Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica D 206 (2005) 213–231
PHYSICA D
www.elsevier.com/locate/physd

Breather solutions in the diffraction managed NLS equation

Panayotis Panayotaros

Depto. Matemáticas y Mecánica, I.I.M.A.S.-U.N.A.M., Apdo. Postal 20-726, 01000 México D.F., México

Received 22 October 2004; received in revised form 27 April 2005; accepted 9 May 2005
Available online 6 June 2005
Communicated by C.K.R.T. Jones

Abstract

We show the existence of localized breather solutions in an averaged version of the discrete nonlinear Schrödinger equation (NLS) with diffraction management, a system that models coupled waveguide arrays with periodic diffraction management geometries. The breather solutions are constrained extrema of the Hamiltonian of the averaged system and their existence is shown by a discrete version of the concentration-compactness principle. The main assumptions are that the averaged diffraction is sufficiently small (compared to strength of the nonlinearity) and that the sign of the nonlinearity corresponds to the focusing case. An interesting feature of the problem is that the nonlinear interaction between neighboring lattice sites can be large and is of infinite range. On the other hand, the interaction decays rapidly at sufficiently large distances, and this plays an important role in the proof. The results also apply to higher dimensional lattices, and to the discrete NLS equation.

MSC: 37J45; 53C22; 70F16

PACS: 42.65Sf; 45.20.Jj

Keywords: Nonlinear Schrödinger equations; Discrete breathers; Variational methods

1. Introduction

The study of the dynamics of nonlinear lattice systems has a long history, e.g. in models of solids, and in recent years discrete systems have found many applications in nonlinear optics (see e.g. [6,15]). In the present work we consider an array of coupled waveguides with the zigzag diffraction management geometry introduced in [8] and study coherent structures using as a starting point the model of [3]

$$\partial_t u_n = \mathrm{i}D(t)(u_{n+1} - 2u_n + u_{n-1}) - 2\mathrm{i}\gamma(g(a))_n, \quad (g(u))_n = |a_n|^2 a_n, \tag{1.1}$$

E-mail address: panos@mym.iimas.unam.mx.

0167-2789/$ – see front matter
doi:10.1016/j.physd.2005.05.006

with u_n the complex amplitude (of the electric field) at the site $n \in \mathbf{Z}$, and D a T-periodic function with average δ and oscillating part $\tilde{d}$. System (1.1) is also a discrete version of the cubic nonlinear Schrödinger (NLS) equation with dispersion management, and as in the continuous equation one of the main technological motivations behind the diffraction management idea is to produce and control low power coherent structures.

Problems of interest for (1.1) include the existence and stability of localized periodic or quasi-periodic solutions (found numerically in [3]), and multi-bump solutions and their interactions. A technical complication is the fact that (1.1) is a non-autonomous system. To obtain some possible insight into the above problems we will approximate (1.1) by the autonomous (but "nonlocal") system

$$\partial_t a_n = \mathrm{i}\delta(\Delta a)_n - 2\mathrm{i}\gamma(\bar{g}_L(a))_n, \quad (\bar{g}_L(a))_n = \frac{1}{T}\int_0^T (L_\tau^{-1} g(L_\tau a))_n \,\mathrm{d}\tau, \quad (L_t u)_n = (\mathrm{e}^{\mathrm{i}\int_0^t \tilde{d}(\sigma)\,\mathrm{d}\sigma \Delta} u)_n, \tag{1.2}$$

with $(\Delta a)_n = a_{n+1} - 2a_n + a_{n-1}$. The relation between the solutions of (1.1) and (1.2) will be discussed in Section 2. We will see that approximating (1.1) by (1.2) involves an averaging argument that is similar to the one used in the continuous NLS with dispersion management (see [10,1]).

We will show that for $|\delta|/|\gamma|$ sufficiently small and $\delta\gamma < 0$, the averaged Eq. (1.2) has localized breather-type solutions. The result also holds for the analogue of (1.2) in integer lattices of arbitrary dimension. The proof we give is variational and rests on the interpretation of the breather solutions as constrained minima (for $\gamma < 0$) of the Hamiltonian of (1.2). The constraint is the l_2 norm. The argument is a discrete version of the concentration-compactness principle (see e.g. [12,5]). A similar strategy was used in [19] for the local discrete NLS equation, and also for the nonlocal nonlinearity in the continuous version of (1.2) ([20], see also [14]). In the nonlocal discrete case the main effort goes into controlling the value of the Hamiltonian in the overlap between distant bumps and thus assuring that minimizing sequences cannot split. The corresponding estimates rely on the fact that the nonlinear interaction between sufficiently distant sites decays rapidly. Once the minimizer is shown to converge the problem is simpler that its continuous counterpart, and the dimension of the lattice only affects the condition on the size of $|\delta|/|\gamma|$. In the limit of constant D the operator L_t becomes the identity and we recover the local discrete NLS. In that case we also expect that for $|\delta|/|\gamma|$ sufficiently small we can obtain the existence of breathers by the continuation arguments of [3,13]. It seems that an advantage of the variational approach here is that it allows us to handle more general nonlinear interactions between neighboring sites.

The paper is organized as follows. In Section 2 we set the notation, outline the steps leading from the original model (1.1) to the averaged Eq. (1.2) and state and discuss the results. In Section 3 we present an outline and the first part the proof of the existence of breathers, examining convergent and vanishing minimizing sequences. In Section 4 we show that a minimizing sequence cannot split, and must therefore have a subsequence converging to the infimum (up to translations).

2. Discrete NLS equations and breathers

We consider the lattice of integers $\mathbf{Z}$, and complex valued functions $u(t)$ on $\mathbf{Z}$ that evolve according to the non-autonomous system

$$\partial_t u = \mathrm{i}D(t)\Delta u - 2\mathrm{i}\gamma g(u), \tag{2.1}$$

where

$$(\Delta u)_j = u_{j+1} - 2u_j + u_{j-1}, \quad g_j(u) = |u_j|^2 u_j, \tag{2.2}$$

and f_j is the value of $f : \mathbf{Z} \to \mathbf{C}$ at the site j. The function D is real valued and γ is a real constant. We further assume that D is T-periodic, and we decompose it as

$$D(t) = \delta + \tilde{d}(t), \quad \text{with} \quad \delta = \frac{1}{T}\int_0^T D(\tau)\,\mathrm{d}\tau \tag{2.3}$$

the average. Physically, the "time" t in (2.1) is the distance along the waveguides, while the "spatial variable" j of (2.1) is the transverse direction (more precisely the index of the waveguide, see e.g. [8]). Model (2.1) was proposed by [3]. Also, u_n is the complex amplitude of (any) one of the components of the electric field at the site n. The initial condition $u(t_0)$ for (2.1) is the emitted light and to simplify the notation we may consider initial data at $t_0 = 0$, shifting $d(t)$ if necessary.

In this work we study an approximate autonomous system derived by an averaging argument. To obtain the averaged equation we rewrite (2.1) using the variable $a(t)$ defined by

$$a(t) = L_t^{-1}u(t), \quad \text{with} \quad L_t = \mathrm{e}^{\mathrm{i}\tilde{\Lambda}(t)\Delta}, \quad \text{and} \quad \tilde{\Lambda}(t) = \int_0^t \tilde{d}(\tau)\,\mathrm{d}\tau. \tag{2.4}$$

By (2.1), the evolution equation for $a(t)$ is then

$$\partial_t a = \mathrm{i}\delta\Delta a - 2\mathrm{i}\gamma L_t^{-1} g(L_t a), \tag{2.5}$$

with the initial condition $a(0) = u(0)$. Note that $L_t^{-1} = \mathrm{e}^{-\mathrm{i}\tilde{\Lambda}(t)\Delta}$ and that the right hand side of (2.5) is T-periodic. Eq. (2.5) is then replaced by the averaged equation

$$\partial_t a = i\delta\Delta a - 2\mathrm{i}\gamma \bar{g}_L(a), \quad \text{with} \quad \bar{g}_L(a) = \frac{1}{T}\int_0^T L_\tau^{-1} g(L_\tau a)\,\mathrm{d}\tau. \tag{2.6}$$

The distance between the solutions of (2.5) and (2.6) can be estimated using averaging arguments and we expect that for $|\gamma|$, $|\delta|$ of $\mathrm{O}(\epsilon)$, $|\epsilon|$ small, and $\omega = 2\pi/T > 0$ of at least O(1), solutions of the two systems that correspond the same O(1) size initial conditions should stay $\mathrm{O}(\epsilon)$ close for a time of $\mathrm{O}(\epsilon^{-1})$ (Distances are measured in the l_2 norm defined below.) The steps are the ones leading to the averaging theorem in [16], modified for flows in l_2, i.e. the requisite Lipshitz continuity properties of the right hand sides are the same. Alternative approaches are e.g. in [17,20,14].

The above definitions and notation can be extended to the case where u, and a are complex valued functions on $\mathbf{Z}^d$, $d \geq 1$. In particular, let $D_{k,+}$ ($D_{k,-}$) denote the forward (backward) first order difference operators along the kth direction, and set $\Delta_k = D_{k,+}D_{k,-}$ and $\Delta = \sum_{k=1}^{d} \Delta_k$. Clearly, for $d = 1$, the new definition of Δ agrees with the one in (2.3).

