POLYNOMIAL UPPER BOUNDS FOR THE ORBITAL INSTABILITY OF THE 1D CUBIC NLS BELOW THE ENERGY NORM

J. COLLIANDER, M. KEEL, G. STAFFILANI, H. TAKAOKA, AND T. TAO

Abstract. We study the long-time behaviour of the focusing cubic NLS on $\mathbf{R}$ in the Sobolev norms H^s for $0 < s < 1$. We obtain polynomial growth-type upper bounds on the H^s norms, and also limit any orbital H^s instability of the ground state to polynomial growth at worst; this is a partial analogue of the H^1 orbital stability result of Weinstein [27], [26]. In the sequel to this paper we generalize this result to other nonlinear Schrödinger equations. Our arguments are based on the "I-method" from earlier papers [9]-[15] which pushes down from the energy norm, as well as an "upside-down I-method" which pushes up from the L^2 norm.

1. Introduction

We consider the long-time behaviour of solutions to the Cauchy problem for the one-dimensional focusing cubic nonlinear Schrödinger equation

$$iu_t + u_{xx} = -F(u); \quad u(x,0) = u_0(x) \tag{1.1}$$

where $u(x,t)$ is a complex-valued function on $\mathbf{R} \times \mathbf{R}$, $F(u)$ is the focusing cubic nonlinearity $F(u) := u\overline{u}u$, and $u_0(x)$ lies in the Sobolev space $H^s(\mathbf{R})$ for some $s \in \mathbf{R}$.

It is known ([25], [1]) that the Cauchy problem (1.1) is globally well-posed in H^s for all[1] $s \geq 0$. Furthermore, due to the many conservation laws of (1.1), we know that if s is an integer and the initial data is in H^s, then the H^s norm stays bounded for all time. For instance, for $s = 1$ one can obtain uniform H^1 bounds by exploiting the conservation of the Hamiltonian

$$H(u) := \int \frac{1}{2}|u_x|^2 - \frac{1}{4}|u|^4 \, dx$$

J.E.C. is supported in part by N.S.F. grant DMS 0100595 and N.S.E.R.C. grant RGPIN 250233-03.

M.K. is supported in part by N.S.F. Grant DMS 9801558.

G.S. is supported in part by N.S.F. Grant DMS 0100375 and by a grant from Hewlett and Packard and the Sloan Foundation.

H.T. is supported in part by J.S.P.S. Grant No. 13740087.

T.T. is a Clay Prize Fellow and is supported in part by a grant from the Packard Foundation.

[1] For $s < 0$ one does not even have local well-posedness, at least if one demands uniform continuity of the solution map; see [21].

and the L^2 norm, combined with the Gagliardo-Nirenberg inequality.

However when s is not an integer, the standard iteration argument only gives bounds on the H^s norms which grow exponentially in time. For $s > 1$ there are polynomial growth bounds[2] in [24], [3], [8].

The first result of this paper is to extend these techniques to s between 0 and 1.

Theorem 1.1. *If* $0 < s < 1$ *and* $u_0 \in H^s$, *then we have*

$$\|u_0(t)\|_{H^s} \leq C(\|u_0\|_{H^s})(1+|t|)^{2s+}.$$

The proof of this theorem proceeds by an "upside-down" version of the "I-method" ([9], [11], [12], [13], [14], [20]; see also [6], [10], [19]), in which one applies a differentiation operator $\mathcal{D} = \mathcal{D}_N$ to the solution instead of a smoothing operator $I = I_N$, and proves that the quantity $\|\mathcal{D}_N u(t)\|_2^2$ is almost conserved in time. One should compare the result in Theorem 1.1 with the ones established in [3], [8] and [24] for the time asymptotic of the H^s norm for smooth ($s > 1$) KdV and Schrödinger type solutions. The method used by Bourgain in [3] is also based on an improved local estimate, but the improvement is not obtained by replacing the norm H^s with a better integral (which is our argument here), but instead by using well-posedness results below the enrgy norm H^1. In [8] and [24] the improvement of the local estimate is obtained by using sharp bilinear estimates in negative Sobolev spaces.

We remark that the same polynomial growth result in Theorem 1.1 also holds for the defocusing cubic NLS (in which $-|u|^2u$ is replaced by $+|u|^2u$) and is slightly easier to obtain. It is likely that one can use the correction term techniques in [10], [14] to improve the exponent $2s+$ substantially[3], perhaps all the way down to $0+$. Certainly one expects to obtain an exponent which goes to 0 as $s \to 1^-$ by exploiting conservation of the Hamiltonian.

One can view Theorem 1.1 as a bound on the possible H^s instability of the 0 solution $u \equiv 0$ to (1.1). It is not strong enough to say that small H^s perturbations to this solution at time zero remain small perturbations for all later time, but it limits the growth of the perturbation to polynomial growth at best.

The next result of this paper concerns the H^s orbital stability of ground states for (1.1). For simplicity we shall only consider the ground states at energy 1 (the other energies can then be recovered by a scaling argument[4]). It is known [7] that

[2]It is an open question whether one has some sort of scattering in H^s for this equation, which would of course imply that the H^s norm remains bounded. Some recent progress in this direction is in [23].

[3]In the particular case of the cubic 1D NLS equation (1.1), one may also be able to exploit the complete integrability of the equation to obtain bounds on the H^s norms which are uniform in time, and perhaps also to obtain global stability bounds for solitons and multisolitons as well. On the other hand, the methods here do not exploit complete integrability and are applicable to a wide range of Hamiltonian evolution equations.

[4]More specifically, for every energy $E > 0$, there is a unique positive even Schwartz function Q_E obeying $(Q_E)_{xx} + |Q_E|^2 Q_E = EQ_E$, but these ground states are linked by the scaling $Q_E(x) = E^{1/2}Q(E^{1/2}x)$. Because the equation (1.1) is L^2-subcritical, all of these ground states have different L^2 mass. Since the L^2 mass is an invariant of the NLS flow, we can thus restrict

there exists a unique even positive Schwartz function $Q(x)$ on $\mathbf{R}$ which solves the equation[5]

$$Q_{xx} + |Q|^2 Q = Q. \tag{1.2}$$

The Cauchy problem (1.1) with initial data $u_0 = Q$ then has an explicit solution $u(t) = e^{it}Q$. More generally, for any $x_0 \in \mathbf{R}$ and $e^{i\theta} \in S^1$, the Cauchy problem with initial data $u_0(x) = e^{i\theta}Q(x - x_0)$ has explicit solution $e^{i(\theta+t)}Q(x - x_0)$. If we thus define the two-dimensional *ground state cylinder*[6] $\Sigma \subset H^1(\mathbf{R})$ by

$$\Sigma := \{e^{i\theta}Q(\cdot - x_0) : x_0 \in \mathbf{R}, e^{i\theta} \in S^1\}$$

we see that the nonlinear flow (1.1) preserves Σ. Also note that each element of Σ obeys (1.2) (though of course most ground states are not even or positive).

In [27] (see also [26]) Weinstein showed that the ground state cylinder Σ was H^1-stable. More precisely, he showed an estimate of the form

$$\mathrm{dist}_{H^1}(u(t), \Sigma) \sim \mathrm{dist}_{H^1}(u_0, \Sigma) \tag{1.3}$$

(when $\mathrm{dist}_{H^1}(u_0, \Sigma)$ is small), for all H^1 solutions $u(t)$ to (1.1) and all times $t \in \mathbf{R}$. In other words, solutions which start close to a ground state in H^1 at time $t = 0$, will stay close to a ground state for all time (though the nearby ground state may itself vary in time[7]).

To prove (1.3), Weinstein employed the *Lyapunov functional*[8]

$$L(u) := 2H(u) + \int |u|^2 = \int |u_x|^2 + |u|^2 - \frac{1}{2}|u|^4 \, dx, \tag{1.4}$$

which is well-defined for all $u \in H^1$. Since this quantity is a combination of the Hamiltonian and the L^2 norm, it is clearly an invariant of the flow (1.1). More explicitly, for sufficiently smooth functions $u(x, t)$ we have the formula

$$\partial_t L(u) = 2\langle u_t, -u_{xx} + u - F(u)\rangle \tag{1.5}$$

which clearly vanishes if u solves (1.1). Here and in the sequel we use $\langle , \rangle$ to denote the *real* inner product

$$\langle u, v\rangle := \mathrm{Re} \int u\overline{v} \, dx.$$

---

our attention to a sphere in L^2, in which case only one energy E is relevant. One can then use the scale invariance of (1.1) to set $E = 1$.

[5]Indeed, we have the explicit formula $Q(x) := 2^{-1/2}/\cosh(x)$, although we will not use this formula in this paper.

[6]Note that the ground state cylinder is the orbit of Q under the phase and translation invariances of NLS. We do not utilize the scaling invariance because, as mentioned earlier, this changes the L^2 norm of Q. Also we do not utilize Gallilean invariance because this does not preserve the Hamiltonian.

[7]For instance, consider the solution $u(x, t) = e^{i(\varepsilon x - \varepsilon^2 t)}e^{it}Q(x - 2\varepsilon t)$ for some small ε; this is a Gallilean transformation of the ground state solution $u(t) = e^{it}Q$, and is close to this solution at time zero. However at later times, the solution slowly drifts away from the original ground state solution, although it remains close to the ground state cylinder Σ.

[8]This is the functional for energy $E = 1$. For other energies it is given by $L(u) = 2H(u) + \int E|u|^2$.

From (1.5) and (1.2) we see that the ground states in Σ are critical points of L. In fact they are minimizers of L; more precisely, we have the fundamental *coercivity estimate*

$$L(u) - L(Q) \sim \mathrm{dist}_{H^1}(u, \Sigma)^2 \text{ whenever } u \in H^1 \text{ and } \mathrm{dist}_{H^1}(u, \Sigma) \ll 1; \tag{1.6}$$

see [27]. The stability estimate (1.3) then follows easily from (1.6) and the conservation of L.

Note that the functional L is invariant under phase rotation $u \to e^{i\theta} u$ and translation $u \mapsto u(\cdot - x_0)$. Thus one cannot expect a coercivity bound like (1.6) in these directions. Of course, this is consistent with (1.6) since the ground state cylinder Σ is itself invariant under these symmetries; the point of (1.6) is that there are no other directions (in the tangent space of H^1 at Q) for which L can be invariant or concave.

