

# Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus

Erwan Faou<sup>\*†</sup>    Ludwig Gauckler<sup>‡</sup>    Christian Lubich<sup>§</sup>

Version of 10 October 2012

## Abstract

It is shown that plane wave solutions to the cubic nonlinear Schrödinger equation on a torus behave orbitally stable under generic perturbations of the initial data that are small in a high-order Sobolev norm, over long times that extend to arbitrary negative powers of the smallness parameter. The perturbation stays small in the same Sobolev norm over such long times. The proof uses a Hamiltonian reduction and transformation and, alternatively, Birkhoff normal forms or modulated Fourier expansions in time.

## 1 Introduction and statement of the result

Consider the cubic nonlinear Schrödinger equation (NLS) on the  $d$ -dimensional torus  $\mathbb{T}^d = \mathbb{R}^d / (2\pi\mathbb{Z})^d$ , for arbitrary dimension  $d \geq 1$ , in the defocusing ( $\lambda = 1$ ) or focusing ( $\lambda = -1$ ) case,

$$i\partial_t u = -\Delta u + \lambda|u|^2 u, \quad x \in \mathbb{T}^d, t \in \mathbb{R}. \quad (1.1)$$

For initial data made of a single Fourier mode,  $u_*(x, 0) = \rho e^{im \cdot x}$ , the equation has the plane-wave solution  $u_*(x, t) = \rho e^{i(m \cdot x - \omega t)}$  with  $\omega = |m|^2 + \lambda\rho^2$ . We show that under

---

<sup>\*</sup>INRIA and ENS Cachan Bretagne, Avenue Robert Schumann, F-35170 Bruz, France (Erwan.Faou@inria.fr).

<sup>†</sup>Département de mathématiques et applications, École normale supérieure, 45 rue d'Ulm, F-75230 Paris Cedex 05, France.

<sup>‡</sup>Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany (gauckler@math.tu-berlin.de).

<sup>§</sup>Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany (lubich@na.uni-tuebingen.de).

generic perturbations of such initial data by functions with small  $H^s$  Sobolev norm, for sufficiently large Sobolev exponent  $s$ , the solution remains essentially localized in the  $m$ th Fourier mode over very long times, and the perturbation remains small in the same  $H^s$  norm. For the precise formulation of the result we decompose the solution in the Fourier basis,  $u(x, t) = \sum_{j \in \mathbb{Z}^d} u_j(t) e^{ij \cdot x}$ .

**Theorem 1.1.** *Let  $\rho_0 > 0$  be such that  $1 + 2\lambda\rho_0^2 > 0$ , and let  $N > 1$  be fixed arbitrarily. There exist  $s_0 > 0$ ,  $C \geq 1$  and a set of full measure  $\mathcal{P}$  in the interval  $(0, \rho_0]$  such that for every  $s \geq s_0$  and every  $\rho \in \mathcal{P}$ , there exists  $\varepsilon_0 > 0$  such that for every  $m \in \mathbb{Z}^d$  the following holds: if the initial data  $u(\bullet, 0)$  are such that*

$$\|u(\bullet, 0)\|_{L^2} = \rho \quad \text{and} \quad \|e^{-im \cdot \bullet} u(\bullet, 0) - u_m(0)\|_{H^s} = \varepsilon \leq \varepsilon_0,$$

then the solution of (1.1) with these initial data satisfies

$$\|e^{-im \cdot \bullet} u(\bullet, t) - u_m(t)\|_{H^s} \leq C\varepsilon \quad \text{for} \quad t \leq \varepsilon^{-N}. \quad (1.2)$$

As will further be shown, this theorem (together with the conservation of  $L^2$  norm) implies long-time *orbital stability* in  $H^s$ : the solution stays in  $H^s$  close to the orbit  $e^{i\varphi} u_m(0) e^{im \cdot x}$ ,  $\varphi \in \mathbb{R}$ , of the nonlinear Schrödinger equation:

$$\inf_{\varphi \in \mathbb{R}} \|e^{-im \cdot \bullet} u(\bullet, t) - e^{i\varphi} u_m(0)\|_{H^s} \leq \sqrt{2}C\varepsilon \quad \text{for} \quad t \leq \varepsilon^{-N}. \quad (1.3)$$

To our knowledge, this is the first long-time orbital stability result as well as long-time high-regularity result for the cubic NLS (1.1) in dimension  $d > 1$ . It is a contrasting counterpart to the instability results for *rough* perturbations given by Christ, Colliander & Tao [11, 12], Carles, Dumas & Sparber [9] and Hani [29].

In the 1D case, much is already known about  $H^1$ -orbital stability of periodic waves for the cubic NLS through the work by Zhidkov [30, Sect. 3.3] and Gallay & Haragus [19, 20]; see also further references therein.

To obtain the  $H^s$ -stability result (with  $s \gg 1$ ) presented above, the techniques used are more closely related to long-time stability and high-regularity results by Bambusi & Grébert [4, 2, 24] and Gauckler & Lubich [21] for small solutions to *modifications* of the periodic cubic NLS (1.1) by the addition of a convolution term  $V \star u$ , which eliminates the resonance of the frequencies of the linearization of (1.1) around 0. Such a resonance-removing modification by a convolution potential was previously studied also by Bourgain [6] and more recently by Eliasson & Kuksin [17]. In this *non-resonant* case,  $H^s$ -stability of small solutions can be proven generically with respect to the external parameter  $V$ , and in the case where the solutions are analytic, the constant can be optimized to some Nekhoroshev-like estimate as shown by Faou & Grébert in [18].

As far as the *resonant* case is concerned, i.e., the cubic NLS (1.1) in dimension  $d > 1$ , the question of the  $H^s$ -behavior of solutions is much more delicate and has recently known many advances. Bourgain [5] gives upper bounds of the form

$$\|u(t)\|_{H^s} \leq t^{\alpha(s-1)} \|u(0)\|_{H^s} \quad \text{for } t > 0$$

in dimension  $d = 2$  or  $d = 3$  for sufficiently large  $s$ , either for the defocusing case  $\lambda > 0$  or for small initial data in  $L^2$  ( $d = 2$ ) or  $H^1$  ( $d = 3$ ). It is conjectured by Bourgain in [7] that the previous bound can be refined to subpolynomial growth in time, that is  $t^\alpha$  for all  $\alpha > 0$ . Colliander, Keel, Staffilani, Takaoka & Tao [14] and Guardia & Kaloshin [25] proved, in dimension  $d = 2$  and in the defocusing case  $\lambda > 0$ , for any  $\varepsilon, M > 0$  and  $s > 1$ , the existence of solutions  $u$  to (1.1) such that  $\|u(0)\|_{H^s} < \varepsilon$  and  $\|u(T)\|_{H^s} > M$  for some time  $T > 0$ . In a similar direction, Carles & Faou [10] proved the existence of initial data which propagate energy from low to high modes in an energy cascade that prevents the existence of Nekhoroshev-like stability results with preservation of the actions for small solutions.

In contrast to these results indicating unstable behavior, Theorem 1.1 above indeed gives examples of  $H^s$ -stability over very long time for the resonant dynamics (1.1), in a way which looks like known results in non-resonant cases. The peculiarity of the situation is that the plane wave solution creates new frequencies in the dynamics of the resonant equation (1.1). These new frequencies depend on the initial data and turn out to be generically non-resonant. This allows us to separate the dynamics between the modes and to prove the preservation of  $H^s$ -regularity over long time. Note that this does not contradict the results and conjectures about the existence of solutions with Sobolev norm growths, but it shows the existence of many *integrable islands* (using an expression due to T. Kappeler) that are  $H^s$ -stable under the dynamics of the resonant equation (1.1). A different resonance-removing effect of the initial value is used by Bourgain [8] to prove long-time  $H^s$ -stability of small solutions to (1.1) in dimension  $d = 1$ .

