Ergod. Th. & Dynam. Sys. (2004), **24**, 1331–1357
DOI: 10.1017/S0143385703000750
© *2004 Cambridge University Press*
Printed in the United Kingdom

Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations

JEAN BOURGAIN

Institute for Advanced Study, School of Mathematics,
Olden Lane, Princeton, NJ 08540, USA
(e-mail: bourgain@math.ias.edu)

(*Received* 4 *March* 2003 *and accepted in revised form* 19 *May* 2003)

Abstract. This is a partial survey of various developments of the classical KAM theory in dynamical systems to infinite-dimensional lattice models and Hamiltonian PDEs. Particular emphasis is given on Nekhoroshev stability and growth of higher Sobolev norms. Some new results are also presented.

1. *Introduction*

This exposé is a summary of some of the author's encounters with the subject mentioned in the title and certainly not a full review of this area of research. The reference list is also fairly restricted. However, we will occasionally point to the various survey works that may provide a more complete picture, both historically and mathematically.

Our basic motivation is to understand qualitatively certain aspects of the long-time behaviour of solutions of Hamiltonian partial differential equations (PDEs) on bounded domains (periodic or Dirichlet boundary conditions etc.). This may be achieved using various approaches but in what follows we will mainly focus on implementing methods from classical dynamical systems.

More precisely, we will discuss the Nekhoroshev stability under an ε-perturbation, possible growth of higher Sobolev norms of classical solutions for $t \to \infty$ and finally the existence and abundance of quasi-periodic and almost periodic motion in phase space (KAM theory for PDEs).

1.1. *Long-time stability of solutions of ε-perturbations of linear or integrable Hamiltonian PDEs.*

Typical examples in this respect are nonlinear Schrödinger (NLS)

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

(or, for waves NLW) equations

$$iu_t - u_{xx} + V(x)u + \varepsilon \frac{\partial H}{\partial \overline{u}} = 0 \tag{1.1}$$

$$y_{tt} - y_{xx} + \rho y + \varepsilon F'(y) = 0 \tag{1.2}$$

$$iu_t + u_{xx} \pm u|u|^2 + \varepsilon \frac{\partial H}{\partial \overline{u}} = 0 \tag{1.3}$$

(Dirichlet or periodic boundary conditions).

In (1.1), (1.3), $H = H(u, \overline{u})$ is a real analytic function of $u, \overline{u}$ and, in (1.2), $F(y)$ is a real analytic function of y.

Roughly speaking, the problem is for how long solutions of perturbed and unperturbed equations will stay 'orbitally close' to each other. By 'non-trivial' stability times T_ε, we mean $T_\varepsilon \gg 1/\varepsilon$.

This is the PDE setting of the well-known Nekhoroshev problem [**N**].

1.2. *Diffusion in higher Sobolev spaces.* Consider, for instance, the two-dimensional (2D) defocusing cubic NLS on $\mathbb{T}^2$:

$$iu_t + \Delta u - u|u|^2 = 0 \tag{1.4}$$

with Hamiltonian

$$H(\phi) = \frac{1}{2}\int |\nabla \phi|^2 + \frac{1}{4}\int |\phi|^4. \tag{1.5}$$

Contrary to its one-dimensional counterpart (1D) (cf. Zakharov–Shabat [**ZS**]), equation (1.4) is non-integrable and no conserved quantities related to Sobolev norms beyond H^1 are at our disposal. For $u(0) \in H^s(\mathbb{T}^2)$, $s \geq 1$, there is a unique global solution satisfying $u(t) \in H^s$ for all time (see [**B1, B3**]). Moreover, it can be shown that

$$\|u(t)\|_{H^s} < Ct^{C(s-1)} \quad \text{for } t \to \infty \tag{1.6}$$

(with C some constant). Because of the absence of *a priori* bounds on higher Sobolev norms, one may ask the following question.

PROBLEM 1. *Fixing $s > 1$, could it happen that, with $u(0) \in H^s$,*

$$\sup_t \|u(t)\|_{H^s} = \infty? \tag{1.7}$$

If so, how fast will $\|u(t)\|_{H^s}$ grow when $t \to \infty$.

For equation (1.4), no examples of smooth solutions satisfying (1.7) are known. The best known upper bounds are power-like in time. Note that, in the linear case,

$$iu_t + \Delta u + V(x,t)u = 0 \tag{1.8}$$

on $\mathbb{T}^d$, with V real and smooth in x, t with uniform bounds in time, V periodic in x but with no specified behaviour in time, one has, for all $s > 0$,

$$\|u(t)\|_{H^s} \ll |t|^\varepsilon \|u(0)\|_{H^s} \quad \text{for all } \varepsilon > 0. \tag{1.9}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

(There are also examples illustrating a slow growth in time, even assuming V to be time periodic.) See [**B2, B3**]. In view of this linear theory, one could hope for improvements over (1.6) in the nonlinear case. At this point, however, we have been unable to implement these methods in the nonlinear setting. We only observe that (1.9) does apply to NLS equations on $\mathbb{T}^d$ of the form

$$iu_t + \Delta u + |u * W|^2 u = 0$$

or

$$iu_t + \Delta u + P(|u * W|^2)u = 0$$

with P a real polynomial and supp $\widehat{W} \subset \mathbb{Z}^d$ finite or $\widehat{W}$ rapidly decaying. Indeed, one may simply put, in (1.8),

$$V(x,t) = P(|u(t) * W|^2)(x)$$

which satisfies uniform smoothness bounds in x, t as a consequence of the equation and the *a priori* bound on $\|u(t)\|_2$.

We will sketch in the appendix a dynamical systems approach to (1.6) leading to slightly better exponents.

Let us also observe here that in the NLW case, there are examples that exhibit growth in the higher Sobolev norms as a power of time (see [**B4**]).

1.3. *Invariant tori for Hamiltonian PDEs.* Over the last 15 years there have been many contributions to the so-called 'KAM theory' for PDEs. Some of these developments led, in fact, to new results in the classical context of finite dimensional phase space. Some survey works on the subject are [**K, Cr, PM**].

We distinguish between

(i) 1D invariant tori $\leftrightarrow$ time periodic solutions,

(ii) finite-dimensional invariant tori $\leftrightarrow$ time quasi-periodic solutions and

(iii) infinite-dimensional invariant tori $\leftrightarrow$ time almost-periodic solutions.

In the context of (iii), one may, in particular, consider almost-periodic solutions on a full set of frequencies (corresponding to 'half-dimensional' invariant tori in phase space). Most of the available results relate to items (i) and (ii). We will only discuss some progress resulting from certain new developments on quasi-periodic localization for lattice Schrödinger operators (such as the extensive use of subharmonic function techniques and the theory of semi-algebraic sets). See [**B5**]. It did permit us to treat KAM problems for PDEs in space dimension $D > 1$, which is significantly harder due to large sets of resonances.

The simplest model is that of a perturbed linear equation with parameters (the perturbation is nonlinear). Thus, in the NLS setting, consider an equation

$$iq_t = \mathcal{L}q + \varepsilon \frac{\partial H_1}{\partial \bar{q}} \tag{1.10}$$

where $H_1 = H_1(q, \bar{q})$ is a real polynomial expression in $q, \bar{q}$ with real coefficients. Here $q = (q_n)_{n \in \mathbb{Z}^d}$. Select a finite set of modes $n_1, \ldots, n_b \in \mathbb{Z}^d$. The operator $\mathcal{L}$ is

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

given by a multiplier

$$\begin{aligned} \mu_{n_j} &= \lambda_j \quad (1 \le j \le b) \\ \mu_n &= |n|^2 \quad \text{for } n \in \mathbb{Z}^d \backslash \{n_1, \dots, n_b\} \end{aligned} \tag{1.11}$$

and where $\lambda = (\lambda_1, \dots, \lambda_b)$ is a parameter (ranging in an open subset of $\mathbb{R}^b$). The choice $\mu_n = |n|^2$ corresponds to the $-\Delta$ operator (which is assumed to be independent of λ for simplicity). Nonlinear problems without parameters may often be reduced to this format, extracting parameters from the nonlinearity and amplitude–frequency modulation.

For $\varepsilon = 0$, (1.10) has the obvious quasi-periodic solution

$$\begin{aligned} q_{n_j}(t) &= a_j e^{i\lambda_j t} \quad (1 \le j \le b) \\ q_n(t) &= 0 \quad \text{for } n \notin \{n_1, \dots, n_b\}. \end{aligned}$$

The persistency result (see [**B5**, ch. XIX]) claims then, roughly, that there is a Cantor subset Ω_ε of the λ-parameter set Ω, $\text{mes}(\Omega \backslash \Omega_\varepsilon) \xrightarrow{\varepsilon \to 0} 0$ and a smooth map $\lambda \mapsto \lambda'$ on Ω, such that, for $\lambda \in \Omega_\varepsilon$, there is a quasi-periodic solution of (1.10),

$$q_n(t) = \sum_{k \in \mathbb{Z}^b} \widehat{q_n}(k) e^{ik\lambda' t},$$

satisfying

$$\begin{aligned} \widehat{q_{n_j}}(e_j) &= a_j \quad (1 \le j \le b) \quad (e_j = j\text{-unit vectors in } \mathbb{Z}^b) \\ |\widehat{q_n}(k)| &\lesssim e^{-c(|n|+|k|)} \\ \sum_{(n,k) \notin \mathcal{R}} |\widehat{q_n}(k)| &< \sqrt{\varepsilon} \quad \text{where } \mathcal{R} = \{(n_j, e_j) \mid j = 1, \dots, b\}. \end{aligned}$$

The result is proven using a Lyapounov–Schmidt decomposition and the Newton iteration method. Thus, the core of the analysis consists in controlling the inverses of (suitable restrictions) of linearized operators of the form

$$(k\lambda' - \mu_n)\delta_{(n,k),(n',k')} + \varepsilon S \tag{1.12}$$

with index set $\mathbb{Z}^{d+b}$ and where S is a Toeplitz-type operator defined from the previous approximate solution obtained along the iteration. Thus, the operators (8.8) are closely related to the quasi-periodic operators arising in the theory of lattice Schrödinger operators.

However, the presence of the normal frequencies $\{\mu_n = |n|^2\}$ leads to rather non-trivial additional difficulties, especially in space dimension $d > 1$ (the case $d = 2$ was treated in [**B6**] but problems for $d \ge 3$ were left open because of the sets of resonant sites were too large).

The method described in [**B5**] permits us to treat this problem in arbitrary dimension d.