The second main result of this paper is to partially extend the H^1 orbital stability result to H^s orbital stability-type result for $0 \leq s < 1$. Unfortunately, as in Theorem 1.1, our estimate loses a polynomial factor in t, so we cannot exclude the possibility of polynomial orbital instability in H^s:

Theorem 1.2. *Let $0 \leq s < 1$, and suppose $\mathrm{dist}_{H^s}(u_0, \Sigma) \ll 1$. Then we have*

$$\mathrm{dist}_{H^s}(u(t), \Sigma) \lesssim t^{1-s+} \mathrm{dist}_{H^s}(u_0, \Sigma) \tag{1.7}$$

whenever

$$1 \leq t \ll \mathrm{dist}_{H^s}(u_0, \Sigma)^{-1/(1-s+)}.$$

In particular, $u(t)$ stays in a bounded subset of H^s for all times $|t| \ll \mathrm{dist}_{H^s}(u_0, \Sigma)^{-1/(1-s+)}$.

Note that a naive application of the local well-posedness theory would lose a factor of the form $\exp(Ct)$ in the estimate (1.7), so that one could only assure $u(t)$ stays in a bounded subset of H^s for times $|t| \lesssim \log(1/\mathrm{dist}_{H^s}(u_0, \Sigma))$. Note that for times t close to $\mathrm{dist}_{H^s}(u_0, \Sigma)^{1/(1-s+)}$, the right-hand side of (1.7) is close to 1, so that we are no longer keeping $u(t)$ close to Σ. At this point one can use Theorem 1.1 to control the further development of $u(t)$.

The proof of (1.7) proceeds via the smoothing operator $I = I_N$ mentioned earlier. Basically, the idea is to show that the modified Lyapunov functional $L(Iu)$ is almost conserved, and then combine this with (1.6) to obtain (1.7). It turns out that a naive implementation of this approach loses an epsilon power of $\mathrm{dist}_{H^s}(u_0, \Sigma)$ because the operator I does not quite preserve the ground state cylinder, but this can be rectified by the standard technique of choosing an approximating ground state to obey specially chosen orthogonality conditions. For expository reasons we have chosen to give the naive versions of the argument first, and only give the full argument at the end of the paper.

The factor t^{1-s+} in (1.7) can probably be reduced, however the factor $\mathrm{dist}_{H^s}(u_0, \Sigma)$ on the right-hand side is necessary (as can be seen even when $t \sim 1$).

In all of these arguments it is crucial that the L^2 and H^1 norms are both subcritical (in the sense that they scale as a negative power of length using the natural scaling of (1.1)). It is because of this that we cannot extend these results to any other NLS with an algebraic nonlinearity[9]. In the sequel to this paper we shall obtain some partial results of the above type in the case when p is not an odd integer; the main new difficulty is to commute the I operator with non-algebraic nonlinearities $F(u)$.

2. Notation

We use $A \lesssim B$ to denote $A \leq CB$, where C is a constant depending on s which may vary from line to line. We use $a+$, $a-$ to denote quantities of the form $a+\varepsilon$, $a-\varepsilon$, where ε is arbitrarily small. We use $\langle\xi\rangle$ to denote $1+|\xi|$.

We define the spatial Fourier transform by

$$\hat{f}(\xi) := \int_{\mathbf{R}} e^{-ix\xi} f(x)\, dx$$

and the spacetime Fourier transform by

$$\tilde{u}(\xi,\tau) := \int_{\mathbf{R}}\int_{\mathbf{R}} e^{-i(x\xi+t\tau)} u(x,t)\, dxdt.$$

Following [2], we define the $X^{s,b}$ spaces by the norm

$$\|u\|_{X^{s,b}} := \|u\|_{s,b} := \|\langle\xi\rangle^s \langle\tau-\xi^2\rangle^b \tilde{u}(\xi,\tau)\|_{L^2_\tau L^2_\xi}.$$

For any time interval I, we define the restricted $X^{s,b}$ spaces by

$$X^{s,b}_I := \{u|_{\mathbf{R}\times I} : u \in X^{s,b}\}$$

with the usual norm

$$\|v\|_{X^{s,b}_I} := \inf\{\|u\|_{s,b} : u|_{\mathbf{R}\times I} = v\}.$$

We shall need the Strichartz estimate

$$\|u\|_{L^6_{x,t}} \lesssim \|u\|_{0,1/2+}; \tag{2.1}$$

this can be obtained by writing an $X^{0,1/2+}$ function as an average of modulated free Schrödinger waves (as in [17]) and then using the $L^6_{x,t}$ Strichartz estimate for free solutions (see e.g. [18] and the references therein).

Interpolating (2.1) with the trivial bound $\|u\|_{L^2_{x,t}} \leq \|u\|_{0,0}$ we obtain

$$\|u\|_{L^4_{x,t}} \lesssim \|u\|_{0,3/8+}. \tag{2.2}$$

We also record the variant estimate

$$\|u\|_{L^8_t L^4_x} \lesssim \|u\|_{0,1/2+}; \tag{2.3}$$

this can be obtained by interpolating (2.1) with the energy estimate $\|u\|_{L^\infty_t L^2_x} \lesssim \|u\|_{0,1/2+}$.

[9]More specifically, the NLS equation is only L^2 subcritical when $p < 1+\frac{4}{n}$, while an algebraic nonlinearity only occurs when p is an odd integer. Since $p > 1$, the only subcritical algebraic equation occurs when $n=1$, $p=3$.

From [6] (see also [13]) we recall the improved bilinear Strichartz estimate (in one spatial dimension)

$$\|u_1 u_2\|_{L^2_{t,x}} \lesssim N^{-1/2} \|u_1\|_{0,1/2+} \|u_2\|_{0,1/2+} \tag{2.4}$$

whenever u_1 has Fourier support in the region $|\xi| \sim N$, and u_2 has Fourier support in the region $|\xi| \ll N$.

Let $n \geq 2$, and let $m(\xi_1, \dots, \xi_n)$ be a function supported on the hyperplane $\{\xi_1 + \dots + \xi_n = 0\}$. We use the notation

$$\Lambda_n(m(\xi_1, \dots, \xi_n); f_1, \dots, f_n)$$

to denote the multilinear form

$$\Lambda_n(m(\xi_1, \dots, \xi_n); f_1, \dots, f_n) := \int_{\xi_1 + \dots + \xi_n = 0} m(\xi_1, \dots, \xi_n) \hat{f}_1(\xi_1) \dots \hat{f}_n(\xi_n).$$

3. Proof of Theorem 1.1

We shall divide the proof of this theorem into several broad steps. These steps will also appear in the proof of Theorem 1.2.

- **Step 0. Preliminaries; introduction of the modified energy.**

Fix $0 < s < 1$, and fix the global H^s solution u. Henceforth all implicit constants may depend on s and the quantity $\|u_0\|_{H^s}$. Our task is to show that

$$\|u(T)\|_{H^s} \lesssim T^{2s+} \tag{3.1}$$

for all $T \gg 1$. By the H^s global well-posedness and regularity theory (see e.g. [1]) and the usual limiting argument it suffices to do this for smooth, rapidly decreasing u.

We now apply an "upside-down" version of the I-method; whereas the strategy of the I method is to mollify the solution at high frequencies to make it smoother (e.g. in the energy class H^1), here we amplify the solution at high frequencies instead to make it rougher (specifically, we place it in L^2). In contrast, the proof of Theorem 1.2 in later sections will proceed via the ordinary "I-method".

Fix T, and let $N \gg 1$ be a large quantity depending on T to be chosen later. Let $\theta(\xi)$ be a smooth even real-valued symbol such that $\theta(\xi) = 1$ for $|\xi| \leq N$ and $\theta(\xi) = |\xi|^s / N^s$ for $|\xi| > 2N$, and let $\mathcal{D}$ be the Fourier multiplier

$$\widehat{\mathcal{D}f}(\xi) := \theta(\xi) \hat{f}(\xi). \tag{3.2}$$

Thus $\mathcal{D}$ is the identity for low frequencies $|\xi| \leq N$, and becomes a differentiation operator of order s for high frequencies $|\xi| \gtrsim N$.

Define the *modified energy* $E_N(t)$ at time t by

$$E_N(t) := \|\mathcal{D}u(t)\|_2^2. \tag{3.3}$$

From Plancherel we have the upper bound

$$E_N(0) \lesssim \|u_0\|_{H^s}^2 \lesssim 1.$$

The heart of the argument shall lie in the following almost conservation law for E_N.

Lemma 3.1. *If $t_0 \in \mathbf{R}$ is such that $E_N(t_0) \leq C$ for some bounded constant $C = O(1)$, then we have*

$$E_N(t_0 + \delta) = E_N(t_0) + O(N^{-1/2+}) \tag{3.4}$$

where $\delta > 0$ is an absolute constant depending only on s and C.

The error bound of $O(N^{-1/2+})$ might not be sharp; any improvement in this estimate will lead to a better polynomial growth bound than (3.1). It may be that one can improve this result by adding suitably chosen "correction terms" to $E_N(t)$, in the spirit of [10], [14].

- **Step 1. Deduction of** (3.1) **from Lemma 3.1.**

If we assume Lemma 3.1, then we may iterate it to obtain $E_N(t) \lesssim 1$ for all $0 \leq t \ll N^{1/2-}$. In particular, if we assume $T \ll N^{1/2-}$, we have from Plancherel and (3.3) that

$$\|u(T)\|_{H^s} \lesssim N^s E_N(T)^{1/2} \lesssim N^s.$$

Optimizing N in terms of T we obtain (3.1) as desired.

It remains to prove Lemma 3.1. This is done in several stages.

- **Step 2. Control u at time t_0.**

By hypothesis we have

$$\|\mathcal{D}u(t_0)\|_2 \lesssim 1. \tag{3.5}$$

As one can see from (3.2), this is essentially an H^s-type bound on u (up to powers of N). However, in order to obtain good polynomial growth bounds it is important that we use the norm $\|\mathcal{D}u\|_2$ throughout rather than $\|u\|_{H^s}$.

- **Step 3. Control u on the time interval $[t_0 - \delta, t_0 + \delta]$.**

We now use (3.5) to claim

$$\|\mathcal{D}u\|_{X^{0,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim 1 \tag{3.6}$$

for any $0 < \varepsilon \ll 1$, if $0 < \delta \ll 1$ is a sufficiently small constant (depending on s and the bound in (3.5), but not on N). This will be achieved by techniques similar to those used to obtain local well-posedness using the $X^{s,b}$ spaces (as in e.g. [2]).

