpha > 0$. Such sequence does not exist if $\epsilon^{-\alpha} \in 2\mathbb{N}\bigcap[2(m+1)^2, +\infty)$.

We suppose that $\epsilon^{-\alpha} \notin \mathbb{Q}$, then $\mathrm{Reson}^{\geq 2m+1} = 0$ is equivalent to a linear system, which has a unique solution

(4.9)
$$\begin{cases}
a_{2m-j,m+k-j,k} = a_{k,m+k-j,2m-j} = \frac{i(k-j)}{(m-j)(1-2(k-m)(k-j)\epsilon^{\alpha})}, & \forall 0 \leq j \leq m-1,\\
a_{j,k,k+m-j} = a_{k+m-j,k,j} = \frac{i(m-k)}{(m-j)(1-2(k-m)(k-j)\epsilon^{\alpha})}, & \forall 0 \leq j \leq m-1,\\
a_{m,k,k} = a_{k,k,m} = \frac{i}{2},\\
a_{k,k+j-m,j} = a_{j,k+j-m,k} = \frac{i}{1-2(j-m)(k-m)\epsilon^{\alpha}}, & \forall j \geq 2m+1,
\end{cases}$$

for all $k \geq 2m+1$. In the case $m = 0$, (4.9) has only the last two formulas. When $\alpha > 0$, the sequence $(a_{j,l,k})_{j-l+k=m}$ can be arbitrarily large, for $0 < \epsilon \ll 1$. We suppose that ϵ^{α} is an irrational algebraic number of degree $d \geq 2$. Then we have the Liouville estimate [23]

$$|a_{j,j+k-m,k}| \leq \frac{1}{c_{\epsilon,\alpha}}(2(j-m)(k-m))^{d-1}, \qquad \forall j,k \geq 2m+1,$$

which loses the regularity of v in the estimates of $X_{F_m}(v)$. It is difficult to find the same kind of estimate for the transcendental numbers, which can preserve the regularity. So we return to the case $\alpha = 0$.

4.3. Long time H^s-stability in the case $\alpha = 0$. For $\alpha = 0$ and every $m \in \mathbb{N}$ and $s \geq 1$, assume that u is the smooth solution of the NLS-Szegő equation

$$i\partial_t u + \partial_x^2 u = \Pi(|u|^2 u), \qquad u(0,x) = e^{imx} + \epsilon f(x), \qquad \|f\|_{H^s} \leq 1,$$

and $u(t,x) = e^{i\arg u_m(t)}(e^{imx}+\epsilon v(t,x))$. Then v is solves the following Hamiltonian equation

$$\begin{aligned}&i\partial_t v + \partial_x^2 v - H_{e^{2imx}}(v) - (1-m^2+\epsilon\varphi'(t))v\\ =&\varphi'(t)e^{imx} + \epsilon\Pi(e^{-imx}v^2 + 2e^{imx}|v|^2) + \epsilon^2\Pi(|v|^2 v).\end{aligned}$$

Its energy functional is

$$\mathcal{H}^{m,\epsilon}(t,v) = H_0^m(v) + \varphi'(t)\mathcal{L}_m(v) + \epsilon(\mathcal{H}_1^m(v) + \frac{\varphi'(t)}{2}\mathcal{N}_2(v)) + \frac{\epsilon^2}{4}\mathcal{N}_4(v),$$

with

$$\begin{cases}\mathcal{H}_0^m(v) = \frac{1}{2}\|\partial_x v\|_{L^2}^2 + \frac{1-m^2}{2}\|v\|_{L^2}^2 + \frac{1}{2}\int_{\mathbb{S}^1}\mathrm{Re}(e^{2imx}\overline{v}^2)\mathrm{d}x,\\ \mathcal{L}_m(v) = \mathrm{Re}v_m,\\ \mathcal{H}_1^m(v) = \mathrm{Re}(\int_{\mathbb{S}^1} e^{-imx}|v|^2 v)\mathrm{d}x,\\ \mathcal{N}_p(v) = \|v\|_{L^p}^p, \qquad p=2 \quad \text{or} \quad 4.\end{cases}$$

We define $\tilde{\mathcal{N}}_2^m(v) := \|v - \mathbb{P}_{2m}v\|_{L^2}^2 = \sum_{n\geq 2m+1}|v_n|^2$ and the following proposition holds.

Proposition 4.4. For every $s > \frac{1}{2}$ and $m\in\mathbb{N}$, there exists a sequence $(a_{j,l,k})_{j-l+k=m}$ such that $a_{j,l,k} = a_{k,l,j}$, $\sup_{m\geq 1}\sup_{j-l+k=m}|a_{j,l,k}| = \frac{1}{2}$ and the functional $\mathcal{F}_m : H_+^s \to \mathbb{R}$, defined by

$$\mathcal{F}_m(v) = \sum_{\substack{j-l+k=m\\ j,k,l\in\mathbb{N}}}\mathrm{Re}(a_{j,l,k}v_j\overline{v}_l v_k), \qquad \forall v\in C_+^\infty,$$

satisfies that $\{\mathcal{F}_m, \mathcal{L}_m\} = -\tilde{\mathcal{N}}_2^m$ and $\mathcal{R}_m := \{\mathcal{F}_m, \mathcal{H}_0^m\} + \mathcal{H}_1^m$ is a finite sum of the Fourier modes $v_1, \cdots, v_{3m}$. Moreover, for all $v, h\in H_+^s$, we have

$$\|X_{\mathcal{F}_m}(v)\|_{H^s} \lesssim_{m,s} \|v\|_{H^s}^2, \qquad \|\mathrm{d}X_{\mathcal{F}_m}(v)h\|_{H^s} \lesssim_{m,s} \|v\|_{H^s}\|h\|_{H^s}.$$