The corresponding model problem for NLW is the equation

$$iq_t = Bq + \varepsilon B^{-1} \frac{\partial H_1}{\partial \bar{q}} \tag{1.13}$$

where B is now a Fourier multiplier $(\mu_n)_{n \in \mathbb{Z}^d}$ defined as

$$\begin{aligned} \mu_{n_j} &= \lambda_j > 0 \quad (1 \le j \le b) \\ \mu_n &= |n| \quad \text{if } n \notin \{n_1, \dots, n_b\} \end{aligned}$$

(corresponding to $\sqrt{-\Delta}$).

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

In this setting, we have the analogue of the NLS persistency theorem stated earlier (see [**B5**, ch. XX]) (only cases $b = 1$ (time-periodic solutions) and $d = 1$ have been treated in previous works).

With respect to item (iii)—'infinite dimensional tori'—we mention the papers [**B7**] and [**P**]. They basically produce a similar result (by different methods). One starts with an unperturbed almost-periodic solution on a full set of frequencies

$$u(x,t) = \sum_{j=1}^{\infty} a_j e^{i\lambda_j t}\varphi_j(x) \tag{1.14}$$

(for a 1D Schrödinger or wave equation with Dirichlet boundary conditions) where the amplitudes $\{a_j\}$ decay extremely rapidly. It is possible in this setting to prove persistency results—as in the quasi-periodic case.

Neither of the arguments [**B7, P**] provides an explicit decay condition on the $\{a_j\}$.

Obtaining such a result is certainly a most interesting problem and may require new ideas. Let us mention in this respect the paper [**FSW**] on disordered systems with infinitely many degrees of freedom where maximal invariant tori are produced with an explicit (but super-exponential) decay rate for the action variables. In this model, the interactions in the perturbation are finite range however, which is significantly easier to deal with.

To summarize, there certainly seems to be little understanding of 'what part' of the phase space for a non-integrable Hamiltonian PDE is filled with KAM tori. The belief is that it is a 'small' part of the phase space and that chaotic diffusion is present in most of it.

2. *Nekhoroshev stability in high dimension*

A first observation is that, when dealing with a large or infinite-dimensional phase space and a Hamiltonian perturbation of an integrable system

$$H(I,\theta) = H_0(I) + \varepsilon H_1(I,\theta) \tag{2.1}$$

(where $I = (I_j), \theta = (\theta_j)$ refer to usual action-angle variables), one needs to specify the following topologies.

(i) To express smallness of the perturbation.

(ii) To express the stability.

The following is also of importance.

(iii) The more precise nature of the perturbation H_1 in (2.1).

In particular, one should expect that models with finite-range interactions should be easier to deal with.

We will distinguish between the following two mechanisms:

(A) perturbation from linear non-resonant systems; and

(B) perturbation from a convex integrable Hamiltonian.

At this point, most of the PDE applications with reasonably large classes of data belong to category (A). Observe (as we will illustrate later) that sometimes a small-data nonlinear problem may be expressed in format (A) by normal form manipulations and modulation of the amplitude and frequencies.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Some further comments. As mentioned earlier, we will view stability times $T_\varepsilon \gg 1/\varepsilon$ as non-trivial. Our notion of a 'very long' stability time will refer to $\log T_\varepsilon \gg \log 1/\varepsilon$ (rather than $T_\varepsilon > \exp 1/\varepsilon^a$ for some $a > 0$, common in finite-dimensional phase space).

Another important point is that to obtain non-trivial results, there are constraints between the 'size of the perturbation' and the 'size of the tori' for which we try to prove long-time near-stability. By the size of the invariant tori, we mean its metrical entropy (in some unspecified topology). If the tori are too large, one may possibly only have stability times $T_\varepsilon \sim 1/\varepsilon$ (see [**B8**]). But these constraints are poorly understood at this point and there is a big gap between positive and negative results. In particular, the available results based on stability mechanism (B) are very restrictive from this respect and require the dimension of the tori to be sublogarithmic in $1/\varepsilon$.

2.1. *Perturbation of linear diophantine systems.* A first basic and well-known idea consists in removing as much as possible of the perturbation $H_1(I, \theta)$ by canonical transformations of the phase space. The final result of this procedure is the conversion of the Hamiltonian to a sum of three components: an integrable part, a nearly resonant part and a 'remainder' which is very small (cf. [**W**]).

Consider a Hamiltonian function $F(I, \theta)$ and let Γ be the time-1 shift for the corresponding Hamiltonian flow

$$\dot{I}_j = -\varepsilon \frac{\partial F}{\partial \theta_j}$$
$$\dot{\theta}_j = \varepsilon \frac{\partial F}{\partial I_j}.$$

Under the transformation Γ, the equations

$$\dot{I}_j = -\frac{\partial H}{\partial \theta_j}$$
$$\dot{\theta}_j = \frac{\partial H}{\partial I_j}$$

are transformed into the corresponding equations with Hamiltonian function $H' = H\Gamma^{-1}$ where, essentially,

$$H' = H + \varepsilon\{H, F\} + 0(\varepsilon^2)$$

and

$$\{H, F\} = \sum_j \left(\frac{\partial H}{\partial I_j} \frac{\partial F}{\partial \theta_j} - \frac{\partial H}{\partial \theta_j} \frac{\partial F}{\partial I_j} \right)$$

is the Poisson bracket.

Since $H = H_0(I) + \varepsilon H_1$, one gets roughly

$$\begin{aligned} H' &= H_0 + \varepsilon H_1 + \varepsilon\{H_0, F\} + 0(\varepsilon^2) \\ &= H_0(I) + \varepsilon H_{1,0}(I) + \varepsilon H_1'(I, \theta) + \varepsilon\{H_0, F\} + 0(\varepsilon^2) \end{aligned} \tag{2.2}$$

where

$$H_1(I, \theta) = \sum_{n \in \mathbb{Z}^d} H_{1,n}(I) e^{i\langle n, \theta \rangle} \quad \text{and} \quad H_1' = H_1 - H_{1,0}(I). \tag{2.3}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Thus, if

$$F = \sum_{n \neq 0} F_n(I) e^{i\langle n, \theta\rangle},$$

we have

$$\{H_0, F\} = \sum_j \frac{\partial H_0}{\partial I_j} \frac{\partial F}{\partial \theta_j} = i \sum_{n \in \mathbb{Z}^d} \langle \nabla_I H_0, n\rangle F_n(I) e^{i\langle n, \theta\rangle}$$

hence,

$$\{H_0, F\} = -H_1'$$

if we choose

$$F_n(I) = i \frac{H_{1,n}(I)}{\langle \nabla_I H_0, n\rangle} \quad (n \neq 0). \tag{2.4}$$

Assume, for simplicity, that $H_1(I, \theta) = \sum' H_{1,n}(I) e^{i\langle n, \theta\rangle}$ with $|n|$ bounded (otherwise cutoffs need to be introduced).

We consider the case where H_0 is given by

$$H_0(I) = \lambda I \tag{2.5}$$

and $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$ satisfies certain diophantine properties.

In this situation, (2.4) becomes

$$F_n(I) = i \frac{H_{1,n}(I)}{\langle \lambda, n\rangle} \quad (n \neq 0) \tag{2.6}$$

with some control on the divisor $\langle \lambda, n\rangle$ since λ is diophantine. After the transformation, one gets for the new Hamiltonian (2.2)

$$H_0(I) + \varepsilon H_{1,0}(I) + \frac{\varepsilon^2}{\eta} H_2(I, \theta) \tag{2.7}$$

with a new perturbation $\varepsilon^2 \eta^{-1} H_2$ and where η is a lower bound on the small divisors $\langle \lambda, n\rangle$ in (2.6). Assume $\eta \gg \varepsilon$. Repeat the procedure. In a phase space of fixed finite dimension d, one may expect a lower bound on the small divisors after M iterations of size

$$\min_{|n| \lesssim M} |\langle n, \lambda\rangle| > M^{-C(d)} \tag{2.8}$$

(because λ is assumed to be diophantine). To have a progressive scheme, we require that, say, $M^{-C(d)} > \varepsilon^{1/4}$, hence $M < \varepsilon^{1/C(d)}$. This allows us, by canonical transformations, to reduce H to a Hamiltonian of the form

$$H'(I, \theta) = \underbrace{H_0(I) + \varepsilon H_0'(I)}_{\text{integrable}} + \underbrace{0\left(\exp \frac{1}{\varepsilon^a}\right)}_{\text{perturbation}}$$

with $a = a(\lambda, d) > 0$. This yields thus stability times of order $\exp(1/\varepsilon^a)$.

Remarks. Previous analysis may be performed in complex canonical variables $q_j = (\operatorname{Re} q_j, \operatorname{Im} q_j)$ as well. There is no need to pass to action-angle variables (I_j, θ_j) which transformation becomes singular for I_j near 0.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

We consider next the situation of a Hamiltonian on an infinite-dimensional phase space. Let

$$H(q,\bar{q}) = \sum_{n=1}^{\infty} \lambda_n |q_n|^2 + \varepsilon H_1(q,\bar{q}).$$

In the PDE context, the $\{q_n\}$ are given by the Fourier modes of the function, thus

$$u(x) = \sum q_n \varphi_n(x)$$

and we denote

$$\|u\|_{H^s} = \Big(\sum |n|^{2s} |q_n|^2 \Big)^{1/2}.$$

We assume again that the frequency vector $\lambda = (\lambda_n)_{n\in\mathbb{Z}_+}$ satisfies certain diophantine properties and also certain asymptotic properties for $n \to \infty$, as will be clear from our future discussion.

The main idea in proving the stability of the unperturbed solution $\{q_n(t) = a_n e^{i\lambda_n t}\}$ with $\{a_n\}$ satisfying appropriate decay conditions is to construct a large finite-dimensional nearly invariant torus with zero-Lyapounov exponents.

Fix N large (depending on ε). Using canonical transformations again, we put H in the form

$$H' = \sum \lambda_n |q_n|^2 + H_0'(|q_n|^2) + 0(\varepsilon^M) + H_1'(q,\bar{q}) \tag{2.9}$$

where, in the (additional) last term H_1', all terms are at least *cubic* in the modes $\{q_n \mid |n| > N\}$. This property ensures us the zero-Lyapounov exponent property. Achieving this clearly requires to satisfy non-resonance conditions of the form

$$\Big| \sum_{j\le N} k_j \lambda_j \pm \lambda_{j_1} + \lambda_{j_2} \Big| > \varepsilon^{1/4} \tag{2.10}$$

where

$$\sum_{j\le N} |k_j| \lesssim M \quad \text{and} \quad j_1, j_2 \text{ are arbitrary.}$$

((2.10) is what is called in [**K**] the '2^e-Melnikov condition'.)