To begin with, by the standard energy estimate for $X^{s,b}$ spaces (see e.g. [2], or [16], page 771; note that the multiplier $\mathcal{D}$ commutes with the Schrödinger operator

$i\partial_t + \partial_{xx}$ and is thus harmless) we have

$$\|\mathcal{D}u\|_{X^{0,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim \|\mathcal{D}u(t_0)\|_{L^2} + \delta^{0+}\|\mathcal{D}(iu_t + u_{xx})\|_{X^{0,-1/2+2\varepsilon}_{[t_0-\delta,t_0+\delta]}}.$$

Since we are allowed to choose δ to be sufficiently small, it suffices from (3.5), (1.1) and standard continuity (or iteration) arguments to show the trilinear estimate

$$\|\mathcal{D}(u_1\overline{u}_2u_3)\|_{0,-1/2+2\varepsilon} \lesssim \|\mathcal{D}u_1\|_{0,1/2+}\|\mathcal{D}u_2\|_{0,1/2+}\|\mathcal{D}u_3\|_{0,1/2+}.$$

We may assume that the u_i have non-negative Fourier transforms. Since $w(\xi+\eta) \lesssim w(\xi) + w(\eta)$ we can obtain a fractional Leibnitz rule, and reduce to showing

$$\begin{aligned}
\|\mathcal{D}(u_1)\overline{u}_2u_3\|_{0,-1/2+2\varepsilon} &\lesssim \|\mathcal{D}u_1\|_{0,1/2+}\|\mathcal{D}u_2\|_{0,1/2+}\|\mathcal{D}u_3\|_{0,1/2+}\\
\|u_1\mathcal{D}(\overline{u}_2)u_3\|_{0,-1/2+2\varepsilon} &\lesssim \|\mathcal{D}u_1\|_{0,1/2+}\|\mathcal{D}u_2\|_{0,1/2+}\|\mathcal{D}u_3\|_{0,1/2+}\\
\|u_1\overline{u}_2\mathcal{D}(u_3)\|_{0,-1/2+2\varepsilon} &\lesssim \|\mathcal{D}u_1\|_{0,1/2+}\|\mathcal{D}u_2\|_{0,1/2+}\|\mathcal{D}u_3\|_{0,1/2+}.
\end{aligned} \tag{3.7}$$

From the dual of (2.2) we see that $\|f\|_{0,-1/2+2\varepsilon} \lesssim \|f\|_{L^{4/3}_{x,t}}$. The claims (3.7) then follow from Hölder's inequality and several applications of (2.2). This gives (3.6).

- **Step 4. Control the increment of the modified energy.**

To prove (3.4), we now apply the fundamental theorem of Calculus to write the increment of the modified energy (3.3) as

$$E_N(t_0+\delta) - E_N(t_0) = \int_{t_0}^{t_0+\delta} \partial_t E_N(t)\, dt. \tag{3.8}$$

A routine integration by parts[10] shows that

$$\begin{aligned}
\partial_t E_N(t) &= 2\langle \mathcal{D}u_t, \mathcal{D}u\rangle\\
&= 2\langle i\mathcal{D}u_{xx} + i\mathcal{D}F(u), \mathcal{D}u\rangle\\
&= 2\langle iu\overline{u}u, \mathcal{D}^2u\rangle\\
&= -2\mathrm{Im}\Lambda_4(\theta(\xi_4)^2; u(t), \overline{u(t)}, u(t), \overline{u(t)})\\
&= \frac{1}{2}\mathrm{Im}\Lambda_4(\theta(\xi_1)^2 - \theta(\xi_2)^2 + \theta(\xi_3)^2 - \theta(\xi_4)^2; u(t), \overline{u(t)}, u(t), \overline{u(t)}),
\end{aligned}$$

where in the last step we exploited the symmetry

$$\Lambda_4(\theta(\xi_j)^2; u(t), \overline{u(t)}, u(t), \overline{u(t)}) = \overline{\Lambda_4(\theta(\xi_k)^2; u(t), \overline{u(t)}, u(t), \overline{u(t)})}$$

for $j = 1, 3$ and $k = 2, 4$.

From the above computations, it suffices to show the estimate

$$\left|\int \chi_{[t_0,t_0+\delta]}(t)\Lambda_4(M_4; u_1(t), \overline{u_2(t)}, u_3(t), \overline{u_4(t)})\, dt\right| \lesssim N^{-1/2+}\prod_{i=1}^4 \|\mathcal{D}u_i\|_{0,1/2+}$$

for all functions $u_1, u_2, u_3, u_4 \in X^{0,1/2+}$, where $M_4 = M_4(\xi_1, \xi_2, \xi_3, \xi_4)$ denotes the symbol

$$M_4(\xi_1,\xi_2,\xi_3,\xi_4) := \theta(\xi_1)^2 - \theta(\xi_2)^2 + \theta(\xi_3)^2 - \theta(\xi_4)^2.$$

---

[10]This is of course related to the usual energy method computations to control the growth of higher order energies; see also [3], [24], [8]. Observe that if $\mathcal{D}$ was the identity then this derivative would vanish, thus giving the standard proof of L^2 norm conservation.

We shall argue similarly to [13]. We apply Littlewood-Paley decompositions, and assume that u_i is supported on the region $\langle \xi_i \rangle \sim N_i$ for some dyadic $N_i \geq 1$; of course, we will eventually have to sum in N_i to recover the general case. Define $\{soprano, alto, tenor, baritone\} = \{1, 2, 3, 4\}$ by requiring

$$N_{soprano} \geq N_{alto} \geq N_{tenor} \geq N_{baritone}.$$

We may assume that $N_{soprano} \sim N_{alto}$ since Λ_4 is integrating over the region $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 0$. We may also assume that $N_{soprano} \gtrsim N$ since the symbol $\theta(\xi_1)^2 - \theta(\xi_2)^2 + \theta(\xi_3)^2 - \theta(\xi_4)^2$ vanishes otherwise.

We divide into two cases.

- **Case (a):** $N_{soprano} \gg N_{baritone}$.

In this case $M_4 = O(\theta(N_{soprano})^2)$, so we can essentially [11] bound this term by

$$\theta(N_{soprano})^2 \int \int_{t_0}^{t_0+\delta} |u_{soprano}||u_{alto}||u_{tenor}||u_{baritone}|\ dxdt.$$

We use Hölder's inequality to take u_{alto} and u_{tenor} in $L^4_{x,t}$, and $u_{soprano}u_{baritone}$ in $L^2_{x,t}$. Using two applications of (2.2) and one application of (2.4) we can bound this by

$$\theta(N_{soprano})^2 N_{soprano}^{-1/2} \prod_{i=1}^4 \|u_i\|_{0,1/2+} \lesssim N_{soprano}^{-1/2} \prod_{i=1}^4 \|\mathcal{D}u_i\|_{0,1/2+}.$$

The claim then follows by summing in the N_i.

- **Case (b):** $N_{soprano} \sim N_{baritone}$.

In this case all the frequencies are comparable. The key observation is that for $|\xi_i| \sim N_{soprano}$ and $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 0$, we have

$$|\frac{M_4}{\xi_1^2 - \xi_2^2 + \xi_3^2 - \xi_4^2}| \lesssim \frac{\theta(N_{soprano})^2}{N_{soprano}^2}. \tag{3.9}$$

To see this, we write the denominator as

$$\begin{aligned} \xi_1^2 - \xi_2^2 + \xi_3^2 - \xi_4^2 &= (\xi_1 - \xi_2)(\xi_1 + \xi_2) + (\xi_3 - \xi_4)(\xi_3 + \xi_4) \\ &= (\xi_1 - \xi_2 - \xi_3 + \xi_4)(\xi_1 + \xi_2) \\ &= 2(\xi_1 + \xi_4)(\xi_1 + \xi_2) \end{aligned}$$

[11]To be more precise, one should replace u_i by Mu_i, the Hardy-Littlewood maximal function of u_i. This is because the symbol M_4 is equal to $\theta(N_{soprano})^2$ times a smooth multiplier which obeys good kernel estimates. We omit the details. An alternate approach is to reduce to the case when all the u_i have non-negative spacetime Fourier transform (which is legitimate since the $X^{s,b}$ norms do not care about the phase of $\tilde{u}_i$), however the time cutoff $\chi_{[t_0,t_0+\delta]}(t)$ then presents an annoying technical difficulty (since its Fourier transform is not non-negative).

and the numerator as

$$\begin{aligned} M_4 =&\theta(\xi_1)^2 - \theta(\xi_2)^2 + \theta(\xi_3)^2 - \theta(\xi_4)^2 \\ =&\theta(\xi_3 + (\xi_1 + \xi_2) + (\xi_1 + \xi_4))^2 - \theta(\xi_3 + (\xi_1 + \xi_4))^2 \\ &+ \theta(\xi_3)^2 - \theta(\xi_3 + (\xi_1 + \xi_2))^2; \end{aligned}$$

the claim then follows from the double mean value theorem and the estimate

$$|\frac{d^2}{d\xi^2}(\theta(\xi)^2)| \lesssim \frac{\theta(N_{soprano})^2}{N_{soprano}^2}$$

for all $\xi = O(N_{soprano})$.

To use (3.9) we must use the spacetime Fourier transform, which requires us to first modify the cutoff $\chi_{[t_0,t_0+\delta]}(t)$. Write $\chi_{[t_0,t_0+\delta]}(t) = a(t) + b(t)$, where $a(t)$ is $\chi_{[t_0,t_0+\delta]}(t)$ convolved with a smooth approximation to the identity of width $N_{soprano}^{-100}$, and $b(t) = \chi_{[t_0,t_0+\delta]}(t) - a(t)$.

Consider the contribution of $b(t)$. This term is essentially bounded by

$$\theta(N_{soprano})^2 \int\int |b(t)||u_1(t)||u_2(t)||u_3(t)||u_4(t)| \; dxdt;$$

by Hölder's inequality and (2.3) we may bound this by

$$N_{soprano}^{-10} \prod_{i=1}^{4} \|u_i\|_{0,1/2+}$$

which easily sums to be acceptable.