A striking fact is that the solutions we consider are not small in  $H^s$ , as is traditionally the case when studying non-resonant situations. In particular, the sign of the nonlinearity enters into the statement of the Theorem: while the solution has to be small enough in the focusing case  $\lambda < 0$ , there is no restriction on the size of the initial data in the defocusing case  $\lambda > 0$ .

The plan of proof is the following: We first separate the dynamics of the plane wave from the other modes, and we perform some reductions and transformations (in Section 2) to obtain a problem in a Hamiltonian form with new frequencies depending on the initial data. We use the gauge invariance of the equation as well as the  $L^2$  norm preservation to eliminate the dynamics of the plane wave itself in the equation. We end up with a situation of a semi-linear Hamiltonian system with non-resonant fre-

quencies, for which there are two known techniques to arrive at Theorem 1.1: *Birkhoff normal forms*, which use a sequence of nonlinear canonical coordinate transforms to transform the system to a form from which the dynamical properties can be read off, and *modulated Fourier expansions*, which embed the system into a larger modulation system having almost-invariants that allow us to infer the desired long-time properties. In Section 3 we reduce Theorem 1.1 to a known abstract result on the long-time near-conservation of super-actions, which was proved by Grébert [24] via Birkhoff normal forms and by Gauckler [22] via modulated Fourier expansions, under differing conditions which we verify for both approaches.

**Remark 1.2.** *Following the approach of [18], we could optimize  $N \simeq |\log \varepsilon|^\beta$  with  $\beta < 1$  in Equation (1.2), if the perturbation of the plane wave is analytic.*

## 2 Reductions and Transformations

### 2.1 Reduction to the case $m = 0$

In terms of Fourier coefficients, (1.1) is given by

$$i\dot{u}_j = |j|^2 u_j + \lambda \sum_{j=j_1-j_2+j_3} u_{j_1} \bar{u}_{j_2} u_{j_3}, \quad (2.1)$$

where  $|j|$  denotes the Euclidean norm of  $j \in \mathbb{Z}^d$ . With the given  $m \in \mathbb{Z}^d$ , we transform to

$$v_j = u_{j+m} e^{it(|m|^2 + 2j \cdot m)}. \quad (2.2)$$

Note that this transformation preserves the  $L^2$  norm. The equation for  $v_j$  is

$$(|m|^2 + 2j \cdot m)v_j + i\dot{v}_j = |j+m|^2 v_j + \lambda \sum_{j=j_1-j_2+j_3} v_{j_1} \bar{v}_{j_2} v_{j_3}$$

or equivalently

$$i\dot{v}_j = |j|^2 v_j + \lambda \sum_{j=j_1-j_2+j_3} v_{j_1} \bar{v}_{j_2} v_{j_3},$$

so that  $v(x, t) = \sum_{j \in \mathbb{Z}^d} v_j(t) e^{ij \cdot x}$  is a solution of (1.1) and is localized in the zero mode if  $u$  is localized in the  $m$ th mode. In other words, up to the transformation (2.2), we can restrict our attention to the case  $m = 0$ .

## 2.2 Elimination of the zero mode

We make the change of variables  $v \mapsto (a, \theta, w)$  with  $v = (v_j)_{j \in \mathbb{Z}^d}$ ,  $w = (w_j)_{0 \neq j \in \mathbb{Z}^d}$ ,  $0 \leq a \in \mathbb{R}$  and  $\theta \in \mathbb{R}$  defined by

$$v_0 = ae^{-i\theta} \quad \text{and} \quad v_j = w_j e^{-i\theta} \quad \text{for} \quad 0 \neq j \in \mathbb{Z}^d.$$

The equation for  $w_j$  reads

$$i\dot{w}_j + \dot{\theta}w_j = |j|^2 w_j + \frac{\partial \widehat{P}}{\partial \bar{w}_j}(a, w, \bar{w}) \quad (2.3)$$

with

$$\widehat{P}(a, w, \bar{w}) = \frac{\lambda}{2} \sum_{j_1 + j_2 - j_3 - j_4 = 0} w_{j_1} w_{j_2} \bar{w}_{j_3} \bar{w}_{j_4},$$

where the sum is over indices  $j_1, \dots, j_4 \in \mathbb{Z}^d$  and where we use the convention  $a = w_0 = \bar{w}_0$ . With the same convention the corresponding equation for  $a$  reads

$$i\dot{a} + \dot{\theta}a = \lambda \sum_{j_1 + j_2 - j_3 = 0} w_{j_1} w_{j_2} \bar{w}_{j_3},$$

and taking the real part shows that

$$\dot{\theta} = \frac{1}{2a} \frac{\partial \widehat{P}}{\partial a}(a, w, \bar{w}).$$

Inserting this formula into (2.3) yields an equation for  $w_j$  which does not depend on  $\theta$  anymore,

$$i\dot{w}_j = |j|^2 w_j + \frac{\partial \widehat{P}}{\partial \bar{w}_j}(a, w, \bar{w}) + \frac{-w_j}{2a} \frac{\partial \widehat{P}}{\partial a}(a, w, \bar{w}). \quad (2.4)$$

Now note that by the conservation of the  $L^2$  norm and Parseval's equality we have for all times  $t$ ,

$$\rho^2 = |v_0|^2 + \sum_{j \neq 0} |v_j|^2 = a^2 + \sum_{j \neq 0} |w_j|^2$$

which means

$$a = \sqrt{\rho^2 - \sum_{j \neq 0} |w_j|^2}. \quad (2.5)$$

Hence we can not only forget the dynamics of  $\theta$  but also of  $a$  : it will be controlled by the  $w_j$ ,  $j \neq 0$ , using (2.5). We arrive at a system of differential equations for a reduced set of variables

$$z = (z_j)_{j \in \mathcal{Z}} := (w_j)_{j \in \mathcal{Z}} \quad \text{with} \quad \mathcal{Z} := \mathbb{Z}^d \setminus \{0\}.$$

### 2.3 The reduced Hamiltonian system

Considering  $a$  as a function of  $w_j$  and  $\bar{w}_j$ ,  $j \in \mathcal{Z}$ , given by (2.5) we have

$$\frac{\partial a}{\partial \bar{w}_j} = \frac{-w_j}{2a},$$

and therefore the equations (2.4) for  $z = (w_j)_{j \in \mathcal{Z}}$  are Hamiltonian,

$$i\dot{z}_j = \frac{\partial \tilde{H}}{\partial \bar{z}_j}(z, \bar{z}), \quad j \in \mathcal{Z}, \quad (2.6)$$

with the real-valued Hamiltonian function

$$\tilde{H}(z, \bar{z}) = \sum_j |j|^2 z_j \bar{z}_j + \tilde{P}(z, \bar{z}) \quad \text{with} \quad \tilde{P}(z, \bar{z}) = \hat{P}(a, w, \bar{w}).$$

This Hamiltonian function takes the form

$$\begin{aligned} \tilde{H}(z, \bar{z}) &= \lambda \rho^4 + \sum_j (|j|^2 + \lambda \rho^2) z_j \bar{z}_j + \frac{\lambda}{2} \rho^2 \sum_j \bar{z}_j \bar{z}_{-j} + \frac{\lambda}{2} \rho^2 \sum_j z_j z_{-j} \\ &+ \frac{\lambda}{2} \sum_{j_1 + j_2 - j_3 - j_4 = 0} z_{j_1} z_{j_2} \bar{z}_{j_3} \bar{z}_{j_4} - \frac{3\lambda}{2} \left( \sum_{j_1} z_{j_1} \bar{z}_{j_1} \right) \left( \sum_{j_2} z_{j_2} \bar{z}_{j_2} \right) \\ &- \frac{\lambda}{2} \left( \sum_{j_1} z_{j_1} z_{-j_1} \right) \left( \sum_{j_2} z_{j_2} \bar{z}_{j_2} \right) - \frac{\lambda}{2} \left( \sum_{j_1} \bar{z}_{j_1} \bar{z}_{-j_1} \right) \left( \sum_{j_2} z_{j_2} \bar{z}_{j_2} \right) \\ &+ \lambda \left( \sum_{j_1 + j_2 - j_3 = 0} z_{j_1} z_{j_2} \bar{z}_{j_3} + \sum_{j_1 - j_2 - j_3 = 0} z_{j_1} \bar{z}_{j_2} \bar{z}_{j_3} \right) \sqrt{\rho^2 - \sum_{j_4} z_{j_4} \bar{z}_{j_4}}. \end{aligned} \quad (2.7)$$