Proof. For the convenience of the reader, the detailed calculus for $\mathcal{R}_m = \{\mathcal{F}_m, \mathcal{H}_0^m\} + \mathcal{H}_1^m$ and formula (4.9) in the case $\alpha = 0$ are postponed in Appendix. We define $a_{j,j+k-m,k} = 0$, for all $0\leq j,k\leq 2m$ and $a_{n,m+1+n,2m+1} = a_{2m+1,m+1+n,n} = 0$, for all $0\leq n\leq m-1$. Combing Proposition 4.3 and (4.9) with $\alpha = 0$, we have

$$\begin{cases}\mathrm{Reson}^{\geq 2m+1}(v)\Big|_{\alpha=0} = 0, \qquad \mathrm{Re}\left(\mathrm{Reson}^{low}(v)\Big|_{\alpha=0}\right) = \mathcal{R}_m(v),\\ \{\mathcal{F}_m, \mathcal{L}_m\}(v) = 2\mathrm{Im}(\sum_{k\geq 0}\overline{a}_{m,k,k}|v_k|^2 + \frac{1}{2}\sum_{j+k=2m}a_{j,m,k}v_j v_k) = -\sum_{n\geq 2m+1}|v_n|^2.\end{cases}$$

By (4.9) with $\alpha = 0$, we have $|a_{j,j+k-m,k}| \leq \frac{1}{2}$, for all $j,k\geq 0$. By the definition of $\mathcal{F}_m$, we have

$$\begin{cases}[\widehat{X_{\mathcal{F}_m}(v)}](n) = -2i\sum_{k-l+n=m}\overline{a}_{k,l,n}\overline{v}_k v_l - i\sum_{k-n+l=m}a_{k,n,l}v_k v_l\\ [\widehat{\mathrm{d}X_{\mathcal{F}_m}(v)h}](n) = -2i\sum_{k-l+n=m}\overline{a}_{k,l,n}(\overline{v}_k h_l + v_l\overline{h}_k) - 2i\sum_{k-n+l=m}a_{k,n,l}v_k h_l,\end{cases}$$

for all $n\geq 0$. The last two estimates are obtained by Young's convolution inequality for $l^1 * l^2 \hookrightarrow l^2$.

□

4.3.1. *The Birkhoff normal form.* Set $\chi_\sigma^m := \exp(\epsilon\sigma X_{\mathcal{F}_m})$ the Hamiltonian flow of $\epsilon\mathcal{F}_m$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\chi_\sigma^m(v) = \epsilon X_{\mathcal{F}_m}(\chi_\sigma^m(v)), \qquad \chi_0^m(v) = v.$$

We perform the canonical transformation $\Psi_{m,\epsilon} := \chi_1^m$. We want to reduce the energy functional $\mathcal{H}^{m,\epsilon}$ to the following normal form

$$\mathcal{H}^{m,\epsilon}(t) \circ \Psi_{m,\epsilon} = \mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m(v) + \epsilon\left(\mathcal{R}_m + \varphi'(t)(-\tilde{\mathcal{N}}_2^m + \frac{\mathcal{N}_2}{2})\right) + \mathcal{O}(\epsilon^2).$$

Since $\mathcal{R}_m$ depends only on low-frequency Fourier modes $v_1, \cdots, v_{3m}$, the high-frequency filtering H^s-norm of $\Psi_{m,\epsilon}^{-1}(v)$ is appropriately estimated by the Birkhoff normal form transformation. The estimate of $\|\mathbb{P}_{3m}(v)\|_{H^s}$ is given by (4.4). The next lemma will give the local existence of χ_σ^m, for $|\sigma| \leq 1$ and estimate the difference between v and $\Psi_{m,\epsilon}^{-1}(v)$.

LEMMA 4.5. For every $s > \frac{1}{2}$ and $m \in \mathbb{N}$, there exist two constants $\gamma_{m,s}, C_{m,s} > 0$ such that for all $v \in H_+^s$, if $\epsilon\|v\|_{H^s} \leq \gamma_{m,s}$, then $\chi_\sigma^m(v)$ is well defined on the interval $[-1,1]$ and the following estimates hold:

$$\begin{aligned}
&\sup_{\sigma\in[-1,1]} \|\chi_\sigma^m(v)\|_{H^s} \leq 2\|v\|_{H^s},\\
&\sup_{\sigma\in[-1,1]} \|\chi_\sigma^m(v) - v\|_{H^s} \leq C_{m,s}\epsilon\|v\|_{H^s}^2,\\
&\|\mathrm{d}\chi_\sigma^m(v)\|_{B(H^s)} \leq \exp(C_{m,s}\epsilon\|v\|_{H^s}|\sigma|), \qquad \forall\sigma \in [-1,1].
\end{aligned}$$

The proof is based on a bootstrap argument, which is similar to Lemma 3.3, given by Lemma 3.1 with $q = 2$. The energy functional $\mathcal{H}^{m,\epsilon}$ is reduced to the normal form in the following lemma.

LEMMA 4.6. For all $s > \frac{1}{2}$, $m \in \mathbb{N}$ and $0 < \epsilon < \epsilon_m^*$, there exists a smooth mapping $\mathcal{Y}_m : \mathbb{R} \times H_+^s \longrightarrow H_+^s$ and a constant $C'_{m,s} > 0$ such that for all $t \in \mathbb{R}$, we have

$$X_{\mathcal{H}^{m,\epsilon}(t)\circ\Psi_{m,\epsilon}} = X_{\mathcal{H}_0^m} + \varphi'(t)X_{\mathcal{L}_m} + \epsilon\left(X_{\mathcal{R}_m} + \varphi'(t)(-X_{\tilde{\mathcal{N}}_2^m} + \frac{1}{2}X_{\mathcal{N}_2})\right) + \epsilon^2\mathcal{Y}_m(t),$$

and $\sup_{t\in\mathbb{R}} \|\mathcal{Y}_m(t,v)\|_{H^s} \leq C'_{m,s}\|v\|_{H^s}^2(1 + \|v\|_{H^s})$, for all $v \in H_+^s$ such that $\epsilon\|v\|_{H^s} \leq \gamma_{m,s}$. Set $w(t) := \Psi_{m,\epsilon}^{-1}(v(t))$, then we have

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|w(t) - \mathbb{P}_{3m}(w(t))\|_{H^s}^2\right| \leq C'_{m,s}\epsilon^2\|w(t)\|_{H^s}^3(1 + \|w(t)\|_{H^s}), \tag{4.10}$$

if $\epsilon\|w(t)\|_{H^s} \leq \gamma_{m,s}$.

