Nonlinearity **12** (1999) 1587–1600. Printed in the UK PII: S0951-7715(99)96754-3

On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi

Jürgen Pöschel
Mathematisches Institut A, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany
E-mail: poschel@mathematik.uni-stuttgart.de

Received 13 August 1998, in final form 10 May 1999
Recommended by J Laskar

Abstract. We consider the nonlinear Schrödinger equation

$$\mathrm{i}u_t = u_{xx} - mu - f(|u|^2)u$$

on a finite x-interval with Dirichlet boundary conditions. Assuming that f is real analytic with $f(0) = 0$ and $f'(0) \neq 0$, we show that the equilibrium solution $u \equiv 0$ enjoys a certain kind of Nekhoroshev stability. If most of the energy is located in finitely many modes and sufficiently small, then the amplitudes of these modes are almost constant over a time interval,which is exponentially long in the inverse of the total energy.

This result is due to Bambusi, but the proof given here is conceptually and technically simpler. It may also apply to a larger class of nonlinear partial differential equations.

AMS classification scheme numbers: 35B35, 35Q55

1. Result

We consider the nonlinear Schrödinger equation

$$\mathrm{i}u_t = u_{xx} - mu - f(|u|^2)u \tag{1}$$

on the finite x-interval $[0, \pi]$ with Dirichlet boundary conditions. The parameter m is real, and f is supposed to be real analytic in some neighbourhood of the origin. Absorbing a constant into m, we may assume that $f(0) = 0$. In addition, we require that

$$f'(0) \neq 0.$$

Thus, we have the equation

$$\mathrm{i}u_t = u_{xx} - mu \mp |u|^2 u + \mathrm{O}_5(u) \tag{2}$$

after rescaling u.

We are going to describe some stability properties of the equilibrium solution $u \equiv 0$. In particular, we show that under certain circumstances the amplitudes of finitely many modes are almost constant over exponentially long time intervals. To state this result, write

$$u = \sum_{k\geqslant 1} q_k \varphi_k(x)$$

with complex coefficients q_k and $\varphi_k = \sqrt{2/\pi}\,\sin kx$, and introduce norms $\|\cdot\|_p$ by

$$\|u\|_p^2 = \sum_{k\geqslant 1} k^{2p}|q_k|^2. \tag{3}$$

0951-7715/99/061587+14$30.00

We are interested in the amplitudes of the basic modes φ_k,

$$I = (I_1, I_2, \ldots), \qquad I_k = \tfrac{1}{2}|q_k|^2,$$

as a function of time along a solution of the Schrödinger equation. The following result is a slight refinement of a theorem due to Bambusi [1].

Theorem 1.1 (Bambusi [1]). *If $\|u^{\mathrm{o}}\|_1$ is sufficiently small, then the solution of the nonlinear Schrödinger equation (2) with initial value u^{o} exists for all times and satisfies*

$$\|u(t)\|_1 \leqslant 2\|u^{\mathrm{o}}\|_1 \qquad \text{for} \quad -\infty < t < \infty.$$

Moreover, if $K \subset \mathbb{Z}$ is a finite index set of cardinality n, if δ is sufficiently small depending on K, and

$$\|u^{\mathrm{o}}\|_1 \leqslant \delta, \qquad \|u^{\mathrm{o}} - \pi_K u^{\mathrm{o}}\|_0 \leqslant \delta^{3/2}, \tag{4}$$

where π_K denotes the orthogonal projection onto span $\{\varphi_k : k \in K\} \subset L^2_{\mathbb{C}}[0, \pi]$, *then*

$$\|I(t) - I^{\mathrm{o}}\|_0 \leqslant c_1\delta^{2+2a} \qquad \text{for} \quad |t| \leqslant \exp\left(\frac{1}{c_2\delta^{2a}}\right),$$

where $a = \frac{1}{2n}$ and $c_i = \sqrt{n}\tilde{c}_i$ with universal constants $\tilde{c}_i$ for $i = 1, 2$.

Remark 1. The two estimates of the theorem may be combined into an estimate of $\|I(t) - I^{\mathrm{o}}\|_1$ over exponentially long time intervals.

Remark 2. We did not try to make explicit how small δ has to be as a function of K and the nonlinearity. This requires more technical estimates. On the other hand, the results are independent of the parameter m.

Remark 3. Condition (4) is satisfied for certain classes of entire functions, as explained by Bambusi [1]. It is an open question, however, whether the result extends to general analytic initial data u^{o}.

Remark 4. The Nekhoroshev estimate is nontrivial only if actually $\|u^{\mathrm{o}}\|_1$ is of the order of δ. If $\|u^{\mathrm{o}}\|_1 \leqslant \delta^{3/2}$, then the first estimate implies that

$$\|I(t)\|_0 \leqslant \|u(t)\|_0^2 \leqslant 4\delta^3$$

and thus $\|I(t) - I^{\mathrm{o}}\|_0 \leqslant 8\delta^3 \leqslant c_1\delta^{2+2a}$ for *all* t.

The rest of this paper is devoted to the proof of theorem 1.1. It is based on Niederman's proof of the Nekhoroshev stability of elliptic equilibria in classical Hamiltonian systems with a certain convexity condition, as described in [5, 6]. This proof employs Lochak's idea [3] of approximating arbitrary orbits by periodic ones and analysing the stability properties of the latter.

The proof given here is conceptually and technically simpler than Bambusi's original proof in [1]. First, there is no need to apply *two* normalizing transformations with exponentially small remainders. Second, by using Niederman's approach, there is no need to introduce action-angles for certain 'large' coordinates. All in all, the proof appears as a rather straightforward generalization from the finite-dimensional case as described in [6].

The result is not restricted to the nonlinear Schrödinger equation (1). Similar results should hold for those partial differential equations for which a KAM theorem is available, such as the nonlinear wave equation on a finite interval and perturbations of the periodic KdV equation. Other classes of equations are desribed in the remarks in [1] at the end of the first section.

In the next section we reduce theorem 1.1 to a stability result about an infinite dimensional Hamiltonian in Birkhoff normal form. The latter is then proven by extending the arguments in [6].