Now consider the contribution of $a(t)$. In light of the estimate

$$\|a(t)u_1\|_{0,1/2+} \lesssim N_{soprano}^{0+} \|u_1\|_{0,1/2+}$$

(see [13]) it suffices to show

$$\int \Lambda_4(M_4; u_1(t), \overline{u_2(t)}, u_3(t), \overline{u_4(t)}) \; dt \lesssim N_{soprano}^{-1} \prod_{i=1}^{4} \|\mathcal{D}u_i\|_{0,1/2+}.$$

We may assume that the spacetime Fourier transforms of u_i are non-negative. By Plancherel and (3.9) it suffices to show that

$$\int_{\xi_1+\xi_2+\xi_3+\xi_4=0} \int_{\tau_1+\tau_2+\tau_3+\tau_4=0} |\xi_1^2 - \xi_2^2 + \xi_3^2 - \xi_4^2| \tilde{u}_1(\xi_1,\tau_1)\tilde{\overline{u_2}}(\xi_2,\tau_2)\tilde{u}_3(\xi_3,\tau_3)\tilde{\overline{u_4}}(\xi_4,\tau_4)$$

$$\lesssim N_{soprano} \prod_{i=1}^{4} \|u_i\|_{0,1/2+}.$$

From the triangle inequality we have

$$|\xi_1^2 - \xi_2^2 + \xi_3^2 - \xi_4^2| \lesssim |\tau_1 - \xi_1^2| + |\tau_2 + \xi_2^2| + |\tau_3 - \xi_3^2| + |\tau_4 + \xi_4^2|;$$

interpolating this with the trivial bound of $N_{soprano}^2$ we obtain

$$|\xi_1^2 - \xi_2^2 + \xi_3^2 - \xi_4^2| \lesssim N_{soprano}(|\tau_1 - \xi_1^2|^{1/2} + |\tau_2 + \xi_2^2|^{1/2} + |\tau_3 - \xi_3^2|^{1/2} + |\tau_4 + \xi_4^2|^{1/2}).$$

By symmetry, it thus suffices to show

$$|\int_{\xi_1+\xi_2+\xi_3+\xi_4=0}\int_{\tau_1+\tau_2+\tau_3+\tau_4=0} |\tau_1-\xi_1^2|^{1/2}\tilde{u}_1(\xi_1,\tau_1)\tilde{\overline{u_2}}(\xi_2,\tau_2)\tilde{u}_3(\xi_3,\tau_3)\tilde{\overline{u_4}}(\xi_4,\tau_4)|$$
$$\lesssim \prod_{i=1}^4 \|u_i\|_{X^{0,1/2+}}.$$

By Plancherel and Hölder's inequality the left-hand side is bounded by

$$\|\mathcal{F}(|\tau_1-\xi_1^2|^{1/2}\tilde{u}_1)\|_{L_t^2L_x^2}\|u_2\|_{L_t^6L_x^6}\|u_3\|_{L_t^6L_x^6}\|u_4\|_{L_t^6L_x^6}$$

which is acceptable by (2.1) and the definition of $X^{0,1/2+}$. This completes the proof of Lemma 3.1, and hence of Theorem 1.1.

4. Proof of Theorem 1.2: first attempt

We now prove Theorem 1.2. We shall begin by giving a simplified version of the argument which does not capture the full power of $\mathrm{dist}_{H^s}(u_0,\Sigma)$ in (1.7), but introduces the main ideas and establishes the key multi-linear estimates (Lemma 4.2). In the next section, we shall give a better version of this argument which recovers most of this power in the next section, and then in the section after that we shall give the most refined version of the argument that gives (1.7) with no loss.

This argument is very similar to the previous one, and we have deliberately given the two arguments a nearly identical structure.

- **Step 0. Preliminaries; introduction of the modified energy.**

By the existing global well-posedness theory and a standard limiting argument we may assume that u is a global smooth solution which is rapidly decreasing in space.

Fix $0 < s < 1$, u, and let $N \gg 1$ be chosen later. Let $m(\xi)$ be a smooth even real-valued symbol such that $m(\xi) = 1$ for $|\xi| \leq N$ and $m(\xi) = |\xi|^{s-1}/N^{s-1}$ for $|\xi| > 2N$, and let I be the Fourier multiplier

$$\widehat{If}(\xi) := m(\xi)\hat{f}(\xi).$$

Thus I is the identity for frequencies $|\xi| \ll N$ and is smoothing of order $1-s$ for high frequencies $|\xi| \gtrsim N$.

In analogy with the quantity (3.3) used in the proof of Theorem 1.1, we define the *modified energy*[12] $E_N(t)$ by

$$E_N(t) := L(Iu(t)), \tag{4.1}$$

where the Lyapunov functional $L()$ was defined in (1.4).

[12]In a more standard application of the I-method, e.g. [9], [11], [12], [20]) we would take $E_N(t) = H(Iu(t))$. The main difference here is thus to add an L^2 component to the modified energy in order to make the ground state cylinder an approximate minimizer of the energy.

We now estimate $E_N(0)$. Write $\sigma := \mathrm{dist}_{H^s}(u_0, \Sigma)$, thus by hypothesis $0 < \sigma \ll 1$ and there exists a ground state $\tilde{Q} \in \Sigma$ such that

$$\|u_0 - \tilde{Q}\|_{H^s} \lesssim \sigma.$$

Applying I, we see that

$$\|Iu_0 - I\tilde{Q}\|_{H^1} \lesssim N^{1-s}\sigma.$$

On the other hand, since $\tilde{Q}$ is smooth, its Fourier transform is rapidly decreasing, and

$$\|I\tilde{Q} - \tilde{Q}\|_{H^1} \lesssim N^{-C} \tag{4.2}$$

for any C. Thus, if we assume

$$N \gtrsim \sigma^{0-}, \tag{4.3}$$

we then have

$$\|Iu_0 - \tilde{Q}\|_{H^1} \lesssim N^{1-s}\sigma.$$

By (1.6) we thus have

$$E_N(0) - L(Q) \lesssim N^{2-2s}\sigma^2.$$

We shall make the assumption that

$$N^{2-2s}\sigma^2 \ll 1. \tag{4.4}$$

The heart of the argument shall lie in the following almost conservation law for the modified energy E_N.

Lemma 4.1. *If $t_0 \in \mathbf{R}$ is such that $E_N(t_0) - L(Q) \ll 1$, then we have*

$$E_N(t_0 + \delta) = E_N(t_0) + O(N^{-1+}) \tag{4.5}$$

where $\delta > 0$ is an absolute constant depending only on s.

As in the previous section, we do not know if the error bound $O(N^{-1+})$ is sharp; any improvement in this error estimate will ultimately lead to an improvement of the t^{1-s+} factor in (1.7) (if one uses the most refined version of the argument below, see Section 6).

- **Step 1. Deduction of a weak form of** (1.7) **from Lemma 4.1.**

If we assume Lemma 4.1, then we may iterate it to obtain $E_N(t) - L(Q) \lesssim N^{2-2s}\sigma^2$ for all

$$1 \leq t \ll N^{1-}N^{2-2s}\sigma^2, \tag{4.6}$$

where we assume

$$N^{1-}N^{2-2s}\sigma^2 \gg 1 \tag{4.7}$$

(note that this automatically implies (4.3)).

Fix t as above. Then by (1.6) we have[13]

$$\|Iu(t) - \tilde{Q}(t)\|_{H^1} \lesssim N^{1-s}\sigma$$

for some ground state $\tilde{Q}(t) \in \Sigma$ depending on t. Using (4.3) as before, we may modify this to

$$\|Iu(t) - I\tilde{Q}(t)\|_{H^1} \lesssim N^{1-s}\sigma$$

which implies that

$$\|u(t) - \tilde{Q}(t)\|_{H^s} \lesssim N^{1-s}\sigma$$

and hence that

$$\mathrm{dist}_{H^s}(u(t), \Sigma) \lesssim N^{1-s}\sigma. \tag{4.8}$$

Optimizing in N subject to (4.4), (4.6), (4.7) we obtain

$$\mathrm{dist}_{H^s}(u(t), \Sigma) \lesssim t^{\frac{1-s}{3-2s-}}\sigma^{\frac{1}{3-2s-}} \tag{4.9}$$

whenever $t \ll \sigma^{-1/(1-s)-}$, which is a weak form of (1.7). In the next two sections we shall use more refined arguments to improve this bound.

It remains to prove Lemma 4.1. This shall be done in several stages.

- **Step 2. Control u at time t_0.**

From the hypothesis $E_N(t_0) - L(Q) \ll 1$ and (1.6) (or the Gagliardo-Nirenberg inequality) we have

$$\|Iu(t_0)\|_{H^1} \lesssim 1. \tag{4.10}$$

Up to powers of N, this is basically an H^s bound on $u(t_0)$, but to obtain good exponents it is important that we work with the $\|Iu\|_{H^1}$ norm rather than the $\|u\|_{H^s}$.

- **Step 3. Control u on the time interval $[t_0 - \delta, t_0 + \delta]$.**

We now claim that (4.10) implies the spacetime estimate

$$\|Iu\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim 1 \tag{4.11}$$

if $0 < \delta \ll 1$ is a sufficiently small constant and $0 < \varepsilon \ll 1$. As before, it suffices to show the estimate

$$\|I(u_1\overline{u}_2u_3)\|_{1,-1/2+\varepsilon} \lesssim \|Iu_1\|_{1,1/2+}\|Iu_2\|_{1,1/2+}\|Iu_3\|_{1,1/2+}. \tag{4.12}$$

We can write the left-hand side as

$$\|I\langle\nabla\rangle(u_1\overline{u}_2u_3)\|_{0,-1/2+\varepsilon},$$

where $I\langle\nabla\rangle$ is the multiplier with symbol $m(\xi)\langle\xi\rangle$. Since $m(\xi+\eta)\langle\xi+\eta\rangle \lesssim m(\xi)\langle\xi\rangle + m(\eta)\langle\eta\rangle$, we can use the fractional Leibnitz rule, together with (2.2) and its dual, to obtain this estimate.

[13]Strictly speaking, to apply (1.6) we must assume beforehand that $Iu(t)$ is close to a ground state. But by hypothesis this is true at time $t = 0$, and so one may proceed by a standard continuity argument which we omit. This continuity argument will also be used later in the paper without further mention.

- **Step 4. Control the increment of the modified energy.**

Now that we have obtained (4.11), the next step is again the fundamental theorem of Calculus (3.8). We introduce the nonlinear functional $\Omega(v)$, defined on smooth (H^2 will do) functions v on $\mathbf{R} \times \mathbf{R}$ by the formula

$$\Omega(v) := \int_{t_0}^{t_0+\delta} \left\langle iI\big(v(t)_{xx} + F(v(t))\big), -Iv(t)_{xx} + Iv(t) - F(Iv(t)) \right\rangle \, dt. \tag{4.13}$$

From (3.8), (1.5) and (1.1) we have

$$E_N(t_0 + \delta) - E_N(t_0) = 2\Omega(u), \tag{4.14}$$

so in view of (4.11) it suffices to prove the following estimate, which we shall re-use in later sections.