Expanding  $\sqrt{\rho^2 - x}$  into a convergent power series for  $|x| < \rho^2$ , we can write the Hamiltonian (2.7) as the infinite sum

$$\tilde{H}(z, \bar{z}) = \lambda \rho^4 + \sum_{r \geq 2} \tilde{H}_r(z, \bar{z})$$

where  $\tilde{H}_r(z, \bar{z})$  is a homogeneous polynomial of degree  $r$  in terms of  $(z_j, \bar{z}_j)$ , which is of the form

$$\tilde{H}_r(z, \bar{z}) = \sum_{p+q=r} \sum_{\substack{(k,l) \in \mathcal{Z}^p \times \mathcal{Z}^q \\ \mathcal{M}(k,l)=0}} \tilde{H}_{kl} z_{k_1} \cdots z_{k_p} \bar{z}_{l_1} \cdots \bar{z}_{l_q}$$

where

$$\mathcal{M}(k, l) = k_1 + \dots + k_p - l_1 - \dots - l_q \quad (2.8)$$

denotes the *momentum* of the multi-index  $(k, l)$ . We note that the Taylor expansion of  $\tilde{H}$  contains only terms with zero momentum, and its coefficients satisfy the bound

$$|\tilde{H}_{kl}| \leq \tilde{M} \tilde{L}^{p+q} \quad \text{for all } (k, l) \in \mathcal{Z}^p \times \mathcal{Z}^q, \quad (2.9)$$

where  $\tilde{M}$  and  $\tilde{L}$  depend on  $\rho$ .

## 2.4 Diagonalization and non-resonant frequencies

We study now the linear part of the system (2.6). As we will see, its eigenvalues are non-resonant for almost all parameters  $\rho$ . Moreover, we can control the diagonalization of this linear operator.

The linear part in the differential equation for  $z_j$  is  $(|j|^2 + \lambda\rho^2)z_j + \lambda\rho^2\bar{z}_{-j}$ . On taking the equation for  $z_j$  together with that for  $\bar{z}_{-j}$ , we are thus led to consider the matrix (for  $n = |j|^2$ )

$$A_n = \begin{pmatrix} n + \lambda\rho^2 & \lambda\rho^2 \\ -\lambda\rho^2 & -n - \lambda\rho^2 \end{pmatrix}.$$

**Lemma 2.1.** *For all  $n \geq 1$ , the matrix  $A_n$  is diagonalized by a  $2 \times 2$  matrix  $S_n$  that is real symplectic and hermitian and has condition number smaller than 2:*

$$S_n^{-1} A_n S_n = \begin{pmatrix} \Omega_n & 0 \\ 0 & -\Omega_n \end{pmatrix} \quad \text{with} \quad \Omega_n = \sqrt{n^2 + 2n\lambda\rho^2}.$$

*Proof.* We obtain

$$S_n = \frac{1}{\sqrt{(\Omega_n + n)(\Omega_n + n + 2\lambda\rho^2)}} \begin{pmatrix} n + \lambda\rho^2 + \Omega_n & -\lambda\rho^2 \\ -\lambda\rho^2 & n + \lambda\rho^2 + \Omega_n \end{pmatrix}$$

and

$$S_n^{-1} = \frac{1}{\sqrt{(\Omega_n + n)(\Omega_n + n + 2\lambda\rho^2)}} \begin{pmatrix} n + \lambda\rho^2 + \Omega_n & \lambda\rho^2 \\ \lambda\rho^2 & n + \lambda\rho^2 + \Omega_n \end{pmatrix},$$

and the statements of the lemma then follow by direct verification.  $\square$

Note that the condition  $1 + 2\lambda\rho^2 > 0$  in Theorem 1.1 ensures that all the eigenvalues  $\Omega_n$  are real, or equivalently, that the linearization of the system (2.6) at 0 is stable. The frequencies  $\Omega_n$  turn out to satisfy Bambusi's non-resonance inequality [1, 4] for almost all norm parameters  $\rho > 0$ .

**Lemma 2.2.** *Let  $r > 1$  and  $\rho_0 > 0$  with  $1 + 2\lambda\rho_0^2 > 0$ . There exist  $\alpha = \alpha(r) > 0$  and a set of full Lebesgue measure  $\mathcal{P} \subset (0, \rho_0]$  such that for every  $\rho \in \mathcal{P}$  there is a  $\gamma > 0$  such*

that the following non-resonance condition is satisfied: for all positive integers  $p, q$  with  $p + q \leq r$  and for all  $m = (m_1, \dots, m_p) \in \mathbb{N}^p$  and  $n = (n_1, \dots, n_q) \in \mathbb{N}^q$ ,

$$|\Omega_{m_1} + \dots + \Omega_{m_p} - \Omega_{n_1} - \dots - \Omega_{n_q}| \geq \frac{\gamma}{\mu_3(m, n)^\alpha}, \quad (2.10)$$

except if the frequencies cancel pairwise. Here,  $\mu_3(m, n)$  denotes the third-largest among the integers  $m_1, \dots, m_p, n_1, \dots, n_q$ .

*Proof.* The proof is similar to the one given in [4, Section 5.1] for the frequencies  $\sqrt{n^2 + \rho}$ ,  $n \in \mathbb{N}$ .

(a) Considering the frequencies as functions of  $\sigma = \rho^2$ ,  $\Omega_n = \Omega_n(\sigma)$ , we have

$$\frac{d^k}{d\sigma^k} \Omega_n = c_k (n\lambda)^k \Omega_n^{1-2k} \quad (2.11)$$

with  $c_k = (3 - 2k)c_{k-1}$  for  $k \geq 1$  and  $c_0 = 1$ . This shows that, for  $0 < n_0 < n_1 < \dots < n_K$ , the matrix

$$\left( \frac{d^k}{d\sigma^k} \Omega_{n_l} \right)_{k,l=0}^K$$

is of Vandermonde form with determinant

$$\begin{aligned} \det \left( \frac{d^k}{d\sigma^k} \Omega_{n_l} \right)_{k,l=0}^K &= c_0 \cdots c_K \Omega_{n_0} \cdots \Omega_{n_K} \prod_{k < l} \left( \frac{n_l \lambda}{\Omega_{n_l}^2} - \frac{n_k \lambda}{\Omega_{n_k}^2} \right) \\ &= c_0 \cdots c_K \Omega_{n_0} \cdots \Omega_{n_K} \prod_{k < l} \frac{(n_k - n_l) \lambda}{(n_k + 2\lambda\sigma)(n_l + 2\lambda\sigma)}. \end{aligned}$$

We infer the lower bound

$$\left| \det \left( \frac{d^k}{d\sigma^k} \Omega_{n_l} \right)_{k,l=0}^K \right| \geq \frac{C}{n_K^{K(K+1)}} \quad (2.12)$$

with a constant  $C$  depending on  $K$ , a lower bound of  $1 + 2\lambda\rho^2$  and an upper bound of  $\rho$  (see also [4, Lemma 5.1]).