Remark 2.0.1. The diffraction management idea is also meaningful for $d = 2$, although the particular "isotropic" diffraction management $D(t)\Delta$ here may be too restrictive. In the special case of the local discrete NLS, the $d = 2$ problem is already of physical interest (see e.g. [2]).

We look for localized solutions of (2.6) that have the form $a(t) = \mathrm{e}^{-\mathrm{i}\lambda t}A$, for some $\lambda \in \mathbf{R}$. We refer to these periodic orbits as breathers. By (2.6), we must then find $A : \mathbf{Z}^d \to \mathbf{C}$ that satisfies

$$\lambda A = -\delta\Delta A + 2\gamma \bar{g}_L(A) \tag{2.7}$$

and decays at infinity.

The strategy will be to use the variational structure of the equation. Consider the standard hermitian inner product $\langle u, v\rangle_h = \sum_{n\in\mathbf{Z}^d} u_n v_n^*$ on pairs of complex valued functions u, v on $\mathbf{Z}^d$. Also let X be $l_2 = l_2(\mathbf{Z}^d, \mathbf{C})$, the real Hilbert space of square summable complex valued functions on $\mathbf{Z}^d$ with the inner product $\langle\cdot,\cdot\rangle$ given by $\langle u, v\rangle = Re\langle u, u\rangle_h$, u, $v \in X$ (i.e. we identify $\mathbf{C}$ with $\mathbf{R}^2$ and use the complex notation for convenience). The norm $\|\cdot\|$ of $u \in X$ is

$\|u\| = \|u\|_{l_2} = (Re\langle u, u\rangle)^{1/2}$, $u \in X$. Similarly, let $\|u\|_{l_4}^4 = \sum_{n\in\mathbf{Z}^d} |u_n|^4$ ($< \infty$ if $u \in X$). We define the functional $\bar{H}$ on X by

$$\bar{H}(v) = \delta \sum_{k=1}^{d} \|D_{k,+}v\|_{l_2}^2 + \gamma \frac{1}{T}\int_0^T \|L_\tau v\|_{l_4}^4 \, \mathrm{d}\tau, \tag{2.8}$$

and look for critical points of $\bar{H}$ in the set $X_c = \{v \in X : \|v\| = c\}$. We will show the following:

Theorem 2.1. *Consider the functional $\bar{H}$ above with $\delta \geq 0$ and $\gamma < 0$. Assume that (i) D is piecewise continuous and bounded on $[0, T]$, and (ii) that given $c > 0$, there exists $\mu > 0$ for which $\delta/|\gamma| < \mu$. Then the infimum of $\bar{H}$ on X_c is attained.*

Remark 2.1.1. The constant μ depends on the dimension d, the function $\tilde{d}$, and c. We will see that that the sign, and relative size assumptions on δ, γ imply that, given any $c > 0$, the infimum of $\bar{H}$ on X_c is strictly negative. This property is crucial for the proof.

Remark 2.1.2. Condition (i) on D here means that either D is continuous in $[0, T]$, or there exists a finite number of points $0 < \tau_1 < \cdots < \tau_n < T$ so that D is continuous in each of the closed intervals $[0, \tau_1], \ldots, [\tau_n, T]$. Thus D can at worst have (finite) jump discontinuities at a finite number of points of $(0, T)$. In the physical examples D is piecewise constant.

The Euler–Lagrange equation corresponding to the above variational problem is precisely (2.7). Note in particular that $L_\tau^{-1} = L_\tau^\dagger$, the adjoint of L_τ in X. We also see that that $\bar{H}$ is C^1 in X. Standard arguments therefore yield:

Proposition 2.2. *Let $\tilde{a} \in X_c$ satisfy $\bar{H}(\tilde{a}) = \inf_{v\in X_c} \bar{H}(v)$. Then there exists a $\lambda \in \mathbf{R}$ for which $A = \tilde{a}$ satisfies (2.7).*

The proof of Theorem (2.1) will be given in the next two sections; we here make some basic observations. First, note that $\bar{H}$ and the l_2 norm are invariant under (i) translations in $\mathbf{Z}^d$, and (ii) the circle action $v \to \mathrm{e}^{\mathrm{i}\phi}v$, $\phi \in \mathbf{R}$. Thus, if $\tilde{a}$ is a minimizer of $\bar{H}$ on X_c, the integer translates of $\tilde{a}$, and points on the circle $\mathrm{e}^{\mathrm{i}\phi}\tilde{a}$, $\phi \in \mathbf{R}$ are also minimizers.

A consequence of the variational characterization of breather solutions $a(t) = \mathrm{e}^{-\mathrm{i}\lambda t}A$ of (2.5) with A a minimizer of $\bar{H}$ on X_c is that they also satisfy a stability property. To make this precise consider the corresponding breather periodic orbit (i.e. invariant circle) $\gamma_c(A)$ (i.e. seen a subset of X_c). Note that the initial value problem for (2.6) has global solutions $a(t)$ in $C^0(\mathbf{R}, l_2)$. This follows from the Lipschitz continuity of the right hand side of (2.6) and the conservation of the l_2 norm. We can therefore consider the evolution of the distance between a breather invariant circle $\gamma_c(A)$ and a solution $a(t)$ that starts near $\gamma_c(A)$. We have the following.

Proposition 2.3. *Let $\Gamma_c(A)$ be an isolated set of breather invariant circles corresponding to a minimizer A of $\bar{H}$ on X_c as above. Also let $a(t)$ be a solution of (2.6) with initial condition a_0. Then given $\epsilon > 0$ there exists $\epsilon_0 > 0$ for which $\sup_{x\in\Gamma_c(A)} \|a_0 - x\|_{l_2} < \epsilon_0$ implies $\sup_{x\in\Gamma_c(A)} \|a(t) - x\|_{l_2} < \epsilon$, $\forall t \in \mathbf{R}$.*

The geometrical ideas behind the proof are simple and we will not give the details here. Note however that the proposition does not imply that individual invariant circles corresponding to minimizers are orbitally stable. For this it would be sufficient to show that such invariant circles are isolated. The numerical evidence in [3] suggests that this is probably the case. We note that in the special case of the local NLS with $|\delta|/|\gamma|$ small where the periodic orbits can be approximated explicitly we expect that we can apply ideas from perturbation theory to obtain Nekhoroshev stability results (see [4]). A further question is the existence of other critical points of $\bar{H}$ on X_c, e.g. saddles. Some interesting suggestions for the discrete NLS are in [11].

We also remark that the proof of Theorem (2.1) in the higher dimensional case is not significantly different from the one-dimensional one (it is only more involved at a few points). On the other hand, the result is interesting from the point of view of possible "higher-dimensional time" extensions of dynamical concepts relevant to one-dimensional lattices (e.g. homoclinic orbits, multi-bump solutions, shadowing, etc.).

Remark 2.3.1. We have not considered here the problem of finding breather solutions when $\delta/\gamma < 0$ is fixed and c varies, e.g. setting $\gamma = -1$ and fixing $\delta > 0$. The existence of breathers of arbitrary l_2 norm c is then subtle since the infimum of $\bar{H}$ on X_c may fail to be strictly negative. In the case of the discrete NLS the question is settled in [19], where it was shown that if $d \geq 2$ then the infimum of $\bar{H}$ on X_c is negative only if c is above a certain threshold.

Remark 2.3.2. Theorem 2.1 can be equivalently restated for the case where $\delta \leq 0$, $\gamma > 0$. Then, given $c > 0$, and assuming that $|\delta|/|\gamma| < \mu$, the functional $\bar{H}$ attains its supremum over X_c. The constant μ is exactly the one in Theorem 2.1 and the alternative statement follows immediately by applying Theorem 2.1 to $-\bar{H}$ and using the conditions of the theorem on $-\delta, -\gamma$. In the case where $\delta, \gamma > 0$ (< 0) we see in the next section (Remark 3.2.2) that the infimum (supremum) of $\bar{H}$ on X_c can not be attained.

3. Minimizing sequences: vanishing and convergent cases

We now outline the steps of the proof of Theorem (2.1). We start with a minimizing sequence $\{a_n\} \in X_c$ (See the remark on notation after the present paragraph.) In Appendix A we show that $\bar{H}$ is continuous in $X = l_2$ and it is therefore sufficient to find a subsequence of $\{a_n\}$ that converges strongly in l_2. To see that this is the case we distinguish three possibilities for the minimizing sequence. First we observe in Proposition 3.2 that the supremum of $|a_n(j)|$ over $j \in \mathbf{Z}^d$ is bounded away from zero. Translating the a_n so that the maxima of $|a_n|$ are at the origin, we define a number $\Gamma > 0$ that indicates the portion of the l_2 norm of a_n that stays concentrated around the origin as $n \to \infty$. In the case $\Gamma = c^2$ we show in Proposition 3.3 that $\{a_n\}$ has a convergent subsequence. It therefore remains to eliminate the case $\Gamma < c^2$, and we do this in Section 4. Intuitively, $\Gamma < c^2$ corresponds to a_n splitting into pieces that carry away to infinity nontrivial portions of the l_2 norm of a_n as $n \to \infty$.