2. Preparations

As is well known, the nonlinear Schrödinger equation (1) can be written as a Hamiltonian system, such that

$$\dot{u} = \mathrm{i}\nabla H(u)$$

with

$$H(u) = \tfrac{1}{2}\langle Au, u\rangle + \tfrac{1}{2}\int_0^\pi g(|u|^2)\,\mathrm{d}x,$$

where $A = \partial_x^2 + m$ and $g' = f$, and the gradient of H is taken with respect to the inner product $\langle u, v\rangle = \operatorname{Re}\int_0^\pi u\bar{v}\,\mathrm{d}x$. The underlying phase space is $H_0^1[0,\pi]$, the Hilbert space of all complex valued L^2-functions on $[0,\pi]$ with an L^2-derivative and vanishing boundary values.

Following [2] we rewrite this system as a classical Hamiltonian system in infinitely many coordinates. Writing

$$u = \sum_{k\geqslant 1} q_k\varphi_k, \qquad \varphi_k = \sqrt{\frac{2}{\pi}}\sin kx,$$

as above, we obtain

$$\begin{aligned} H &= \Lambda + G \\ &= \tfrac{1}{2}\sum_{k\geqslant 1}\lambda_k|q_k|^2 + \tfrac{1}{2}\int_0^\pi g(|u|^2)\,\mathrm{d}x \end{aligned}$$

with $\lambda_k = k^2 + m$. We view H as a Hamiltonian in the real and imaginary part of $q_k = x_k - \mathrm{i}y_k$. Thus, $|q_k|^2 = x_k^2 + y_k^2$, the equations of motion are

$$\dot{x}_k = \frac{\partial H}{\partial y_k}, \qquad \dot{y}_k = -\frac{\partial H}{\partial x_k},$$

and

$$X_H = \sum_{k\geqslant 1}\left(\frac{\partial H}{\partial y_k}\frac{\partial}{\partial x_k} - \frac{\partial H}{\partial x_k}\frac{\partial}{\partial y_k}\right)$$

is the Hamiltonian vector field associated with H.

We study this Hamiltonian on a phase space

$$\mathcal{S}_p = \ell_p^2 \times \ell_p^2 \ni (x, y) = z$$

with squared norm $\|z\|_p^2 = \langle z, z\rangle_p = \|x\|_p^2 + \|y\|_p^2$, where ℓ_p^2 is the Hilbert space of all real sequences $x = (x_1, x_2, \ldots)$ with squared norm

$$\|x\|_p^2 = \langle x, x\rangle_p = \sum_{k\geqslant 1} k^{2p}x_k^2 < \infty. \tag{5}$$

The symplectic structure is the usual one,

$$\sum_{k\geqslant 1}\mathrm{d}x_k\wedge\mathrm{d}y_k = \frac{1}{2\mathrm{i}}\sum_{k\geqslant 1}\mathrm{d}q_k\wedge\mathrm{d}\bar{q}_k.$$

In the case of the nonlinear Schrödinger equation we may fix $p = 1$.

The following lemma is proven in [2].

Lemma 2.1. *For $p > \frac{1}{2}$, the Hamiltonian vector field X_G is real analytic as a map from a neighbourhood of the origin in $\mathcal{S}_p$ into $\mathcal{S}_p$, with*

$$\|X_G\|_p = \mathrm{O}(\|z\|_p^3).$$

Thus, X_G is a genuine vector field on $\mathcal{S}_p$. On the other hand, the linear vector field X_Λ is unbounded on $\mathcal{S}_p$, since it takes values in $\mathcal{S}_{p-2}$.

From now on we consider the setting of equation (2). The sign and higher-order terms being irrelevant for our arguments, we may even assume the nonlinearity to be $|u|^2 u$. We then find

$$G = \tfrac{1}{4}\int_0^\pi |u|^4 \,\mathrm{d}x = \tfrac{1}{4}\sum_{i,j,k,l} G_{ijkl} q_i q_j \bar{q}_k \bar{q}_l$$

with $G_{ijkl} = \int_0^\pi \varphi_i\varphi_j\varphi_k\varphi_l \,\mathrm{d}x$, and an easy calculation gives $G_{ijij} = (2+\delta_{ij})/2\pi$. It turns out that all other nonzero coefficients can be removed by a symplectic transformation, thus putting the Hamiltonian in a Birkhoff normal form up to order four. The following result is proven in [2].

Lemma 2.2. *For the Hamiltonian $H = \Lambda + G$ there exists a real analytic, symplectic change of coordinates Γ in a neighbourhood of the origin in $\mathcal{S}_p$, $p > \frac{1}{2}$, that takes H into its Birkhoff normal form up to order four. That is,*

$$H \circ \Gamma = \Lambda + \bar{G} + K,$$

such that $X_{\bar{G}}$ and X_K are real analytic vector fields defined in a neighbourhood of the origin in $\mathcal{S}_p$ with $\|X_K\|_p = \mathrm{O}(\|z\|_p^5)$ and

$$\bar{G} = \tfrac{1}{2}\sum_{i,j\geqslant 1} \bar{G}_{ij}|q_i|^2|q_j|^2$$

with uniquely determined coefficients $\bar{G}_{ij} = (4-\delta_{ij})/4\pi$. Moreover, $\|\Gamma - \mathrm{id}\|_p = \mathrm{O}(\|z\|_p^3)$.

Thus, the Hamiltonian H is brought into the infinite-dimensional analogue of the classical fourth-order Birkhoff normal form:

$$H \circ \Gamma = \Lambda + \bar{G} + K = \langle \lambda, I\rangle + \tfrac{1}{2}\langle AI, I\rangle + K,$$

where $\lambda = (\lambda_1, \lambda_2, \ldots)$, $I = (I_1, I_2, \ldots)$ with

$$I_k = \tfrac{1}{2}|q_k|^2 = \tfrac{1}{2}(x_k^2 + y_k^2),$$

and $A_{ij} = 4\bar{G}_{ij} = (4-\delta_{ij})/\pi$.

Further transformations. It is elementary to observe that

$$\|u\|_0^2 = \int_0^\pi |u|^2 \,\mathrm{d}x$$

is an integral of the nonlinear Schrödinger equation as well as of its linear and nonlinear part separately. Accordingly,

$$\Phi = \tfrac{1}{2}\sum_{k\geqslant 1} |q_k|^2 = \sum_{k\geqslant 1} I_k$$

is an integral of $H = \Lambda + G$, which also commutes with Λ and G individually.