PROOF. For every $t \in \mathbb{R}$, we expand the energy $\mathcal{H}^{m,\epsilon}(t) \circ \Psi_{m,\epsilon} = \mathcal{H}^{m,\epsilon}(t) \circ \chi_1^m$ with Taylor's formula at time $\sigma = 1$ around 0. Since $\chi_0^m = \mathrm{Id}_{H_+^s}$, we have

$$
\begin{aligned}
&(\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m) \circ \chi_1^m \\
=&\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m + \frac{\mathrm{d}}{\mathrm{d}\sigma}[(\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m) \circ \chi_\sigma^m]|_{\sigma=0} \\
&+ \int_0^1 (1-\sigma)\frac{\mathrm{d}^2}{\mathrm{d}\sigma^2}[(\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m) \circ \chi_\sigma^m]\mathrm{d}\sigma \\
=&\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m + \epsilon\{\mathcal{F}_m, \mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m\} \\
&+ \epsilon^2 \int_0^1 (1-\sigma)\{\mathcal{F}_m, \{\mathcal{F}_m, \mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m\}\} \circ \chi_\sigma^m \mathrm{d}\sigma
\end{aligned}
$$

and

$$
\begin{aligned}
\left(\mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2\right) \circ \chi_1^m =&\mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2 + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\sigma}[(\mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2) \circ \chi_\sigma^m]\mathrm{d}\sigma \\
=&\mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2 + \epsilon \int_0^1 \{\mathcal{F}_m, \mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2\} \circ \chi_\sigma^m \mathrm{d}\sigma.
\end{aligned}
$$

Since $\mathcal{F}_m$ solves the homological system $\begin{cases} \{\mathcal{F}_m, \mathcal{H}_0^m\} + \mathcal{H}_1^m = \mathcal{R}_m \\ \{\mathcal{F}_m, \mathcal{L}_m\} + \tilde{\mathcal{N}}_2^m = 0 \end{cases}$ in Proposition 4.4, we have

$$
\begin{aligned}
&\mathcal{H}^{m,\epsilon}(t) \circ \chi_1^m \\
=&\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m + \epsilon\left(\{\mathcal{F}_m, \mathcal{H}_0^m\} + \mathcal{H}_1^m + \varphi'(t)(\{\mathcal{F}_m, \mathcal{L}_m\} + \frac{1}{2}\mathcal{N}_2)\right) \\
&+ \epsilon^2\Bigg[\int_0^1 \{\mathcal{F}_m, (1-\sigma)\{\mathcal{F}_m, \mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m\} \\
&+ \mathcal{H}_1^m + \frac{\varphi'(t)}{2}\mathcal{N}_2\} \circ \chi_\sigma^m \mathrm{d}\sigma + \frac{\mathcal{N}_4 \circ \chi_1^m}{4}\Bigg] \\
=&\mathcal{H}_0^m + \varphi'(t)\mathcal{L}_m + \epsilon\left(\mathcal{R}_m + \varphi'(t)(-\tilde{\mathcal{N}}_2^m + \frac{1}{2}\mathcal{N}_2)\right) \\
&+ \epsilon^2\left[\int_0^1 \mathcal{G}_m(t,\sigma) \circ \chi_\sigma^m \mathrm{d}\sigma + \frac{\mathcal{N}_4 \circ \chi_1^m}{4}\right],
\end{aligned}
$$

where $\mathcal{G}_m(t,\sigma) = \{\mathcal{F}_m, (1-\sigma)\mathcal{R}_m + \sigma\mathcal{H}_1^m + \varphi'(t)((\sigma-1)\tilde{\mathcal{N}}_2^m + \frac{1}{2}\mathcal{N}_2)\}$. We set

$$
\mathcal{Y}_m(t,v) := \int_0^1 X_{\mathcal{G}_m(t,\sigma)\circ\chi_\sigma^m}(v)\mathrm{d}\sigma + \frac{1}{4}X_{\mathcal{N}_4\circ\chi_1^m}(v),
$$

then we get

$$
X_{\mathcal{H}^{m,\epsilon}(t)\circ\chi_1^m} = X_{\mathcal{H}_0^m} + \varphi'(t)X_{\mathcal{L}_m} + \epsilon\left(X_{\mathcal{R}_m} + \varphi'(t)(-X_{\tilde{\mathcal{N}}_2^m} + \frac{1}{2}X_{\mathcal{N}_2})\right) + \epsilon^2\mathcal{Y}_m(t).
$$

Since $\mathcal{F}_m$, $\mathcal{H}_m^1$ and $\mathcal{R}_m$ are homogeneous series of order 3 with uniformly bounded coefficients, $\mathcal{N}_2$ and $\tilde{\mathcal{N}}_2^m$ are homogeneous series of order 2 with uniformly bounded coefficients, we have

$$
\begin{cases} \|X_{\{\mathcal{F}_m,\mathcal{H}_m^1\}}(v)\|_{H^s} + \|X_{\{\mathcal{F}_m,\mathcal{R}_m\}}(v)\|_{H^s} \lesssim_s \|v\|_{H^s}^3, \\ \|X_{\{\mathcal{F}_m,\mathcal{N}_2\}}(v)\|_{H^s} + \|X_{\{\mathcal{F}_m,\tilde{\mathcal{N}}_2^m\}}(v)\|_{H^s} \lesssim_s \|v\|_{H^s}^2, \end{cases}
$$

because for $\mathcal{J}_m(v) = \sum_{j-l+k=m} \mathrm{Re}(b_{j,l,k} v_j \overline{v}_l v_k)$ with $\sup_{j-l+k=m} |b_{j,l,k}| < +\infty$, we have

$$\begin{aligned}
&\{\mathcal{F}_m, \mathcal{J}_m\}(v) \\
=&4\mathrm{Im} \sum_{n\geq 0} \partial_{\overline{v}_n} \mathcal{F}_m(v) \partial_{v_n} \mathcal{J}_m(v) \\
=& \sum_{k_1+k_2=l_1+l_2} \mathrm{Im}(4\overline{a}_{k_1,l_1,m+l_1-k_1} b_{l_2,k_2,m+k_2-l_2} \\
&+ a_{l_1,l_1+l_2-m,l_2} \overline{b}_{k_1,k_1+k_2-m,k_2}) \overline{v}_{k_1} \overline{v}_{k_2} v_{l_1} v_{l_2} \\
&+ \sum_{k_1+k_2+l_1-l_2=2m} 2\mathrm{Im}(a_{k_1,k_1+l_1-m,l_1} b_{k_2,l_2,m+l_2-k_2} \\
&- a_{k_2,l_2,m+l_2-k_2} b_{k_1,k_1+l_1-m,l_1}) v_{k_1} v_{k_2} v_{l_1} \overline{v}_{l_2}
\end{aligned}$$