Lemma 4.2. *We have*

$$|\Omega(v)| \lesssim N^{-1+}$$

whenever

$$\|Iv\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta, t_0+\delta]}} \lesssim 1.$$

We now prove Lemma 4.2. By repeated integration by parts and a symmetrization (cf. the computations after (3.8)) we can expand the integrand in (4.13) as

$$\begin{aligned}
&\langle iIv_{xx}, -F(Iv)\rangle + \langle iIF(v), -Iv_{xx} + Iv - F(Iv)\rangle \\
=&\langle iIv_{xx}, -F(Iv) + IF(v)\rangle + \langle iF(v), I^2 v\rangle + \langle iIF(v), -F(Iv)\rangle \\
=&\mathrm{Re}\big(i\Lambda_4(-\xi_1^2 m(\xi_1)(-m(\xi_2)m(\xi_3)m(\xi_4) + m(\xi_2+\xi_3+\xi_4)); v, \overline{v}, v, \overline{v}) \\
&\quad + i\Lambda_4(m(\xi_4)^2, v, \overline{v}, v, \overline{v}) \\
&\quad - i\Lambda_6(m(\xi_1+\xi_2+\xi_3)m(\xi_4)m(\xi_5)m(\xi_6); v, \overline{v}, v, \overline{v}, v, \overline{v})\big) \\
=&\mathrm{Re}\big(C_1\Lambda_4(M_4'; v, \overline{v}, v, \overline{v}) + C_2\Lambda_4(M_4''; v, \overline{v}, v, \overline{v}) + C_3\Lambda_6(M_6; v, \overline{v}, v, \overline{v}, v, \overline{v})\big)
\end{aligned}$$

where C_1, C_2, C_3 are imaginary constants and

$$\begin{aligned}
M_4' &:= \xi_1^2 m(\xi_1)\big(m(\xi_2)m(\xi_3)m(\xi_4) - m(\xi_2+\xi_3+\xi_4)\big) \\
M_4'' &:= m(\xi_1)^2 - m(\xi_2)^2 + m(\xi_3)^2 - m(\xi_4)^2 \\
M_6 &:= m(\xi_1+\xi_2+\xi_3)m(\xi_4)m(\xi_5)m(\xi_6) - m(\xi_1)m(\xi_2)m(\xi_3)m(\xi_4+\xi_5+\xi_6).
\end{aligned}$$

It thus suffices to show the bounds

$$\Big|\int_{t_0}^{t_0+\delta} \Lambda_4(M_4'; u_1(t), \overline{u}_2(t), u_3(t), \overline{u}_4(t)) \, dt\Big| \lesssim N^{-1} \prod_{i=1}^4 \|Iu_i\|_{1,1/2+} \tag{4.15}$$

$$\Big|\int_{t_0}^{t_0+\delta} \Lambda_4(M_4''; u_1(t), \overline{u}_2(t), u_3(t), \overline{u}_4(t)) \, dt\Big| \lesssim N^{-1} \prod_{i=1}^4 \|Iu_i\|_{1,1/2+} \tag{4.16}$$

$$\Big|\int_{t_0}^{t_0+\delta} \Lambda_6(M_6; u_1(t), \overline{u}_2(t), u_3(t), \overline{u}_4(t), u_5(t), \overline{u}_6(t)) \, dt\Big| \lesssim N^{-1} \prod_{i=1}^6 \|Iu_i\|_{1,1/2+}. \tag{4.17}$$

As before, we first restrict the Fourier support of u_i to the region $\langle \xi_i \rangle \sim N_i$ for some $N_i \geq 1$, and promise to sum in the N_i later.

- **Step 4(a): Proof of** (4.15)**.**

We first prove (4.15). We shall assume that $N_2 \geq N_3 \geq N_4$; the other cases are similar. We may then assume $N_2 \gtrsim N$ since the symbol vanishes otherwise.

We split into two cases: $N_2 \gg N_3$ and $N_2 \sim N_3$.

- **Case 4(a).1:** $N_2 \gg N_3$**.**

We may assume that $N_1 \sim N_2$ since $\xi_1+\xi_2+\xi_3+\xi_4 = 0$. We then bound the symbol M_4' by $N_1^2 m(N_1)m(N_2) \sim N_1 m(N_1) N_2 m(N_2)$, and estimate this contribution by

$$N_1 m(N_1) N_2 m(N_2) \int\int |u_1||u_2||u_3||u_4|.$$

Applying Cauchy-Schwartz followed by two applications of (2.4) we obtain a bound of

$$N_1^{1/2} m(N_1) N_2^{1/2} m(N_2) \prod_{i=1}^{4} \|u_i\|_{0,1/2+} \lesssim N_1^{-1/2} N_2^{-1/2} \prod_{i=1}^{4} \|Iu_i\|_{1,1/2+}.$$

Summing in the N_i we see that this case is acceptable (we lose some logarithms of N_1 and N_2 from the N_3 and N_4 summation, which is why we only end up with a bound of N^{-1+} instead of N^{-1}).

- **Case 4(a).2:** $N_2 \sim N_3$**.**

Since $\langle \xi_2+\xi_3+\xi_4 \rangle \sim N_1$, we may bound the symbol M_4' by $N_1^2 m(N_1)^2$, and estimate the contribution by

$$N_1^2 m(N_1)^2 \int\int |u_1||u_2||u_3||u_4|.$$

We bound $N_1^2 m(N_1)^2 \lesssim N_1 m(N_1) N_2 m(N_2)$ and apply (2.2) four times to bound this by

$$N_1 m(N_1) N_2 m(N_2) \prod_{i=1}^{4} \|u_i\|_{0,1/2+} \lesssim \frac{1}{N_3 m(N_3)} \prod_{i=1}^{4} \|Iu_i\|_{1,1/2+}.$$

Summing in the N_i we see that this case is acceptable (again, the N_1 and N_4 summations cost us some logarithms).

- **Step 4(b): Proof of** (4.16) **and** (4.17)**.**

Now we prove (4.16), which is rather easy due to the lack of derivatives in the symbol. We may assume that $N_1 \geq N_2, N_3, N_4$. We may assume that $N_1 \gtrsim N$ since the symbol vanishes otherwise. Then if we bound the symbol M_4'' by $O(1)$ and use (2.2) we estimate this term by

$$\prod_{i=1}^{4} \|u_i\|_{0,1/2+} \lesssim \frac{1}{N_1 m(N_1)} \prod_{i=1}^{4} \|Iu_i\|_{1,1/2+}.$$

Summing in the N_i we see that this case is acceptable, again losing some logarithms in the N_2, N_3, N_4 summations. (Indeed, it is clear one could extract far more decay from this term if desired). Finally, the estimate (4.17) is similar to (4.16) (just use (2.1) instead of (2.2)).

This concludes the proof of Lemma 4.2, hence of Lemma 4.1, which gives a weak version of Theorem 1.2.

5. Proof of Theorem 1.2: second attempt

In the previous section we gave a partial proof of Theorem 1.2, but with the wrong power of $\mathrm{dist}_{H^s}(u_0, \Sigma)$. The problem was that we were not really exploiting the fact that $u(t)$ was close to the ground state cylinder Σ; for instance, the bound (4.10) would still be true if u was at a distance ~ 1 from the ground state cylinder in the $\|Iu\|_{H^1}$ metric. To improve upon these results we must consider not just $u(t)$, but also the difference $w(t) := u(t) - Q(t)$ between u and an appropriate ground state $Q(t) \in \Sigma$. In particular we wish to exploit the fact that w has small norm (with a bound which depends linearly on σ).

In this section we use the above ideas to refine the argument of the previous section, and obtain a near miss to (1.7). Unfortunately our power of $\mathrm{dist}_{H^s}(u_0, \Sigma)$ will still be off by an epsilon, mainly because the smoothing operator I does not quite preserve the ground state cylinder (so that $IQ(t)$ is not a ground state). To fix this problem and get the sharp power of $\mathrm{dist}_{H^s}(u_0, \Sigma)$ requires some further refinements which we delay until the next section in order to simplify the exposition.

We now turn to our second attempt at proving Theorem 1.2, deliberately repeating much of the structure of the arguments from previous sections.

- **Step 0. Preliminaries; introduction of the modified energy.**

We make the same reductions as the previous section, and leave the definition of the modified energy $E_N(t)$ from (4.1) unchanged. The main difference is that we sharpen Lemma 4.1 to

Lemma 5.1. *If $t_0 \in \mathbf{R}$ is such that $E_N(t_0) \leq L(Q) + \tilde\sigma^2$ for some $N^{-C} < \tilde\sigma \ll 1$ for some arbitrary constant C, then we have*

$$E_N(t_0 + \delta) = E_N(t_0) + O(N^{-1+}\tilde\sigma^2) \tag{5.1}$$

where $\delta > 0$ is an absolute constant depending only on s.

As in the previous section, we do not know if the factor $O(N^{-1+})$ can be improved. The quadratic exponent $\tilde\sigma^2$ is probably sharp, however, since the derivative of $E_N(t_0)$ will contain terms which are quadratic in the difference $w(t) = u(t) - Q(t)$.

- **Step 1. Deduction of a slightly weakened form of (1.7) from Lemma 5.1.**

Assume Lemma 5.1 for the moment. We substitute this lemma, with $\tilde{\sigma} \sim N^{1-s}\sigma$, in place of Lemma 4.1 in Step 1 of the previous section; note that $\tilde{\sigma}$ is admissible by (4.3). The argument then proceeds as before, with (4.6) replaced by

$$1 \leq t \ll N$$

(and (4.7) is no longer needed). For this range of t we obtain (4.8). Optimizing in N subject to (4.3), (4.4), (4.6) we find that we may improve (4.9) to

$$\mathrm{dist}_{H^s}(u(t), \Sigma) \lesssim t^{1-s+}\sigma^{1-} \tag{5.2}$$

whenever

$$t \ll \sigma^{-\frac{1}{1-s}-}.$$

This is within an epsilon of (1.7). To remove this last epsilon we shall need a more refined argument, presented in the next section.

It remains to prove Lemma 5.1. This shall be done in the usual sequence of stages. The main difference is that, instead of controlling u, we shall control a difference $w(t) = u(t) - Q(t)$ between $u(t)$ and a suitably chosen ground state $Q(t)$. This will let us recover the powers of $\tilde{\sigma}$ in (5.1).