The normalizing transformation Γ of the preceding lemma preserves this property for the integral $\Phi\circ\Gamma$. Moreover, it also preserves Φ:

$$\Phi \circ \Gamma = \Phi.$$

This may seen by inspecting the proof of lemma 2.2 in [2]. Alternatively, one may interpret the normalizing transformation as an averaging transformation with respect to the Hamiltonian

$$\langle \lambda, I\rangle|_{m=0} = \langle \alpha, I\rangle = \sum_{k\geqslant 1}\alpha_k I_k, \qquad \alpha_k = \lambda_k|_{m=0} = k^2,$$

whose flow is periodic, and apply the reasoning of the proof of lemma 4.1 below.

Thus, Φ is also an integral of $H \circ \Gamma = \Lambda + \bar{G} + K$, which in addition commutes with Λ, $\bar{G}$ and K individually. Moreover, Bambusi made the observation that

$$\langle \lambda, I \rangle = \langle \alpha, I \rangle + m\Phi,$$
$$\pi \langle AI, I \rangle = -\langle I, I \rangle + 4\Phi^2.$$

Thus we have

$$H \circ \Gamma = \langle \alpha, I \rangle - \frac{1}{2\pi} \langle I, I \rangle + K + S$$

with

$$S = m\Phi + \frac{2}{\pi}\Phi^2,$$

where S commutes with each constituent of the Hamiltonian $H \circ \Gamma$. For this Hamiltonian we prove the following estimates.

Estimates 2.3. *Consider the Hamiltonian $H \circ \Gamma$ on $\mathcal{S}_1$. If $\|z^{\mathrm{o}}\|_1$ is sufficiently small, then the orbit with initial value z^{o} exists for all times and satisfies*

$$\|z(t)\|_1 \leqslant 2\|z^{\mathrm{o}}\|_1 \qquad \text{for} \quad -\infty < t < \infty.$$

Moreover, if $K \subset \mathbb{Z}$ is a finite index set of cardinality n, if δ is sufficiently small depending on K, and

$$\|z^{\mathrm{o}}\|_1 \leqslant \delta, \qquad \|z^{\mathrm{o}} - \pi_K z^{\mathrm{o}}\|_0 \leqslant \delta^{3/2}$$

where π_K denotes the orthogonal projection onto the space spanned by the coordinates with index in K, then

$$\|I(t) - I^{\mathrm{o}}\|_0 \leqslant c\gamma\delta^{2+2a} \qquad \text{for} \quad |t| \leqslant \exp\left(\frac{1}{d\gamma\delta^{2a}}\right)$$

with $a = \frac{1}{2n}$, $\gamma = \sqrt{n}$, and universal constants c, d.

Since $\|u\|_p = \|z\|_p$ in view of (3) and (5) and $\Gamma = \mathrm{id} + \mathrm{O}_3(z)$, this will prove theorem 1.1.

We reduce these statements to a more standard setting. Performing the scaling

$$x = \delta\tilde{x}, \qquad y = \delta\tilde{y}$$

and dividing the resulting Hamiltonian by δ^4, which amounts to an additional rescaling of time by δ^2, we obtain

$$\tilde{H} = \frac{1}{\varepsilon}\langle \alpha, \tilde{I} \rangle - \frac{1}{2\pi}\langle \tilde{I}, \tilde{I} \rangle + \varepsilon\tilde{K}(\tilde{z}, \varepsilon) + \tilde{S}(\tilde{z}, \varepsilon)$$

with

$$\varepsilon = \delta^2.$$

If δ is sufficiently small, $\tilde{K}$ and $\tilde{S}$ are real analytic on some fixed large ball around the origin in the complexification of $\mathcal{S}_p$, and $\|X_{\tilde{K}}\|_p$ is bounded there uniformly for all small δ. In these coordinates we prove the following estimates.

Estimates 2.4. *Consider the Hamiltonian $\tilde{H}$ on $\mathcal{S}_1$. If ε is sufficiently small, then every orbit with initial value z^{o} satisfying $\|z^{\mathrm{o}}\|_1 \leqslant 1$ exists for all time and satisfies*

$$\|z(t)\|_1 \leqslant 2\|z^{\mathrm{o}}\|_1 \qquad \text{for} \quad -\infty < t < \infty.$$

Moreover, if $K \subset \mathbb{Z}$ is a finite index set of cardinality n, and in addition

$$\|z^{\mathrm{o}} - \pi_K z^{\mathrm{o}}\|_0 \leqslant \varepsilon^{1/4},$$

then

$$\|I(t) - I^{\mathrm{o}}\|_0 \leqslant c\gamma\varepsilon^a \qquad \text{for} \quad |t| \leqslant \varepsilon \exp\left(\frac{1}{d\gamma\varepsilon^a}\right)$$

with $a = \frac{1}{2n}$, $\gamma = \sqrt{n}$, *and universal constants* c, d.

These estimates are proven in the following two sections for a slightly more general Hamiltonian. They clearly imply estimates 2.3 and thus prove theorem 1.1.

3. General stability results

Set-up. On some fixed phase space $\mathcal{S}_p$, we now consider a Hamiltonian

$$H = \frac{1}{\varepsilon}\langle\alpha, I\rangle + \frac{1}{2}\langle AI, I\rangle + \varepsilon F(z, \varepsilon) + S(z, \varepsilon) \tag{6}$$

with $\varepsilon > 0$ small, where $I = (I_1, I_2, \ldots)$ with $I_k = \frac{1}{2}(x_k^2 + y_k^2)$. We make the following assumptions.

Assumptions A.

(1) There are positive constants $a \leqslant b$ such that

$$a^2\|z\|_p^2 \leqslant \langle\alpha, I\rangle \leqslant b^2\|z\|_p^2.$$

(2) There are positive constants $m \leqslant M$ such that for some q with $0 \leqslant q \leqslant p$,

$$m\|I\|_q^2 \leqslant \langle AI, I\rangle \leqslant M\|I\|_q^2.$$

(3) The vector field X_F is real analytic on the complex ball

$$B_4 = \{z \in \mathcal{S}_{p,\mathbb{C}} : \|z\|_p < 4R\},$$

where $R = b/a \geqslant 1$, and

$$\|X_F\|_{p;4} = \sup_{B_4} \|X_F\|_p \leqslant E,$$

uniformly for all small ε. Here, *real analytic* means *analytic* in each argument and *real* for real arguments.