and $\{\mathcal{F}_m, \mathcal{N}_2\}(v) = -2\mathrm{Im}(\sum_{j-l+k=m} a_{j,l,k} v_j \overline{v}_l v_k)$. Recall that

$$\sup_{0<\epsilon<\epsilon_m^*} \sup_{t\in\mathbb{R}} |\varphi'(t)| < +\infty$$

and $X_{\mathcal{N}_4}(v) = -4i\Pi(|v|^2 v)$, then we have

$$\sup_{0\leq\sigma\leq 1} \sup_{t\in\mathbb{R}} \|X_{\mathcal{G}_m(t,\sigma)}(v)\|_{H^s} + \|X_{\mathcal{N}_4}(v)\|_{H^s} \lesssim_{m,s} \|v\|_{H^s}^2 (1 + \|v\|_{H^s}).$$

By using Lemma 4.5 and (3.11), for all $v \in H^s_+$ such that $\epsilon\|v\|_{H^s} \leq \gamma_{m,s}$, we have

$$\begin{aligned}
\sup_{0\leq\sigma\leq 1} \sup_{t\in\mathbb{R}} \|X_{\mathcal{G}_m(t,\sigma)\circ\chi_\sigma^m}(v)\|_{H^s} &\leq \sup_{0\leq\sigma\leq 1} \sup_{t\in\mathbb{R}} \|\mathrm{d}\chi_{-\sigma}^m(\chi_\sigma^m(v))\|_{B(H^s)} \|X_{\mathcal{G}_m(t,\sigma)}(\chi_\sigma^m(v))\|_{H^s} \\
&\lesssim_{m,s} \sup_{0\leq\sigma\leq 1} e^{C_{m,s}\epsilon\|\chi_\sigma^m(v)\|_{H^s}} \|\chi_\sigma^m(v)\|_{H^s}^2 (1 + \|\chi_\sigma^m(v)\|_{H^s}) \\
&\lesssim_{m,s} \|v\|_{H^s}^2 (1 + \|v\|_{H^s})
\end{aligned}$$

and $\sup_{t\in\mathbb{R}} \|X_{\mathcal{N}_4\circ\chi_1^m}(v)\|_{H^s} \lesssim_{m,s} \|v\|_{H^s}^2 (1 + \|v\|_{H^s})$. Then we obtain the estimate of $\mathcal{Y}_m$.

Since $w(t) = \chi_{-1}^m(v(t))$, $w = \sum_{n\geq 0} w_n e^{inx}$ solves the following infinite dimensional Hamiltonian system of Fourier modes:

$$i\partial_t w_n(t) = (1 + n^2 - m^2 - \epsilon\varphi'(t)) w_n(t) + i\epsilon^2 \widehat{\mathcal{Y}_m(t, w(t))}(n), \quad \forall n \geq 3m+1.$$

because for all $n \geq 3m+1$, we have

$$\begin{cases} \widehat{H_{e^{2imx}}(w(t))}(n) = \widehat{X_{\mathcal{R}_m}(w(t))}(n) = \widehat{X_{\mathcal{L}_m}(w(t))}(n) = 0, \\ \widehat{X_{\mathcal{N}_2}(w(t))}(n) = \widehat{X_{\tilde{\mathcal{N}}_2^m}(w(t))}(n) = -2iw_n(t). \end{cases}$$

Consequently, if $\epsilon\|w(t)\|_{H^s} \leq \gamma_{m,s}$, then we have

$$\begin{aligned}
\left|\partial_t \|w(t) - \mathbb{P}_{3m}(w(t))\|_{H^s}^2\right| &\leq 2\epsilon^2 \sum_{n\geq 3m+1} (1+n^2)^s |\widehat{\mathcal{Y}_m(t, w(t))}(n)||w_n(t)| \\
&\leq 2\epsilon^2 \|\mathcal{Y}_m(t, w(t))\|_{H^s} \|w(t)\|_{H^s} \\
&\lesssim_{m,s} \epsilon^2 \|w(t)\|_{H^s}^3 (1 + \|w(t)\|_{H^s}).
\end{aligned}$$

□

4.3.2. *End of the proof of Theorem* 1.8. The proof is completed by a bootstrap argument and estimate (4.10), obtained by the Birkhoff normal form transformation. It suffices to prove the case $u(0) \in C_+^\infty$ by the same density argument in the proof of Proposition 1.7.

Proof. For all $m \in \mathbb{N}$ and $s \geq 1$, we recall that $\partial_t v(t) = X_{\mathcal{H}^{m,\epsilon}(t)}(v(t))$ and $w(t) = \chi_{-1}^m(v(t))$. In Proposition 4.2, there exists $A_{m,s} \geq 1$ such that $\sup_{0<\epsilon<1} \|v(0)\|_{H^s} \leq A_{m,s}$. By using (4.4), we have

$$\sup_{0<\epsilon<1} \sup_{t\in\mathbb{R}} \|\mathbb{P}_{3m}(v(t))\|_{H^s} \leq (1+9m^2)^{\frac{s}{2}} I_m.$$

Set $K_{m,s} := \max(6A_{m,s}, 6(1+9m^2)^{\frac{s}{2}} I_m)$, $\epsilon_{m,s} := \min(\epsilon_m^*, \frac{\gamma_{m,s}}{3K_{m,s}}, \frac{1}{48C_{m,s}K_{m,s}})$ and

$$T_{m,s} := \sup\{t \geq 0 : \sup_{0\leq\tau\leq t} \|v(\tau)\|_{H^s} \leq 2K_{m,s}\}.$$

We choose $\epsilon \in (0, \epsilon_{m,s})$. Since $\epsilon = \epsilon\|v(0)\|_{H^s} \leq \epsilon A_{m,s} \leq \gamma_{m,s}$, Lemma 4.5 yields that

$$\|v(0) - w(0)\|_{H^s} = \|v(0) - \chi_{-1}^m(v(0))\|_{H^s} \leq \epsilon A_{m,s}^2 C_{m,s} \leq A_{m,s}.$$