Notation: Let U be a subset of $\mathbf{R}^d$, and a function $f : U \to \mathbf{C}$. In this and the next section $f(j)$ will denote the value of f at the point j. Subscripts will from now on denote indices in a sequence, e.g. as in a_n above.

As a preliminary step we give an explicit expression and a basic estimate for the operator $L_t = \mathrm{e}^{\mathrm{i}\tilde{\Lambda}(t)\Delta}$ of (2.4).

Lemma 3.1. *Let $L_t : X \to X$ be the operator defined in (2.4), and let $v \in X$. Then*

$$(L_t v)(n) = \sum_{m \in \mathbf{Z}^d} G_t(n-m)v(m), \quad n \in \mathbf{Z}^d, \quad with \tag{3.1}$$

$$G_t(y) = \mathrm{e}^{-2\mathrm{i}\tilde{\Lambda}(t)} \prod_{j=1}^{d} (\mathrm{i})^{|y_j|} \mathcal{J}_{|y_j|}(2\tilde{\Lambda}(t)), \quad y = [y_1, \ldots, y_d] \in \mathbf{Z}^d, \tag{3.2}$$

and $\mathcal{J}_p$ the Bessel function of integer order $p \geq 0$. (The dependence of G_t on the function $\tilde{\Lambda}$ is suppressed from the notation.) Also,

$$|G_t(y)| \leq \prod_{j=1}^{d} \frac{|\tilde{\Lambda}(t)|^{|y_j|}}{|y_j|!}, \quad \forall y = [y_1, \ldots, y_d] \in \mathbf{Z}^d. \tag{3.3}$$

Proof. Let $I_\pi^d = [-\pi, \pi]^d$ and define the semi discrete Fourier transform $\mathcal{F}: X \to L_2(I_\pi^d, \mathbf{C})$ by

$$\hat{u}(k) = \sum_{n\in\mathbf{Z}^d} \mathrm{e}^{-\mathrm{i}\langle k,n\rangle} u(n), \quad k \in I_\pi^d, \tag{3.4}$$

i.e. we use the notation $\hat{u} = \mathcal{F}u$. $\mathcal{F}$ is an isometry with inverse $\mathcal{F}^{-1}$ defined by

$$u(n) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}\langle k,n\rangle} \hat{u}(k)\, \mathrm{d}k_1 \cdots \mathrm{d}k_d, \quad n \in \mathbf{Z}^d. \tag{3.5}$$

We compute

$$(\hat{\Delta u})(k) = -4 \left(\sum_{j=1}^{d} \sin^2 \frac{k_j}{2} \right) \hat{u}(k), \quad k = [k_1, \ldots, k_d], \tag{3.6}$$

hence

$$(L_t u)(n) = \sum_{m\in\mathbf{Z}^d} G_t(n-m)u(m), \quad \text{with} \tag{3.7}$$

$$G_t(y) = \frac{1}{(2\pi)^d} \int_{I_\pi^d} \prod_{j=1}^{d} (\mathrm{e}^{-\mathrm{i}k_j y_j}\, \mathrm{e}^{-\mathrm{i}4\tilde{\Lambda}(t)\sin^2(k_j/2)})\, \mathrm{d}k_1 \cdots \mathrm{d}k_d, \quad y = [y_1, \ldots, y_d]. \tag{3.8}$$

But

$$\begin{aligned}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k_j y_j}\, \mathrm{e}^{-\mathrm{i}4\tilde{\Lambda}(t)\, \sin^2(k_j/2)}\, \mathrm{d}k_j &= \mathrm{e}^{-\mathrm{i}2\tilde{\Lambda}(t)} \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k_j y_j}\, \mathrm{e}^{\mathrm{i}2\tilde{\Lambda}(t)\, \cos k_j}\, \mathrm{d}k_j \\ &= \frac{\mathrm{e}^{-\mathrm{i}2\tilde{\Lambda}(t)}}{\pi} \int_0^{\pi} \mathrm{e}^{\mathrm{i}2\tilde{\Lambda}(t)\, \cos k_j} \cos |y_j| k_j\, \mathrm{d}k_j = (\mathrm{i})^{|y_j|}\, \mathrm{e}^{-\mathrm{i}2\tilde{\Lambda}(t)} \mathcal{J}_{|y_j|}(2\tilde{\Lambda}(t))\end{aligned} \tag{3.9}$$

by the definition of $\mathcal{J}_p$, the Bessel function of integer order p. By (3.8), (3.9) and Fubini we immediately obtain (3.2). The bound on $|G_t(y)|$ follows from the definition of $G_t(y)$ and the basic inequality

$$|\mathcal{J}_p(z)| \le \frac{|z/2|^p}{p!}, \quad \forall z \in \mathbf{C}, \quad p = 0, 1, 2, \ldots,$$

for the Bessel functions (see e.g. [18]). □

Proposition 3.2. *Consider $\bar{H}$ as in (2.7) with $\delta \ge 0$, $\gamma < 0$. Let $\{u_n\}$ be sequence in X_c possessing a subsequence $\{v_n\}$ satisfying $\lim_{n\to\infty} \sup_{j\in\mathbf{Z}^d} |v_n| = 0$. Then there exists $\delta_0 > 0$ for which $\delta < \delta_0$ implies that $\{u_n\}$ cannot be minimizing sequence for $\bar{H}$* in X_c.

Proof. Let $v(0) = c$, $v(j) = 0$ for $j \ne 0$. We have $v \in X_c$. Using (3.9), and the facts that $\tilde{\Lambda}(0) = 0$, $|\mathcal{J}_0(0)| = 1$, we see that if $\gamma < 0$, then

$$\gamma \frac{1}{T} \int_0^T \|L_t v\|_{l_4}^4\, \mathrm{d}t \le \gamma \frac{c^4}{T} \int_0^T |\mathcal{J}_0(\tilde{\Lambda}(t))|^4\, \mathrm{d}t < 0. \tag{3.10}$$

Then $\bar{H}(v) < 0$ for $\delta > 0$ sufficiently small and therefore the infimum of $\bar{H}$ in X_c is strictly negative. On the other hand, consider the subsequence $\{v_n\}$ whose maximum over $\mathbf{Z}^d$ goes to zero as above. We observe that $\|L_t v_n\|_{l_4}^4 \to 0$ as $n \to \infty$, uniformly in $t \in [0, T]$. This is because

$$\|L_t v_n\|_{l_4}^4 \leq \sup_{j\in\mathbf{Z}^d} |(L_t v_n)(j)|^2 \|L_t v_n\|_{l_2}^2 \leq c^2 (\sup_{j\in\mathbf{Z}^d} |(L_t v_n)(j)|)^2 \tag{3.11}$$

since L_t preserves the l_2 norm, $\forall t$. From (3.3), the kernel G_t is absolutely summable $\forall t$, and by (3.11) we have

$$\frac{1}{T}\int_0^T \|L_t v_n\|_{l_4}^4 \, \mathrm{d}t \leq c^2 K(\tilde{\lambda})(\sup_{j\in\mathbf{Z}^d} |v_n(j)|)^2 \to 0 \quad \text{as} \quad n \to \infty, \tag{3.12}$$

with the constant $K(\tilde{\Lambda})$ depending on $\sup_{t\in[0,T]} |\tilde{\Lambda}(t)|$. Thus the quartic part of $\bar{H}$ vanishes and since the quadratic part is nonnegative by $\delta \geq 0$ we have that $\inf_{u\in X_c} \bar{H}(u) \geq 0$, a contradiction. □

Remark 3.2.1. The above proof makes it clear that the constant δ_0 depends on c, the dimension d, and the function $\tilde{\delta}$. For $\delta = 0$ ($\gamma < 0$) the fact that the infimum of $\bar{H}$ on X_c is strictly negative is automatic. In the case where $\frac{\delta}{\gamma}$ is fixed however, finding a trial function $v \in X_c$ with $\bar{H}(v) < 0$ could be difficult or impossible (i.e. if there is a threshold), and one would require an approach similar to the one in [19].

Remark 3.2.2. The arguments of Proposition 3.2 also show that if $\delta\gamma > 0$, $\delta > 0$ then the infimum of $\bar{H}$ on X_c cannot be attained. To see this define the sequence $\{a_n\}$ by $a_n(k) = (2n+1)^{-(d/2)}c$, if $|k_1| \leq n, \ldots, |k_d| \leq n$, and $a_n(k) = 0$ otherwise ($k = [k_1, \ldots, k_d] \in \mathbf{Z}^d$). We have that $\|a_n\| = c$, for all n, moreover $\sup_{j\in\mathbf{Z}^d} |a_n(j)| \to 0$ as $n \to \infty$. Computing the quadratic part of $\bar{H}$ and using the argument in (3.12) we then see that $\bar{H}(a_n) \to 0$ as $n \to \infty$. Similarly, if $\delta\gamma > 0$, $\delta < 0$ the supremum of $\bar{H}$ on X_c is not attained.