(b) Having established the lower bound (2.12) we can proceed as in [4, Lemmas/Corollaries 5.2–5.6] to show that for  $\alpha$  sufficiently large compared to  $r$  the set

$$\begin{aligned} \mathcal{P}_\gamma = \left\{ \sigma \in [0, \rho_0^2] : \left| \sum_{n=1}^N k_n \Omega_n + m \right| \geq \frac{\gamma}{N^\alpha} \text{ for all } N \geq 1, \text{ all } m \in \mathbb{Z} \right. \\ \left. \text{and all } 0 \neq k \in \mathbb{Z}^N \text{ with } \sum_{n=1}^N |k_n| \leq r \right\} \end{aligned}$$



has large Lebesgue measure,

$$|[0, \rho_0^2] \setminus \mathcal{P}_\gamma| \leq C\gamma^{1/r}.$$

In fact, for fixed  $k \in \mathbb{Z}^n$ , the lower bound (2.12) and equation (2.11) imply that  $|d^l/d\sigma^l \sum_{n=1}^N k_n \Omega_n| \geq CN^{-K(K+1)}$  for at least one derivative  $l \leq \sum_n |k_n|$  (see Lemma 5.2 and Corollary 5.3 in [4]). This estimate of a derivative can be lifted to a corresponding estimate of  $|\sum_{n=1}^N k_n \Omega_n + m|$  with  $m \in \mathbb{Z}$  for many values of  $\sigma$  (see Lemma 5.4 and Corollary 5.5 in [4]). Finally one considers the intersection over all  $k \in \mathbb{Z}^N$ , all  $N \geq 1$  and all  $m \in \mathbb{Z}$  (see Lemma 5.6 in [4]).

(c) Part (b) establishes the non-resonance condition stated in the lemma with the largest integer among  $m_1, \dots, m_p, n_1, \dots, n_q$  instead of the third-largest. In order to get the third-largest index instead we proceed as in Lemma 5.7 of [4] using the asymptotic behaviour

$$\Omega_n = n + \lambda\sigma - \frac{\sigma^2}{2(n + \lambda\sigma)} + \mathcal{O}\left(\frac{1}{n^2}\right) \quad (2.13)$$

of the frequencies. Indeed, considering a linear combination of frequencies as in (2.10) we distinguish three cases depending on the size of the largest index  $\mu_1(m, n)$  and of the second largest index  $\mu_2(m, n)$  among the integers  $m_1, \dots, m_p, n_1, \dots, n_q$  in comparison to the third-largest index  $\mu_3(m, n)$ : Either we can bound  $\mu_1(m, n)$  in terms of  $\mu_3(m, n)$  and apply part (b) directly, or we can consider  $\Omega_{\mu_1(m, n)} - \Omega_{\mu_2(m, n)}$  by (2.13) as a small perturbation of an integer number (taking the role of  $m$  in  $\mathcal{P}_\gamma$  of part (b)), or in the case that  $\Omega_{\mu_1}$  and  $\Omega_{\mu_2}$  have the same sign in (2.10) and  $\mu_2(m, n)$  is large compared to  $\mu_3(m, n)$  we can exclude any near-resonance.  $\square$

## 2.5 The transformed Hamiltonian system

Applying the symplectic<sup>1</sup> linear transformation

$$\begin{pmatrix} z_j \\ \bar{z}_{-j} \end{pmatrix} = S_n \begin{pmatrix} \xi_j \\ \bar{\xi}_{-j} \end{pmatrix} \quad \text{for } j \in \mathcal{Z} \text{ and } n = |j|^2,$$

with the matrices  $S_n$  of Lemma 2.1, to the Hamiltonian system (2.6) of equations for  $z_j$ , we end up with a Hamiltonian system

$$i \frac{d}{dt} \xi_j(t) = \frac{\partial H}{\partial \bar{\xi}_j}(\xi(t), \overline{\xi(t)}), \quad j \in \mathcal{Z} = \mathbb{Z}^d \setminus \{0\},$$

---

<sup>1</sup> The transformation is symplectic with respect to the standard form  $\omega(\xi, \xi') = \sum_j \text{Im}(\xi_j \bar{\xi}'_j)$  as a direct calculation using the real symplecticity of  $S_n$  ( $\det(S_n) = 1$ ) shows.

with the real-valued Hamilton function

$$H(\xi, \bar{\xi}) = \tilde{H}(z, \bar{z}),$$

with  $\tilde{H}$  of (2.7). This Hamiltonian is of the form

$$H(\xi, \bar{\xi}) = \sum_{j \in \mathcal{Z}} \omega_j |\xi_j|^2 + P(\xi, \bar{\xi}), \quad (2.14)$$

where the frequencies are  $\omega_j = \Omega_n$  for  $|j|^2 = n$  with  $\Omega_n = \sqrt{n^2 + 2n\lambda\rho^2}$ , and the non-quadratic term  $P$  is of the form

$$P(\xi, \bar{\xi}) = \sum_{p+q \geq 3} \sum_{\substack{k \in \mathcal{Z}^p, l \in \mathcal{Z}^q \\ \mathcal{M}(k,l)=0}} H_{kl} \xi_{k_1} \cdots \xi_{k_p} \bar{\xi}_{l_1} \cdots \bar{\xi}_{l_q}, \quad (2.15)$$

where the sum is still only over multi-indices with zero momentum (2.8), since the transformation mixes only terms that give the same contribution to the momentum. From (2.9) and Lemma 2.1 we obtain the following bound for the Taylor coefficients.

**Lemma 2.3.** *There exist  $M > 0$  and  $L > 0$  such that for all positive integers  $p, q$  with  $p + q \geq 3$  the coefficients in (2.15) are bounded by*

$$|H_{kl}| \leq M L^{p+q} \quad \text{for all } k \in \mathcal{Z}^p, l \in \mathcal{Z}^q.$$

The Hamiltonian equations of motion are now

$$i \frac{d}{dt} \xi_j(t) = \omega_j \xi_j(t) + \frac{\partial P}{\partial \bar{\xi}_j}(\xi(t), \bar{\xi}(t)), \quad j \in \mathcal{Z}, \quad (2.16)$$

where the nonlinearity is of the form

$$\frac{\partial P}{\partial \bar{\xi}_j}(\xi, \bar{\xi}) = \sum_{p+q \geq 2} \sum_{\substack{k \in \mathcal{Z}^p, l \in \mathcal{Z}^q \\ \mathcal{M}(k,l)=j}} P_{j,k,l} \xi_{k_1} \cdots \xi_{k_p} \bar{\xi}_{l_1} \cdots \bar{\xi}_{l_q} \quad (2.17)$$

with  $P_{j,k,l}$  an integral multiple (at most  $(q+1)$  times) of  $H_{k,(l,j)}$ .

Note that after the change of variables, the weighted  $\ell^2$ -norm

$$\|\xi\|_s = \left( \sum_{j \in \mathcal{Z}} |j|^{2s} |\xi_j|^2 \right)^{\frac{1}{2}}$$

of the sequence  $\xi$  is equivalent to the Sobolev norm  $\|e^{-im \cdot \bullet} u - u_m\|_{H^s}$  of the corresponding function  $u$ ,

$$\hat{c} \|\xi\|_s \leq \|e^{-im \cdot \bullet} u - u_m\|_{H^s} \leq \hat{C} \|\xi\|_s \quad (2.18)$$

with positive constants depending on  $\lambda$  and  $\rho$ . In particular, under the assumptions of Theorem 1.1, the system (2.16) has small initial values whose  $\ell_s^2$  norm is of order  $\varepsilon$ .

### 3 Long-time near-conservation of super-actions

In this section we give the proof of Theorem 1.1. After the transformations of the previous section, we can verify that the conditions required to apply existing results on the long-time near-conservation of so-called super-actions are fulfilled. A transformation back to the original variables then gives Theorem 1.1.

#### 3.1 Super-actions

Without the nonlinearity  $\frac{\partial P}{\partial \bar{\xi}_j}$  in (2.16), the *actions*

$$I_j(\xi, \bar{\xi}) = |\xi_j|^2, \quad j \in \mathcal{Z},$$

would be exactly conserved along solutions of (2.16). In the presence of the nonlinearity and in view of the partial resonance  $\omega_j = \Omega_n$  for all  $j \in \mathcal{Z}$  with  $|j|^2 = n$  and the non-resonance of the  $\Omega_n$  as given by Lemma 2.2, there remains long-time near-conservation of *super-actions*

$$J_n(\xi, \bar{\xi}) = \sum_{|j|^2=n} I_j(\xi, \bar{\xi}), \quad n \in \mathbb{N}, \quad (3.1)$$

along solutions of (2.16) provided that the initial value is small. The precise result in our situation is the following.