So we have $\|w(0)\|_{H^s} \leq \frac{K_{m,s}}{3}$. For all $t \in [0, T_{m,s}]$, we have $\epsilon\|v(t)\|_{H^s} \leq \gamma_{m,s}$. Hence Lemma 4.5 gives the following estimate

$$\begin{aligned}
\|w(t)\|_{H^s} \leq& \|v(t)\|_{H^s} + \|\chi_{-1}^m(v(t)) - v(t)\|_{H^s} \\
\leq& \|v(t)\|_{H^s} + C_{m,s}\epsilon\|v(t)\|_{H^s}^2 \\
\leq& 2K_{m,s} + 4C_{m,s}K_{m,s}^2\epsilon \\
\leq& 3K_{m,s}.
\end{aligned}$$

So we have $\epsilon \sup_{0\leq t\leq T_{m,s}} \|w(t)\|_{H^s} \leq \gamma_{m,s}$, which implies that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|w(t) - \mathbb{P}_{3m}(w(t))\|_{H^s}^2\right| \leq C'_{m,s}\epsilon^2\|w(t)\|_{H^s}^3(1 + \|w(t)\|_{H^s}),$$

in Lemma 4.6. Set $d_{m,s} := \frac{1}{486K_{m,s}^2 C'_{m,s}}$. We can obtain the following estimate:

$$\begin{aligned}
&\|w(t) - \mathbb{P}_{3m}(w(t))\|_{H^s}^2 \\
\leq& \|w(0)\|_{H^s}^2 + C'_{m,s}|t|\epsilon^2 \sup_{0\leq\tau\leq T_{m,s}} \|w(\tau)\|_{H^s}^3 (1 + \sup_{0\leq\tau\leq T_{m,s}} \|w(\tau)\|_{H^s}) \\
\leq& \frac{K_{m,s}^2}{9} + 162C'_{m,s}K_{m,s}^4|t|\epsilon^2 \\
\leq& \frac{4K_{m,s}^2}{9},
\end{aligned}$$

for all $0 \leq t \leq \min(T_{m,s}, \frac{d_{m,s}}{\epsilon^2})$. We use Lemma 4.5 again to obtain that

$$\begin{aligned}
\|v(t)\|_{H^s} \leq& 2\|w(t) - v(t)\|_{H^s} + \|w(t) - \mathbb{P}_{3m}(w(t))\|_{H^s} + \|\mathbb{P}_{3m}(v(t))\|_{H^s} \\
\leq& 2C_{m,s}\epsilon\|v(t)\|_{H^s}^2 + \frac{2K_{m,s}}{3} + \frac{K_{m,s}}{6} \\
\leq& K_{m,s},
\end{aligned}$$

for all $t \in [0, \frac{d_{m,s}}{\epsilon^2}]$. In the case $t < 0$, we use the same procedure and we replace t by $-t$. Consequently, we have

$$\sup_{|t|\leq\frac{d_{m,s}}{\epsilon^2}} \|u(t) - e^{i(\cdot + \arg u_m(t))}\|_{H^s} = \epsilon \sup_{|t|\leq\frac{d_{m,s}}{\epsilon^2}} \|v(t)\|_{H^s} \leq K_{m,s}\epsilon.$$

□

5. Comparison to the NLS equation

Although we have some similar results for the NLS equation, there are still some differences between the NLS equation and the NLS-Szegő equation. We denote by $u = u(t,x) = \sum_{n\geq 0} u_n(t)e^{inx}$ the solution of the NLS equation

$$i\partial_t u + \partial_x^2 u = |u|^2 u. \tag{5.1}$$

In Fourier modes, we have $i\partial_t u_n = n^2 u_n + \sum_{k_1-k_2+k_3=n} u_{k_1}\overline{u}_{k_2}u_{k_3}$. Fix $m \in \mathbb{Z}$, for every $n \in \mathbb{Z}$, we define $v_n := u_{n+m}e^{i(m^2+2nm)t}$. Then $\|v(t)\|_{L^2} = \|u(t)\|_{L^2}$ and we have

$$i\partial_t v_n = n^2 v_n + \sum_{k_1-k_2+k_3=n} v_{k_1}\overline{v}_{k_2}v_{k_3}. \tag{5.2}$$

If u is localized in the m-th Fourier mode, then v is localized on the zero mode. Thus the orbital stability of the traveling wave $\mathbf{e}_m$ can be reduced to the case $m = 0$. In Faou–Gauckler–Lubich [8], long time H^s-orbital stability of plane wave solutions is established by limiting the mass of the initial data to a certain full measure subset of $(0, +\infty)$ for the defocusing cubic Schrödinger equation with the time interval $[-\epsilon^{-N}, \epsilon^{-N}]$, for all $N \geq 1$ and $s \gg 1$. However, this above transformation $u \mapsto v$ does not preserve the L^2 norm for the NLS-Szegő equation and formula (5.2) fails too.

On the other hand, the approach that we use to prove Theorem 1.8 does not work for (5.1). In fact, the negative high frequency Fourier modes in the term $\mathcal{H}_1^m(v) = \operatorname{Re}\int_{\mathbb{S}^1} e^{-imx}|v|^2 v$ can not be cancelled by the homological equation. The energy functional of the equation of v can not be reduced as $\mathcal{H}_0^m + \mathcal{O}(\epsilon^2)$ by using the same method in this paper.

The Szegő filtering to (5.1) makes it possible to cancel all the high frequency resonances in the term $\mathcal{R}_m = \{\mathcal{F}_m, \mathcal{H}_0^m\} + \mathcal{H}_1^m$. Then we use a bootstrap argument to deal with the equation $\partial_t w(t) = X_{\mathcal{H}^{m,\epsilon}(t)\circ\chi_1^m}(w(t))$ after the Birkhoff normal form transformation.

6. Appendix

We give the details of the homological equation in Proposition 4.4 and prove (4.9) and Proposition 4.3 in the case $\alpha = 0$.