(4) S is an integral of H, not necessarily analytic, which also commutes with each of its constituents individually.

Note that the constant m in assumption (2) is unrelated to the parameter m of equation (1), which is no longer needed. In the following, an orbit is always understood to be a real orbit, and its initial value is denoted by z^{o}. Along an orbit, we write for example $H(t)$ for $H(z(t))$.

Lyapunov stability. The following result is standard and based on the fact that $\langle\alpha, I\rangle$ is definite. The convexity of A is not needed.

Lemma 3.1. *Suppose assumptions A hold, except for the convexity of the operator* A. *If* ε *is sufficiently small, then for every initial position* z^{o} *in* $\|z\|_p \leqslant 1$ *the orbit exists for all times and satisfies*

$$\|z(t)\|_p \leqslant 2R\|z^{\mathrm{o}}\|_p + \sqrt{R}$$

for all t. *In particular, if* X_F *vanishes at the origin, then* $\|z(t)\|_p \leqslant 2R\|z^{\mathrm{o}}\|_p$ *for all times* t.

Proof. Along an orbit, $H(t) - S(t) = H(0) - S(0)$ by energy conservation and symmetry. As long as the orbit stays inside B_4 we then have

$$\frac{a^2}{\varepsilon}\|z(t)\|_p^2 - M\|z(t)\|_p^4 - 4RE\varepsilon \leqslant \frac{b^2}{\varepsilon}\|z^{\rm o}\|_p^2 + M\|z^{\rm o}\|_p^4 + 4RE\varepsilon.$$

Hence, for ε sufficiently small,

$$\frac{a^2}{2}\|z(t)\|_p^2 \leqslant 2b^2\|z^{\rm o}\|_p^2 + \frac{a^2}{2}R$$

and thus, with $R = b/a$,

$$\|z(t)\|_p \leqslant 2R\|z^{\rm o}\|_p + \sqrt{R}.$$

This implies that with $\|z^{\rm o}\|_p \leqslant 1$ the orbit cannot escape from B_4. So this estimate holds for all times.

If $X_F(0) = 0$, then we may use $|F(z)| \leqslant C\|z\|_p^2$ to avoid the R-terms in the preceding estimates for ε sufficiently small, giving $\|z(t)\|_p \leqslant 2R\|z^{\rm o}\|_p$. □

Normal form. Consider the integrable Hamiltonian

$$H_0 = \langle\tilde{\alpha}, I\rangle + \tfrac{1}{2}\langle AI, I\rangle, \qquad \tilde{\alpha} = \varepsilon^{-1}\alpha.$$

In this section the dependence of H_0 on ε and the definiteness of $\langle\tilde{\alpha}, I\rangle$ is irrelevant.

Along an orbit we have $I(t) = I(0) = I(x^{\rm o}) =: I^{\rm o}$, and the motion is almost-periodic with frequencies

$$\omega^{\rm o} = \omega(I^{\rm o}) = \tilde{\alpha} + AI^{\rm o}.$$

Consider the case of T-periodic frequencies:

$$T\omega^{\rm o} \in 2\pi\mathbb{Z}^\infty, \qquad T\omega^{\rm o} \neq 0,$$

for some real T. The set of all real initial positions z with these frequencies $\omega^{\rm o}$ forms a torus

$$\mathcal{T}(I^{\rm o}) = \{z : I(z) = I^{\rm o}\} \subset \mathcal{S}_p,$$

whose dimension equals the number of positive components of $I^{\rm o}$. Later on, we will assume $\mathcal{T}(I^{\rm o})$ to be finite dimensional, but this assumption is not needed here. The case $I^{\rm o} = 0$ is also admitted.

Our aim is to obtain a normal form for H in a neighbourhood of $\mathcal{T}(I^{\rm o})$. To simplify notation we introduce

$$J = I - I^{\rm o} \Leftrightarrow I = I^{\rm o} + J,$$

keeping in mind that these are just real analytic expressions in z. No coordinate transformation is involved. Up to an irrelevant additive constant, we then can write

$$H_0 = \langle\omega^{\rm o}, I\rangle + \tfrac{1}{2}\langle AJ, J\rangle.$$

The total Hamiltonian H is written as

$$H = h(I) + g(I) + f(z) + S(z),$$

where $h = \langle\omega^{\rm o}, I\rangle$, $g = \frac{1}{2}\langle AJ, J\rangle$, and $f = \varepsilon F(z, \varepsilon)$, while the dependence of S on ε is dropped.

We study this Hamiltonian H on complex domains

$$D_{r,s} = \{z : \|I(z) - I^{\rm o}\|_q < r, \|z\|_p < sR\} \subset \mathcal{S}_{p,\mathbb{C}}.$$

For $\|z^{\rm o}\|_p \leqslant 1$ and $r > 0$, $s > 1$, these are nonempty neighbourhoods of $\mathcal{T}(I^{\rm o})$. We restrict ourselves to $s \leqslant 4$.

In the following we do not attempt to obtain estimates with particularly sharp constants. Indeed, we suppress all constants, and use the notations $u \lessdot v$ and $u \cdot\!< v$ to indicate that $u < cv$ and $cu < v$, respectively, with some constant $c \geqslant 1$, which depends only on a, b, m and M. Similarly for $u =\!\cdot\, v$ later on. For example, on $D_{s,r}$ we have

$$\|X_g\|_p \leqslant \|AJ\|_\infty \|z\|_p \lessdot \|AJ\|_q \lessdot \|J\|_q < r$$

by assumption A(2).

Let $\|\cdot\|_{p;r,s} = \sup_{D_{r,s}} \|\cdot\|_p$. The following lemma is proven in the next section.