Corollary 3.3. *Let $\delta < \delta_0$, with δ_0 as in Proposition 3.2. Then if $\{u_n\} \subset X_c$ is a minimizing sequence for $\bar{H}$ in X_c, we have that for every subsequence of $\{v_n\}$ of $\{u_n\}$ there exists $a > 0$ for which $\sup_{j\in\mathbf{Z}^d} |v_n(j)| \geq a$, $\forall n$.*

Let $\{a_n\}$ be a subsequence of a minimizing sequence as in Corollary 3.3. For each index n we can translate a_n so that the supremum of $a_n(j)$ over $j \in \mathbf{Z}^d$ is attained at $j = 0$. We refer to this new sequence by $\{a_n\}$ (by a slight abuse of notation). We therefore have $|a_n(0)| \geq c$, $\forall n$. We will construct a notion of "asymptotic mass in a bounded region" to characterize the behavior of $\{a_n\}$. We first establish some basic notation.

Let $B_r = \{k \in [k_1, \ldots, k_d] \in \mathbf{Z}^d : |k_j| \leq r,\ j = 1, \ldots, d\}$, $m(a_n, r) = \sum_{k\in B_r} |a_n(k)|^2$. Let $r = 1$. We can extract from $\{a_n\}$ a subsequence $\{a_n^1\}$ for which $m(a_n^1, r)$ converges. We proceed inductively, increasing r: given $\{a_n^r\}$ with $m(a_n^r, r)$ convergent we extract a subsequence $\{a_n^{r+1}\}$ of $\{a_n^r\}$ for which $m(a_n^r, r)$ is convergent. We thus obtain a sequence of subsequences $\{a_n^1\} \supset \cdots \supset \{a_n^r\} \supset \cdots$ of the minimizing sequence $\{a_n\} \in X_c$. We also let $m_r = \lim_{n\to\infty} m(a_n^r, r)$, $r > 0$. The sequence $\{m_r\}$ is bounded above by c^2, and it easy to check that is also increasing. We therefore define Γ by

$$\Gamma = \lim_{r\to\infty} m_r. \tag{3.13}$$

The definition and the above construction implies that $0 < \Gamma \leq c^2$. In Proposition 3.5 we see that if $\Gamma = c^2$ then the original minimizing sequence $\{a_n\}$ has a convergent subsequence. On the other hand, in the next section we show that $\Gamma < c^2$ implies that $\{a_n\}$ can not be a minimizing sequence.

Remark 3.3.1. Note that the definition of Γ depends on the subsequences $\{a_n^r\}$ we choose as we increase r. It will become clear however that the conclusions following from the value of Γ are independent of this choice.

Proposition 3.4. *Consider* $\{a_n\}$ in X_c, *and* Γ *as above. Suppose that* $\Gamma = c^2$. *Then, there exists a subsequence* $\{\tilde{a}_n\}$ *of* $\{a_n\}$ *satisfying that for every* $\epsilon > 0$, *there exist* $R, N > 0$ *for which* $n > N$ *implies*

$$c^2 - \sum_{k \in B_R} |\tilde{a}_n(k)|^2 < \epsilon. \tag{3.14}$$

Proof. Consider the subsequences $\{a_n^1\} \supset \cdots \supset \{a_n^r\} \supset \cdots$ of $\{a_n\}$ and let $d_n = a_n^n$, $n \geq 1$ define the diagonal sequence $\{d_n\}$. Clearly, $\{d_n\}$ is a subsequence of $\{a_n\}$. By the definition of Γ, and the assumption $\Gamma = c^2$, given any $\epsilon > 0$ we can find R for which

$$|c^2 - m_R| < \frac{\epsilon}{2}. \tag{3.15}$$

By the definition of d_n we can also find $N > 0$ for which $n > N$ implies

$$\left| \sum_{k \in B_R} |d_n(k)|^2 - m_R \right| < \frac{\epsilon}{2}. \tag{3.16}$$

Combining (3.15), (3.16) we have (3.14) with $\tilde{a}_n = d_n$. □

Proposition 3.5. *Consider* $\{a_n\}$ in X_c, Γ *as above. Suppose that* $\Gamma = c^2$. *Then* $\{a_n\}$ *has a subsequence that converges to an element* $\tilde{a} \in X_c$.

Proof. We consider the diagonal $\{d_n\}$ defined in Proposition 3.4, and

$$\|d_{n_1} - d_{n_2}\|_{l_2}^2 = \sum_{k \in B_R} |d_{n_1}(k) - d_{n_2}(k)|^2 + \sum_{k \in \mathbf{Z}^d \setminus B_R} |d_{n_1}(k) - d_{n_2}(k)|^2. \tag{3.17}$$

Let $\epsilon > 0$. By Proposition 3.4 we can choose $R, N' > 0$ so that the second term can be bounded as

$$\sum_{k \in \mathbf{Z}^d \setminus B_R} |d_{n_1}(k) - d_{n_2}(k)|^2 \leq \sum_{k \in \mathbf{Z}^d \setminus B_R} |d_{n_1}(k)|^2 + \sum_{k \in \mathbf{Z}^d \setminus B_R} |d_{n_2}(k)|^2 \tag{3.18}$$

$$= (c^2 - \sum_{k \in B_R} |\tilde{a}_{n_1}(k)|^2) + (c^2 - \sum_{k \in B_R} |\tilde{a}_{n_1}(k)|^2) < 2\epsilon, \tag{3.19}$$

for $n_1, n_2 > N'$. Since the restriction of d_n to any B_r is convergent, we can chose $N'' > 0$ for which $n_1, n_2 > N''$ implies.

$$\sum_{k \in B_R} |d_{n_1}(k) - d_{n_2}(k)|^2 < \epsilon. \tag{3.20}$$

Combining (3.17)–(3.20) we see that $\{d_n\}$ is Cauchy and hence convergent in l_2. Since $\|d_n\|_{l_2} = c$, $\forall n$, the limit belongs to X_c. □

4. Minimizing sequences: splitting case

We now consider the case where $\Gamma \neq c^2$, and show that the sequence $\{a_n\}$ can not be a minimizing sequence. The sequences $\{a_n\}$, $\{a_n^r\}$, and Γ are as defined after Corollary 3.3 and we also use the diagonal sequence $\{d_n\}$ defined in the proof of Proposition 3.4. $\overline{H}$ is as in Theorem 2.1 and we also let $\Lambda = \max_{t \in [0,T]} \tilde{\Lambda}(t)$. By assumption (i) of Theorem 2.1, Λ is finite.

Proposition 4.1. *Consider the sequences* $\{a_n\}$, $\{a_n^r\}$, $r \geq 1$, *and the diagonal* $\{d_n\}$ *as above, with* $\Gamma < c^2$. *Also let* $p = p(R, R_1) = (R_1 - R - 2)/2$. *Then, for every* $\epsilon > 0$ *there exist* $R_1 > R > 0$ *satisfying*

$$\frac{\Lambda^p p^{d+1}}{p!} < 1, \quad 3R + 1 \leq R_1, \quad \frac{1}{R_1 - R} < \epsilon, \tag{4.1}$$

and $N > 0$ *for which* $n > N$ *implies*

$$\left| \Gamma - \sum_{k \in B_R} |d_n(k)|^2 \right| < 2\epsilon, \tag{4.2}$$

$$\left| \Gamma - \sum_{k \in B_{R_1}} |d_n(k)|^2 \right| < \epsilon, \tag{4.3}$$

$$\sum_{k \in B_{R_1} \setminus B_R} |d_n(k)|^2 < 5\epsilon. \tag{4.4}$$

Proof. Let $\epsilon > 0$. By the definitions of m_r, Γ (after Corollary 3.3) we can choose $R > 0$ for which $|m_R - \Gamma| < \epsilon$. Also, if $n \geq R$ then $\{d_n\}$ is a subsequence of $\{a_n^R\}$ which converges to m_R, i.e. we can choose $N_2 \geq R$ for which $n > N_2$ implies

$$\left| \sum_{k \in B_R} |d_n(k)|^2 - m_R \right| < \epsilon. \tag{4.5}$$

We thus obtain (4.2). Also, choose $R_1 > R$ satisfying (4.1) and $|m_R - \Gamma| < \epsilon/2$. As before we can find $N_3 > R_1$ for which $n > N_3$ implies

$$\left| \sum_{k \in B_{R_1}} |d_n(k)|^2 - m_{R_1} \right| < \frac{\epsilon}{2}. \tag{4.6}$$

We thus obtain (4.3). By (4.3), (4.2) we have

$$|a - b| < \epsilon, \quad |2a| < 2\epsilon, \quad b > 0, \text{ with} \tag{4.7}$$

$$a = \Gamma - \sum_{k \in B_R} |d_n(k)|^2, \quad b = \sum_{k \in B_{R_1} \setminus B_R} |d_n(k)|^2. \tag{4.8}$$

We easily see that (4.5) implies $b < 5\epsilon$, i.e. (4.4). □

Lemma 4.2. *Consider the sequences* $\{a_n\}$, *and* $\{a_n^r\}$, $r \geq 1$ *above with* $\Gamma < c^2$ *and let* $\{d_n\}$ *be the diagonal as in Proposition 4.1. For every* $\epsilon > 0$, *there exist sequences* $\{v_n\}$, $\{w_n\}$ *of complex valued functions on* $\mathbf{Z}^d$ *and* $N > 0$ *for which* $n > N$ *implies*

$$v_n(k) + w_n(k) = d_n(k), \quad \forall k \in \mathbf{Z}^d, \tag{4.9}$$

$$\left| \Gamma - \|v_n\|_{l_2}^2 \right| < 7\epsilon, \quad \left| (c^2 - \Gamma) - \|w_n\|_{l_2}^2 \right| < 6\epsilon, \tag{4.10}$$

$$\bar{H}(d_n) = \bar{H}(v_n) + \bar{H}(w_n) + h_n, \text{ with} \quad |h_n| < k\sqrt{\epsilon}, \tag{4.11}$$

and k a constant depending on c, γ, δ *and* d.