Achieving (2.10) requires at least that (unless the λ_n are growing rapidly)

$$M \log N \lesssim \log \frac{1}{\varepsilon}. \tag{2.11}$$

The additional issue of the j_1, j_2 modes (not subject to a restriction) is taken care of by using the additional asymptotic properties of the λ_j for $j \to \infty$. For instance, if $\lambda_j = j^2 + 0(1)$, as in the Schrödinger case, one may exploit the fact that $|\lambda_{j_1} - \lambda_{j_2}| \to \infty$ for $j_1 \neq j_2 \to \infty$ (in one dimension).

Denote

$$E_s = \sum |n|^{2s} |q_n|^2 = \|u\|_{H^s}^2.$$

According to the H' flow,

$$\dot{E}_s = \operatorname{Im} \sum |n|^{2s} \bar{q}_n \frac{\partial H'}{\partial \bar{q}_n}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

hence,

$$|\dot{E}_s| < 0(\varepsilon^M E_s + N^{-s} E_s^{3/2}).$$

This leads to the condition

$$|t| < T \sim \min(\varepsilon^{-M}, N^s). \tag{2.12}$$

For such times, $E_s(t)$ remains bounded (assuming $E_s(0)$ bounded) and the solution remains near the torus $\prod_{j\leq N}[|q_j| = |q_j(0)|] \times \{0\}$.

In view of (2.11), (2.12), we need, at least, to impose a time restriction

$$T \lesssim \min(\varepsilon^{-M}, \varepsilon^{-cs/M}). \tag{2.13}$$

Thus, to obtain a non-trivial stability time, the Sobolev exponent s should be taken to be large enough. More precise statements involving NLS and NLW equations will appear in the next section.

Remark. Let $(\lambda_n)_{n=0,1,2,\ldots}$ be independently chosen in $[0, 1]$ and consider the analytic case, thus $|q_n(0)| < e^{-\rho n}$ are exponentially decaying. We then obtain, instead of (2.12), a condition

$$T \sim \min(\varepsilon^{-M}, e^{(1/2)\rho N})$$

hence, from (2.11) and the optimal choice of M,

$$T \sim \min\left(\varepsilon^{-M}, \exp\left(\frac{1}{\varepsilon}\right)^{c/M}\right) \sim \exp c\frac{(\log(1/\varepsilon))^2}{\log\log(1/\varepsilon)}. \tag{2.14}$$

Time scale (2.14) is thus much shorter than the exponential $\exp(1/\varepsilon^a)$ $(a > 0)$ in the finite-dimensional case. In fact, it is of the same order as obtained in [**BFG**] for infinite systems of coupled harmonic oscillators with finite-range couplings.

Note also that in the PDE examples discussed next, the non-resonance properties of the (λ_n) will not be as good as in the independent case considered earlier and the stability times will be shorter than (2.14).

2.2. *Some PDE applications.* We mention some of the applications of the previously described method to NLS and NLW.

THEOREM 1. [**B7**] *Consider the NLS*

$$iu_t - u_{xx} + V(x)u + \varepsilon\frac{\partial H}{\partial \bar{u}} = 0 \tag{2.15}$$

or the NLW

$$y_{tt} - y_{xx} + \rho y + \varepsilon F'(y) = 0 \tag{2.16}$$

with Dirichlet boundary conditions on $[0, 1]$. *In (2.15), V is a 'typical' smooth odd potential, H is a real polynomial in $|u|^2$. In (2.16), $\rho \in \mathbb{R}_+$ is 'typical' and $F(y)$ an even polynomial.*

Then, given a positive integer A, there is a Sobolev exponent $s = s(A)$ such that any bounded initial data $H^s[0, 1]$ will stay $\sqrt{\varepsilon}$-close in $H^{s/2}$ to a quasi-periodic solution (in fact, the solution of the unperturbed = linear equation with appropriate frequency adjustment) for times $|t| < 1/\varepsilon^A$. For such times, the H^s-norm of the solution remains bounded.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

A few comments about the 'typicality' assumptions.

In (2.16), we have $\lambda_n = \sqrt{n^2 + \rho}$. For typical ρ, these satisfy non-resonance properties to a sufficient extent that the method described in the previous section can be applied. Regarding asymptotic behaviour, $\lambda_n = |n| + 0(1/|n|)$. This fact is also important in establishing (2.10). Concerning (2.15), we use the fact that the Dirichlet spectrum of a smooth potential V admits asymptotics

$$\lambda_n = \pi^2 n^2 + c_0(V) + c_1(V)\frac{1}{n} + \cdots + c_R(V)\frac{1}{n^R} + 0(n^{-R-1}) \tag{2.17}$$

up to any order R and where the $c_s(V)$ coefficients are given by multi-linear expressions in V (see [**PT**]). The 'typicality' of V refers here to the non-vanishing of these coefficients.

Let us also point out that, in Theorem 1, the assumption of H^s-boundedness for large s forces the 'compactness' of the unperturbed torus as brought up earlier.

Reduction of the nonlinear problem to a perturbation of a non-resonant linear system may sometimes be achieved by normal forms and amplitude–frequency modulation. Here is an example (see [**B8**]).

THEOREM 2. *Consider an NLS*

$$iu_t + u_{xx} + v u|u|^2 + \frac{\partial H_1}{\partial \bar{u}} = 0 \tag{2.17}$$

(1D periodic boundary conditions) where $v \neq 0$ and $H_1(u, \bar{u}) = P(|u|^2)$, P a real polynomial satisfying $P(0) = P'(0) = P''(0) = 0$.

Given a positive integer A, there is again an exponent $s = s(A)$ such that for $s > s(A)$ and 'typical' initial data $u(0) = \phi \in H^s(\mathbb{T})$, $\|\phi\|_{H^s} < \varepsilon$, one has

$$\sup_{|t|<T} \|u(t)\|_{H^s} \lesssim \varepsilon \quad \textit{for} \quad T = \left(\frac{1}{\varepsilon}\right)^A.$$

Starting from our Hamiltonian

$$H(q, \bar{q}) = \sum n^2 |q_n|^2 + v \sum_{n_1 - n_2 + n_3 - n_4 = 0} q_{n_1} \bar{q}_{n_2} q_{n_3} \bar{q}_{n_4} + \text{higher order } (\geq 6), \tag{2.18}$$

we make an initial reduction of the quartic term in (2.18) using a canonical transformation going back to [**KP**]. The key observation here is that

$$\begin{cases} n_1 - n_2 + n_3 - n_4 = 0 \\ n_1^2 - n_2^2 + n_3^2 - n_4^2 = 0 \text{ (resonance)} \end{cases} \Leftrightarrow \{n_1, n_3\} = \{n_2, n_4\}.$$

This permits us to transform H in the form

$$H'(q, \bar{q}) = \sum (n^2 + \rho)|q_n|^2 + v \sum |q_n|^4 + \text{(higher order)}. \tag{2.19}$$

We also exploit the conservation of the L^2-norm $\sqrt{\sum |q_n|^2}$ here. The first two terms in (2.19) define an integrable Hamiltonian satisfying the twist condition and for typical values of the action variables $I_n = |q_n|^2$, we obtain a perturbation of a linear non-resonant system. The main difficulty in this analysis comes from the fact that if we consider data in H^s with large s, which makes it easier to estimate the contribution of the tail modes q_n, $|n| > N$, the $\sum |q_n|^4$-term in (2.19) produces clearly less amplitude–frequency modulation and, hence, the $(\lambda_n)_{n \in \mathbb{Z}}$ frequencies will have worse non-resonance properties.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

2.3. *Perturbation of a convex integrable Hamiltonian.* Consider a real analytic Hamiltonian expressed in action-angle variables

$$H(I,\theta) = H_0(I) + \varepsilon H_1(I,\theta) \quad I = (I_1,\dots,I_d), \quad \theta = (\theta_1,\dots,\theta_d) \tag{2.20}$$

with $H_0(I)$ strictly convex, say

$$H_0(I) = \lambda I + \nu|I|^2 + \cdots \quad (\nu \neq 0).$$

In this situation, Nekhoroshev's theorem states that, for times $t < T = \exp c(1/\varepsilon^a)$ ($a = a(d) > 0$), there will be little change in the action variables. We will first recall Lochak's simple proof of this fact [**Lo**].

Replacing λ by $\lambda + 2\nabla H_0(I_0)$, we may replace I by $J = I - I_0$. Write, for simplicity,

$$H(J,\theta) = \lambda J + |J|^2 + \varepsilon H_1(J,\theta) \tag{2.21}$$

(there are no assumptions on λ).

Following the usual route, we first bring by canonical transformations H in the form of a sum of three components:

$$H'(J,\theta) = H_0(J) + \varepsilon H_1'(J,\theta) + 0(\varepsilon^A) \tag{2.22}$$

where H_1' consists of nearly resonant monomials. Here we need to be more precise.

Define for some fixed $\delta \gg \varepsilon$ to be specified later,

$$n \in \mathbb{Z}^d \begin{cases} \delta & \text{non-resonant if } |n\lambda| > \delta, \quad (*) \\ \delta & \text{resonant if } |n\lambda| \leq \delta. \quad (**) \end{cases}$$

Thus,

$$H_1'(J,\theta) = \sum_{|n|<B}^{(**)} c_n(J) e^{in\theta} \tag{2.23}$$

Here B is some cutoff related to A, say $A \sim B$.

The key idea in Lochak's argument is to invoke a rational approximation of $\lambda = (\lambda_1,\dots,\lambda_d)$ by a vector $\omega = (a_1/q,\dots,a_d/q)$. Take

$$Q \sim \varepsilon^{-1/10} \tag{2.24}$$

and $Q \leq q \leq Q^2$ satisfying $\max_{1\leq i\leq d} \|q\lambda_i\| \leq Q^{-1/d}$. Hence, there are $a_1,\dots,a_d \in \mathbb{Z}$ satisfying

$$|\lambda - \omega| = \max_i \left|\lambda_i - \frac{a_i}{q}\right| < \frac{1}{qQ^{1/d}}. \tag{2.25}$$

Take

$$\delta = \frac{1}{100q}. \tag{2.26}$$

If now $n \in \mathbb{Z}^d$, $|n| < B$ is δ-resonant, we get

$$|n\omega| \lesssim |n\lambda| + B|\lambda - \omega| < \delta + \frac{B}{qQ^{1/d}}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

and, therefore, either $n\omega = 0$ or

$$1 \leq \left| \sum_{i=1}^{d} n_i a_i \right| < \delta q + BQ^{-1/d} = \frac{1}{100} + BQ^{-1/d}.$$

Assuming

$$B \ll Q^{1/d}, \quad \text{hence } A \ll \left(\frac{1}{\varepsilon}\right)^{1/10d} \tag{2.27}$$

the conclusion is that

$$H_1'(J, \theta) = \sum_{\substack{|n|<B \\ n\omega=0}} c_n(J) e^{in\theta}. \tag{2.28}$$

Next, we estimate $|I(t) - I_0| = |J(t)|$ for $|t| < T$.