Lemma 3.2 (normal form lemma). *Consider a real analytic Hamiltonian $H = h + g + f$ on $D_{r,s}$, where $h = \langle \omega, I \rangle$ and g are integrable and in involution, the flow of h is T-periodic, and*

$$\|X_g\|_{p;4r,4} < \delta, \qquad \|X_f\|_{p;4r,4} < \varepsilon.$$

If

$$\delta \lessdot r, \qquad mT\varepsilon \cdot\!< r, \qquad mTr \cdot\!< 1 \tag{7}$$

with some integer $m \geqslant 1$ and $0 < r \cdot\!< 1$, then there exists a real analytic, symplectic transformation $\Psi : D_{3r,3} \to D_{4r,4}$ with $\|\Psi - \mathrm{id}\|_{p;3r,3} \lessdot T\varepsilon$, such that

$$H \circ \Psi = h + \tilde{g} + \hat{f}$$

with $\{h, \tilde{g}\} = 0$ and

$$\|X_{\tilde{g}} - X_g\|_{p;3r,3} < 2\varepsilon, \qquad \|X_{\hat{f}}\|_{p;3r,3} < 2^{-m}\varepsilon. \tag{8}$$

Moreover, if S is an integral of H, not necessarily analytic, which commutes also with h, g and f, then S is invariant under Ψ, hence

$$(H + S) \circ \Psi = h + \tilde{g} + \hat{f} + S,$$

and S also commutes with $\tilde{g}$ and $\hat{f}$.

Remark. *An inspection of equation (14) in the proof of lemma 3.2 shows that the implicit constant in $\delta \lessdot r$ may be arbitrary large, influencing only the implicit constant in $mTr \cdot\!< 1$.*

Local stability. We use the normal form to prove a stability result near periodic solutions which is analogous to theorem 1 of [4].

Lemma 3.3 (local stability). *Consider the Hamiltonian (6) of the set-up and suppose assumptions A hold. Let z^{o} be an initial position in $\|z\|_p \leqslant 1$ with a T-periodic frequency vector $\omega^{\mathrm{o}} = \varepsilon^{-1}\alpha + AI^{\mathrm{o}}$. If*

$$\varepsilon \cdot\!< r^2, \qquad mTr \cdot\!< 1$$

with an integer $m \geqslant 1$ and $0 < r \cdot\!< 1$, then for every initial position with amplitudes $I(0)$ satisfying $\|I(0) - I^{\mathrm{o}}\|_q \cdot\!< r$ one has

$$\|I(t) - I^{\mathrm{o}}\|_q < r \qquad \text{for} \quad |t| < 2^m\varepsilon.$$

Proof. Rewriting the Hamiltonian in terms of $I = I^{\mathrm{o}} + J$ as in the previous section we obtain $H = h + g + f + S$ with $g = \frac{1}{2}\langle AJ, J\rangle$ and

$$\|X_g\|_{p;4r,4} \lessdot r, \qquad \|X_f\|_{p;4r,4} < \varepsilon.$$

Since $mT\varepsilon \cdot\!< mTr^2 \cdot\!< r$ also, the normal form lemma applies.

The normalizing transformation Ψ satisfies

$$\|\Psi - \mathrm{id}\|_{p;3r,3} <\cdot T\varepsilon \cdot< Tr^2 \cdot< r.$$

So the image of $D_{3r,3}$ under Ψ covers $D_{2r,2}$, and the shift of $\|I\|_q$ under Ψ is bounded by a fraction of r. It therefore suffices to prove the claim for the Hamiltonian in normal form on $D_{3r,3}$.

Let $H = h + \tilde{g} + \hat{f} + S$, where $h = \langle \omega^{\mathrm{o}}, I\rangle$ commutes with $\tilde{g} = g + \hat{g}$, the function S commutes with H, and the estimates (8) hold. Along an orbit, consider $h(t) = h(z(t))$. If we first assume that ω^{o} is T-periodic and bounded, then we may differentiate along orbits and obtain

$$\dot{h} = \frac{\mathrm{d}}{\mathrm{d}t}h = \{h, H\} = \{h, \hat{f}\} = dh(X_{\hat{f}}).$$

As long as the orbit stays inside the domain of the normal form we may use assumptions A(1) and A(2) to obtain $\varepsilon\omega_k^{\mathrm{o}} <\cdot k^{2p}$ and thus

$$|\dot{h}| <\cdot \frac{1}{\varepsilon}\|z\|_p \|X_{\hat{f}}\|_{p;3r,3} <\cdot 2^{-m}.$$

For $|t| < 2^m\varepsilon$ and as long as the orbit stays inside the domain of the normal form, we thus have

$$|h(t) - h(0)| <\cdot \varepsilon.$$

This estimate, however, does *not* depend on a bound on the frequencies ω_k^{o}, but only on the constants appearing in assumption A. By a straightforward approximation argument, it then extends to unbounded, T-periodic frequencies ω^{o} as well.

By energy conservation and symmetry, $H(t) = H(0)$ and $S(t) = S(0)$ for all times t, so

$$|g(t)| \leqslant |g(0)| + |h(t) - h(0)| + |\hat{g}(t) - \hat{g}(0)| + |\hat{f}(t) - \hat{f}(0)|.$$

The estimates for h as well as for $\hat{g} = \tilde{g} - g$ and $\hat{f}$ then give

$$|g(t)| <\cdot |g(0)| + \varepsilon \qquad \text{for} \quad |t| < 2^m\varepsilon,$$

and with the convexity of g by assumption A(2),

$$m\|J(t)\|_q^2 <\cdot M\|J(0)\|_q^2 + \varepsilon \qquad \text{for} \quad |t| < 2^m\varepsilon.$$

Together with the hypotheses $\|J(0)\|_q \cdot< r$ and $\varepsilon \cdot< r^2$ we thus arrive at

$$\|J(t)\|_q < r \qquad \text{for} \quad |t| < 2^m\varepsilon,$$

provided the orbit stays inside the domain of the normal form. But this is guaranteed just by this estimate and, in the case $q < p$, the Lyapunov stability of lemma 3.1. □

Global stability. To extend the local result of the preceding section to a global one, we need to make another assumption. To keep things simple we just cover the case of the nonlinear Schrödinger equation.