Remark 4.2.1. The splitting of d_n into v_n, w_n depends on ϵ.

Lemma 4.2 contains the main computation and we will give its proof below. To conclude that $\{a_n\}$ with $\Gamma < c^2$ is not a minimizing sequence we will use (4.11) and the subadditivity of the functional $\bar{H}$, given by Lemma 4.3 below. In particular, define P_α by $P_\alpha = \inf_{\|u\|^2_{l_2}=\alpha} \bar{H}(u)$, with $\bar{H}$ as in (2.7). We have:

Lemma 4.3. *Consider $\bar{H}$ as in (2.7) with $\delta \geq 0$, $\gamma < 0$ and let $\alpha, \beta > 0$. Then*

$$P_{\alpha+\beta} < P_\alpha + P_\beta. \tag{4.12}$$

The subadditivity property for the quartic NLS on $\mathbf{R}^d$ is shown in [5,20]. The proof uses a scaling argument that applies to the present case almost verbatim.

Proposition 4.4. *Consider the sequences $\{a_n\}$, and $\{a_n^r\}$, $r \geq 1$ above with $\Gamma < c^2$ and let $\{d_n\}$ be the diagonal sequence as in Proposition 4.1. Then $\{a_n\}$ cannot be a minimizing sequence.*

Proof. Given $\epsilon > 0$, we can use (4.10) to approximate v_n, w_n by $\tilde{v}_n \in X_\Gamma$, $\tilde{w}_n \in X_{c^2-\Gamma}$ to within an l_2 error of $O(\epsilon)$. By the continuity of $\bar{H}$ in l_2, and (4.11) in Lemma 4.2, we can choose $N > 0$ so that $n > N$ implies

$$\bar{H}(d_n) + |\tilde{h}_n| \geq \bar{H}(\tilde{v}_n) + \bar{H}(\tilde{w}_n), \text{ with } \quad |\tilde{h}_n| < \tilde{k}\sqrt{\epsilon}, \tag{4.13}$$

and $\tilde{k}$ independent of ϵ ($\tilde{k}$ will depend on c and Γ), hence

$$\bar{H}(d_n) + |\tilde{h}_n| \geq P_\Gamma + P_{c^2-\Gamma}. \tag{4.14}$$

Assuming that $\{d_n\}$ is a subsequence of a minimizing sequence, taking the limit $n \to \infty$ in (4.14) and using the bound on $\tilde{h}_n$ in (4.13), we have

$$P_{c^2} + \tilde{k}\sqrt{\epsilon} \geq P_\Gamma + P_{c^2-\Gamma}, \quad \forall \epsilon > 0. \tag{4.15}$$

By Lemma 4.3 we therefore have $P_{c^2} \geq P_\Gamma + P_{c^2-\Gamma} > P_{c^2}$, a contradiction. □

In proving Lemma 4.2 we will use the estimates in Lemma 4.5 below. The proof is somewhat lengthy but follows from elementary arguments and we give a sketch at the end of the section.

Lemma 4.5. *Let $r_1 > r > 0$ with $2r + 1 \leq r_1 - r$, and $d \geq 1$. Also consider the kernel G_t defined by (3.7)–(3.9). Let $\Lambda = \max_{t\in[0,T]} |\tilde{\Lambda}(t)|$. Then*

$$\sup_{k\in B_r} \left(\sum_{m\in\mathbf{Z}^d\setminus B_{r_1}} |G(k-m)| \right)^2 \leq \sum_{j=1}^{d} K_j(d,\Lambda) \left(\frac{\Lambda^{(r_1-r)}}{(r_1-r)!} \right)^{2j}, \tag{4.16}$$

$$\sum_{k\in\mathbf{Z}^d\setminus B_{r_1}} \left(\sum_{m\in B_r} |G(k-m)| \right)^2 \leq (r_1-r)^d \sum_{j=1}^{d} \tilde{K}_j(d,\Lambda) \left(\frac{\Lambda^{(r_1-r)}}{(r_1-r)!} \right)^{2j}, \tag{4.17}$$

with constants K_j, $\tilde{K}_j$ that depend on d, Λ.

Proof of Lemma 4.2. We start by constructing the functions v_n, w_n. We consider $R_1 > R > 0$ as in Proposition 4.1, and functions $\phi, \phi_1 : \mathbf{R}^d \to [0, 1] \subset \mathbf{R}$ that satisfy

$$\phi(k) = \begin{cases} 1, & \text{if } k \in B_R \\ 0, & \text{if } k \in \mathbf{Z}^d \setminus B_{R_1} \end{cases}, \tag{4.18}$$

$$\phi_1(k) = \begin{cases} 1, & \text{if } k \in \mathbf{Z}^d \setminus B_{R_1} \\ 0, & \text{if } k \in B_R \end{cases} . \tag{4.19}$$

Further, we require that ϕ, ϕ_1 satisfy

$$\phi(k) + \phi_1(k) = 1, \quad \forall k \in \mathbf{R}^d. \tag{4.20}$$

Such functions exist by partition of unity arguments. To simplify notation we denote the functions in $\mathbf{R}^d$ and their restriction to $\mathbf{Z}^d$ by the same symbol. We define v_n, w_n by

$$v_n(k) = \phi(k)d_n(k), \quad w_n(k) = \phi_1(k)d_n(k), \quad k \in \mathbf{Z}^d. \tag{4.21}$$

Property (4.9) is immediate, while (4.10) follows easily from Proposition 4.1, and (4.18)–(4.20) and we omit the details. It remains to show (4.11).

Notation: We use the abbreviations $v = v_n$, $w = w_n$, and $v_t = L_t v$, $w_t = L_t w$. Also, $K(\alpha, \beta, \ldots)$ will denote a constant depending on $\alpha, \beta, \ldots$.

Using the abbreviated notation above, and $d_n = v + w$, we write

$$\bar{H}(d_n) = \bar{H}(v) + \bar{H}(w) + q(v, w) + Q_1(v, w) + Q_2(v, w) + Q_3(v, w), \text{ with} \tag{4.22}$$

$$q(v, w) = \delta \sum_{j=1}^{d} \sum_{k \in \mathbf{Z}^d} 2Re[(D_j v)(k)(D_j w)^*(k)], \tag{4.23}$$

$$Q_1(v, w) = \gamma \frac{1}{T} \int_0^T \sum_{k \in \mathbf{Z}^d} 2|v_t(k)|^2 |w_t(k)|^2 \, dt, \tag{4.24}$$

$$Q_2(v, w) = \gamma \frac{1}{T} \int_0^T \sum_{k \in \mathbf{Z}^d} 4Re\left(v_t(k) w_t^*(k)\right)^2 \, dt, \tag{4.25}$$

$$Q_3(v, w) = \gamma \frac{1}{T} \int_0^T \sum_{k \in \mathbf{Z}^d} 4(|v_t(k)|^2 + |w_t(k)|^2) Re\left(v_t(k) w_t^*(k)\right) \, dt. \tag{4.26}$$

We will estimate the "overlap" terms q, Q_1, Q_2, Q_3. We start with the quadratic term q of (4.24) and use the notation $(T_j u)(k) = u(k_1, \ldots, k_j + 1, \ldots, k_d)$, $j \in \{1, \ldots, d\}$. We have

$$q = 2\delta \sum_{j=1}^{d} \sum_{k \in \mathbf{Z}^d} Re[2v(k)w^*(k) - (T_j v)(k)w^*(k) - (v)(k)(T_j w)^*(k)]. \tag{4.27}$$

For the first term in the sum we observe that

$$\left| \sum_{j=1}^{d} \sum_{k \in \mathbf{Z}^d} Re[2v(k)w^*(k)] \right| \leq 2d \sum_{k \in B_{R_1} \setminus B_R} |v(k)w^*(k)| \tag{4.28}$$

$$\leq 2d \sum_{k \in B_{R_1} \setminus B_R} |d_n(k)|^2 \leq 10 d\epsilon. \tag{4.29}$$

The last two inequalities follow from the definition of ϕ, ϕ_1 in (4.18), (4.19), and (4.4) of Proposition 4.1. To bound the second and third terms in (4.27) we further restrict ϕ, ϕ_1 by requiring

$$\phi(k)=\begin{cases}1, & \text{if } k\in B_{R+1}\\ 0, & \text{if } k\in \mathbf{Z}^d\setminus B_{R_1-1}\end{cases}, \tag{4.30}$$

$$\phi_1(k)=\begin{cases}1, & \text{if } k\in \mathbf{Z}^d\setminus B_{R_1-1}\\ 0, & \text{if } k\in B_{R+1}\end{cases}, \tag{4.31}$$

and (4.20). Clearly, such functions exist and satisfy (4.18), (4.19). Using (4.30), (4.31) we bound the second term of (4.27) as

$$\left|\sum_{j=1}^{d}\sum_{k\in\mathbf{Z}^d} Re\left[(T_jv)(k)w^*(k)\right]\right| \le \sum_{j=1}^{d}\sum_{k\in B_{R_1}\setminus B_R}|(T_jv)(k)w^*(k)| \tag{4.32}$$

$$\le d\|v\|_{l_2}\left(\sum_{k\in B_{R_1}\setminus B_R}|w(k)|^2\right)^{1/2} \le \sqrt{(\Gamma+7\epsilon)5\epsilon}. \tag{4.33}$$

The last inequality follows from (4.10). The third term of (4.27) is estimated very similarly and we obtain the bound of (4.33). We therefore have that

$$|q(v,w)| < K(c,d,\delta)\sqrt{\epsilon}. \tag{4.34}$$

(Note that we can also obtain an O(ϵ) bound by a few extra steps, but this is not necessary.)