- **Step 2. Control w at time t_0.**

Fix t_0. By (1.6) there exists a ground state $Q_{t_0} \in \Sigma$ depending on t_0 such that

$$\|Iu(t_0) - Q_{t_0}\|_{H^1} \lesssim \tilde{\sigma}.$$

The above estimate asserts that $u(t_0)$ is in some sense close to Q_{t_0}. If $u(t_0)$ was in fact equal to Q_{t_0}, then the evolution of $u(t)$ would follow the curve of ground states[14] $Q(t) : \mathbf{R} \to \Sigma$ defined by

$$Q(t) := e^{i(t-t_0)}Q_{t_0}. \tag{5.3}$$

Thus it is natural to define $w(t) := u(t) - Q(t)$. By (4.2) and the previous we have

$$\|Iw(t_0)\|_{H^1} \lesssim \tilde{\sigma}, \tag{5.4}$$

where we have used the hypothesis that $\tilde{\sigma} \gtrsim N^{-C}$ for some C.

- **Step 3. Control w on the time interval $[t_0 - \delta, t_0 + \delta]$.**

The idea is now to run a local well posedness argument for w instead of u in order to gain the powers of $\tilde{\sigma}$ in (5.1). Specifically, we wish to obtain the spacetime estimate

$$\|Iw\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim \tilde{\sigma} \tag{5.5}$$

for some small absolute constants $0 < \delta \ll 1$ and $0 < \varepsilon \ll 1$.

To obtain this, we observe from (1.1) and (1.2) that w obeys the difference equation

$$iw_t + w_{xx} = -G(w(t), Q(t)) \tag{5.6}$$

[14]This heuristic is a little inaccurate because of the presence of the I, and also because the error $Iu(t_0) - Q_{t_0}$ can affect the modulation parameters of the approximating ground state. We address these issues in the next section, when we remove the epsilon losses which arise from the arguments in this section.

where G is the nonlinear expression

$$G(w(t), Q(t)) := F(Q(t) + w(t)) - F(Q(t)).$$

The exact form of G is not important, save for the fact that G is cubic in $Q(t)$, $\overline{Q(t)}$, w, $\overline{w}$, and is always at least linear in w, $\overline{w}$. By (5.4), (5.6) and the standard $X^{s,b}$ energy estimate (as in previous sections), we thus have

$$\|Iw\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim \tilde{\sigma} + \delta^{0+}\|IG(w)\|_{X^{1,-1/2+2\varepsilon}_{[t_0-\delta,t_0+\delta]}}.$$

By several applications of (4.12), and the observation that $Q(t)$, $\overline{Q(t)}$ are Schwartz in x and thus live in every (time-localized) $X^{s,b}$ space, we therefore have

$$\|Iw\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim \tilde{\sigma} + \delta^{0+}\|Iw\|_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} + \delta^{0+}\|Iw\|^3_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}}.$$

The claim (5.5) then follows if δ is small enough by standard continuity (or iteration) arguments.

- **Step 4. Control the increment of the modified energy.**

We now prove (5.1). By (4.14) it suffices to show that

$$\Omega(Q(t) + w(t)) = O(N^{-1+}\tilde{\sigma}^2).$$

We could do this by direct computation, but we present instead a simple argument (based on isolating the terms in $\Omega(Q(t)+w(t))$ which are linear in w) which obtains this bound as a nearly automatic consequence of Lemma 4.2.

To prove the above estimate it suffices by the hypothesis $\tilde{\sigma} \gtrsim N^{-C}$ to prove the more general bound

$$\Omega(Q(t) + k\frac{w(t)}{\tilde{\sigma}}) = O(N^{-1+}|k|^2) + O(N^{-C-1}|k|) \tag{5.7}$$

for any real number k such that $|k| \lesssim 1$.

To prove (5.7), first observe from (4.13) that the left-hand side of (5.7) is a polynomial $P(k)$ in k of degree at most 6 (with the coefficients depending on $Q(t)$, $w(t)$, σ, of course). From (4.13), (1.2), and integration by parts we see that the constant term of P is zero. Also, from Lemma 4.2 and (5.5) we see that the left-hand side of (5.7) is $O(N^{-1+})$ for all $|k| \lesssim 1$. Thus all the coefficients of $P(k)$ are $O(N^{-1+})$. To finish the argument it suffices to show that the linear term of $P(k)$ is $O(N^{-C-1})$. Equivalently, it suffices to show that the portion of $\Omega(Q(t) + w(t))$ which is linear in w, $\overline{w}$ is $O(N^{-C-1}\tilde{\sigma})$.

To show this we return to (4.14), and exploit the fact that $IQ(t)$ is nearly a minimizer of L. If we compute $E_N(t) = L(IQ(t) + Iw(t))$ using (1.4) and suppress terms which are not linear in w, $\overline{w}$, we obtain

$$\begin{aligned} E_N(t) =& 2\langle IQ(t)_x, Iw(t)_x\rangle + 2\langle IQ(t), Iw(t)\rangle \\ &- 2\langle F(IQ(t)), Iw(t)\rangle + \text{nonlinear terms} \end{aligned}$$

Integrating by parts and using (1.2), we obtain

$$E_N(t) = 2\langle w(t), R(t)\rangle + \text{nonlinear terms}$$

where

$$R(t) := I(IF(Q(t)) - F(IQ(t))).$$

Differentiating this in time using (5.6), and again suppressing nonlinear terms, we obtain

$$\begin{aligned}\partial_t E_N(t) =& 2\big\langle iw_{xx}(t) + 2iw(t)\overline{Q(t)}Q(t) + iQ(t)\overline{w}(t)Q(t), R(t)\big\rangle \\ &+ 2\langle w(t), \partial_t R(t)\rangle + \text{nonlinear terms}\end{aligned}$$

We can bound the linear terms by (for instance)

$$\lesssim \|w(t)\|_2(\|R(t)\|_{H^2} + \|\partial_t R(t)\|_2).$$

Because $Q(t)$ is Schwartz, the operator I is almost the identity on $Q(t)$ or $F(Q(t))$, and so it is easy to see that[15]

$$\|R(t)\|_{H^2} + \|\partial_t R(t)\|_2 \lesssim N^{-C-1}.$$

Also from (5.5) we have $\|w(t)\|_2 \lesssim \tilde{\sigma}$. Thus the linear part of $\Omega(Q(t) + w(t))$ is $O(N^{-C-1}\tilde{\sigma})$ as desired. This proves (5.1), which then almost gives (1.7).

6. Proof of Theorem 1.2: final argument

In the previous sections one had to assume (4.3) in order to use the approximation $Q \approx IQ$. Ultimately this assumption caused us to miss (1.7) by an epsilon. In order to avoid this loss we shall need to avoid using the approximation $Q \approx IQ$, at least in situations in which one does not have enough powers of σ in the estimates.

- **Step 0. Preliminaries; introduction of the modified energy.**

This step is the same as in the previous section, except that we need to replace the modified energy $L(Iu)$ from (4.1) by a slightly different quantity (because IQ is not quite a minimizer of L).

To motivate the argument we first recall the more precise statement of Weinstein's estimate (1.6):

Lemma 6.1. [27], [26] *Let $\tilde{Q} \in \Sigma$ be a ground state, and let $w \in H^1$ obey the two orthogonality conditions*[16]

$$\langle w, iF(\tilde{Q})\rangle = \langle w, \partial_x F(\tilde{Q})\rangle = 0.$$

Then, if the H^1 norm of w is sufficiently small, we have the coercivity estimate

$$L(\tilde{Q} + w) - L(\tilde{Q}) = L(\tilde{Q} + w) - L(Q) \sim \|w\|_{H^1}^2.$$

---

[15] Note that the time derivative ∂_t on $R(t)$ is not dangerous because the time-dependence of $Q(t)$ and $R(t)$ are given by a simple phase rotation e^{it}. In the next section however we shall need a more sophisticated time evolution for the approximating ground state $Q(t)$.

[16] If one writes $u := \tilde{Q} + w$, then these orthogonality conditions have the geometric interpretation that $\tilde{Q}$ is the closest ground state in Σ to u, as measured in H^1 norm.

We shall need to apply this lemma to estimate quantities of the form $L(Q(t)+Iw(t))$ (which will be our substitute for $L(Iu)$). Thus we shall require w to obey the orthogonality conditions[17]

$$\langle w(t), iIF(Q(t))\rangle = \langle w(t), \partial_x IF(Q(t))\rangle = 0 \tag{6.1}$$

for all times t.

To use Lemma 6.1 we thus need a decomposition $u = Q(t) + Iw(t)$ of the solution u which obeys (6.1). This is the purpose of the following Lemma.

Lemma 6.2. *If $u \in H^s(\mathbf{R})$ is such that $\mathrm{dist}_{H^s}(u, \Sigma) \ll N^{s-1}$, and N is sufficiently large depending on s, then we may decompose $u = \tilde{Q} + w$, where $\tilde{Q} \in \Sigma$ is a ground state, w obeys* (6.1)*, and*

$$\|Iw\|_{H^1} \lesssim N^{1-s} \mathrm{dist}_{H^s}(u, \Sigma). \tag{6.2}$$

Proof Define the metric d on H^s by

$$d(u,v) := \|I(u-v)\|_{H^1} = \left(\langle I(u-v), I(u-v)\rangle + \langle \partial_x I(u-v), \partial_x I(u-v)\rangle\right)^{1/2}.$$

Clearly we have

$$d(u, \Sigma) \lesssim N^{1-s} \mathrm{dist}_{H^s}(u, \Sigma) \ll 1.$$

We now claim that there exists a ground state Q' in Σ which minimizes $d(u, Q')$, so that $d(u, Q') = d(u, \Sigma)$. To see this, first observe (since u was assumed to be smooth and rapidly decreasing) that $\langle Iu, Ie^{i\theta}Q(x - x_0)\rangle \to 0$ as $x_0 \to \pm\infty$, and similarly for $\langle \partial_x Iu, \partial_x Ie^{i\theta}Q(x-x_0)\rangle$. By orthogonality we thus have

$$d(u, e^{i\theta}Q(x-x_0)) \to (\|Iu\|_{H^1}^2 + \|IQ\|_{H^1}^2)^{1/2} \gtrsim 1 \gg d(u, \Sigma)$$

as $x_0 \to \pm\infty$. Thus, in order to minimize $d(u, Q')$ on Σ, it suffices to restrict Q' to a compact subset of the cylinder Σ. By compactness and smoothness we then see that a minimizer Q' must exist.