Assumption B. In the Hamiltonian (6) of the set-up, the frequencies α are T_0-periodic, and A is a diagonal operator:

$$T_0\alpha \in 2\pi\mathbb{Z}^\infty, \qquad A = \mathrm{diag}(A_1, A_2, \ldots).$$

Lemma 3.4 (global stability). *Consider the Hamiltonian (6) of the set-up, and suppose assumptions A and B hold. Fix an index set $K \subset \mathbb{Z}$ of finite cardinality n. If ε is sufficiently small, then for every initial position z with*

$$\|z\|_p \leqslant 1, \qquad \|z - \pi_K z\|_q \leqslant \varepsilon^{1/4}, \tag{9}$$

one has

$$\|I(t) - I(0)\|_q <\cdot\ \gamma \varepsilon^a \qquad \text{for} \quad |t| \leqslant \varepsilon \exp\left(\frac{1}{c_0 \gamma \varepsilon^a}\right),$$

with $a = \frac{1}{2n}$, $\gamma = \sqrt{n}$, and a constant c_0 depending only on a, b, m and M.

Proof. Fix $K \subset \mathbb{Z}$ and an initial position z satisfying (9). Let

$$\omega = \omega(I) = \varepsilon^{-1}\alpha + AI, \qquad I = I(z).$$

The simple idea, due to Lochak [3], is to approximate this initial position by a periodic orbit of the integrable reference Hamiltonian H_0 and to apply the local stability result. The basic ingredient of this approach is Dirichlet's theorem on simultaneous approximations, stating that for every $\omega \in \mathbb{R}^n$ and every integer $Q \geqslant 1$,

$$\min_{\substack{q \in \mathbb{Z} \\ 1 \leqslant q \leqslant Q}} \min_{p \in \mathbb{Z}^n} |q\omega - p|_\infty \leqslant \frac{1}{Q^{1/n}}. \tag{10}$$

In our case, it suffices to approximate only *finitely* many frequencies, since we suppose α to be T_0-periodic and $\|z - \pi_K z\|_q$ to be small. Let $\bar{z} = (\bar{x}, \bar{y})$, where $\bar{x}$ denotes the n-vector obtained from x by deleting all components with an index not in K. Similarly, let

$$\bar{\upsilon} = \bar{\omega} - \varepsilon^{-1}\bar{\alpha} = \bar{A}\bar{I}, \qquad \bar{I} = \bar{I}(\bar{z}).$$

Given $\varepsilon > 0$, choose

$$Q = \varepsilon^{-a(n-1)}.$$

Scaling down $\bar{\upsilon}$ slightly until one of its components becomes an integer, approximating the remaining $n-1$ components by a rational frequency vector according to (10) and then scaling back, we obtain a T-periodic frequency vector $\bar{\upsilon}^{\mathrm{o}}$ such that

$$|\bar{\upsilon} - \bar{\upsilon}^{\mathrm{o}}|_\infty \leqslant \frac{2\pi}{T Q^{1/(n-1)}} = 2\pi \frac{\varepsilon^a}{T}, \qquad 1 \leqslant T \leqslant 2\pi Q. \tag{11}$$

Scaling $\bar{\upsilon}^{\mathrm{o}}$ by another factor, which is ε-close to 1, we can also arrange that T is an integer multiple of εT_0, while the estimate (11) deteriorates by a factor, which only depends on $|\bar{A}|$ and hence only on M.

It follows that $\bar{\omega}^{\mathrm{o}} = \varepsilon^{-1}\bar{\alpha} + \bar{\upsilon}^{\mathrm{o}}$ is also T-periodic and satisfies

$$|\bar{\omega} - \bar{\omega}^{\mathrm{o}}|_\infty <\cdot\ \frac{\varepsilon^a}{T}.$$

This frequency vector $\bar{\omega}^{\mathrm{o}}$ in turn corresponds to an amplitude vector $\bar{I}^{\mathrm{o}}$ with

$$\|\bar{I} - \bar{I}^{\mathrm{o}}\|_q <\cdot\ \gamma \frac{\varepsilon^a}{T},$$

due to assumptions A(2) and B.

Now let I^{o} be the vector $\bar{I}^{\mathrm{o}}$ 'filled up with zeros' so that $\pi_K I^{\mathrm{o}} = \bar{I}^{\mathrm{o}}$. By construction, $\omega^{\mathrm{o}} = \varepsilon^{-1}\alpha + AI^{\mathrm{o}}$ is also T-periodic, and

$$\begin{aligned} \|I - I^{\mathrm{o}}\|_q &\leqslant \|I - \pi_K I\|_q + \|\pi_K I - I^{\mathrm{o}}\|_q \\ &\leqslant \|z - \pi_K z\|_q^2 + \|\bar{I} - \bar{I}^{\mathrm{o}}\|_q <\cdot\ \sqrt{\varepsilon} + \gamma \frac{\varepsilon^a}{T} <\cdot\ \gamma \frac{\varepsilon^a}{T}, \end{aligned}$$

as we have $\sqrt{\varepsilon} = \varepsilon^a/Q \lessdot \varepsilon^a/T$. If we set

$$r =\cdot \gamma \frac{\varepsilon^a}{T}$$

with a *sufficiently large* implicit constant, then we have $\|I - I^{\circ}\|_q \cdot\!< r$ and $\varepsilon \cdot\!< r^2$ as required.

Finally, choosing m by $mTr \cdot\!= 1$, or $m \cdot\!= \gamma^{-1}\varepsilon^{-a}$, we may apply the local stability result around I° and obtain

$$\|I(t) - I(0)\|_q < r \qquad \text{for} \quad |t| < 2^m \varepsilon.$$

With $r \lessdot \gamma\varepsilon^a$ the final result follows. □

4. Proof of the normal form lemma

We consider a Hamiltonian $H = h + g + f$, where $h = \langle \omega, I \rangle$ and g are integrable and in involution, the flow of h is T-periodic, and X_f is small. The aim is to transform H into some normal form with respect to h up to an exponentially small remainder.

As usual, this lemma is proven by iterating an averaging transformation a finite number of times—namely, m times. The general step of this procedure is described in the next lemma.