**Theorem 3.1** (Long-time near-conservation of super-actions). *Fix  $N > 1$  arbitrarily. For every  $\rho_0 > 0$  such that  $1 + 2\lambda\rho_0^2 > 0$  there exist  $s_0 > 0$  and a set of full measure  $\mathcal{P}$  in the interval  $(0, \rho_0]$  such that for all  $s \geq s_0$  and  $\rho \in \mathcal{P}$  the following holds: There exist  $\varepsilon_0 > 0$  and  $C$  such that for small initial data satisfying*

$$\|\xi(0)\|_s \leq \varepsilon \leq \varepsilon_0,$$

*the super-actions of the solution of (2.16) starting with  $\xi(0)$  at  $t = 0$  are nearly conserved,*

$$\sum_{n \geq 1} n^s \frac{|J_n(\xi(t), \bar{\xi}(t)) - J_n(\xi(0), \bar{\xi}(0))|}{\varepsilon^2} \leq C\varepsilon^{\frac{1}{2}},$$

*over long times*

$$0 \leq t \leq \varepsilon^{-N}.$$

Since  $\|\xi\|_s^2 = \sum_{n \geq 1} n^s J_n(\xi, \bar{\xi})$ , this theorem implies that  $\|\xi(t)\|_s$  stays of order  $\varepsilon$  over long times  $t \leq \varepsilon^{-N}$ . When we transform this result back to the solution  $u$  of (1.1), we immediately get Theorem 1.1 on using (2.18).

There are two entirely different approaches to prove Theorem 3.1, *Birkhoff normal forms* and *modulated Fourier expansions*. Both approaches will be outlined in the following subsections. Each proof relies on a non-resonance condition on the frequencies  $\omega_j$  describing the linear part in (2.16), a regularity condition on the nonlinearity in (2.16) and a condition on the interaction of modes (zero-momentum condition). Based on Lemmas 2.2 and 2.3, these assumptions will be verified in the following subsections, separately for each approach since they are not exactly the same for both proofs. Once the conditions are verified, we can directly apply results from Bambusi & Grébert [2, 24] (using Birkhoff normal forms) and Gauckler [22] (using modulated Fourier expansions) to obtain Theorem 3.1.

## 3.2 Proof of Theorem 3.1 via Birkhoff normal forms

We follow the Birkhoff normal form approach as developed in [1, 2, 4, 24]. We verify that the assumptions of [24, Theorem 7.2] are fulfilled by the system (2.16).

### 3.2.1 Regularity of the nonlinearity

For multi-indices  $k = (k_1, \dots, k_p) \in \mathcal{Z}^p$  and  $l = (l_1, \dots, l_q) \in \mathcal{Z}^q$  we denote by  $\mu_i(k, l)$  the  $i$ -th largest integer among  $|k_1|, \dots, |k_p|, |l_1|, \dots, |l_q|$ , so that  $\mu_1(j) \geq \mu_2(j) \geq \mu_3(j) \geq \dots$ . Moreover, for a given positive radius  $r$ , we set

$$B_s(r) = \{(\xi, \bar{\xi}) \in \mathbb{C}^{\mathcal{Z}} \times \mathbb{C}^{\mathcal{Z}} : \|\xi\|_s \leq r\}.$$

To apply Theorem 7.2 in [24], the Hamilton function

$$H = H_0 + P$$

with

$$\begin{aligned} H_0(\xi, \bar{\xi}) &= \sum_{j \in \mathcal{Z}} \omega_j |\xi_j|^2 \\ P(\xi, \bar{\xi}) &= \sum_{p+q \geq 3} \sum_{\substack{k \in \mathcal{Z}^p, l \in \mathcal{Z}^q \\ \mathcal{M}(k, l) = 0}} H_{kl} \xi_{k_1} \cdots \xi_{k_p} \bar{\xi}_{l_1} \cdots \bar{\xi}_{l_q}, \end{aligned}$$

needs to satisfy a non-resonance condition on the frequencies, as provided by Lemma 2.2, and the following conditions on the non-quadratic part (see [24, Definition 4.4]):

(H1) There exists  $s_0 \geq 0$  such that for all  $s \geq s_0$ , there exists  $r > 0$  such that  $P \in \mathcal{C}^\infty(B_s(r), \mathbb{C})$ .

(H2) The Taylor coefficients  $H_{kl}$  satisfy the following property: for all  $(k, l) \in \mathcal{Z}^p \times \mathcal{Z}^q$  we have  $\overline{H}_{kl} = H_{lk}$ , and for all  $p, q$  with  $p + q \geq 3$ , there exists  $\nu \geq 0$  such that for every  $N \in \mathbb{N}$ , there exists  $C$  depending on  $N, p$  and  $q$  such that for all  $(k, l) \in \mathcal{Z}^p \times \mathcal{Z}^q$ ,

$$|H_{kl}| \leq C \mu_3(k, l)^\nu \left( \frac{\mu_3(k, l)}{\mu_3(k, l) + \mu_1(k, l) - \mu_2(k, l)} \right)^N. \quad (3.2)$$

The bound (3.2), used in many works on normal forms applied to nonlinear PDEs - see [15, 16, 1, 24, 3]- implies in particular that the nonlinearity acts on the ball  $B_s(r)$ . Moreover, it is preserved by the Poisson bracket of two functions and by the normal form construction under a non-resonance condition implying a control of the small denominator by the third largest integer.

We now show that (H1) and (H2) are implied by the coefficient estimates of Lemma 2.3 together with the fact that the Hamiltonian has only terms with zero momentum. Let us consider a fixed multi-index  $(k, l)$  satisfying  $\mathcal{M}(k, l) = 0$ . Following the proof of [24, Lemma 5.2], we see that we always have

$$|\mu_1(k, l) - \mu_2(k, l)| \leq |\mathcal{M}(k, l)| + \sum_{n=3}^{p+q} \mu_n(k, l) \leq (p + q - 2)\mu_3(k, l).$$

From this relation, we infer

$$\left( \frac{\mu_3(k, l)}{\mu_3(k, l) + \mu_1(k, l) - \mu_2(k, l)} \right)^N \geq (p + q - 1)^{-N}.$$

Using Lemma 2.3, we thus see that the coefficients  $H_{kl}$  satisfy the bound (3.2) with the constant  $C = ML^{p+q}(p + q - 1)^N$  and  $\nu = 0$ . This yields (H2). The assertion (H1) results from the fact that  $\sqrt{\rho^2 - \sum_{j \in \mathcal{Z}} z_j \bar{z}_j}$  is analytic on  $B_s(r)$  for  $r < \rho$ , and that the monomials with zero momentum terms define a smooth Hamiltonian as soon as  $s_0 > d/2$  (see for instance [4]). This ensures that  $P \in \mathcal{C}^\infty(B_s(r), \mathbb{C})$ .

### 3.2.2 A normal form result

For a given Hamiltonian  $K \in \mathcal{C}^\infty(B_s(r), \mathbb{C})$  satisfying  $K(\xi, \bar{\xi}) \in \mathbb{R}$  we denote by  $X_K(\xi, \bar{\xi})$  the Hamiltonian vector field

$$X_K(\xi, \bar{\xi})_j = \left( i \frac{\partial K}{\partial \xi_j}, -i \frac{\partial K}{\partial \bar{\xi}_j} \right), \quad j \in \mathcal{Z},$$

associated with the Poisson bracket

$$\{K, G\} = i \sum_{j \in \mathcal{Z}} \frac{\partial K}{\partial \xi_j} \frac{\partial G}{\partial \bar{\xi}_j} - \frac{\partial K}{\partial \bar{\xi}_j} \frac{\partial G}{\partial \xi_j},$$

which is well defined for Hamiltonian functions  $K$  and  $G$  in the class of Hamiltonians defined above.