To estimate Q_1 of (4.24) we write

$$Q_1=\gamma\frac{1}{T}\int_0^T 2\,(Q_1(\mathrm{I})+Q_1(\mathrm{II})+Q_1(\mathrm{III}))\,\mathrm{d}t, \quad \text{with} \tag{4.35}$$

$$Q_1(\mathrm{I})=\sum_{k\in B_R}|v_t(k)|^2|w_t(k)|^2, \tag{4.36}$$

$$Q_1(\mathrm{II})=\sum_{k\in B_{R_1}\setminus B_R}|v_t(k)|^2|w_t(k)|^2, \tag{4.37}$$

$$Q_1(\mathrm{III})=\sum_{k\in \mathbf{Z}^d\setminus B_{R_1}}|v_t(k)|^2|w_t(k)|^2. \tag{4.38}$$

We further restrict ϕ, ϕ_1 by requiring

$$\phi(k)=\begin{cases}1, & \text{if } k\in B_{R+((R_1-R)/2)-1}\\ 0, & \text{if } k\in \mathbf{Z}^d\setminus B_{R+((R_1-R)/2)+1}\end{cases}, \tag{4.39}$$

$$\phi_1(k)=\begin{cases}1, & \text{if } k\in \mathbf{Z}^d\setminus B_{R+((R_1-R)/2)+1}\\ 0, & \text{if } k\in B_{R+((R_1-R)/2)-1}\end{cases}, \tag{4.40}$$

and (4.20). Clearly, such functions exist and satisfy the properties (4.18), (4.19), (4.30), (4.31) used above. To bound $Q_1(I)$ of (4.36) we first note that

$$|Q_1(I)| \le \left(\sup_{k\in B_R} |v_t(k)|^2\right) \sum_{k\in B_R} |w_t(k)|^2 \le c^2 \sum_{k\in B_R} |w_t(k)|^2. \tag{4.41}$$

We also have

$$\begin{aligned}\sum_{k\in B_R} |w_t(k)|^2 &= \sum_{k\in B_R} \left(\sum_{m\in \mathbf{Z}^d} G(k-m)w(m)\right)^2 \\ &= \sum_{k\in B_R} \left(\sum_{m\in \mathbf{Z}^d\setminus B_{R+((R_1-R)/2)}} G(k-m)w(m)\right)^2 \\ &\le \left(\sup_{m\in \mathbf{Z}^d} |w(m)|\right)^2 \sum_{k\in B_R} \left(\sum_{m\in \mathbf{Z}^d\setminus B_{R+((R_1-R)/2)}} |G(k-m)|\right)^2 \\ &\le c^2 \sum_{k\in B_R} \left(\sum_{m\in \mathbf{Z}^d\setminus B_{R+(R_1-R)/2}} |G(k-m)|\right)^2 \\ &\le c^2(2R+1)^d \sup_{k\in B_R} \left(\sum_{m\in \mathbf{Z}^d\setminus B_{R+(R_1-R)/2}} |G(k-m)|\right)^2. \end{aligned} \tag{4.42}$$

Applying (4.16) in Lemma 4.5 to $r = R$, $r_1 = R + (R_1 - R)/2$, and (4.1) in Proposition 4.1, (4.42) becomes

$$\sum_{k\in B_R} |w_t(k)|^2 \le K(c, d, \Lambda)(R_1 - R)^d \left(\frac{\Lambda^{(R_1-R)/2}}{((R_1-R)/2)!}\right)^2 < \frac{K(c, d, \Lambda)}{(R_1 - R)}. \tag{4.43}$$

By (4.41), (4.43) we therefore have

$$|Q_1(\mathrm{I})| < K(c, d, \Lambda)\epsilon. \tag{4.44}$$

To bound $Q_1(\mathrm{III})$ of (4.37) we use

$$|Q_1(\mathrm{III})| \le \left(\sup_{k\in \mathbf{Z}^d} |w_t(k)|^2\right) \sum_{k\in \mathbf{Z}^d\setminus B_{R_1}} |v_t(k)|^2 \le c^2 \sum_{k\in \mathbf{Z}^d\setminus B_{R_1}} |v_t(k)|^2. \tag{4.45}$$

We have

$$\sum_{k\in\mathbf{Z}^d\setminus B_R}|v_t(k)|^2 \le \sum_{k\in\mathbf{Z}^d\setminus B_{R_1}}\left(\sum_{m\in\mathbf{Z}^d}G(k-m)v(m)\right)^2$$
$$= \sum_{k\in\mathbf{Z}^d\setminus B_{R_1}}\left(\sum_{m\in B_{R+((R_1-R)/2)-1}}G(k-m)v(m)\right)^2$$
$$\le \left(\sup_{m\in\mathbf{Z}^d}|v(m)|^2\right)^2\sum_{k\in\mathbf{Z}^d\setminus B_{R_1}}\left(\sum_{m\in B_{R+(R_1-R)/2-1}}|G(k-m)|\right)^2$$
$$\le c^2\sum_{k\in\mathbf{Z}^d\setminus B_{R_1}}\left(\sum_{m\in B_{R+(R_1-R)/2-1}}|G(k-m)|\right)^2. \tag{4.46}$$

Applying (4.17) in Lemma 4.5 to $r = R + (R_1 - R)/2 - 1$, $r_1 = R_1$, and (4.1) in Proposition 4.1 we have

$$\sum_{k\in\mathbf{Z}^d\setminus B_{R_1}}\left(\sum_{m\in B_{R+(R_1-R)/2-1}}|G(k-m)|\right)^2 < \frac{K(c,d,\Lambda)}{(R_1-R)}. \tag{4.47}$$

By (4.45)–(4.47) we therefore have

$$|Q_1(\mathrm{III})| < K(c,d,\Lambda)\epsilon. \tag{4.48}$$

To bound $Q_1(\mathrm{II})$ of (4.36) we let

$$V = \{k\in\mathbf{Z}^d : k\in B_{R+(R_1-R)/2}\setminus B_R\}, \tag{4.49}$$

$$V_1 = \{k\in\mathbf{Z}^d : k\in B_{R_1-1}\setminus B_{R+(R_1-R)/2}\},$$

i.e. $V\cup V_1 = B_{R_1}\setminus B_R$. We can then write

$$Q_1(\mathrm{II}) = Q(\mathrm{IIA}) + Q(\mathrm{IIB}),\quad \text{with} \tag{4.50}$$

$$Q(\mathrm{IIA}) = \sum_{k\in V}|v_t|^2|w_t|^2,\quad Q(\mathrm{IIB}) = \sum_{k\in V_1}|v_t|^2|w_t|^2. \tag{4.51}$$

To estimate $Q_1(\mathrm{IIA})$ we use

$$Q_1(\mathrm{IIA}) = \sum_{k\in V}|v_t(k)|^2\left(\sum_{k\in\mathbf{Z}^d\setminus B_{R+((R_1-R)/2)-1}}G(k-m)w(m)\right)^2, \tag{4.52}$$

and further

$$|Q_1(\mathrm{IIA})| \le Q_1(\mathrm{IIA1}) + Q_1(\mathrm{IIA2}),\quad \text{with} \tag{4.53}$$

$$Q_1(\mathrm{IIA1}) = \sum_{k\in V} |v_t(k)|^2 \left(\sum_{m\in \mathbf{Z}^d \setminus B_{R_1-1}} G(k-m)w(m) \right)^2 . \tag{4.54}$$

$$Q_1(\mathrm{IIA2}) = \sum_{k\in V} |v_t(k)|^2 \left(\sum_{m\in B_{R_1-1} \setminus B_{R+(R_1-R)/2-1}} G(k-m)w(m) \right)^2 , \tag{4.55}$$