Lemma 4.1 (iterative lemma). *Suppose the Hamiltonian $H = h + g + f$ is real analytic on $D_{r,s}$, where the flow of $h = \langle \omega, I \rangle$ is T-periodic and*

$$\|X_g - X_{g_0}\|_{p;r,s} < \gamma, \qquad \|X_{g_0}\|_{p;r,s} < \delta, \qquad \|X_f\|_{p;r,s} < \varepsilon$$

with an integrable Hamiltonian g_0. If

$$T\varepsilon \cdot\!< \rho \cdot\!< \sigma$$

with $0 < \rho < r$ and $0 < \sigma < s$, then there exists a real analytic, symplectic transformation $\Phi : D_{r-\rho,s-\sigma} \to D_{r,s}$, such that $H \circ \Phi = h + g_+ + f_+$ with

$$\|X_{g_+} - X_g\|_{p;r,s} < \varepsilon, \qquad \|X_{f_+}\|_{p;r-\rho,s-\sigma} \lessdot T\left(\frac{\delta}{\sigma} + \frac{\gamma+\varepsilon}{\rho}\right)\varepsilon,$$

and $\{h, g_+ - g\} = 0$. Also, $\|\Phi - \mathrm{id}\|_{p;r-\rho,s-\sigma} < T\varepsilon$.

Moreover, if S is an integral of H, which also commutes with h, g and f, then S is invariant under Φ and also commutes with g_+ and f_+.

Proof. We first assume that ω is a bounded frequency vector. Then the flow X_h^t is smooth, and we may differentiate along its flow lines as usual. If the frequencies in ω are unbounded, then X_h^t is only continuous, as X_h is not a *bona fide* tangent vector to $\mathcal{S}_p$. This case is handled below by an approximation argument.

So assume ω is bounded. As usual, Φ is written as the time-1-map X_ϕ^1 of a Hamiltonian vector field X_ϕ, where ϕ solves the homological equation

$$\{\phi, h\} = f - \bar{f}.$$

Standard calculations then lead to $H \circ \Phi = h + g_+ + f_+$ with $g_+ = g + \bar{f}$ and

$$f_+ = \int_0^1 \{g + f_t, \phi\} \circ X_\phi^t \,\mathrm{d}t, \qquad f_t = tf + (1-t)\bar{f}.$$

The operator $\{\cdot, h\}$ is semi-simple, so $\bar{f}$ is determined as the projection of f onto the null space of this operator. Whence $\{g_+ - g, h\} = \{\bar{f}, h\} = 0$. Then ϕ is the preimage of $f - \bar{f}$ in its range. Since the flow of h is T-periodic, these solutions are given by

$$\bar{f} = \frac{1}{T}\int_0^T f \circ X_h^t \,\mathrm{d}t, \qquad \phi = \frac{1}{T}\int_0^T t(f - \bar{f}) \circ X_h^t \,\mathrm{d}t. \tag{12}$$

Suppose S commutes with h, g and f. Then in particular $S \circ X_h^t = S$ for all t and thus

$$\{f \circ X_h^t, S\} = \{f \circ X_h^t, S \circ X_h^t\} = \{f, S\} \circ X_h^t = 0.$$

Consequently,

$$\{\bar f, S\} = \frac{1}{T}\int_0^T \{f \circ X_h^t, S\}\,\mathrm{d}t = 0,$$

$$\{\phi, S\} = \frac{1}{T}\int_0^T t\{(f - \bar f) \circ X_h^t, S\}\,\mathrm{d}t = 0.$$

We conclude that $\{g_+, S\} = 0$ and $S \circ \Phi = S$. Similarly, $\{f_+, S\} = 0$.

The vector fields of $\bar f$ and ϕ are

$$X_{\bar f} = \frac{1}{T}\int_0^T (X_h^t)^* X_f\,\mathrm{d}t, \qquad X_\phi = \frac{1}{T}\int_0^T t(X_h^t)^*(X_f - X_{\bar f})\,\mathrm{d}t.$$

Since the flow X_h^t consists of orthogonal transformations, $\|(X_h^t)^* X_f\|_p = \|X_f\|_p$ and thus

$$\|X_{\bar f}\|_{p;r,s} < \varepsilon, \qquad \|X_\phi\|_{p;r,s} < 2T\varepsilon,$$

which also gives the first of the postulated estimates.

To control the mapping of domains we observe that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|z\|_p^2\right| \leqslant 2|\langle z, \dot z_p\rangle| \leqslant 2\|z\|_p\|\dot z\|_p,$$

and similarly $|\frac{\mathrm{d}}{\mathrm{d}t}\|I\|_q^2| \leqslant 2\|I\|_q\|\dot I\|_q$ with

$$\|\dot I\|_q^2 \leqslant \sum_{k\neq 0} |k|^{2q}|z_k|^2|\dot z_k|^2 \leqslant \|z\|_0^2\|\dot z\|_p^2 \leqslant 16\|X_\phi\|_p^2$$

on $D_{r,s}$. Together with the hypotheses $T\varepsilon \cdot\!< \rho \cdot\!< \sigma$ we thus have

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|z\|_p\right| \leqslant \|X_\phi\|_p \cdot\!< \sigma, \qquad \left|\frac{\mathrm{d}}{\mathrm{d}t}\|I\|_q\right| \leqslant 4\|X_\phi\|_p \cdot\!< \rho$$

for almost all t on $D_{r,s}$. Consequently, $X_\phi^t : D_{r-\rho,s-\sigma} \to D_{r,s}$ for $-1 \leqslant t \leqslant 1$, and $\|X_\phi^t - \mathrm{id}\|_{p;r-\rho,s-\sigma} <\!\cdot\, T\varepsilon$.

It remains to estimate f_+, or rather

$$X_{f_+} = \int_0^1 (X_\phi^t)^*[X_g + X_{f_t}, X_\phi]\,\mathrm{d}t.$$

It suffices to estimate the Lie bracket itself, hence $[X_{f_t} X_\phi]$, $[X_g - X_{g_0}, X_\phi]$ and $[X_{g_0}, X_\phi]$. As to the first bracket, we write

$$[X_{f_t}, X_\phi] = \frac{\mathrm{d}}{\mathrm{d}\tau}(X_\phi^\tau)^* X_{f_s}\bigg|_{\tau=0},$$

and observe that the flow of X_ϕ^τ starting in $D_{r-\rho,s-\sigma}$ exists within $D_{r,s}$ for *complex* times τ with

$$|\tau| \cdot\!< \rho\|X_\phi\|_{p;r,s}^{-1}.$$

Hence, with Cauchy's estimate we obtain

$$\|[X_{f_t}, X_\phi]\|_{p;r-\rho,s-\sigma} <\!\cdot\, \frac{1}{\rho}\|X_{f_t}\|_{p;r,s}\|X_\phi\|_{p;r,s} <\!\cdot\, \frac{T\varepsilon}{\rho}\cdot\varepsilon.$$

The estimate of $\|[X_g - X_{g_0}, X_\phi]\|_{p;r-\rho,s-\sigma}$ is similar, using $\|X_g - X_{g_0}\|_{p;r,s} < \gamma$.