We are now ready to apply Theorem 7.2 of [24] to the Hamiltonian (2.14). We obtain the following result:

**Theorem 3.2.** *Let  $\rho$  be in the set  $\mathcal{P}$  of full measure as given by Lemma 2.2 for some  $N \geq 3$ . There exists  $s_0$  and for any  $s \geq s_0$  there exist two neighborhoods  $\mathcal{U}$  and  $\mathcal{V}$  of the origin in  $B_s(\rho)$  and an analytic canonical transformation  $\tau : \mathcal{V} \rightarrow \mathcal{U}$  which puts  $H = H_0 + P$  in normal form up to order  $N$ , i.e.,*

$$H \circ \tau = H_0 + Z + R$$

where

- (i)  $Z$  is a polynomial of degree  $N$  which commutes with all the  $J_n$ ,  $n \geq 1$ , i.e.,  $\{Z, J_n\} = 0$  for all  $n \geq 1$ ,
- (ii)  $R \in \mathcal{C}^\infty(\mathcal{V}, \mathbb{R})$  and  $\|X_R(\xi, \bar{\xi})\|_s \leq C_s \|\xi\|_s^N$  for  $\xi \in \mathcal{V}$ ,
- (iii)  $\tau$  is close to the identity:  $\|\tau(\xi, \bar{\xi}) - (\xi, \bar{\xi})\|_s \leq C_s \|\xi\|_s^2$  for all  $\xi \in \mathcal{V}$ .

Let us recall the principle underlying the proof of this result: the construction of the transformation  $\tau$  is made by induction by cancelling iteratively the polynomials of growing degree in the Hamiltonian  $H_0 + P$ . As  $\tau$  is determined as the flow at time one of a polynomial Hamiltonian  $\chi = \sum_{n \geq 3} \chi_n$ , where  $\chi_n$  are homogeneous polynomials of degree  $n$ , we are led to solve by induction homological equations of the form

$$\{H_0, \chi_n\} = Z_n + Q_n$$

where  $Z_n$  is the  $n$ -th component of the normal form term, and  $Q_n$  a homogeneous polynomial of degree  $n$  depending on  $P$  and on the terms constructed at the previous iterations. Writing the equation in terms of coefficients, this equation can be written in the form

$$(\omega_{k_1} + \dots + \omega_{k_p} - \omega_{l_1} - \dots - \omega_{l_q}) \chi_{kl} = Z_{kl} + Q_{kl}$$

where  $(k, l) \in \mathcal{Z}^p \times \mathcal{Z}^q$  with  $p + q = n$ . Using the non-resonance condition (2.10), we see that we can solve this equation for  $\chi_{kl}$  and set  $Z_{kl} = 0$  without losing too much regularity (i.e.,  $\chi$  will satisfy (H2) for some  $\nu$ ), except for the multi-indices  $(k, l)$  having

equal length  $p = q$  and after permutation,  $|k_1|^2 = |l_1|^2, \dots, |k_p|^2 = |l_p|^2$ . This yields that the normal form term  $Z$  contains only terms of the form  $\xi_{k_1} \cdots \xi_{k_p} \bar{\xi}_{l_1} \cdots \bar{\xi}_{l_q}$  with  $p = q$  and  $|k_1|^2 = |l_1|^2, \dots, |k_p|^2 = |l_p|^2$ . We then check that these terms Poisson-commute with  $J_n$  for all  $n$ , and hence  $\{J_n, Z\} = 0$  for all  $n$ .

The proof of Theorem 3.1 can then be done using  $\|\xi\|_s^2 = \sum_{n \geq 1} n^s J_n(\xi, \bar{\xi})$  and following the proof of Corollary 4.9 in [24].

### 3.3 Proof of Theorem 3.1 via modulated Fourier expansions

Modulated Fourier expansions in time have been developed as a technique for analysing weakly nonlinear oscillatory systems over long times, both continuous and discrete systems, in finite and infinite dimensions. There are two ingredients:

- a solution approximation over short time (the modulated Fourier expansion properly speaking)
- almost-invariants of the modulation system.

The technique can be viewed as embedding the original system in a larger modulation system that turns out to have a Hamiltonian/Lagrangian structure with an invariance property from which results on the long-time behaviour can be inferred.

Modulated Fourier expansions were originally introduced in [26] to explain the long-time behaviour of numerical methods for oscillatory ordinary differential equations; see also [28, Chap. XIII]. In [13, 23] and [21] they were used to study long-time properties of small solutions of nonlinear wave equations and nonlinear Schrödinger equations with an external potential, respectively, and in [22] for general classes of Hamiltonian partial differential equations. For a recent review of the technique and various of its uses see [27].

Let us assume for the moment that there are no resonances among the frequencies, in particular  $\omega_j \neq \omega_{j'}$  also for  $|j| = |j'|$ . A modulated Fourier expansion of the solution  $\xi$  of (2.16) is an approximation of  $\xi$  in terms of products of propagators  $e^{-i\omega_j t}$  for the linear equation with slowly varying coefficient functions,

$$\xi_j(t) \approx \tilde{\xi}_j(t) = \sum_{\mathbf{k}} z_j^{\mathbf{k}}(\varepsilon t) e^{-i(\mathbf{k} \cdot \boldsymbol{\omega})t}, \quad j \in \mathcal{Z}, \quad (3.3)$$

where the sum runs over a finite set of sequences of integers  $\mathbf{k} = (\mathbf{k}(\ell))_{\ell \in \mathcal{Z}} \in \mathbb{Z}^{\mathcal{Z}}$  with finitely many nonzero entries, and where  $\mathbf{k} \cdot \boldsymbol{\omega} = \sum_{\ell \in \mathcal{Z}} \mathbf{k}(\ell) \omega_\ell$ .

Inserting the ansatz (3.3) into the equations of motion (2.16) and equating terms with the same exponential  $e^{-i(\mathbf{k}\cdot\boldsymbol{\omega})t}$  leads to the *modulation system*

$$\begin{aligned} & i\varepsilon \dot{z}_j^{\mathbf{k}} + (\mathbf{k} \cdot \boldsymbol{\omega}) z_j^{\mathbf{k}} \\ &= \omega_j z_j^{\mathbf{k}} + \sum_{p+q \geq 2}^{\infty} \sum_{\substack{\mathbf{k}^1 + \dots + \mathbf{k}^p \\ -\mathbf{l}^1 - \dots - \mathbf{l}^q = \mathbf{k}}} \sum_{\substack{k \in \mathcal{Z}^p, l \in \mathcal{Z}^q \\ \mathcal{M}(k,l)=j}} P_{j,k,l} z_{k_1}^{\mathbf{k}^1} \cdots z_{k_p}^{\mathbf{k}^p} \bar{z}_{l_1}^{\mathbf{l}^1} \cdots \bar{z}_{l_q}^{\mathbf{l}^q} \end{aligned}$$

with the coefficients  $P_{j,k,l}$  of (2.17). Modulated Fourier expansions can hence be seen as embedding the original system of equations in a larger system. The nonlinearity is the partial derivative with respect to  $\bar{z}_j^{\mathbf{k}}$  of the modulation potential

$$\mathcal{P}(\mathbf{z}, \bar{\mathbf{z}}) = \sum_{p+q \geq 3} \sum_{\substack{\mathbf{k}^1 + \dots + \mathbf{k}^p \\ -\mathbf{l}^1 - \dots - \mathbf{l}^q = \mathbf{0}}} \sum_{\substack{k \in \mathcal{Z}^p, l \in \mathcal{Z}^{q+1} \\ \mathcal{M}(k,l)=0}} H_{kl} z_{k_1}^{\mathbf{k}^1} \cdots z_{k_p}^{\mathbf{k}^p} \bar{z}_{l_1}^{\mathbf{l}^1} \cdots \bar{z}_{l_{q+1}}^{\mathbf{l}^{q+1}}$$

with  $\mathbf{z} = (z_j^{\mathbf{k}})_{j,\mathbf{k}}$  and with the coefficients  $H_{kl}$  of the non-quadratic part of the Hamiltonian  $H(\xi, \bar{\xi})$ . The modulation potential is invariant under transformations  $z_j^{\mathbf{k}} \mapsto e^{i\mathbf{k}(\ell)\theta} z_j^{\mathbf{k}}$  for  $\theta \in \mathbb{R}$  and fixed  $\ell \in \mathcal{Z}$ . The modulation system thus inherits the Hamiltonian structure from the original equations of motions for  $\xi$ , and its transformation invariance leads to formal invariants

$$\mathcal{I}_\ell(\mathbf{z}, \bar{\mathbf{z}}) = \sum_{j,\mathbf{k}} \mathbf{k}(\ell) |z_j^{\mathbf{k}}|^2, \quad \ell \in \mathcal{Z},$$

of the modulation system, see [22, Sect. 3.1]. These formal invariants form the cornerstone for the study of long time intervals.