Since H' given by (2.22) is conserved, we have

$$\lambda . J(t) + |J(t)|^2 = 0(\varepsilon)$$

and again by (2.25)

$$|J(t)|^2 < |\omega J(t)| + 0(\varepsilon + \delta\varepsilon^{1/10d}|J(t)|). \tag{2.29}$$

The first term in (2.29) is estimated using the evolution equations. Thus

$$J_j(t) = \int_0^t \dot{I}_j(t) = \int_0^t \frac{\partial H'}{\partial \theta_j} = \varepsilon \int_0^t \frac{\partial H_1'}{\partial \theta_j} + 0(\varepsilon^A T)$$

and

$$\omega J(t) = \varepsilon \int_0^t (\nabla_\theta H_1' \omega) + 0(\varepsilon^A T).$$

But, in view of (2.28),

$$\nabla_\theta H_1' \omega = i \sum_{\substack{|n|<B \\ n\omega=0}} (n\omega) c_n(J) e^{in\theta} = 0.$$

Thus, $|\omega J(t)| = 0(\varepsilon^A T)$ and (2.29) implies for

$$|t| < T = \left(\frac{1}{\varepsilon}\right)^{A/2} \tag{2.30}$$

that

$$|J(t)| = |I(t) - I_0| < 0(\varepsilon^{1/10d}\delta) < \varepsilon^{1/10}. \tag{2.31}$$

In view of the restriction (2.27) on A, we may indeed reach a stability time $T_{\varepsilon,d} \sim \exp c/\varepsilon^a$ with $a = a(d) > 0$.

This proves the result.

Remark 1. Previous argument does require the dimension of phase space $d \ll \log 1/\varepsilon$, as is clear from restriction (2.27).

Remark 2. The stability times obtained from previous considerations get enhanced if λ is near a rational vector. Thus, amazingly, long-time stability gets favoured by two opposite scenarios.

Remark 3. There are some PDE applications of this method (see [**Bam**]). One fixes a few Fourier modes $n_1, \ldots, n_d$ and require $I_n = |q_n|^2$ to be very small for $n \notin \{n_1, \ldots, n_d\}$. Since $d \ll \log 1/\varepsilon$, the conditions on the initial data are unfortunately very restrictive.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

2.4. *Systems with finite-range interactions.* Considering infinite-dimensional systems, one can do more if the perturbation is finite range. Thus, consider

$$H(q,\bar{q}) = \sum_j (V_j I_j + I_j^2) + \varepsilon \sum_{\bar{n},\overline{n'}\in\mathbb{Z}_+^\infty} c(n,n') \operatorname{Re}\Big[\prod q_j^{n_j} \bar{q}_j^{n'_j}\Big] \tag{2.32}$$

denoting $I_j = |q_j|^2$. For the monomials in the perturbation, we make the following assumptions

$$\sum (n_j - n'_j) = 0 \tag{2.33}$$

$$2 \le \sum |n_j| + |n'_j| < C \tag{2.34}$$

$$\operatorname{diam}(\operatorname{supp}\bar{n} \cup \operatorname{supp}\overline{n'}) < C \tag{2.35}$$

where

$$\operatorname{supp}\bar{n} = \{j \in \mathbb{Z} \mid n_j \neq 0\}.$$

Condition (2.35) refers to the 'finite-range' assumption. From (2.33), we get L^2-norm conservation, thus $\sum I_j(t) = \sum I_j^0$ for all time (the property will be preserved under the symplectic transformations later on).

Typical examples (cf. [**FSW, BFG**]) are infinite systems of coupled harmonic oscillators in a random media (we restrict ourselves to a 1D lattice here). For instance,

$$H(q,\bar{q}) = \sum_j (V_j I_j + I_j^2) + \varepsilon \sum (\bar{q}_j q_{j+1} + q_j \bar{q}_{j+1}) \tag{2.36}$$

with $\{V_j\}$ a random potential, corresponds to the nonlinear lattice Schrödinger operator

$$iq_t + Vq + \varepsilon\Delta q + |q|^2 q = 0 \quad \text{where } q = (q_j)_{j\in\mathbb{Z}}. \tag{2.37}$$

Recall that, in the linear case,

$$iq_t + Vq + \Delta q = 0 \tag{2.38}$$

where the $\{V_j\}_{j\in\mathbb{Z}}$ are independent and identically distributed (i.i.d.) variables (thus on the 1D lattice), the solutions of (2.38) are almost periodic in time, since the lattice Schrödinger operator

$$V_j \delta_{j,j'} + \Delta$$

satisfies dynamical localization (see [**DJLS**]).

At this point, no nonlinear counterpart of this result seems known.

However, for (2.36) and, more generally, (2.32), with say $\{V_j\}_{j\in\mathbb{Z}}$ independently chosen in [0, 1], one may obtain stability times T_ε with $\log T_\varepsilon \gg \log(1/\varepsilon)$ under rather general assumptions. Such result appears in [**BFG**]. We present here a different approach that permits us to relax some of the assumptions from [**BFG**] (a remark is made about this at the end).

What follows is only a sketch of the argument. More precise formulations and details will appear elsewhere.

Fix some $\sigma > 0$ and take $q = (q_j)_{j\in\mathbb{Z}}$ to satisfy an H^σ-bound at $t = 0$, thus

$$\Big(\sum (1+j^2)^\sigma |q_j|^2\Big)^{1/2} < C \tag{2.39}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

implying, in particular, that $I_j^0 = |q_j(0)|^2 \lesssim |j|^{-2\sigma}$. Denoting $J = I - I^0$, we get

$$H = (2I^0 + V)J + |J|^2 + \varepsilon H_1(q, \bar{q})$$

where H_1 is as in (2.32). Put the Hamiltonian in the form (2.22) by canonical transformations. Thus,

$$H' = \lambda J + |J|^2 + \varepsilon H_0'(J) + \varepsilon H_1'(q, \bar{q}) + 0(\varepsilon^A) \tag{2.40}$$

where

$$\lambda = 2I^0 + V. \tag{2.41}$$

Here A denotes some large integer and H_1' consists of nearly resonant monomials with respect to the frequency vector λ. Thus,

$$H_1'(q, \bar{q}) = \sum c(n, n') \prod q_j^{n_j} \bar{q}_j^{n_j'} \tag{2.42}$$

with monomials in (2.42) satisfying

$$\sum (n_j - n_j') = 0$$

$$\sum |n_j| + |n_j'| \lesssim A \quad \text{and} \quad \operatorname{diam}(\operatorname{supp} \bar{n} \cup \operatorname{supp} \overline{n'}) \lesssim A \tag{2.43}$$

$$\left| \sum (n_j - n_j')\lambda_j \right| < \varepsilon^{1/100}. \tag{2.44}$$

Note again that the consistency of (2.40) requires J to satisfy some bound, say

$$\max_{j \in \mathbb{Z}} |J_j| < \varepsilon^{1/20} \tag{2.45}$$

and that will have to be verified at the end.

We now make the following observation. Fix an interval $S \subset \mathbb{Z}$ of length $L_0 \sim A$ and such that $\min_{j \in S} |j| > C(A)$. Since the $\{V_j\}$ are independently chosen in $[0, 1]$, the probability of S to intersect $\operatorname{supp} \bar{n} \cup \operatorname{supp} \overline{n'}$ for some $(n, n', n \neq n')$ in (2.42) is, at most,

$$C(A)\Big(\max_{j \in S} |j|^{-2\sigma} + \varepsilon^{1/100} \Big) < |j|^{-\sigma} + \varepsilon^{1/200}. \tag{2.46}$$

The first term in (2.46) comes from the $2I^0$ contribution in (2.41) satisfying (2.39). Again using independence, this observation implies that basically the nearly resonant monomials from H_1' are supported by separated 'islands' of bounded size. More precisely, there is a collection of disjoint intervals $\mathfrak{S}_\alpha \subset \mathbb{Z}$ of size, at most, $C(A)$ such that

$$\operatorname{dist}(\mathfrak{S}_\alpha, \mathfrak{S}_{\alpha'}) > L_0 \quad \text{for } \alpha \neq \alpha' \tag{2.47}$$

and for any monomial in (2.42), either $\operatorname{supp} \bar{n} \cup \operatorname{supp} \overline{n'} \subset \mathfrak{S}_\alpha$ for some α or $\operatorname{supp} \bar{n} \cup \operatorname{supp} \overline{n'} \subset \overline{\mathfrak{S}}$ where $\overline{\mathfrak{S}} = \{j \in \mathbb{Z} \mid |j| > j_0 = \varepsilon^{-A/20}\}$ is the tail (we make essential use of the 1D structure here).

This gives a decomposition of H_1' in (2.40) as

$$H_1' = \sum_\alpha H_{1,\alpha}' + \overline{H_1'} \tag{2.48}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

where $H'_{1,\alpha}$ (respectively $\overline{H'_1}$) depends on modes q_j with $j \in \mathfrak{S}_\alpha$ (respectively $j \in \overline{\mathfrak{S}}$). It thus suggests a decoupling of the system in subsystems of bounded dimension and a tail. Our aim is to apply the stability result for strictly convex Hamiltonian, as described in the previous section, to these subsystems.

From (2.40) and the preceding, we obtain, for $j \in \mathfrak{S}_\alpha$,

$$i\dot{q}_j = \frac{\partial H'}{\partial \bar{q}_j} = \lambda_j q_j + 2J_j q_j + \varepsilon \frac{\partial H'_0}{\partial J_j} q_j + \varepsilon \frac{\partial H'_{1,\alpha}}{\partial \bar{q}_j} + 0(\varepsilon^A) \tag{2.49}$$

and, if $j \notin \bigcup_\alpha \mathfrak{S}_\alpha \cup \overline{\mathfrak{S}}$,

$$i\dot{q}_j = \lambda_j q_j + 2J_j q_j + \varepsilon \frac{\partial H'_0}{\partial J_j} q_j + 0(\varepsilon^A). \tag{2.50}$$

In particular, (2.50) implies for $j \notin \bigcup \mathfrak{S}_\alpha \cup \bar{\mathfrak{S}}$ that

$$\dot{J}_j = 0(\varepsilon^A) \quad \text{hence } |J_j(t)| < 0(\varepsilon^A T) \text{ for } |t| < T. \tag{2.51}$$

For each α, assign a neighborhood $\tilde{\mathfrak{S}}_\alpha$ of $\mathfrak{S}_\alpha$ by including the $\frac{1}{4}L_0$ neighboring sites. Thus $\tilde{\mathfrak{S}}_\alpha \cap \mathfrak{S}_{\alpha'} = \phi$ for $\alpha' \neq \alpha$ and also $\tilde{\mathfrak{S}}_\alpha \cap \overline{\mathfrak{S}} = \phi$. Define the sub-Hamiltonian

$$H'_\alpha = \sum_{j \in \mathfrak{S}_\alpha} [\lambda_j J_j + J_j^2] + \varepsilon H'_{0,\alpha}(J) + \varepsilon H'_{1,\alpha} \tag{2.52}$$

where $H'_{0,\alpha}$ consists of monomials $\prod J_j^{m_j}$ in $H'_0(J)$ for which $\operatorname{supp} \bar{m} = \{j \mid m_j \neq 0\}$ intersects $\mathfrak{S}_\alpha$. Since $\operatorname{diam} \operatorname{supp} \bar{m} \lesssim A$, the choice of L_0 will ensure that $\operatorname{supp} \bar{m} \subset \tilde{\mathfrak{S}}_\alpha$ and, hence, H'_α depends only on modes q_j with $j \in \tilde{\mathfrak{S}}_\alpha$.