For $Q_1(\mathrm{IIA1})$ we have

$$\begin{aligned} |Q_1(\mathrm{IIA1})| &\le \left(\sup_{m\in \mathbf{Z}^d} |w(m)| \right)^2 \sum_{k\in V} |v_t(k)|^2 \left(\sum_{m\in \mathbf{Z}^d \setminus B_{R_1-1}} |G(k-m)| \right)^2 \\ &\le c^2 \left(\sup_{k\in V} \left(\sum_{m\in \mathbf{Z}^d \setminus B_{R_1-1}} |G(k-m)| \right)^2 \right) \left(\sum_{k\in V} |v_t(k)|^2 \right). \end{aligned} \tag{4.56}$$

Using (4.16) in Lemma 4.5 with $r = R + (R_1 - R)/2$, $r_1 = R_1 - 1$, and (4.1) in Proposition 4.1, (4.56) becomes

$$|Q_1(\mathrm{IIA1})| < K(c, d, \Lambda)\epsilon. \tag{4.57}$$

Also,

$$\begin{aligned} |Q_1(\mathrm{IIA2})| &\le \sum_{k\in V} |v_t(k)|^2 \left(\sum_{m\in V_1} |G(k-m)|^2 \sum_{m\in V_1} |w(m)|^2 \right) \\ &\le K\epsilon \sup_{k\in V} \left(\sum_{m\in V_1} |G(k-m)|^2 \right) \left(\sum_{k\in V} |v_t(k)|^2 \right) < K(c)\epsilon. \end{aligned} \tag{4.58}$$

We have here used the definition of w, and (4.4) in Proposition 4.1. For $Q_1(\mathrm{IIB})$ we have

$$|Q_1(\mathrm{IIB})| \le |Q_1(\mathrm{IIB1})| + |Q_1(\mathrm{IIB2})|, \quad \text{with} \tag{4.59}$$

$$Q_1(\mathrm{IIB1}) = \sum_{k\in V_1} |w_t(k)|^2 \left(\sum_{m\in B_R} G(k-m)v(m) \right)^2 , \tag{4.60}$$

$$Q_1(\mathrm{IIB2}) = \sum_{k\in V_1} |w_t(k)|^2 \left(\sum_{m\in B_{R+(R_1-R)/2} \setminus B_R} G(k-m)v(m) \right)^2 . \tag{4.61}$$

We have

$$|Q_1(\mathrm{IIB1})| \le \left(\sup_{m\in\mathbf{Z}^d}|v(m)|\right)^2 \sum_{k\in V_1}|w_t(k)|^2\left(\sum_{m\in B_R}|G(m-k)|\right)^2$$
$$\le c^2\left(\sup_{m\in\mathbf{Z}^d}|w_t(m)|\right)^2 \sum_{\mathbf{Z}^d\setminus B_{R+(R_1-R)/2}}\left(\sum_{m\in B_R}|G(m-k)|\right)^2. \tag{4.62}$$

Using (4.16) in Lemma 4.5 with $r_1 = R+(R_1-R)/2$, $r = R$, and (4.1) in Proposition 4.1, we therefore have

$$|Q_1(\mathrm{IIB1})| < K(c,d,\Lambda)\epsilon. \tag{4.63}$$

Also,

$$|Q_1(\mathrm{IIB2})| \le \sum_{k\in V_1}|w_t(k)|^2\left(\sum_{m\in V}|G(m-k)|\right)^2\left(\sum_{m\in V}|v(m)|\right)^2 < K(c,d,\Lambda)\epsilon, \tag{4.64}$$

where in the last inequality we have used (4.4) in Proposition 4.1. By (4.56)–(4.58), and (4.57), (4.63), (4.64) we therefore have

$$|Q_1(\mathrm{II})| < K(c,d,\Lambda)\epsilon, \tag{4.65}$$

Collecting (4.44), (4.48), (4.65) we have

$$|Q_1(v,w)| < K(c,d,\Lambda)\epsilon. \tag{4.66}$$

The estimate for Q_2 in (4.25) is immediate: $(Re(v_t(k)w_t(k)))^2 \le |v_t(k)|^2|w_t(k)|^2$ and (4.66) yield

$$|Q_2(v,w)| \le \tfrac{1}{2}|Q_1(v,w)| < K(c,d,\Lambda)\epsilon. \tag{4.67}$$

Also, by (4.26) and (4.66) we have

$$|Q_2(v,w)| \le \left[\sum_{k\in\mathbf{Z}^d}\left(|v_t(k)|^2+|w_t(k)|^2\right)^2\right]^{1/2}\left[\sum_{k\in\mathbf{Z}^d}(\mathrm{Re}(v_t(k)w_t(k)))^2\right]^{1/2} < K(c,d,\Lambda)\sqrt{\epsilon}. \tag{4.68}$$

Adding the estimates for q, Q_1, Q_2, and Q_3 we obtain (4.11) in the lemma.

Note that in the (local) discrete NLS case, where L_t is the identity, the Proof of Lemma 4.2 is significantly shorter. Since $v_t = v$, $w_t = w$, the definition of ϕ, ϕ_1 in (4.39)–(4.40), and Proposition 4.1 readily imply that the overlap terms $Q(\mathrm{I})$–$Q(\mathrm{III})$ in (4.35)–(4.38) are of $\mathrm{O}(\epsilon)$. □

Proof of Lemma 4.5. The details are somewhat laborious but elementary, and we will only give an outline, stressing the decomposition of the set $B^c_{r_1} = \mathbf{Z}^d\setminus B_{r_1}$, $d>1$ appearing in the sums of (4.16), (4.17) (The case $d=1$ is simple and we omit the details.) In particular, we let $d>1$ and write $B^c_{r_1} = B^c_{r_1}(d)\cup B^c_{r_1}(d-1)\cup\cdots\cup B^c_{r_1}(1)$, where $B^c_{r_1}(j)$ is the set of all multi-indices $[k_1,\ldots,k_d]\in B^c_{r_1}$ that have j components with absolute value greater than r_1. We first consider (4.16). We will break the inner sum into sums over the $B^c_{r_1}(j)$. Note that by using appropriate combinatorial constants that depend on j, d, we can replace estimates of sums over the $B^c_{r_1}(j)$ by estimates of sums over the sites with multi-indices satisfying $|k_1|, |k_2|, \ldots, |k_j| > r_1$. We see that

$$\sup_{k\in B_r}\left(\sum_{m\in B^c_{r_1}(j)}|G(k-m)|\right)^2 \le K(d,j,\Lambda)C^{d-j}D^j, \quad \text{with} \tag{4.69}$$

$$C = \sup_{|k_1| \le r_1} \left(\frac{\Lambda^{|k_1 - r_1|}}{|k_1 - r_1|!} + \frac{\Lambda^{|k_1 + r_1|}}{|k_1 + r_1|!} \right)^2, \tag{4.70}$$

$$D = \sup_{|k_d| \le r_1} \left(\sum_{|m_d| > r_1} \frac{\Lambda^{|k_d - m_d|}}{|k_d - m_d|!} \right)^2. \tag{4.71}$$

($K(\alpha, \beta, \ldots)$ is a constant depending on $\alpha, \beta, \ldots$. Also, the indices k_1, k_d, m_d in C, D are integer; the same notation is used in A, B below.) We can easily see that

$$C \le K(d) \left(\frac{\Lambda^{|r_1 - r|}}{|r_1 - r|!} \right)^2, \quad D \le K(\Lambda), \tag{4.72}$$

so that collecting (4.69)–(4.72) we obtain the bound in (4.16). For (4.17), we use the same decomposition of $B^c_{r_1}$ and similar combinatorial considerations in estimating sums over the $B^c_{r_1}(j)$. We see that

$$\sum_{k \in B^c_{r_1}(j)} \left(\sum_{m \in B_r} |G(k - m)| \right)^2 \le K(d, j) A^j B^{d-j}, \quad \text{with} \tag{4.73}$$

$$A = \sum_{|k_1| > r_1} \left(\sum_{|m_1| > r} \frac{\Lambda^{|k_1 - m_1|}}{|k_1 - m_1|!} \right)^2, \tag{4.74}$$

$$B = \sum_{|k_d| \le r_1} \left(\sum_{|m_d| \le r} \frac{\Lambda^{|k_d - m_d|}}{|k_d - m_d|!} \right)^2. \tag{4.75}$$

We have that

$$A \le K(d, \Lambda) \left(\frac{\Lambda^{|r_1 - r|}}{|r_1 - r|!} \right)^2. \tag{4.76}$$

To estimate B in (4.75) we split the inner sum into sums over *(i)* $|k_d| \le r_1$, estimated by $K(\Lambda)(2r + 1)$, and *(ii)* $r < |k_d| \le r_1$, estimated by $K(\Lambda)$. Then,

$$A^j B^{d-j} \le K(\Lambda, d) \left(\frac{\Lambda^{|r_1 - r|}}{|r_1 - r|!} \right)^{2j} \Big(K(\Lambda, d, j)(2r + 1)^{d-j} + \cdots + K(\Lambda, d, 0) \Big). \tag{4.77}$$

Collecting (4.73)–(4.77) and using $2r + 1 \le r_1 - r$, $r_1 - r \ge 1$ we obtain (4.17). □

Acknowledgement

The author would like to acknowledge helpful discussions with P. Padilla.