Observe that the statement and conclusion of Lemma 6.2 is invariant under translations and modulations of u (and hence of $\tilde{Q}$ and w). By these invariances we may thus assume that the minimizer Q' is attained at $Q' = Q$, thus

$$d(u, Q) = d(u, \Sigma) \lesssim N^{1-s} \mathrm{dist}_{H^s}(u, \Sigma). \tag{6.3}$$

The tangent space of Σ at Q is spanned by iQ and Q_x. Differentiating

$$d(u,Q)^2 = \langle I(u-Q), I(u-Q)\rangle + \langle \partial_x I(u-Q), \partial_x I(u-Q)\rangle$$

in these directions, we thus see that

$$\langle I(u-Q), iIQ\rangle + \langle \partial_x I(u-Q), iIQ\rangle = 0;$$
$$\langle I(u-Q), IQ_x\rangle + \langle \partial_x I(u-Q), \partial_x IQ_x\rangle = 0.$$

Integrating by parts and using (1.2) we obtain

$$\langle u - Q, iI^2F(Q)\rangle = \langle u - Q, \partial_x I^2 F(Q)\rangle = 0. \tag{6.4}$$

[17]This trick of choosing $Q(t)$ to obey carefully selected orthogonality conditions is very common in stability analysis, see e.g. [22].

This is almost what we want, except that we have I^2 instead of I. To rectify this we shall need to use perturbation theory to shift and modulate the ground state slightly. In other words, we set $\tilde{Q} := e^{i\theta} Q(x - x_0)$ for some $|\theta|, |x_0| \ll 1$. Setting $w := u - \tilde{Q}$, we see from (6.4) that (6.1) becomes

$$\begin{aligned} \langle \tilde{Q} - Q, iIF(Q) \rangle &= \langle u - Q, i(I - I^2)F(Q) \rangle \\ \langle \tilde{Q} - Q, \partial_x IF(Q) \rangle &= \langle u - Q, \partial_x (I - I^2)F(Q) \rangle. \end{aligned} \tag{6.5}$$

Since $F(Q)$ is Schwartz, the right-hand sides are $O(N^{-100} d(u, Q)) = O(N^{-99} \mathrm{dist}_{H^s}(u, \Sigma))$. Expanding the left-hand sides to first order in θ, x_0, we see that

$$\langle \tilde{Q} - Q, iIF(Q) \rangle = \theta \langle iQ, iIF(Q) \rangle + x_0 \langle -Q_x, iIF(Q) \rangle + O(|\theta|^2 + |x_0|^2)$$

and

$$\langle \tilde{Q} - Q, \partial_x IF(Q) \rangle = \theta \langle iQ, \partial_x IF(Q) \rangle + x_0 \langle -Q_x, \partial_x IF(Q) \rangle + O(|\theta|^2 + |x_0|^2).$$

However an easy integration by parts using (4.2) gives the estimates

$$\begin{aligned} \langle iQ, iIF(Q) \rangle &= \|Q\|_4^4 + O(N^{-100}) \\ \langle -Q_x, iIF(Q) \rangle, \langle iQ, \partial_x IF(Q) \rangle &= O(N^{-100}) \\ \langle -Q_x, \partial_x IF(Q) \rangle &= -3 \int Q^2 Q_x^2 + O(N^{-100}). \end{aligned}$$

Thus if N is large enough, the left-hand side of (6.5) has a non-degenerate Jacobian at the point $(\theta, x_0) = (0, 0)$. By the inverse function theorem (or the contraction mapping theorem) we can thus solve (6.5) for some $\theta, x_0 = O(N^{-99} \mathrm{dist}_{H^s}(u, \Sigma))$. The condition (6.2) then follows from (6.3) (since the distance between Q and $\tilde{Q}$ is $O(N^{-99} \mathrm{dist}_{H^s}(u, \Sigma))$ in any reasonable norm). ■

Applying this Lemma at each time t we thus have ground states $Q(t)$ and a function $w(x, t)$ such that $u(t) = Q(t) + w(t)$ for all times t for which $\mathrm{dist}_{H^s}(u, \Sigma) \ll N^{s-1}$. (We can ignore this latter condition by standard continuity arguments, since we will eventually verify this hypothesis (with some room to spare) at the end of the argument).

We now redefine the modified energy $E_N(t)$ as

$$E_N(t) := L(Q(t) + Iw(t)). \tag{6.6}$$

This should be compared to the energy $L(IQ(t) + Iw(t))$ from (4.1) used in the previous section. From Lemma 6.1 we have the analogue of (1.6)

$$E_N(t) - L(Q) \sim \|Iw(t)\|_{H^1}^2 \tag{6.7}$$

whenever $\|Iw(t)\|_{H^1}$ is sufficiently small. In particular, at time zero we see from Lemma 6.2 that

$$E_N(0) - L(Q) \sim \|Iw(0)\|_{H^1}^2 \lesssim N^{2-2s} \sigma^2.$$

The analogue of Lemma 5.1 is

Lemma 6.3. *If $t_0 \in \mathbf{R}$ is such that $E_N(t_0) \leq L(Q) + \tilde{\sigma}^2$ for some $0 < \tilde{\sigma} \ll 1$, then (5.1) holds for some $\delta > 0$ which is an absolute constant depending only on s.*

- **Step 1. Deduction of** (1.7) **from Lemma 6.3.**

We repeat Step 1 of the previous section, using (6.7) as a substitute for (1.6). The main differences are that there is no hypothesis of the form $\tilde{\sigma} > N^{-C}$ in Lemma 6.3, so that there is no need to assume (4.3). We omit the details.

Note that while there is some logarithmic losses in the $O(N^{-1+})$ factor of (5.1), there are none in the $\tilde{\sigma}^2$ factor. This is what allows us to get the sharp power of $\mathrm{dist}_{H^s}(u_0, \Sigma)$ in (1.7), although the power t^{1-s+} is probably not sharp.

It remains to prove Lemma 6.3, which we divide into the now-familiar sequence of steps.

- **Step 2. Control w at time t_0.**

We now prove Lemma 6.3. Fix t_0. By (6.7) we again obtain (5.4).

- **Step 3. Control w on the time interval $[t_0 - \delta, t_0 + \delta]$.**

As before, the next step is to obtain (5.5). It will be convenient to write $Q(t)$ more explicitly as

$$Q(x,t) =: e^{i\theta(t)} e^{it} Q(x - x_0(t)) \tag{6.8}$$

where the modulation parameters $\theta(t)$, $x_0(t)$ are real-valued; we shall shortly derive equations for the evolution of these parameters. This ansatz should be compared with (5.3); note that the standard ground state $u(t) = e^{it}Q$ occurs when θ, x_0, w all identically vanish.

From (1.1), (1.2), (6.8) we see that w obeys the difference equation

$$\begin{aligned} iw_t + w_{xx} &= -G(w(t), Q(t)) - Q(t) - iQ(t)_t \\ &= -G(w(t), Q(t)) + \dot{\theta}(t)Q(t) + i\dot{x}_0(t)\partial_x Q(t) \end{aligned} \tag{6.9}$$

where G was defined in the previous section. We may thus repeat Step 3 of the previous section, provided that we can show the estimate

$$\|\dot{\theta}(t)Q(t) + i\dot{x}_0(t)\partial_x Q(t)\|_{X^{1,-1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}} \lesssim \|Iw\|_{X^{1,1/2+}_{[t_0-\delta,t_0+\delta]}}.$$

Since $Q(t)$ is Schwartz, it suffices to show that

$$|\dot{\theta}(t)|, |\dot{x}_0(t)| \lesssim \|w(t)\|_{H^s} \lesssim \|Iw(t)\|_{H^1}. \tag{6.10}$$

To do this, we argue in a manner reminiscent of the proof of Lemma 6.2. We first introduce the renormalized function $\tilde{w}(t)$ defined by

$$\tilde{w}(x,t) := e^{-i\theta(t)} e^{-it} w(x + x_0(t)).$$

From (6.9) and (6.8) we see that $\tilde{w}$ evolves according to the equation

$$i\tilde{w}_t + \tilde{w}_{xx} = -G(\tilde{w}, Q) + \dot{\theta}(t)(Q + \tilde{w}) + \tilde{w} + \dot{x}_0(t)\partial_x(Q + \tilde{w}). \tag{6.11}$$

On the other hand, from (6.1) we have the orthogonality relations

$$\langle \tilde{w}(t), A_j \rangle = 0 \text{ for } j = 0, 1,$$

where A_0, A_1 are the Schwartz functions

$$A_0 := I(iQ^3); \quad A_1 := I\partial_x(Q^3).$$

Let $j = 0, 1$. Differentiating the previous in time and then applying (6.11) we obtain

$$-\langle i\tilde{w}_{xx}, A_j \rangle = -\dot{\theta}(t)\langle i(Q+\tilde{w}), A_j \rangle + \langle i\tilde{w}, A_j \rangle + \dot{x}_0(t)\langle \partial_x(Q+\tilde{w})A_j \rangle - \langle iG(\tilde{w}, Q), A_j \rangle$$

which we rewrite as

$$\langle i(Q+\tilde{w}), A_j \rangle \dot{\theta}(t) - \langle \partial_x(Q+\tilde{w}), A_j \rangle \dot{x}_0(t) = -\langle \tilde{w}, i(A_j)_{xx} \rangle + \langle G(\tilde{w}, Q), iA_j \rangle$$

This is a linear system of two equations $j = 0, 1$ and two unknowns $\dot{\theta}(t)$, $\dot{x}_0(t)$. To invert this system we first observe that the right hand side is $O(\|w\|_{H^s})$ since the A_j are Schwartz. Also, since $IQ - Q$ has norm $O(N^{-100})$ in any reasonable space, we have the coefficient estimates

$$\begin{aligned}
\langle i(Q+\tilde{w}), A_0 \rangle &= \langle iQ, iQ^3 \rangle + \langle i(IQ - Q), iQ^3 \rangle + O(\|\tilde{w}\|_{H^s}) \\
&= \|Q\|_4^4 + O(N^{-100}) + O(\|\tilde{w}\|_{H^s}) \\
\langle i(Q+\tilde{w}), A_1 \rangle &= \langle i\tilde{w}, A_1 \rangle + O(N^{-100}) \\
&= O(\|\tilde{w}\|_{H^s}) + O(N^{-100}) \\
\langle \partial_x(Q+\tilde{w}), A_0 \rangle &= -\langle \tilde{w}, \partial_x A_0 \rangle + O(N^{-100}) \\
&= O(\|\tilde{w}\|_{H^s}) + O(N^{-100}) \\
\langle \partial_x(Q+\tilde{w}), A_1 \rangle &= \langle \partial_x Q, \partial_x(Q^3) \rangle + \langle \partial_x(IQ - Q), \partial_x(Q^3) - \langle w, \partial_x A_1 \rangle \\
&= -\langle Q_{xx}, Q^3 \rangle + O(N^{-100}) + O(\|\tilde{w}\|_{H^s}).
\end{aligned}$$

Observe that the absolute constants $\|Q\|_4^4$ and $-\langle Q_{xx}, Q^3 \rangle = 3\int Q^2 Q_x^2$ are both non-zero. Thus if N is sufficiently large and $\|\tilde{w}\|_{H^s} = \|w\|_{H^s}$ is sufficiently small, we can invert the above 2×2 linear system and obtain the desired bounds (6.10). Note that this argument shows that one can make $\theta(t)$ and $x_0(t)$ differentiable in t.