Equation (2.49) may be rewritten as

$$i\dot{q}_j = \frac{\partial H'_\alpha}{\partial \bar{q}_j} + 0(\varepsilon^A) \tag{2.53}$$

(the $0(\varepsilon^A)$ term also depends of course on modes q_j with $j \notin \tilde{\mathfrak{S}}_\alpha$).

It follows from (2.52), (2.53) and (2.51) that

$$\begin{aligned} \frac{d}{dt} H'_\alpha &= \sum_{j \in \mathfrak{S}_\alpha} \left(\frac{\partial H'_\alpha}{\partial q_j} \dot{q}_j + \frac{\partial H'_\alpha}{\partial \bar{q}_j} \dot{\bar{q}}_j \right) + \sum_{j \in \tilde{\mathfrak{S}}_\alpha \setminus \mathfrak{S}_\alpha} \varepsilon \frac{\partial H'_{0,\alpha}}{\partial J_j} \dot{J}_j \\ &= -i \sum_{j \in \mathfrak{S}_\alpha} \left(\frac{\partial H'_\alpha}{\partial q_j} \frac{\partial H'_\alpha}{\partial \bar{q}_j} - \frac{\partial H'_\alpha}{\partial \bar{q}_j} \frac{\partial H'_\alpha}{\partial q_j} \right) + 0(\varepsilon^A) + 0(\varepsilon^A) = 0(\varepsilon^A). \end{aligned}$$

Thus, for $|t| < T < \varepsilon^{-A/50}$, we obtain

$$|J_j(t)| < \varepsilon \quad \text{for } j \notin \bigcup \mathfrak{S}_\alpha \cup \overline{\mathfrak{S}} \tag{2.54}$$

and

$$|H'_\alpha(t) - H'_\alpha(0)| < \varepsilon. \tag{2.55}$$

For $j \in \overline{\mathfrak{S}}$,

$$i\dot{q}_j = \lambda_j q_j + 2J_j q_j + \varepsilon \frac{\partial H'_0}{\partial J_j} q_j + \varepsilon \frac{\partial \bar{H}'_1}{\partial \bar{q}_j} + 0(\varepsilon^A)$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

and since, again by (2.33),

$$\sum_{j\in\overline{\mathfrak{S}}}\left(q_j\frac{\partial H_1'}{\partial q_j}-\bar{q}_j\frac{\partial H_1'}{\partial \bar{q}_j}\right)=0,$$

we have

$$\frac{d}{dt}\left(\sum_{j\in\overline{\mathfrak{S}}}|q_j|^2\right)=0(\varepsilon^A).$$

Hence, for $|t|<T$,

$$\sum_{|j|>j_0}|q_j|^2<\sum_{|j|>j_0}I_j^0+0(\varepsilon^A T)<0(j_0^{-2\sigma}+\varepsilon^A T)<\varepsilon$$

$$\sum_{|j|>j_0}|J_j(t)|<\varepsilon. \tag{2.56}$$

Our aim is to derive from (2.55) an estimate on $(J_j)_{j\in\mathfrak{S}_\alpha}$ using the strictly convex Nekhoroshev as described in the previous section.

Thus first, for each α, consider the frequency vector $(\lambda_j)_{j\in\tilde{\mathfrak{S}}_\alpha}$ and choose an integer $\varepsilon^{-1/10}<r_\alpha<\varepsilon^{-1/5}$ and a rational vector $\omega^\alpha=(a_j/r_\alpha)_{j\in\tilde{\mathfrak{S}}_\alpha}$ satisfying

$$\max_{j\in\tilde{\mathfrak{S}}_\alpha}|\lambda_j-\omega_j|<\frac{1}{r_\alpha}\varepsilon^\gamma \tag{2.57}$$

where $\gamma=\gamma(A)>0$ (the dimensions $d<C(A)$ here).

Take

$$\delta_\alpha=\frac{1}{100r_\alpha}.$$

Partitioning the H_α' monomials into δ_α-resonant and non-resonant, further canonical transformations convert $H_{1,\alpha}'$ to $H_{1,\alpha}''$ of the form

$$H_{1,\alpha}'=\sum_{\substack{(\mathrm{supp}\,\bar{n}\cup\mathrm{supp}\,\overline{n'})\cap\mathfrak{S}_\alpha\neq\phi\\ \sum(n_j-n_j')\omega_j=0}}c(n,n')\prod q_j^{n_j}\prod\bar{q}_j^{n_j'}. \tag{2.58}$$

The full Hamiltonian H' in (2.40) is transformed into

$$H''=\lambda J+|J|^2+\varepsilon H_0''(J)+\varepsilon\left(\sum_\alpha H_{1,\alpha}''+\overline{H_1'}\right)+0(\varepsilon^A) \tag{2.59}$$

(where we pre-assume a bound $|J_j|<\delta_\alpha\varepsilon^\kappa$ for $j\in\tilde{\mathfrak{S}}_\alpha$ and some $\kappa>0$).

Note that the monomials in $H_{1,\alpha}''$ are supported by a set which may be slightly larger than $\mathfrak{S}_\alpha$ (certainly contained in $\tilde{\mathfrak{S}}_\alpha$) but we will keep the notation $\mathfrak{S}_\alpha$ for it.

Statements (2.49)–(2.56) remain preserved.

In particular, from (2.52), (2.55) and (2.57), it follows that

$$\sum_{j\in\mathfrak{S}_\alpha}\lambda_j J_j+J_j^2=0(\varepsilon)$$

$$\sum_{j\in\mathfrak{S}_\alpha}J_j^2=0\left(\varepsilon+\left|\sum_{j\in\mathfrak{S}_\alpha}\omega_j J_j\right|+\frac{1}{r_\alpha}\varepsilon^\gamma\sum_{j\in\mathfrak{S}_\alpha}|J_j|\right)$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

and, using (2.58),

$$\left|\frac{d}{dt}\left[\sum_{j\in\mathfrak{S}_\alpha}\omega_j J_j\right]\right| \le \left|\sum_{j\in\mathfrak{S}_\alpha}\omega_j\left(\operatorname{Im} q_j\frac{\partial H''_{1,\alpha}}{\partial q_j}\right)\right| + 0(\varepsilon^A) = 0(\varepsilon^A).$$

It follows that, for $|t| < T$,

$$\sum_{j\in\mathfrak{S}_\alpha}|J_j| = 0\left(\frac{\varepsilon^\gamma}{r_\alpha}\right) < \varepsilon^{1/10}. \tag{2.60}$$

Recall also (2.54), (2.56). This justifies our pre-assumption on J, in particular (2.45).

Finally, we make estimates related to the H^σ-norm (assuming $0 < \sigma \le 2$ say).

We have

$$\frac{d}{dt}\left[\sum_{j\in\mathfrak{S}_\alpha} I_j\right]\Bigg| < \left|\sum_{j\in\mathfrak{S}_\alpha}\left(q_j\frac{\partial H''_{1,\alpha}}{\partial q_j} - \bar{q}_j\frac{\partial H''_{1,\alpha}}{\partial \bar{q}_j}\right)\right| + 0(\varepsilon^A) = 0(\varepsilon^A)$$

since the first term vanishes by (2.33).

Thus,

$$\sum_{j\in\mathfrak{S}_\alpha} I_j(t) < \sum_{j\in\mathfrak{S}_\alpha} I_j^0 + 0(\varepsilon^A T).$$

Recalling also (2.51) for $j \notin \bigcup \mathfrak{S}_\alpha \cup \overline{\mathfrak{S}}$,

$$I_j(t) < I_j^0 + 0(\varepsilon^A T).$$

Therefore, clearly

$$\sum_{|j|<j_0}|j|^{2\sigma} I_j(t)$$

$$\lesssim \sum |j|^{2\sigma} I_j^0 + 0(\varepsilon^A T j_0^{2\sigma+1}) < C + 0\left(\varepsilon^A\left(\frac{1}{\varepsilon}\right)^{(A/50)+[(2\sigma+1)/20]A}\right) < C.$$

Again from (2.59),

$$\left|\frac{d}{dt}\left[\sum_{|j|>j_0}|j|^{2\sigma} I_j\right]\right| \le \left|\sum_{|j|>j_0}|j|^{2\sigma}\left(q_j\frac{\partial\overline{H'_1}}{\partial q_j} - \bar{q}_j\frac{\partial\overline{H'_1}}{\partial\bar{q}_j}\right)\right| + 0\left(\varepsilon^{A/2}\left(\sum_j |j|^{2\sigma} I_j\right)\right). \tag{2.61}$$

Again invoking (2.33), one may estimate

$$|(2.61)| < \sum |c(n,n')|\left|\sum_{|j|>j_0}(n_j - n'_j)|j|^{2\sigma}\right|\prod |q_j|^{n_j+n'_j} = 0\left(\frac{1}{j_0}\sum |j|^{2\sigma} I_j\right).$$

Therefore, collecting previous estimates, for $|t| < T$,

$$\sum |j|^{2\sigma} I_j(t) < C + 0\left(\frac{T}{j_0} + T\varepsilon^{A/2}\right)\max_{t'<T}\left[\sum |j|^{2\sigma} I_j(t')\right]$$

where

$$\frac{T}{j_0} + T\varepsilon^{A/2} < \left(\frac{1}{\varepsilon}\right)^{A(1/50-1/20)} < \varepsilon^{A/40}.$$

Consequently,

$$\sup_{|t|<T}\|q(t)\|_{H^\sigma} < C.$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

THEOREM 3. *Consider the 1D Hamiltonian (2.32) with finite-range interactions as described and* $\{V_j\}_{j\in\mathbb{Z}}$ *i.i.d. with smooth distribution. Let* $\sigma > 0$ *and* $q|_{t=0} = q(0) \in H^\sigma$, $\sum |j|^{2\sigma}|q_j|^2 < C$. *Then* $\sup_j |I_j(t) - I_j(0)| < \sqrt{\varepsilon}$ *and* $\|q(t)\|_{H^\sigma} < C+1$ *for time scales* T_ε *with* $\log T_\varepsilon \gg \log(1/\varepsilon)$.