Appendix A

We outline the proof of the continuity of $\bar{H} : X \to \mathbf{R}$. This property is used in Theorem 2.1. In Proposition 2.2 we also use that $\bar{H}$ is (Frechet) C^1 in X; the proof uses similar standard tools and is omitted. (The notation here is as in Section 2.)

A useful observation for handling the quartic part of $\bar{H}$ is that $\tau \mapsto L_\tau$ is norm-continuous in $[0, T]$. This allows us to view the integral $\int_0^T$ as a Riemann integral and pass norms inside the integral. In particular, recall that by the assumptions on D we have a finite number of intervals $I_0 = [0, \tau_1]$, $\bar{I}_1 = (\tau_1, \tau_2), \ldots, I_n = (\tau_n, T]$ so that where $\tilde{d}$ is continuous in each $\bar{I}_j$. By the fact that Δ is bounded, the maps $t \mapsto \tilde{d}(t)\Delta \in \mathcal{L}(X)$ are therefore strongly continuous in all the intervals $\bar{I}_j$. ($\mathcal{L}(X)$ is the set of bounded linear operators on X.) We can then use the following.

Lemma A.1. *Consider the non-autonomous initial value problem $u_t = A(t)u$, $u(t_0) = u_0 \in X$, with $A(t) \in \mathcal{L}(X)$, $\forall t \in [t_0, t_1]$. Assume that $t \mapsto A(t) \in \mathcal{L}(X)$ is strongly continuous in $[t_0, t_1]$. Then there exists a unique 1-parameter norm-continuous semigroup of bounded linear operators $U(t, t_0) \in \mathcal{L}(X)$, $t \in [t_0, t_1]$, with $U(t, t_0)u_0$ satisfying the initial value problem in $[t_0, t_1]$.*

Lemma A.1 is proved in [9], chap. 7.1 (see also [7], chap. 5.9). Applying the lemma to the initial value problem for the equation $u_t = \tilde{d}(t)\Delta u$ in the intervals $\bar{I}_j$, i.e. with $A(t) = \tilde{d}(t)\Delta u$, we have that $U(\tau, \tau_j) = L_{\tau,\tau_j} = \exp(\mathrm{i} \int_{\tau_j}^{\tau} \tilde{d}(s)\, \mathrm{d}s \Delta)$, $\tau \in \bar{I}_j$, $j = 0, \ldots, n$. It is easy to check that if $\tau \in \bar{I}_k$ then $L_\tau = L_{\tau,\tau_k} L_{\tau_k,\tau_{k-1}} \ldots L_{\tau_1,0}$, $\forall k \in \{0, \ldots, n\}$, and that therefore $\tau \mapsto L_\tau$ is norm-continuous, for all $t \in [0, T]$. We then have:

Lemma A.2. *The function $\bar{H} : X \to \mathbf{R}$ is Lipschitz continuous on every bounded subset of X.*

Proof. First consider $\bar{H}_0(u) = \sum_{k=1}^{d} \|D_{k,+}u\|^2$, the quadratic part of $\bar{H}$. Consider the case $d = 1$ and let $(T_+u)_n = u_{n+1}$, $n \in \mathbf{Z}$. Expanding $|u_{n+1} - u_n|^2 - |v_{n+1} - v_n|^2$ and letting $u, v \in X$ we compute that

$$\begin{aligned} |\bar{H}_0(u) - \bar{H}_0(v)| &\le 2(|\langle u - v, u\rangle| + |\langle u - v, v\rangle| + |\langle T_+(u - v), v\rangle| + |\langle u - v, T_+v\rangle|) \\ &\le 4(\|u\| + \|v\|)\|u - v\|. \end{aligned} \tag{A.1}$$

For $d \ge 1$, $u, v \in X$ imply

$$|\bar{H}_0(u) - \bar{H}_0(v)| \le \left| \sum_{k=1}^{d} (\|D_{k,+}u\|^2 - \||D_{k,+}v\|^2) \right| \le 4d(\|u\| + \|v\|)\|u - v\|, \tag{A.2}$$

where in the last inequality we have repeated the steps leading to the $d = 1$ estimate.

Also, consider the quartic term $\bar{H}_2$ defined by $(\gamma/T)\bar{H}_2 = \bar{H} - \delta\bar{H}_0$. Given $u \in X$, define the map $h_u : [0, T] \to \mathbf{R}$ by $h_u(\tau) = \|L_\tau u\|_{l_4}^4$. The map h_u is the composition of $A_u : [0, T] \to X$, defined by $A_u(\tau) = L_\tau u$, and $G : X \to \mathbf{R}$, defined by $G(\psi) = \|\psi\|_{l_4}^4$. Applying Lemma A.1 to L_τ as above, the map A_u is continuous in $[0, T]$. To see the continuity of G, we calculate that for any $\psi, \chi \in X$

$$\begin{aligned} |G(\psi) - G(\chi)| &= \left| \sum_{n \in \mathbf{Z}^d} (|\psi_n|^2 + |\chi_n|^2)[(\psi_n - \chi_n)\psi_n^* - \chi_n(\psi_n^* - \chi_n^*)] \right| \\ &\le (\|\psi\|^2 + \|\chi\|^2)(\|\psi\| + \|\chi\|)\|\psi - \chi\|. \end{aligned} \tag{A.3}$$

Therefore, the composition h_u is also continuous, $\forall u \in X$. Letting $v, w \in X$ and using the triangle inequality for the Riemann integral we then have

$$|\bar{H}_2(v) - \bar{H}_2(w)| = \left| \int_0^T (h_v(\tau) - h_w(\tau)) d\tau \right| \le \int_0^T |(h_v(\tau) - h_w(\tau))| d\tau. \tag{A.4}$$

Estimating the last term as in (A.3) and using the fact that L_τ is an isometry, $\forall\tau$, we then obtain

$$|\bar{H}_2(v) - \bar{H}_2(w)| \leq (\|v\|^2 + \|w\|^2)(\|v\| + \|w\|)\|v - w\|. \tag{A.5}$$

The statement follows by combining (A.2), (A.5). □

References

[1] M.R. Ablowitz, G. Biondini, Multiscale pulse dynamics in communication systems with strong dispersion management, Opt. Lett. 23 (1998) 1668–1670.
[2] F.K. Abdulaev, B.B. Baizakov, M. Salerno, Stable two dimensional dispersion managed soliton, Phys. Rev. E 68 (2003) 066605.
[3] M.R. Ablowitz, Z.H. Musslimani, Discrete diffraction managed solitons, Phys. Rev. Lett. 87 (2001) 254102.
[4] D. Bambusi, Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators, Nonlinearity 9 (1996) 433–457.
[5] T. Cazenave, Semilinear Schrödinger equations, A.M.S., Providence, 2003.
[6] D.N. Christodoulides, R.I. Joseph, Discrete self-focusing in nonlinear arrays of coupled waveguides, Opt. Lett. 13 (1988) 794–796.
[7] K. Engel, R. Nagel, One-parameter semigroups for linear evolution equations, Springer, New York, 2000.
[8] H.S. Eisenberg, Y. Silberberg, R. Morandotti, J.S. Aitchison, Diffraction management, Phys. Rev. Lett. 85 (2000) 1863.
[9] H.O. Fattorini, The Cauchy problem, Addison-Wesley, London, 1983.
[10] I.R. Gabitov, S.K. Turytsin, Averaged pulse dynamics in a cascaded transmission line with passive dispersion compensation, Opt. Lett. 21 (1996) 327–329.
[11] Y.S. Kivshar, D.K. Campbell, Peierls-Nabarro barrier for highly localized nonlinear modes, Phys. Rev. E 48 (1993) 3077–3081.
[12] P.L. Lions, The concentration-compactnes principle in the calculus of variations. The locally compact case I., Ann. Inst. H. Poincare Anal. Non Lin. 1 (1984) 109–145.
[13] R.S. MacKay, S. Aubry, Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators, Nonlinearity 7 (1994) 1623–1643.
[14] J. Moeser, C.K.R.T. Jones, V. Zharnitsky, Stable pulse solutions for the nonlinear Schrödinger equations with higher order dispersion management, SIAM J. Appl. Anal. 35 (2004) 1486–1511.
[15] R. Morandotti, U. Peschel, J.S. Aitchison, H.S. Eisenberg, Y. Silberberg, Dynamics of discrete solitons in waveguide arrays, Phys. Rev. Lett. 83 (1999) 2726.
[16] P. Panayotaros, Averaging and Benjamin-Feir instabilities in a parametrically forced nonlinear Schrödinger equation, Phys. Lett. A 323 (2004) 403–414.
[17] P. Panayotaros, Quartic normal forms for the periodic nonlinear Schrödinger equation with dispersion management, Physica D 191 (2004) 219–237.
[18] M.A. Pinsky, Introducción al analisis de Fourier y las ondeletas, International Thomson Eds., México, 2003.
[19] M.I. Weinstein, Excitation thresholds for nonlinear localized modes on lattices, Nonlinearity 12 (1999) 673–691.
[20] V. Zharnitsky, E. Grenier, C.K.R.T. Jones, S.K. Turytsin, Stabilizing effect of dispersion management, Physica D 152–3 (2001) 794–817.