On a short time interval of length  $\varepsilon^{-1}$  it is possible to construct an approximate solution of the modulation system in an iterative way, such that—under certain assumptions to be verified below—the ansatz (3.3) describes the solution  $\xi$  up to a small error,

$$\|\xi(t) - \tilde{\xi}(t)\|_s \leq C\varepsilon^{N+3} \quad \text{for } 0 \leq t \leq c\varepsilon^{-1}$$

with  $\tilde{\xi}$  given by (3.3) with the approximate solution of the modulation system, see [22, Sects. 3.2–3.4]. The constants depend on  $N$  and  $s$  but not on  $\varepsilon$ . Along this approximate solution  $\mathbf{z}$  of the modulation system, the formal invariants  $\mathcal{I}_\ell$  then become almost-invariants,

$$\sum_{\ell \in \mathcal{Z}} |\ell|^{2s} \left| \frac{d}{dt} \mathcal{I}_\ell(\mathbf{z}(\varepsilon t), \overline{\mathbf{z}(\varepsilon t)}) \right| \leq C\varepsilon^{N+2},$$



which are close to the actions  $I_\ell(\xi, \bar{\xi}) = |\xi_\ell|^2$ ,

$$\sum_{\ell \in \mathcal{Z}} |\ell|^{2s} |\mathcal{I}_\ell(\mathbf{z}(\varepsilon t), \overline{\mathbf{z}(\varepsilon t)}) - I_\ell(\xi(t), \overline{\xi(t)})| \leq C\varepsilon^{\frac{5}{2}}.$$

These almost-invariants allow us to repeat the construction of modulated Fourier expansions on short time intervals of length  $\varepsilon^{-1}$  and patch  $\varepsilon^{-N+1}$  of those short intervals together, see [22, Sect. 3.5]. On a long time interval of length  $\varepsilon^{-N}$  we then get near-conservation of actions as stated in Theorem 3.1 (with actions instead of super-actions).

Compared to the above description, the modulated Fourier expansion for our problem (2.16) has some subtleties that are caused by the partial resonances  $\omega_j = \omega_{j'}$  for  $|j| = |j'|$ . Since all sums in the nonlinearity of (2.16) involve only products of the form  $\xi_{k_1} \cdots \xi_{k_p} \bar{\xi}_{l_1} \cdots \bar{\xi}_{l_q}$  with  $k_1 + \cdots + k_p - l_1 - \cdots - l_q = j$  (by the zero momentum condition in the Hamiltonian), only modulation functions  $z_j^{\mathbf{k}}$  with

$$j = j(\mathbf{k}) = \sum_{\ell \in \mathcal{Z}} \mathbf{k}(\ell)\ell$$

can be different from zero. Moreover, since the frequencies  $\omega_j$  in (2.16) are partially resonant,  $\omega_j = \omega_{j'}$  for  $|j| = |j'|$ , we can distinguish exponentials  $e^{-i(\mathbf{k}^1 \cdot \boldsymbol{\omega})t}$  and  $e^{-i(\mathbf{k}^2 \cdot \boldsymbol{\omega})t}$  only if

$$\mathbf{k}^1 - \mathbf{k}^2 \notin \left\{ \mathbf{k} : \sum_{|\ell|^2=n} \mathbf{k}(\ell) = 0 \text{ for all } n \in \mathbb{N} \right\}.$$

For this reason, the sum in (2.16) is in our situation only over a set of representatives of sequences  $\mathbf{k}$  where  $j(\mathbf{k})$  or  $\mathbf{k} \cdot \boldsymbol{\omega}$  are distinguishable (in the above sense). The main consequence is that the quantities  $\mathcal{I}_\ell$  from above are no longer invariants of the modulation system, but only certain sums of them:

$$\mathcal{J}_n(\mathbf{z}, \bar{\mathbf{z}}) = \sum_{\ell \in \mathcal{Z}: |\ell|^2=n} \mathcal{I}_\ell(\mathbf{z}, \bar{\mathbf{z}}), \quad n \in \mathbb{N}.$$

Along the approximate solution of the modulation system, they are close to the corresponding sums of the actions  $I_\ell$ , the super-actions  $J_n$ . In this way we get long-time near-conservation of super-actions as in Theorem 3.1.

We finally state and verify the assumptions needed for the iterative construction of modulation functions. The first lemma below summarises the assumptions on the nonlinearity in (2.16), whereas the second lemma below deals with the non-resonance condition on the frequencies describing the linear part of (2.16). The properties stated in these lemmas are precisely the assumptions under which Theorem 3.1 has been shown in [22, Theorem 2.7].

**Lemma 3.3.** *The expansion (2.17) of the nonlinearity in (2.16) has the following properties.*

(i) *It fulfills the zero momentum condition*

$$P_{j,k,l} = 0 \quad \text{if } j \neq \mathcal{M}(k,l)$$

for  $j \in \mathcal{Z}$ ,  $k \in \mathcal{Z}^p$  and  $l \in \mathcal{Z}^q$ .

(ii) *There exist constants  $C_{p,q,s}$  depending only on  $p$ ,  $q$ ,  $s$  and  $\rho$  such that for*

$$|P|_j^{p,q}(\xi^1, \dots, \xi^p, \bar{\xi}^1, \dots, \bar{\xi}^q) = \sum_{k \in \mathcal{Z}^p, l \in \mathcal{Z}^q} |P_{j,k,l}| \xi_{k_1}^1 \cdots \xi_{k_p}^p \bar{\xi}_{l_1}^1 \cdots \bar{\xi}_{l_q}^q$$

the estimate

$$\| |P|^{p,q}(\xi^1, \dots, \xi^p, \bar{\xi}^1, \dots, \bar{\xi}^q) \|_s \leq C_{p,q,s} \|\xi^1\|_s \cdots \|\xi^p\|_s \|\bar{\xi}^1\|_s \cdots \|\bar{\xi}^q\|_s \quad (3.4)$$

holds for  $\xi^1, \dots, \xi^p, \bar{\xi}^1, \dots, \bar{\xi}^q \in \ell_s^2$  if  $s > \frac{d}{2}$ .

(iii) *There exist  $r_0 > 0$  depending only on  $\rho$ , and  $C_s$  depending in addition on  $s > \frac{d}{2}$  such that*

$$\sum_{p+q \geq 2} C_{p,q,s} |z|^{p+q-2} \leq C_s \quad \text{for all } z \in \mathbb{C} \text{ with } |z| \leq r_0.$$

*Proof.* Property (i) is obvious, see (2.17).

For property (ii) we recall that (3.4) was verified in [22, Subsect. 2.6] in the situation  $P_{j,k,l} = 0$  for  $j \neq \mathcal{M}(k,l)$  and  $P_{j,k,l} = 1$  else. The proof is just a repeated application of the Cauchy-Schwarz inequality, and the corresponding constants  $C_{p,q,s}$  are given by  $C^{p+q}$  with  $C$  depending only on  $s$ . In our situation here, the coefficients  $P_{j,k,l}$  vanish for  $j \neq \mathcal{M}(k,l)$  and can be bounded with Lemma 2.3. This implies that the second property (ii) is satisfied with constants  $C_{p,q,s} = (q+1)ML^{p+q+1}C^{p+q}$ .