Remark. This result is in a similar vein to the one proven in [**BFG**] but there are some significant differences. First, in [**BFG**], the perturbation in (2.32) is at least cubic in $\{q_j\}$ and does not cover (2.36). Second, the condition on σ is more restrictive there—more precisely, the authors assume $|q_j(0)| < |j|^{-2}$ in one dimension. However, [**BFG**] applies to arbitrary dimension, while ours is specifically 1D. Also our method is different, since we allow some separated islands with possible resonances in the perturbation of the final Hamiltonian. Their effect is then controlled using the Nekhoroshev stability in the strictly convex case and we exploit the first term $\sum I_j^2$ in (2.32).

PROBLEM. *What may be said about the stability of* $\{I_j\}$ *for the Hamiltonian (2.36) without any assumption on the* $\{V_j\}$*?*

Same problem with a more general short-range perturbation.

3. *An example*

We mention an example from [**B8**] that illustrates, to some extent, why 'too large' tori may have no stability, even in the weakest sense, beyond the obvious $T_\varepsilon \sim 1/\varepsilon$.

Consider the Hamiltonian on phase space $\mathbb{C}^{N+1}$;

$$H = \frac{1}{2}|q_0|^4 + \sum_1^N \lambda_n |q_n|^2 + \frac{1}{2}\sum_1^N |q_n|^4 + \varepsilon H_1 \tag{3.1}$$

where

$$H_1(q,\bar{q}) = \sum_1^N \sigma_k \left| \frac{1}{\sqrt{N}} \sum_1^N e^{ik\lambda_n} q_n \right|^6 (q_0 + \bar{q}_0). \tag{3.2}$$

The vector $\lambda = \{\lambda_n\}_{1\le n\le N}$ consists of randomly chosen numbers in $[1,2]$ according to a probability distribution with Fourier transform ρ, satisfying $\rho(y) < (1+y^2)^{-1}$ for $y \to \infty$.

The numbers $\{\sigma_k\}_{1\le k\le N}$ satisfy $|\sigma_k| \le 1$ and are appropriately chosen, as indicated later. Note that the $\sum_0^N I_n^2$ times in (3.1) provide a strictly convex $H_0(I)$.

For random $\{\lambda_n\}_{1\le n\le N}$, one may show that

$$\sum_1^N \left| \frac{1}{\sqrt{N}} \sum_1^N e^{ik\lambda_n} \xi_n \right|^6 \lesssim \left(\sum |\xi_n|^2 \right)^3$$

hence,

$$|H_1(q,\bar{q})| \lesssim \left(\sum_1^N |q_n|^2 \right)^3 |q_0|. \tag{3.3}$$

Therefore, endowing the phase space with the natural symplectic topology $= \ell^2_{N+1}(\mathbb{C})$, H is, indeed, an ε-perturbation of H_0.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

We take $\varepsilon \sim 1/N$. For the initial conditions, we take

$$q_0(0) = \gamma \quad \text{and} \quad q_n(0) \sim \frac{1}{\sqrt{N}} \ (1 \le n \le N) \tag{3.4}$$

where γ is a small constant.

We restrict time $t < T \sim N \sim 1/\varepsilon$. For an appropriate choice of $\{\sigma_k\}$ as before, one may ensure a drift

$$|I_0(T) - I_0(0)| = ||q_0(T)|^2 - |q_0(0)|^2| \gtrsim \gamma^4$$

and, hence, no stability beyond $1/\varepsilon$.

We briefly sketch the argument and refer to [**B8**] for details.

From (3.1) and (3.2),

$$i\dot{q}_0 = |q_0|^2 q_0 + \varepsilon \sum_{k=1}^{N} \sigma_k \left| \frac{1}{\sqrt{N}} \sum_{1}^{N} e^{ik\lambda_n} q_n \right|^6$$

hence,

$$|q_0(T)|^2 - |q_0(0)|^2 = \varepsilon \sum_{k=1}^{N} \sigma_k \int_0^T \left| \frac{1}{\sqrt{N}} \sum_{1}^{N} e^{ik\lambda_n} q_n \right|^6 \operatorname{Im} q_0(t)\, dt. \tag{3.5}$$

For t as before and $n = 1, \ldots, N$, we may write, up to the perturbation,

$$q_n(t) \sim \frac{1}{\sqrt{N}} e^{-i\lambda_n t}$$

and (3.5) approximately becomes

$$\begin{aligned} &\varepsilon \sum_{k=1}^{N} \sigma_k \int_0^T \left| \frac{1}{N} \sum_{1}^{N} e^{i\lambda_n (k-t)} \right|^6 \operatorname{Im} q_0(t)\, dt \\ &\qquad \approx \varepsilon \sum_{k=1}^{N} \sigma_k \int_0^T |\rho(k-t)|^6 \operatorname{Im} q_0(t)\, dt + 0\left(\varepsilon N T \left(\frac{\log T}{N} \right)^3 \right). \end{aligned} \tag{3.6}$$

The last term in (3.6) comes from standard probabilistic considerations and is bounded by $0((\log N)^3/N^2) = o(1)$.

Taking γ sufficiently small in (3.4), we may ensure that

$$\int_0^T |\rho(k-t)|^6 \operatorname{Im} q_0(t) \sim \operatorname{Im} q_0(k)$$

(invoking some tail estimates on ρ).

Thus, (3.6) becomes

$$\varepsilon \sum_{k=1}^{N} \sigma_k \operatorname{Im} q_0(k). \tag{3.7}$$

The next point is that, invoking a Schauder fixed point argument, one may ensure that

$$\sigma_k = \operatorname{Im} q_0(k) \quad \text{for } 1 \le k \le N \tag{3.8}$$

(observe that q_0 does depend on $\{\sigma_k\}$!). If (3.8) holds, one gets that

$$(3.5) \sim (3.8) = \varepsilon \sum_{k=1}^{N} |\operatorname{Im} q_0(k)|^2 \sim \varepsilon N \sim \varepsilon T \sim 1$$

as desired.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Remark. If now in (3.1) we let ε be sufficiently small ($\ll 1/N$), H may be transformed to an ε^A perturbation of an integrable Hamiltonian (since λ is non-resonant). But we may add further terms of the same form to H_1, coupling q_0 to subsequent sets of tail modes (taking N much larger) to reach a similar effect for a sequence of time scales $T_\varepsilon \sim 1/\varepsilon \to \infty$. The resulting perturbation H_1 remains real analytic.

A. *Appendix. Growth of Sobolev norms in the NLS*

Consider a defocusing NLS equation

$$iu_t + \Delta u - \frac{\partial H_1}{\partial \bar{u}} = 0$$

on $\mathbb{T}^d$, where $H_1(u, \bar{u}) = P(|u|^2)$, P a polynomial with positive leading coefficient. In one dimension, two dimensions and subcritical three dimensions (thus det $P \leq 2$), H^s solutions exist globally for $s \geq 1$.

Moreover,

$$\|u(t)\|_{H^s} < |t|^{C(s-1)} \quad \text{for } |t| \to \infty \tag{A.1}$$

for some constant C (see [**B4**] and also [**St**]).

Our aim here is to sketch an argument based on normal form reduction and which leads to better values of the constant C in (A.1).

Let us consider the 1D quintic NLS

$$iu_t + \Delta u - u|u|^4 = 0 \tag{A.2}$$

thus,

$$iu_t = \frac{\partial H}{\partial \bar{u}}$$

where

$$\begin{aligned} H(\phi) &= \int [|\nabla \phi|^2 + |\phi|^6] \\ &= \sum j^2 |q_j|^2 + \sum_{j_1 - j_2 + j_3 - j_4 + j_5 - j_6 = 0} q_{j_1} \bar{q}_{j_2} q_{j_3} \bar{q}_{j_4} q_{j_5} \bar{q}_{j_6} \end{aligned} \tag{A.3}$$

denoting

$$q_j = \hat{\phi}(j) \quad \text{for } j \in \mathbb{Z}.$$

By normal form reduction, we mean phase space transformations that will transform

$$H_1(q, \bar{q}) = \sum_{j_1 - j_2 + j_3 - j_4 + j_5 - j_6 = 0} q_{j_1} \bar{q}_{j_2} q_{j_3} \bar{q}_{j_4} q_{j_5} \bar{q}_{j_6}$$

in expressions involving only monomials of the form

$$q_{j_1} \bar{q}_{j_2} q_{j_3} \cdots q_{j_{2r-1}} \bar{q}_{j_{2r}} \tag{A.4}$$

where

$$j_1 - j_2 + \cdots - j_{2r-1} - j_{2r} = 0 \tag{A.5}$$

and

$$|j_1^2 - j_2^2 + \cdots - j_{2r}^2| < K \tag{A.6}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

with K a large number. Fix a time interval $[0, T]$. We will take

$$K = T^{\delta} \tag{A.7}$$

with $\delta > 0$ an arbitrarily small fixed constant. The degree $2r$ of the monomials (A.4) will then be bounded depending on δ. The main idea is to remain up to a negligible error with only 'nearly-resonant' monomials in the sense of (A.6). This gives us additional information on the frequency configurations that will permit better bounds on the H^s-norms. We assume our initial data to be bounded in $H^1(\mathbb{T})$. From conservation of the Hamiltonian (A.3), there is an *a priori* bound

$$\Big(\sum j^2 |q_j|^2\Big)^{1/2} < C \tag{A.8}$$

for all time.

An important issue is to understand how our symplectic transformations act on the phase space endowed with various H^σ topologies. In one dimension, as observed in [**KP**], these are easily seen to act boundedly on a neighborhood of 0 in $H^\sigma(\mathbb{T})$, if $\sigma > \frac{1}{2}$. Only the algebra property in one dimension of these spaces is involved here. Recall that our symplectic transformations are obtained as time-1 shift Γ_F under the flow

$$i q_t = \frac{\partial F}{\partial \bar{q}} \tag{A.9}$$

where the Hamiltonians F are obtained as follows.

Let

$$H_1(q, \bar{q}) = \sum_{j_1 - j_2 \cdots - j_{2r} = 0} c(j_1, \ldots, j_{2r}) q_{j_1} \bar{q}_{j_2} \cdots \bar{q}_{j_{2r}} \tag{A.10}$$

be (part of) the NLS Hamiltonian obtained at some stage of the process. More specifically, (A.10) will be restricted to monomials satisfying

$$|j_1^2 - j_2^2 + \cdots - j_{2r}^2| > K \tag{A.11}$$

with K as in (A.7). We define then

$$F(q, \bar{q}) \sim \sum \frac{c(j_1, \ldots, j_{2r})}{j_1^2 - j_2^2 + \cdots - j_{2r}^2} q_{j_1} \bar{q}_{j_2} \cdots \bar{q}_{j_{2r}} \tag{A.12}$$

so that

$$\{F, H_0\} = H_1 \quad \text{where } H_0 = \sum j^2 |q_j|^2.$$

Since we impose (A.11) with K large, Γ_F will act boundedly on bounded subsets of H^σ, $\sigma > \frac{1}{2}$ and, hence, $H^1(\mathbb{T})$ (on which we have an *a priori* bound).