For all $v \in \bigcup_{n\in\mathbb{N}} \mathbb{P}_n(C_+^\infty)$, $\mathcal{H}_0^m(v) = \frac{1}{2}\|\partial_x v\|_{L^2}^2 + \frac{1-m^2}{2}\|v\|_{L^2}^2 + \frac{1}{2}\int_{\mathbb{S}^1} \operatorname{Re}(e^{2imx}\overline{v}^2)\mathrm{d}x$ and $\mathcal{F}_m(v) = \sum_{\substack{j-l+k=m\\ j,k,l\in\mathbb{N}}} \operatorname{Re}(a_{j,l,k}v_j\overline{v}_l v_k)$. With the convention $v_n = 0$, for all $n < 0$, we have

$$\begin{aligned}
\partial_{v_n}\mathcal{H}_0^m(v) =& \overline{\partial_{\bar{v}_n}\mathcal{H}_0^m(v)} = \tfrac{1+n^2-m^2}{2}\overline{v}_n + \tfrac{1}{2}v_{2m-n},\\
\partial_{\bar{v}_n}\mathcal{F}_m(v) =& \overline{\partial_{v_n}\mathcal{F}_m(v)} = \sum_{k-l+n=m} \bar{a}_{k,l,n}\overline{v}_k v_l + \tfrac{1}{2}\sum_{k-n+l=m} a_{k,n,l}v_k v_l.
\end{aligned}$$

Combing (3.8), we have the Poisson bracket of $\mathcal{F}_m$ and $\mathcal{H}_0^m$,

$$\begin{aligned}
&\{\mathcal{F}_m, \mathcal{H}_0^m\}(v) \\
=& -2i \sum_{n\geq 0} \left(\partial_{\overline{v}_n} \mathcal{F}_m \partial_{v_n} \mathcal{H}_0^m - \partial_{\overline{v}_n} \mathcal{H}_0^m \partial_{v_n} \mathcal{F}_m\right)(v) \\
=& \mathrm{Im}\Big(\sum_{\substack{k-l+n=m \\ k,l,n\in\mathbb{N}}} 2(1+n^2-m^2)\overline{a}_{k,l,n} \overline{v}_k v_l \overline{v}_n + \sum_{\substack{k-n+l=m \\ k,l,n\in\mathbb{N}}} (1+n^2-m^2) a_{k,n,l} v_k \overline{v}_n v_l\Big) \\
&+ \mathrm{Im}\Big(\sum_{\substack{k-l+n=m \\ k,l,n\in\mathbb{N}, n\leq 2m}} 2\overline{a}_{k,l,n} v_l \overline{v}_k v_{2m-n} + \sum_{\substack{k-n+l=m \\ k,l,n\in\mathbb{N}, n\leq 2m}} a_{k,n,l} v_k v_l v_{2m-n}\Big) \\
=& \mathrm{Im}(\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_4).
\end{aligned}$$

The term $\mathcal{A}_4 = \sum_{\substack{k-n+l=m \\ k,l,n\in\mathbb{N}, n\leq 2m}} a_{k,n,l} v_k v_l v_{2m-n}$ has only finite terms depending on low-frequency Fourier modes $v_1, \cdots, v_{3m}$. We divide $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ into two parts. The first part consists of low-frequency resonances, the second part consists of high-frequency resonances.

$$\mathcal{A}_1 = -\sum_{\substack{k-l+n=m \\ k,l,n\in\mathbb{N}}} 2(1+n^2-m^2) a_{k,l,n} v_k \overline{v}_l v_n = \mathcal{A}_1^{low} + \mathcal{A}_1^{\geq 2m+1},$$

where $\mathcal{A}_1^{low}$ consists of all the resonances $v_j \overline{v}_{j+k-m} v_k$ such that $j, k \leq 2m$,

$$\begin{aligned}
\mathcal{A}_1^{low} =& -\Big(\sum_{j=0}^{\lfloor \frac{m-1}{2} \rfloor} \sum_{k=m-j}^{2m} + \sum_{j=\lfloor \frac{m+1}{2} \rfloor}^{2m-1} \sum_{k=j+1}^{2m} \Big) 2(2+j^2+k^2-2m^2) a_{j,j+k-m,k} v_j \overline{v}_{j+k-m} v_k \\
&- \sum_{j=0}^{m-1} \sum_{k=j+m+1}^{2m} 2(2+j^2+(k+m-j)^2-2m^2) a_{j,k,k+m-j} v_j \overline{v}_k v_{k+m-j} \\
&- \sum_{k=\lfloor \frac{m+1}{2} \rfloor}^{2m} 2(1+k^2-m^2) a_{k,2k-m,k} v_k^2 \overline{v}_{2k-m},
\end{aligned}$$

and $\mathcal{A}_1^{\geq 2m+1}$ contain other resonances $v_j \overline{v}_{j+k-m} v_k$ such that at least one of j, k is strictly larger than $2m$.

$$\begin{aligned}
\mathcal{A}_1^{\geq 2m+1} =& -\sum_{j=0}^{m} \sum_{k\geq 2m+1} 2(2+j^2+(k+m-j)^2-2m^2) a_{j,k,k+m-j} v_j \overline{v}_k v_{k+m-j} \\
&- \sum_{j=m+1}^{2m} \sum_{k\geq 2m+1} 2(2+j^2+k^2-2m^2) a_{j,j+k-m,k} v_j \overline{v}_{j+k-m} v_k \\
&- \sum_{k\geq 2m+1} \sum_{j\geq k+1} 2(2+k^2+j^2-2m^2) a_{k,k+j-m,j} v_k \overline{v}_{k+j-m} v_j \\
&- \sum_{k\geq 2m+1} 2(1+k^2-m^2) a_{k,2k-m,k} v_k^2 \overline{v}_{2k-m}.
\end{aligned}$$

Then, we calculate $\mathcal{A}_2 = \sum_{\substack{j-l+k=m\\ j,l,k\in\mathbb{N}}} (1+l^2-m^2)a_{j,l,k}v_j\overline{v}_l v_k = \mathcal{A}_2^{low} + \mathcal{A}_2^{\geq 2m+1}$. $\mathcal{A}_2^{low}$ consists of all the resonances $v_j\overline{v}_l v_k$ such that $j,k \leq 2m$ or $l = j+k-m \leq 2m$.

$$\begin{aligned}
\mathcal{A}_2^{low} = & \sum_{\substack{j-l+k=m\\ j,l,k\in\mathbb{N}, l\leq 2m}} (1+l^2-m^2)a_{j,l,k}v_j\overline{v}_l v_k \\
& + \sum_{\substack{j-l+k=m\\ m+1\leq j\leq 2m-1\\ j+1\leq k\leq 2m, l\geq 2m+1}} 2(1+l^2-m^2)a_{j,l,k}v_j\overline{v}_l v_k \\
& + \sum_{k=\lfloor \frac{3m}{2}\rfloor+1}^{2m} (1+(2k-m)^2-m^2)a_{k,2k-m,k}v_k^2\overline{v}_{2k-m};
\end{aligned}$$