This completes the proof of (5.5). In particular we have the bounds

$$\|w(t)\|_{H^s} \lesssim \|Iw(t)\|_{H^1} \lesssim \sigma \tag{6.12}$$

on the interval $[t_0 - \delta, t_0 + \delta]$, so from (6.10) we have

$$|\dot{\theta}(t)|, |\dot{x}_0(t)| \lesssim \tilde{\sigma}. \tag{6.13}$$

- **Step 4. Control the increment of the modified energy.**

We now prove (5.1) again (or more precisely, we prove (5.1) for the new definition (6.6) of the modified energy). We write

$$E_N(t) = L(e^{-it}Q(t) + e^{-it}Iw(t)).$$

We now use (1.5) to obtain

$$\partial_t E_N(t) = 2\langle Q(t)_t - iQ(t) + Iw_t - iIw, -\partial_{xx}(Q(t)+Iw) + (Q(t)+Iw) - F(Q(t)+Iw) \rangle.$$

We simplify the right factor using (1.2) to obtain

$$\partial_t E_N(t) = 2\langle Q(t)_t - iQ(t) + Iw_t - iIw, -Iw_{xx} + Iw - G(Iw, Q(t))\rangle.$$

From (6.8), (6.9) we have

$$\begin{aligned} Q(t)_t - iQ(t) + Iw_t - iIw &= i\dot\theta(t)(Q(t) - IQ(t)) - \dot x_0(t)\partial_x(Q(t) - IQ(t)) \\ &\quad + I(iw_{xx} - iw + iG(w, Q(t))). \end{aligned}$$

To show (5.1), it thus suffices by (3.8), (6.13) to show the bounds

$$\int_{t_0}^{t_0+\delta} \tilde\sigma |\langle Q(t) - IQ(t), -Iw_{xx} + Iw - G(Iw, Q(t))\rangle|\, dt \lesssim N^{-1+}\tilde\sigma^2 \tag{6.14}$$

$$\int_{t_0}^{t_0+\delta} \tilde\sigma |\langle \partial_x(Q(t) - IQ(t)), -Iw_{xx} + Iw - G(Iw, Q(t))\rangle|\, dt \lesssim N^{-1+}\tilde\sigma^2 \tag{6.15}$$

$$|\int_{t_0}^{t_0+\delta} \langle I(iw_{xx} - iw + iG(w, Q(t))), -Iw_{xx} + Iw - G(Iw, Q(t))\rangle\, dt| \lesssim N^{-1+}\tilde\sigma^2. \tag{6.16}$$

To prove (6.14), (6.15) we use integration by parts to move all the derivatives onto $Q(t) - IQ(t)$. This function has a norm of N^{-100} in any reasonable space. By (6.12) we thus see that these terms are acceptable.

We now prove (6.16). Again we could do this by direct computation, but we shall instead just use the work that we have already done in previous sections. By (5.5) it suffices to show

$$|\int_{t_0}^{t_0+\delta} \langle iIw_{xx} - iIw + iIG(w, Q(t)), -Iw_{xx} + Iw - G(Iw, Q(t))\rangle| \lesssim N^{-1+}\|Iw\|^2_{X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}}.$$

The left-hand side consists of multilinear expressions of order between 2 and 6 in w. Thus (cf. Step 4 of the previous section) it suffices to prove the estimate

$$|\int_{t_0}^{t_0+\delta} \langle iI\tilde w_{xx} - iI\tilde w + iIG(\tilde w, Q(t)), -I\tilde w_{xx} + I\tilde w - G(I\tilde w, Q(t))\rangle| \lesssim N^{-1+}$$

for all spacetime functions $\tilde w$ in the unit ball of $X^{1,1/2+\varepsilon}_{[t_0-\delta,t_0+\delta]}$. Note that the parameter $\tilde\sigma$ has totally disappeared, and so we are now able to lose factors of $O(N^{-100})$ as necessary.

We now convert the above expression into one which can be dealt with by Lemma 4.2. Fix $\tilde w$, and define the functions $v(t)$ by

$$v(t) := \tilde w(t) + Q(t).$$

From (1.2) we have

$$iI\tilde w_{xx} - iI\tilde w + iIG(\tilde w, Q(t)) = I(iv_{xx} - iv + iF(v)).$$

Similarly we have

$$-I\tilde w_{xx} + I\tilde w - G(I\tilde w, Q(t)) = -Iv_{xx} + Iv - F(Iv) + B(t)$$

where the error $B(t)$ is given by

$$B(t) := [F(I\tilde{w} + IQ(t)) - F(I\tilde{w} + Q(t))] - [IF(Q(t)) - F(Q(t))].$$

Recall that $IQ(t)$ is within $O(N^{-100})$ to $Q(t)$ in H^1 norm (say), and similarly for $IF(Q(t))$ and $F(Q(t))$. Since $I\tilde{w}$ is also bounded in H^1, we thus have

$$\|B(t)\|_{H^1} \lesssim N^{-100},$$

and so the contribution of $B(t)$ is easily seen to be acceptable. Thus it remains to show that

$$|\int_{t_0}^{t_0+\delta} \langle I(iv_{xx} + iv - iF(v), -Iv_{xx} + Iv - F(Iv)\rangle| \lesssim N^{-1+}.$$

From integration by parts we have the identity

$$\langle iIv, -Iv_{xx} + Iv - F(Iv)\rangle = 0$$

so it suffices to show

$$|\int_{t_0}^{t_0+\delta} \langle I(iv_{xx} - iF(v), -Iv_{xx} + Iv - F(Iv)\rangle| \lesssim N^{-1+}.$$

But this is immediate from Lemma 4.2 and (4.13). This finishes the proof of (5.1), and thus Theorem 1.2 is (finally!) completely proved.

References

[1] T. Cazenave, F.B. Weissler, *Critical nonlinear Schrödinger Equation*, Non. Anal. TMA, **14** (1990), 807–836.

[2] J. Bourgain, *Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations*, Part I, Geometric and Funct. Anal. **3** (1993), 107-156.

[3] J. Bourgain, *On the growth in time of higher order Sobolev norms of smooth solutions of Hamiltonian PDE*, IMRN **6** (1996), 277-304.

[4] J. Bourgain, *Refinements of Strichartz Inequality and Applications to 2d-NLS With Critical Nonlinearity*, Inter. Math. Res. Not., (1998), p. 253–284.

[5] J. Bourgain, *Scattering in the energy space and below for 3D NLS*, J. Anal. Math. **75** (1998), 267-297.

[6] J. Bourgain, *New global well-posedness results for non-linear Schrödinger equations*, AMS Publications, 1999.

[7] C.V. Coffman, *Uniqueness of the ground state solution for $\Delta u - u + u^3 = 0$ and a variational characterization of other solutions*, Arch. Rat. Mech. Anal. **46** (1972), 81–95.

[8] J. Colliander, J. Delort, C. Kenig, G. Staffilani, *Bilinear estimates and applications to 2D NLS*, Trans. Amer. Math. Soc. **353** (2001), 3307-3325.

[9] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Almost conservation laws and global rough solutions to a nonlinear Schrodinger equation*, submitted, Math. Res. Letters.

[10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Sharp global well-posedness for periodic and non-periodic KdV and mKdV on* **R** *and* **T**, to appear, J. Amer. Math. Soc.

[11] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Multilinear estimates for periodic KdV equations, and applications*, submitted, J. Funct. Anal.

[12] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Global well-posedness result for KdV in Sobolev spaces of negative index*, EJDE **2001** (2001) No 26, 1-7.

[13] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Global well-posedness for the Schrodinger equations with derivative*, Siam J. Math. Anal. **33** (2001), 649–669.

[14] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *A refined global well-posedness result for the Schrodinger equations with derivative*, to appear, Siam J. Math.

[15] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Polynomial growth and orbital instability bounds for L^2-subcritical NLS below the energy norm* , in preparation.

[16] J. Colliander, G. Staffilani, H. Takaoka, *Global wellposedness for KdV below L^2*, Math. Res. Lett. **6** (1999), no. 5-6, 755–778.

[17] J. Ginibre, *Le probléme de Cauchy pour des EDP semi-linéaires périodiques en variables d'espace*, Séminaire Bourbaki 1994/1995, Astérisque **237** (1996), Exp. 796, 163–187.

[18] M. Keel, T. Tao, *Endpoint Strichartz Estimates*, Amer. Math. J. 120 (1998), 955–980.

[19] M. Keel, T. Tao, *Local and global well-posedness of wave maps on $\mathbf{R}^{1+1}$ for rough data*, IMRN **21** (1998), 1117–1156.

[20] M. Keel, T. Tao, *Global well-posedness of the Maxwell-Klein-Gordon equation below the energy norm*, to appear.

[21] C. Kenig, G. Ponce, L. Vega, *On the ill-posedness of some canonical dispersive equations*, Duke Math. J. **3** (2001), 617–633.

[22] Y. Martel, F. Merle, *Stability of the blowup profile and lower bounds for blowup rate for the critical generalized KdV equation*, to appear, Annals of Math.

[23] K. Nakanishi, *Energy scattering for non-linear Klein-Gordon and Schrodinger equations in spatial dimensions 1 and 2*, JFA **169** (1999), 201–225.

[24] G. Staffilani, *On the growth of high Sobolev norms of solutions for KdV and Schrodinger equations*, Duke Math J. **86** (1997), 109-142.

[25] Y. Tsutsumi, *L^2 solutions for nonlinear Schrodinger equations and nonlinear groups*, Funk. Ekva. 30 (1987), 115-125.

[26] M. Weinstein, *Modulational stability of ground states of nonlinear Schrodinger equations*, SIAM J. Math. Anal. **16** (1985), 472-491.

[27] M. Weinstein, *Lyapunov stability of ground states of nonlinear dispersive equations*, CPAM **39** (1986), 51-68.

University of Toronto

University of Minnesota

Stanford University and Brown University

Hokkaido University

University of California, Los Angeles