Moreover, because of (A.7), the resulting Hamiltonian will only invoke monomials of bounded degree and an additional term that can be made smaller than T^{-C} for any C.

Next, we will need to make a more careful study of the normal form analysis and show that the Γ_F (with F of bounded degree) also act boundedly on the ball of H^σ, for any $\sigma > 0$. At this stage, the more specific structure of the divisors $j_1^2 - j_2^2 - \cdots - j_{2r}^2$ will be essential.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Observe that in two dimensions, the 'obvious' estimate on the symplectic transformations would require $H^\sigma(\mathbb{T}^2)$ with $\sigma > 1$, on which we do not have an *a priori* bound. Justifying the symplectic transformations Γ_F on bounded subsets of H^1 is thus less obvious than in one dimension and already at this stage it requires more refined analysis we are going to describe next. Returning to (A.2), the key ingredient in the local Cauchy problem is the following periodic Strichartz inequality:

$$\|e^{it\Delta}\phi\|_{L^6(\mathbb{T}\times\mathbb{T})} \leq \left(\exp C\frac{\log N}{\log\log N}\right)\|\phi\|_2 \tag{A.13}$$

where

$$e^{it\Delta}\phi = \sum_{j\in\mathbb{Z}} \hat{\phi}(j)e^{i(jx+j^2t)} \tag{A.14}$$

is the solution operator to the linear equation $iu_t + \Delta u = 0$ and we assume in (A.13) that

$$\operatorname{supp}\hat{\phi} \subset B(0, N).$$

(In the periodic case, $e^{it\Delta}$ does not satisfy the space–time L^6-estimate!)

Also in our considerations here, inequality (A.13) will be essential. Define the following norms on polynomials on $\mathbb{T}\times\mathbb{T}$:

$$u(x,t) = \sum_{j,m\in\mathbb{Z}} \hat{u}(j,m)e^{i(jx+mt)}$$

$$\|u\|_{0,p} = \left[\sum_m \left(\sum_j |\hat{u}(j, j^2+m)|^2\right)^{p/2}\right]^{1/p} \tag{A.15}$$

where $1 \leq p \leq \infty$. Note that $\|u\|_{L_t^\infty L_x^2} \leq \|u\|_{0,1}$.

Remark. $\|\cdot\|_{0,p}$ is a Besov-type norm associated to the linear group (A.14) and is related to the '$X_{s,b}$-spaces' used in the study of the Cauchy problem for NLS (see [**B1**]). The $X_{s,b}$-spaces do not seem to be suited for our purpose here. In the notation $\|\cdot\|_{0,p}$, the 0-index refers to zeroth-order derivatives and we may introduce $\|\cdot\|_{\sigma,p}$ similarly.

Thus, if $\operatorname{supp}\hat{u} \subset B(0, N)$, (A.13) easily implies that

$$\|u\|_{L^6(\mathbb{T}^2)} \leq \left(\exp C\frac{\log N}{\log\log N}\right)\|u\|_{0,1}. \tag{A.16}$$

Let us use in the following the symbol $\ll$ for inequalities involving factors N^ε with ε arbitrary small. These factors will afterwards be absorbed at the cost of a bit of the smoothness of $\phi = \sum q_j e^{ijx}$. With this notation, by (A.16), our initial Hamiltonian

$$H_1(q) = \sum_{j_1-j_2\cdots=0} q_{j_1}q_{j_2}\cdots\bar{q}_{j_6}$$

satisfies

$$\int_{\mathbb{T}} H_1(q(t))\,dt \ll \|q\|_{0,1}^6 \tag{A.17}$$

if we replace q by a function of time, identifying

$$q = \sum q_j(t)e^{ijx}.$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

As the recursive assumption, we require that all Hamiltonians (A.10) involved in the process satisfy the bound

$$\mathcal{H}_1 = \int_{\mathbb{T}} H_1(q(t))\,dt \ll 1 \tag{A.18}$$

assuming $\|q\|_{0,1}$ to be bounded (this should hold also replacing the $c(J)$-coefficients in (A.10) by $|c(J)|$). If (A.18), duality implies that

$$\left\| \frac{\partial \mathcal{H}_1}{\partial q} \right\|_{0,\infty} \ll 1. \tag{A.19}$$

Consider, next, the Hamiltonian introduced in (A.12).

Thus, replacing q by $q(t)$, we have

$$\left\| \frac{\partial \mathcal{F}}{\partial q} \right\|_{0,1} = \sum_m \left[\sum_j \left| \left(\sum_{j_2 - j_3 + \cdots + j_{2r} = j} \frac{c(j, j_2, \ldots, j_{2r})}{j^2 - j_2^2 \cdots - j_{2r}^2} q_{j_2} \bar{q}_{j_3} \cdots q_{j_{2r}} \right)^{\wedge} (j^2 + m) \right|^2 \right]^{1/2}. \tag{A.20}$$

Here the factors $q^{(\ell)} (\ell = 2, \ldots, 2r)$ are bounded in $\| \|_{0,1}$-norm; hence, by convexity, they may be replaced by

$$q^{(\ell)}(t) = (e^{it\Delta} \phi^{(\ell)}) e^{it m_\ell} \tag{A.21}$$

with $\phi^{(\ell)}$ bounded in $L^2(\mathbb{T})$ and $m_\ell \in \mathbb{Z}$. If (A.21) holds, then

$$(q_{j_2}^{(2)} \overline{q_{j_3}^{(3)}} \cdots q_{j_{2r}}^{(2r)})^{\wedge}(m') \neq 0$$

only if $m' = j_2^2 - j_3^2 + \cdots + j_{2r}^2 + \bar{m}$ where $\bar{m} = m_2 - m_3 + \cdots + m_{2r}$. Therefore ($\bar{m}$ is fixed here),

$$\begin{aligned} (\text{A.20}) &= \sum_{|m - \bar{m}| > K} \frac{1}{|m - \bar{m}|} \\ &\quad \times \left[\sum_j \left| \left(\sum_{j_2 - j_3 + \cdots + j_{2r} = j} c(j, j_2, \ldots, j_{2r}) q_{j_2} \bar{q}_{j_3} \cdots q_{j_{2r}} \right)^{\wedge} (j^2 + m) \right|^2 \right]^{1/2} \\ &= \sum_{|m - \bar{m}| > K} \frac{1}{|m - \bar{m}|} \left[\sum_j \left| \left(\frac{\partial \mathcal{H}_1}{\partial q} \right)^{\wedge} (j, j^2 + m) \right|^2 \right]^{1/2} \end{aligned} \tag{A.22}$$

and, since (A.19) and $\sum_m 1/|m - \bar{m}|$ is only logarithmically divergent, we may write

$$\left\| \frac{\partial \mathcal{F}}{\partial q} \right\|_{0,1} \ll 1. \tag{A.23}$$

Also, if H_1 satisfies (A.18) and F (A.23), then the Hamiltonian obtained by the Poisson bracket

$$\{H_1, F\} = i \sum_j \left[\frac{\partial H_1}{\partial q_j} \frac{\partial F}{\partial \bar{q}_j} - \frac{\partial H_1}{\partial \bar{q}_j} \frac{\partial F}{\partial q_j} \right]$$

will still satisfy (A.18). Indeed, for $q = q(t)$ bounded in $\| \|_{0,1}$-norm,

$$\left| \int_{\mathbb{T}} \left[\sum_j \frac{\partial \mathcal{H}_1}{\partial \bar{q}_j} \frac{\partial \mathcal{F}}{\partial q_j} \right] dt \right| = \left| \left\langle \frac{\partial \mathcal{F}_1}{\partial q}, \frac{\partial \mathcal{H}_1}{\partial q} \right\rangle \right| \leq \left\| \frac{\partial \mathcal{H}_1}{\partial q} \right\|_{0,\infty} \left\| \frac{\partial \mathcal{F}}{\partial q} \right\|_{0,1} \ll 1$$

from (A.19) and (A.23).

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Finally, returning to Γ_F obtained from (A.9), we estimate the vector field $\partial F/\partial \bar{q}$ say in L^2-norm as follow. Denote, for $j \in \mathbb{Z}$,

$$q_j(t) = q_j e^{ij^2 t}.$$

Thus, $\|q\|_{0,1} = \|q\|_2$ and, with further abuse of notation,

$$\frac{\partial F}{\partial \bar{q}} = \frac{\partial \mathcal{F}}{\partial q}\bigg|_{t=0}$$

so that, again by (A.23),

$$\left\| \frac{\partial F}{\partial \bar{q}} \right\|_2 \leq \left\| \frac{\partial \mathcal{F}}{\partial q} \right\|_{L_t^\infty L_x^2} \leq \left\| \frac{\partial \mathcal{F}}{\partial q} \right\|_{0,1} \ll 1.$$

Thus, the preceding gives an approach to obtaining bounds $\|\partial F/\partial \bar{q}\|_2 \ll N^\varepsilon$ for $q = \{q_j\}_{|j|<N}$ bounded in $\ell^2(\mathbb{Z})$. A bit more careful analysis shows that if we fix any $\sigma > 0$, then, assuming $\|q\|_{H^\sigma}$ bounded, we have

$$\left\| \frac{\partial F}{\partial \bar{q}} \right\|_{H^\sigma} < K^{-\gamma}, \quad \gamma = \gamma(\sigma) > 0 \tag{A.24}$$

with K the lower bound (A.11) on the divisors.

Moreover, we may write, still assuming $\|q\|_{H^\sigma}$ to be bounded, for any $s \geq 0$,

$$\left\| \frac{\partial F}{\partial \bar{q}} \right\|_{H^s} < K^{-\gamma} \|q\|_{H^s}. \tag{A.25}$$

Consequently, Γ_F acts boundedly on $H^s \cap [\|\phi\|_{H^\sigma} < C]$.

Coming back to the problem of possible growth of higher Sobolev norms in (A.2), the preceding thus permits us to convert the equation to

$$i\dot{q}_j = j^2 q_j + \frac{\partial H_1}{\partial \bar{q}_j} + 0(T^{-C}) \quad (j \in \mathbb{Z}) \tag{A.26}$$

with $H_1 = \sum c(\bar{j}) q_{j_1} \bar{q}_{j_2} \cdots \bar{q}_{j_{2r}}$ (r bounded) satisfying (A.18) and involving only monomials such that

$$|j_1^2 - j_2^2 - \cdots - j_{2r}^2| < K = T^\delta$$

($C = C(r, \delta)$ is a large constant).