$\mathcal{A}_2^{\geq 2m+1}$ plays the same role as $\mathcal{A}_1^{\geq 2m+1}$.

$$\begin{aligned}
\mathcal{A}_2^{\geq 2m+1} = & \sum_{j=0}^{m}\sum_{k\geq 2m+1} 2(1+k^2-m^2)a_{j,k,m+k-j}v_j\overline{v}_k v_{m+k-j} \\
& + \sum_{j=m+1}^{2m}\sum_{k\geq 2m+1} 2(1+(j+k-m)^2-m^2)a_{j,j+k-m,k}v_j\overline{v}_{j+k-m}v_k \\
& + \sum_{k\geq 2m+1}\sum_{j\geq k+1} 2(1+(k+j-m)^2-m^2)a_{k,k+j-m,n}v_k\overline{v}_{k+j-m}v_j \\
& + \sum_{k\geq 2m+1} (1+(2k-m)^2-m^2)a_{k,2k-m,k}v_k^2\overline{v}_{2k-m}.
\end{aligned}$$

At last,

$$\mathcal{A}_3 = \sum_{\substack{k-l+n=m\\ k,l,n\in\mathbb{N}, n\leq 2m}} 2\overline{a}_{k,l,n}v_l\overline{v}_k v_{2m-n} = \sum_{\substack{j-l+k=m\\ j,k,l,n\in\mathbb{N}, j\leq 2m}} 2\overline{a}_{l,k,2m-j}v_k\overline{v}_l v_j.$$

Using the same idea, we have $\mathcal{A}_3 = \mathcal{A}_3^{low} + \mathcal{A}_3^{\geq 2m+1}$, with

$$\begin{aligned}
\mathcal{A}_3^{low} = & \sum_{j=0}^{m}\sum_{k=0}^{2m} 2\overline{a}_{2m-j,m+k-j,k}v_j\overline{v}_k v_{m+k-j} \\
& + \sum_{j=m+1}^{2m}\sum_{k=0}^{2m} 2\overline{a}_{2m-j,k,k+j-m}v_j\overline{v}_{k+j-m}v_k;
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{A}_3^{\geq 2m+1} = & \sum_{j=0}^{m}\sum_{k\geq 2m+1} 2\overline{a}_{2m-j,m+k-j,k}v_j\overline{v}_k v_{m+k-j} \\
& + \sum_{j=m+1}^{2m}\sum_{k\geq 2m+1} 2\overline{a}_{2m-j,k,k+j-m}v_j\overline{v}_{k+j-m}v_k.
\end{aligned}$$

Recall that $\mathcal{H}_1^m(v) = \mathrm{Re}(\int_{\mathbb{S}^1} e^{-imx}|v|^2 v) = \sum_{k-l+n=m} \mathrm{Re}(v_k \overline{v}_l v_n)$. A similar calculus as in the case of $\mathcal{A}_1$ shows that $\mathcal{H}_1^m(v) = \mathrm{Re}(B^{low} + B^{\geq 2m+1})$ with

$$\begin{aligned} & B^{low} \\ =& \Big(\sum_{j=0}^{\lfloor \frac{m-1}{2} \rfloor} \sum_{k=m-j}^{2m} + \sum_{j=\lfloor \frac{m+1}{2} \rfloor}^{2m-1} \sum_{k=j+1}^{2m} \Big) 2 v_j \overline{v}_{j+k-m} v_k \\ &+ \sum_{k=\lfloor \frac{m+1}{2} \rfloor}^{2m} v_k^2 \overline{v}_{2k-m} + \sum_{j=0}^{m-1} \sum_{k=j+m+1}^{2m} 2 v_j \overline{v}_k v_{k+m-j}, \end{aligned}$$

$$\begin{aligned} B^{\geq 2m+1} =& \sum_{k \geq 2m+1} \Bigg(\sum_{j=0}^{m} 2 v_j \overline{v}_k v_{k+m-j} \\ &+ \sum_{j=m+1}^{2m} 2 v_j \overline{v}_{j+k-m} v_k + v_k^2 \overline{v}_{2k-m} + \sum_{j \geq k+1} 2 v_k \overline{v}_{k+j-m} v_j \Bigg). \end{aligned}$$

At last we define $\mathrm{Reson}^{low}(v)\Big|_{\alpha=0} = -i(\mathcal{A}_1^{low} + \mathcal{A}_2^{low} + \mathcal{A}_3^{low} + \mathcal{A}_4) + B^{low}$ and

$$\mathrm{Reson}^{\geq 2m+1}(v)\Big|_{\alpha=0} = -i(\mathcal{A}_1^{\geq 2m+1} + \mathcal{A}_2^{\geq 2m+1} + \mathcal{A}_3^{\geq 2m+1}) + B^{\geq 2m+1}.$$

Then we have $\{\mathcal{F}_m, \mathcal{H}_0^m\}(v) + \mathcal{H}_1^m(v) = \mathrm{Re}(\mathrm{Reson}^{low}(v)\Big|_{\alpha=0} + \mathrm{Reson}^{\geq 2m+1}(v)\Big|_{\alpha=0})$. Since $\mathrm{Reson}^{low}(v)\Big|_{\alpha=0}$ contains only finite terms and depends only on $v_1, \cdots, v_{3m}$, so is $\mathcal{R}_m = \mathrm{Re}(\mathrm{Reson}^{low}(v)\Big|_{\alpha=0})$. For high-frequency resonances, we compute

$$\begin{aligned} & \mathrm{Reson}_{\geq 2m+1}(v)\Big|_{\alpha=0} \\ =& \sum_{k \geq 2m+1} \sum_{j=0}^{m-1} 2\left(-i\overline{a}_{2m-j,m+k-j,k} + (1 + 2(m-j)(k-j)) i a_{j,k,k+m-j} + 1\right) \\ & v_j \overline{v}_k v_{k+m-j} \\ &+ \sum_{k \geq 2m+1} \sum_{j=m+1}^{2m} 2((1 - 2(k-m)(j-m)) i a_{j,j+k-m,k} - i\overline{a}_{2m-j,k,k+j-m} + 1) \\ & v_j \overline{v}_{j+k-m} v_k \\ &+ \sum_{k \geq 2m+1} [2(i a_{m,k,k} - i\overline{a}_{m,k,k} + 1) v_m |v_k|^2 + ((1 - 2(k-m)^2) i a_{k,2k-m,k} + 1) \\ & v_k^2 \overline{v}_{2k-m} t] \\ &+ \sum_{k \geq 2m+1} \sum_{j \geq k+1} 2((1 - 2(j-m)(k-m)) i a_{k,j+k-m,j} + 1) v_k \overline{v}_{j+k-m} v_j. \end{aligned}$$