These constants satisfy (iii) with  $r_0$  and  $C_s$  depending only on  $\rho$  and  $s$ .  $\square$

**Lemma 3.4.** *The frequencies  $\Omega_n$  grow like  $n$ ,  $c_1 n \leq \Omega_n \leq C_1 n$  with positive constants  $c_1$  and  $C_1$  depending only on  $\rho$ .*

*Moreover, for all  $\rho_0 > 0$  such that  $1 + 2\lambda\rho_0^2 > 0$  and for all positive integer  $N$ , there exist  $s_0$  and a set of full measure  $\mathcal{P}$  in the interval  $(0, \rho_0]$  such that for all  $s \geq s_0$  and all  $\rho \in \mathcal{P}$  the following non-resonance condition holds: There exist  $\varepsilon_0 > 0$  and  $C_0$  such that for all  $r \leq 2N + 2d + 6$  and all  $0 < \varepsilon \leq \varepsilon_0$ ,*

$$\left( \frac{n}{n_1 \cdots n_r} \right)^{s - \frac{d+1}{2}} \varepsilon^r \leq C_0 \varepsilon^{2N+2d+8}$$

whenever a near-resonance

$$|\Omega_n \pm \Omega_{n_1} \pm \cdots \pm \Omega_{n_r}| < \varepsilon^{\frac{1}{2}}$$

occurs with frequencies that do not cancel pairwise.

*Proof.* The asymptotic growth behaviour of the frequencies is obvious, and the non-resonance condition is implied by the non-resonance condition of Lemma 2.2 as shown in [13, Lemma 1].  $\square$

### 3.4 Proof of the orbital stability (1.3)

We have

$$\inf_{\varphi \in \mathbb{R}} \|e^{-im \cdot \bullet} u(\bullet, t) - e^{i\varphi} u_m(0)\|_{H^s}^2 = \left| |u_m(t)| - |u_m(0)| \right|^2 + \|e^{-im \cdot \bullet} u(\bullet, t) - u_m(t)\|_{H^s}^2,$$

and by the conservation of the  $L^2$  norm

$$\begin{aligned} \left| |u_m(t)| - |u_m(0)| \right|^2 &\leq \left| |u_m(t)|^2 - |u_m(0)|^2 \right| = \left| \sum_{m \neq j \in \mathbb{Z}^d} |u_j(0)|^2 - \sum_{m \neq j \in \mathbb{Z}^d} |u_j(t)|^2 \right| \\ &\leq \max(\|e^{-im \cdot \bullet} u(\bullet, 0) - u_m(0)\|_{L^2}^2, \|e^{-im \cdot \bullet} u(\bullet, t) - u_m(t)\|_{L^2}^2). \end{aligned}$$

The estimate (1.3) thus follows from Theorem 1.1.

## References

- [1] D. Bambusi: Birkhoff normal form for some nonlinear PDEs. *Comm. Math. Phys.* **234**, 253–283 (2003).
- [2] D. Bambusi: A Birkhoff normal form theorem for some semilinear PDEs. *Hamiltonian Dynamical Systems and Applications*, Springer, 2007, pp. 213–247.
- [3] D. Bambusi, J.-M. Delort, B. Grébert and J. Szeftel: Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds. *Comm. Pure Appl. Math.* **60**, 1665–1690 (2007).
- [4] D. Bambusi and B. Grébert: Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.* **135**, 507–567 (2006).
- [5] J. Bourgain: On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices* **6**, 277–304 (1996).

- [6] J. Bourgain: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math. (2)* **148**, 363–439 (1998).
- [7] J. Bourgain: Problems in Hamiltonian PDE's. *Geom. Funct. Anal.*, Special Volume, Part I, 32–56 (2000).
- [8] J. Bourgain: On diffusion in high-dimensional Hamiltonian systems and PDE. *J. Anal. Math.* **80**, 1–35 (2000).
- [9] R. Carles, E. Dumas and Ch. Sparber: Multiphase weakly nonlinear geometric optics for Schrödinger equations. *SIAM J. Math. Anal.* **42**, 489–518 (2010).
- [10] R. Carles and E. Faou: Energy cascades for NLS on the torus. *Discrete Contin. Dyn. Syst.* **32**, 2063–2077 (2012).
- [11] M. Christ, J. Colliander and T. Tao: Ill-posedness for nonlinear Schrödinger and wave equations. *Annales de l'IHP*, to appear. arXiv:math/0311048v1
- [12] M. Christ, J. Colliander and T. Tao: Instability of the periodic nonlinear Schrödinger equation. Preprint, 2003. arXiv:math/0311227v1
- [13] D. Cohen, E. Hairer and Ch. Lubich: Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions. *Arch. Ration. Mech. Anal.* **187**, 341–368 (2008).
- [14] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.* **181**, 39–113 (2010).
- [15] J. M. Delort and J. Szeftel: Long-time existence for small data nonlinear Klein-Gordon equations on tori and spheres. *Internat. Math. Res. Notices* **37**, 1897–1966 (2004).
- [16] J. M. Delort and J. Szeftel: Long-time existence for semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds. *Amer. J. Math.* **128**, 1187–1218 (2006).
- [17] L. H. Eliasson and S. B. Kuksin: KAM for the nonlinear Schrödinger equation. *Ann. of Math. (2)* **172**, 371–435 (2010).
- [18] E. Faou and B. Grébert: A Nekhoroshev type theorem for the nonlinear Schrödinger equation on the d-dimensional torus. Preprint. arXiv:1003.4845v1

- [19] T. Gallay and M. Haragus: Stability of small periodic waves for the nonlinear Schrödinger equation. *J. Diff. Equations* **234**, 544–581 (2007).
- [20] T. Gallay and M. Haragus: Orbital stability of periodic waves for the nonlinear Schrödinger equation. *J. Dyn. Diff. Eqns.* **19**, 825–865 (2007).
- [21] L. Gauckler and Ch. Lubich: Nonlinear Schrödinger equations and their spectral semi-discretizations over long times. *Found. Comput. Math.* **10**, 141–169 (2010).
- [22] L. Gauckler: *Long-time analysis of Hamiltonian partial differential equations and their discretizations*. Dissertation (doctoral thesis), Univ. Tübingen, 2010. <http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-47540>.
- [23] L. Gauckler, E. Hairer, Ch. Lubich and D. Weiss: Metastable energy strata in weakly nonlinear wave equations. *Comm. Partial Differential Equations* **37**, 1391–1413 (2012).
- [24] B. Grébert: Birkhoff normal form and Hamiltonian PDEs. *Sémin. Cong.* **15**, 1–46 (2007).
- [25] M. Guardia and V. Kaloshin: Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. Preprint, 2012. arXiv:1205.5188v1
- [26] E. Hairer and Ch. Lubich: Long-time energy conservation of numerical methods for oscillatory differential equations. *SIAM J. Numer. Anal.* **38**, 414–441 (2000).
- [27] E. Hairer and Ch. Lubich: Modulated Fourier expansions for continuous and discrete oscillatory systems. In *Foundations of Computational Mathematics, Budapest 2011*. Cambridge Univ. Press, 2012.
- [28] E. Hairer, Ch. Lubich and G. Wanner: *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*. 2nd ed., Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2006.
- [29] Z. Hani: *Global and dynamical aspects of nonlinear Schrödinger equations on compact manifolds*. PhD thesis, University of California Los Angeles, 2011.
- [30] P. Zhidkov: *Korteweg-de Vries and nonlinear Schrödinger equations: qualitative theory*. Lecture Notes in Mathematics 1756, Springer-Verlag, Berlin, 2001.