Define, for $s \geq 1$,

$$I_s(q) = \|q\|_{H^s}^2 = \sum_j |j|^{2s} |q_j|^2. \tag{A.27}$$

From (A.26),

$$\dot{I}_s \sim \sum |j|^{2s} \left(q_j \frac{\partial H_1}{\partial q_j} - \bar{q}_j \frac{\partial H_1}{\partial \bar{q}_j} \right) + 0(T^{-C} I_s). \tag{A.28}$$

The first term in (A.28) equals

$$\sum_{\substack{\sum_\ell (-1)^\ell j_\ell = 0 \\ |\sum_\ell (-1)^\ell j_\ell^2| < K}} \left(\sum (-1)^\ell j_\ell^{2s} \right) c(\bar{j}) q_{j_1} \bar{q}_{j_2} \cdots \bar{q}_{j_{2r}}. \tag{A.29}$$

The next main point is the following simple arithmetic lemma.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

LEMMA. *Assume* $\{j_\ell\}_{1\le\ell\le 2r}$ *numbers satisfying*

$$\sum(-1)^\ell j_\ell = 0 \tag{A.30}$$

$$\left|\sum(-1)^\ell j_\ell^2\right| < K. \tag{A.31}$$

Denoting by $j_1^* \ge j_2^* \ge \cdots$ *the decreasing rearrangement of* $\{|j_\ell|\}$, *we have*

$$\left|\sum(-1)^\ell j_\ell^{2s}\right| \lesssim K(j_1^*)^{2(s-1)} j_3^* j_4^*. \tag{A.32}$$

Proof. We may clearly assume $j_1^* = |j_1|$. We distinguish two cases.

Case I: $j_2^* = |j_\ell|$ with ℓ even. Take $j_2^* = |j_2|$. From (A.31)

$$j_1^2 = j_2^2 + 0((j_3^*)^2 + K)$$

hence,

$$|j_1^{2s} - j_2^{2s}| \lesssim j_1^{2(s-1)}((j_3^*)^2 + K)$$

$$\left|\sum(-1)^\ell j_\ell^{2s}\right| \lesssim j_1^{2(s-1)}((j_3^*)^2 + K). \tag{A.33}$$

If $j_4^* \sim j_3^*$, (A.33) implies (A.32). Assume $j_4^* = o(1) j_3^*$. Then (A.30) and (A.31) imply

$$|j_1 - j_2| = (1 + o(1)) j_3^* \tag{A.34}$$

$$|j_1^2 - j_2^2| = (1 + o(1))(j_3^*)^2. \tag{A.35}$$

From (A.34), j_1 and j_2 clearly need to have the same sign. Hence,

$$(1 + o(1))(j_3^*)^2 = ||j_1| - |j_2||(|j_1| + |j_2|) = (1 + o(1)) j_3^*(j_1^* + j_2^*).$$

Unless $j_3^* = 0$, in which case $j_1 = j_2$ and the left-hand side of (A.32) vanishes, this gives

$$(1 + o(1)) j_3^* = j_1^* + j_2^*$$

an obvious contradiction.

Case II: $j_2^* = |j_\ell|$ with ℓ odd.

Take $j_2^* = |j_3|$. From (A.31)

$$(j_1^*)^2 + (j_2^*)^2 \le (j_3^*)^2 + (j_4^*)^2 + \cdots + K$$

hence,

$$(j_1^*)^2 \le (j_4^*)^2 + (j_5^*)^2 + \cdots + K \lesssim (j_4^*)^2 + K,\ j_1^* \lesssim j_4^* + K$$

and (A.32) is obvious. This proves the lemma. □

We now return to (A.29), which, in view of (A.32), we bound by

$$\sum_{\substack{\sum(-1)^\ell j_\ell = 0 \\ |\sum(-1)^\ell j_\ell^2| < K}} |c(j)|(|j_1|^{s-1}|q_{j_1}|)(|j_2|^{s-1}|q_{j_2}|)(|j_3||q_{j_3}|)(|j_4||q_{j_4}|) \prod_{\ell\ge 5} |q_{j_\ell}| \tag{A.36}$$

assuming $|j_1| = j_1^*$, $|j_2| = j_2^*$, $|j_3| = j_3^*$, $|j_4| = j_4^*$ say.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

Define then

$$Q_j^{(1)}(t) = Q_j^{(2)}(t) = |j|^{s-1}|q_j|e^{ij^2t}$$
$$Q_j^{(3)}(t) = Q_j^{(4)}(t) = |j||q_j|e^{ij^2t}$$
$$Q_j^{(\ell)}(t) = |q_j|e^{ij^2t}.$$

Since

$$H_1(q,\bar{q}) = \sum |c(j)|q_{j_1}\bar{q}_{j_2}\cdots\bar{q}_{j_{2r}}$$

satisfies (A.18), it follows that

$$\sup_{m\in\mathbb{Z}}\left|\int_{\mathbb{T}} H_1(Q^{(1)}(t), Q^{(2)}(t),\ldots)e^{-imt}\,dt\right| \tag{A.37}$$
$$\lesssim J^{\varepsilon}\prod_{\ell=1}^{2r}\|Q^{(\ell)}\|_{0,1} \tag{A.38}$$
$$\lesssim \Big(\sum |j|^{2(s-1)+\varepsilon}|q_j|^2\Big)\Big(\sum |j|^2|q_j|^2\Big)\Big(\sum |q_j|^2\Big)^{2(r-2)}. \tag{A.39}$$

(In (A.38) we denoted by $J \sim \max(|j_1|, |j_2|, \ldots, |j_{2r}|) = |j_1|$ and $\varepsilon > 0$ arbitrary small.)

Since (A.36) is obtained by summing (A.37) for $|m| < K$, the bound (A.39) implies

$$\text{(A.29), (A.36)} \lesssim KI_{s-1+\varepsilon}(q)\|q\|_{H^1}^2\|q\|_2^{2(r-2)}$$
$$< T^{\delta}I_{s-1+\varepsilon}(q)\|q\|_{H^1}^{2(r-1)}. \tag{A.40}$$

Invoking interpolation between H^1 and H^s.

$$I_{s-1+\varepsilon}(q) \le I_s(q)^{1-(1-\varepsilon)/(s-1)}\|q\|_{H^1}^{2(1-\varepsilon)/(s-1)}.$$

(A.28) and (A.40) imply that

$$|\dot{I}_s| \lesssim T^{\delta}I_s(q)^{1-(1-\varepsilon)/(s-1)}C(\|q\|_{H^1}) + 0(T^{-C}I_s).$$

Thus, integrating for $|t| \le T$, it follows that

$$I_s(t) \lesssim T^{(1+\delta)(s-1)/(1-\varepsilon)}$$

hence,

$$\Big(\sum |j|^{2s}|q_j(t)|^2\Big)^{1/2} = \|u(t)\|_{H^s} \ll T^{(s-1)/2+\varepsilon}. \tag{A.41}$$

The conclusion is thus

THEOREM A.1. *Consider the 1D defocusing quintic NLS*

$$iu_t + u_{xx} - u|u|^4 = 0$$

with periodic boundary conditions. Then, for $s \ge 1$, $u(0) \in H^s(\mathbb{T})$, we have, for all $\varepsilon > 0$,

$$\|u(t)\|_{H^s} < |t|^{(s-1)/2+\varepsilon} \quad \text{for } |t| \to \infty. \tag{A.42}$$

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press

REFERENCES

[Bam] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Math. Z.* **230** (1999), 345–387.

[B1] B. Bourgain. *Nonlinear Schrödinger Equations (Park City Lectures)*. American Mathematical Society, Providence, RI 1999.

[B2] J. Bourgain. Growth of Sobolev norms in linear Schrödinger equations with quasi-periodic potential. *Comm. Math. Phys.* **204**(1) (1999), 207–257.

[B3] J. Bourgain. On the growth of Sobolev norms in linear Schrödinger equations with smooth time dependent potential. *J. Math.* **77** (1999), 315–348.

[B4] J. Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices* **6** (1996), 277–304.

[B5] J. Bourgain. Greens function estimates for lattice Schrödinger operators and applications. *Ann. Math. Studies* to appear.

[B6] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. Math.* **148** (1998), 463–439.

[B7] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.* **6** (1996), 201–230.

[B8] J. Bourgain. On diffusion in high-dimensional Hamiltonian Systems and PDE. *J. Anal. Math.* **80** (2000), 1–35.

[BFG] G. Benettin, J. Fröhlich and A. Giorgili. A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom. *Comm. Math. Phys.* **119** (1989), 95–108.

[Cr] W. Craig. Problèmes de petits diviseurs dans les equations aux dérivées partielles. *Panoramas et Syntheses*. Société Mathématique de France, Paris, 2000.

[DJLS] R. del Rio, S. Jitomirskaya, Y. Last and B. Simon. Operators with singular continuous spectrum, IV. Hausdorff dimensions, rank one perturbations, and localization. *J. Anal. Math.* **69** (1996), 153–200.

[FSW] J. Fröhlich, T. Spencer and E. Wayne. Localization in disordered, nonlinear dynamical systems *J. Stat. Phys.* **42** (1986), 257–275.

[K] S. Kuksin. *Analysis of Hamiltonian PDE (Oxford Lecture Series in Mathematics and Applications, 19)*. Oxford University Press, Oxford, 2000.

[KP] S. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. Math.* (2) **143** (1996), 149–179.

[Lo] P. Lochak. Canonical perturbation theory via simultaneous approximation. *Usp. Mat. Nauk* **47**(6) (1992), 59–140.

[N] N. Nekhoroshev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Usp. Mat. Nauk* **32**(1) (1977), 5–66.

[P] J. Pöschel. On the construction of almost periodic solutions for a nonlinear Schrödinger equation. *Ergod. Th. & Dynam. Sys.* **22** (2002), 1537–1549.

[PM] R. Perez-Marco. *KAM Methods in PDE*. Sem. Bourbaki, Paris, June 2002.

[PT] J. Pöschel and E. Trubowitz. *Inverse Spectral Theory*. Academic Press, Boston, MA, 1987.

[St] G. Staffilani. Quadratic forms for a 2D semilinear Schrödinger equation. *Duke Math. J.* **86** (1997), 79–108.

[W] C. Wayne. On the elimination of nonresonant harmonics. *Comm. Math. Phys.* **103** (1986), 351–386.

[ZS] V. Zakharov and A. B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Sov. Phys.–JETP* **34** (1972), 62–69.

https://doi.org/10.1017/S0143385703000750 Published online by Cambridge University Press