The bracket $[X_{g_0}, X_\phi]$, however, is estimated slightly differently. We write

$$[X_{g_0}, X_\phi] = -\frac{\mathrm{d}}{\mathrm{d}\tau}(X^\tau_{g_0})^* X_\phi \Big|_{\tau=0}.$$

Since g_0 is integrable, the flow $X^\tau_{g_0}$ leaves I invariant. So the existence of this flow is only restricted by the condition that $\|X^\tau_{g_0}(z)\| < s$ for $\|z\| < s - \sigma$. Whence this flow exists for complex times τ with

$$|\tau| \cdot\!< \sigma \|X_{g_0}\|^{-1}_{p;r,s},$$

and by Cauchy's estimate we obtain

$$\|[X_{g_0}, X_\phi]\|_{p;r-\rho,s-\sigma} <\!\cdot \frac{1}{\sigma}\|X_{g_0}\|_{p;r,s}\|X_\phi\|_{p;r,s} <\!\cdot \frac{T\delta}{\sigma} \cdot \varepsilon.$$

This proves the lemma for bounded frequencies.

Now assume that $\omega = (\omega_1, \omega_2, \ldots)$ is unbounded. Let

$$\omega^n = (\omega_1, \ldots, \omega_n, \omega_n, \ldots), \qquad n \geqslant 1.$$

Each ω^n is bounded, the flow of $h_n = \langle \omega^n, I\rangle$ is T-periodic, and as before we obtain

$$\bar{f}_n = \frac{1}{T}\int_0^T f \circ X^t_{h_n}\,\mathrm{d}t, \qquad \phi_n = \frac{1}{T}\int_0^T t(f - \bar{f}) \circ X^t_{h_n}\,\mathrm{d}t, \tag{13}$$

and similarly $f_{n,+}$ and so on. The pertaining estimates hold *uniformly* in n, as they do not involve any bounds on the frequencies, and as $DX^t_{h_n}$ consists of orthogonal transformations for all n.

In (13) as well as the other identities we may then pass to the limit $n \to \infty$ by the Lebesgue-dominated convergence theorem, thus recovering (12) and the other identities. The same applies to the vector fields $X_{\bar{f}_n}$ and X_{ϕ_n}, which shows that $\bar{f}$ and ϕ are differentiable in the complex and hence also analytic in the unbounded case. The estimates continue to hold, as they hold uniformly in n. This also proves the lemma in the unbounded case. □

Proof of lemma 3.2. We apply the iterative lemma m times, so that

$$H_i \circ \Phi_i = H_{i+1} = h + g_{i+1} + f_{i+1}, \qquad 0 \leqslant i \leqslant m-1,$$

starting with $H_0 = h + g_0 + f_0 = h + g + f = H$, the given Hamiltonian, and choose uniformly

$$\rho = \frac{r}{m}, \qquad \sigma = \frac{1}{m}$$

for each step. We claim that at each step we obtain $\{h, g_i\} = 0$ and

$$\|X_{g_i} - X_{g_0}\|_{p;r-i\rho,s-i\sigma} < 2(1 - 2^{-i})\varepsilon, \qquad \|X_{f_i}\|_{p;r-i\rho,s-i\sigma} < 2^{-i}\varepsilon,$$

as well as $\{g_i, S\} = \{f_i, S\} = 0$ and $S \circ \Phi_i = S$. This clearly holds for $i = 0$, so we can proceed by induction.

To apply the iterative lemma for some $i \geqslant 0$ we observe that

$$T\varepsilon = mT\varepsilon \cdot \frac{1}{m} \cdot\!< \frac{r}{m} = \rho \cdot\!< \sigma$$

by the hypothesis (7) and $r \cdot\!< 1$. So the lemma applies. The estimate for $X_{g_{i+1}}$ is immediate. To estimate $X_{f_{i+1}}$ we note that with $\delta <\!\cdot r$, $\gamma < 2\varepsilon$ we have

$$T\frac{\delta}{\sigma} <\!\cdot mTr, \qquad T\frac{\gamma + \varepsilon}{\rho} <\!\cdot \frac{mT\varepsilon}{r}, \tag{14}$$

so with hypothesis (7) we can arrange that

$$T\left(\frac{\delta}{\sigma}+\frac{\gamma+\varepsilon}{\rho}\right)\cdot\!<\frac{1}{2}.$$

This gives the estimate for $X_{f_{i+1}}$. Moreover, g_{i+1} and f_{i+1} commute with S.

Finally we note that $S\circ\Phi=S$ and $\|\Phi_i-\mathrm{id}\|_{p;r-(i+1)\rho,s-(i+1)\sigma}<T\varepsilon_i=2^{-i}T\varepsilon$, and the estimate for $\Psi=\Phi_0\circ\cdots\circ\Phi_{m-1}$ follows by routine arguments, as well as $S\circ\Psi=S$. This completes the proof of lemma 3.2. □

References

[1] Bambusi D Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations *Math. Z.* to appear

[2] Kuksin S and Pöschel J 1996 Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation *Ann. Math.* **143** 149–79

[3] Lochak P 1992 Canonical perturbation theory via simultaneous approximation *Russ. Math. Surveys* **47** 57–133

[4] Lochak P 1993 Hamiltonian perturbation theory: periodic orbits, resonances and intermittency *Nonlinearity* **6** 885–904

[5] Niederman L 1998 Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system *Nonlinearity* **11** 1465–79

[6] Pöschel J 1999 On Nekhoroshev's estimate at an elliptic equilibrium *Int. Math. Res. Not.* **4** 203–15