After replacing j by $2m-j$ in the sum $m+1 \leq j \leq 2m$, we have the equivalence between $\mathrm{Reson}_{\geq 2m+1}\Big|_{\alpha=0} = 0$ and (4.9) in the case $\alpha = 0$.

References

[1] Bambusi D., *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Physics 234 (2003), 253–283.

[2] Bourgain, J., *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations*, Geometric and Functional Analysis, March 1993, Volume 3, Issue 2, pp 107–156

[3] Boyer, F., Fabrie, P., *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models.* Applied Mathematical Sciences 183. New York: Springer. pp. 102–106. (2013).

[4] Brezis, H., Gallouët, T., *Nonlinear Schrödinger evolution equations.* Nonlinear Anal. 4, 677-681

[5] Colliander, J., Keel, M., Staffilani, G.,Takaoka, H., Tao, T., *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math.181 (2010), 39–113.

[6] Eliasson H., Kuksin S.,*KAM for the nonlinear Schrödinger equation*, Ann. Math. 172(1), 371–435 (2010)

[7] Faddeev, L.D., Takhtajan, L.A.,*Hamiltonian Methods in the Theory of Solitons*, Springer series in Soviet Mathematics, Springer, Berlin, 1987.

[8] Faou E., Gauckler L., Lubich C.,*Sobolev Stability of Plane Wave Solutions to the Cubic Nonlinear Schrödinger Equation on a Torus*, Comm. Partial Differential Equations 38 (2013), no. 7, 1123–1140

[9] Gallay, T., Haragus, M., *Stability of small periodic waves for the nonlinear Schrödinger equation*, J. Diff. Eqns 234, 544-581 (2007)

[10] Gallay, T., Haragus, M., *Orbital stability of periodic waves for the nonlinear Schrödinger equation*, J. Diff. Eqns 19, 825-865 (2007)

[11] Gérard, P. *On the conservation laws of the defocusing cubic NLS equation*, 2015, http://www.math.u-psud.fr/ pgerard/integrales-NLScubique.pdf

[12] Gérard, P., Grellier, S. *The cubic Szegő equation*, Ann. Sci. l'Éc. Norm. Supér. (4) 43 (2010), 761-810

[13] Gérard, P., Grellier, S. *Effective integrable dynamics for a certain nonlinear wave equation*, Anal. PDE, 5(2012), 1139-1155

[14] Gérard, P., Grellier, S. *Invariant tori for the cubic Szegő equation*, Invent. Math. 187:3(2012),707-754. MR 2944951 Zbl 06021979

[15] Gérard, P., Grellier, S. *On the growth of Sobolev norms for the cubic Szegő equation*, text of a talk at IHES, seminar Laurent Schwartz, January 6, 2015.

[16] Gérard, P., Grellier, S., An explicit formula for the cubic Szegő equation, Trans. Amer. Math. Soc. 367 (2015), 2979-2995

[17] Gérard, P., Grellier, S., The cubic Szegő equation and Hankel operators, volume 389 of Astérisque. Soc. Math. de France, 2017.

[18] Grébert, B. *Birkhoff Normal Form and Hamiltonian PDEs*, pp.1-46 in Partial differential equations and applications, edited by X.P.Wang and C.Zhong, Sémin. Congr.15, Soc.Math.France, Paris, 2007. MR2009d:37130 Zbl 1157.37019

[19] Grébert, B., Kappeler, T. *The Defocusing NLS Equation and Its Normal Form*, Series of Lectures in Mathematics, European Mathematical Society, 2014.

[20] Grébert, B., Thomann, L. *Resonant dynamics for the quintic nonlinear Schrödinger equation*, Ann. Inst. H. Poincaré Anal. Non Linéaire 29(3) (2012) 455-477.

[21] Haus, E., Procesi, M. *KAM for beating solutions of the quintic NLS*, Comm. Math. Phys. 354 (2017), no.3, 1101-1132

[22] Lax, P. *Integrals of Nonlinear Equations of Evolution and Solitary Waves*, Comm. Pure Appl. Math. Volume 21, 1968, Pages 467–490

[23] Liouville, J. *Sur des classes très–étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques*, J.Math. pures et app. 16(1851), 133-142

[24] Ogawa, T. *A proof of Trudinger's inequality and its application to nonlinear Schrödinger equations*, Nonlinear Anal. 14 (1990), 765–769.

[25] Peller, V. V. *Hankel operators of class $\mathcal{S}_p$ and their applications (rational approximation, Gaussian processes, the problem of majorization of operators)*, Math. USSR Sb.41(1982), 443–479

[26] Vladimirov, M. V. *On the solvability of a mixed problem for a nonlinear equation of Schrödinger type*, Sov. Math. Dokl. 29 (1984), 281–284.

[27] Yudovich, V. I . *Non-stationary flow of an ideal incompressible liquid*, USSR Comput. Math. Math. Phys. 3 (1963), 1407–1456 (english), Zh. Vuch. Mat. 3 (1963), 1032– 1066 (russian).

[28] Zakharov, V.E., Shabat, A.B., *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media*, Soviet Physics JETP 34-62, 1972.

[29] Zhidkov, P. *Korteweg–de Vries and nonlinear Schrödinger equations: qualitative theory*, Lecture Notes in Mathematics 1756, Springer-Verlag, Berlin, 2001.

Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud XI, CNRS, Université Paris-Saclay, F-91405 Orsay, France

Email address: ruoci.sun@math.u-psud.fr

URL: https://www.researchgate.net/profile/Ruoci_Sun