J. Eur. Math. Soc. 17, 1107–1198

DOI 10.4171/JEMS/528

Yu Deng

Invariance of the Gibbs measure for the Benjamin–Ono equation

Dedicated to Sekai Saionji

Received December 5, 2012 and in revised form April 25, 2013

Abstract. In this paper we consider the periodic Benjamin–Ono equation. We establish the invariance of the Gibbs measure associated to this equation, thus answering a question raised in Tzvetkov [28]. As an intermediate step, we also obtain a local well-posedness result in Besov-type spaces rougher than L^2, extending the L^2 well-posedness result of Molinet [20].

Keywords. Benjamin–Ono equation, Gibbs measure, measure invariance, global well-posedness

1. Introduction

In this paper we study the periodic Benjamin–Ono equation

$$u_t + Hu_{xx} = uu_x, \qquad (t, x) \in I \times \mathbb{T}, \tag{1.1}$$

where I is a time interval and $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Also the Hilbert transform H is defined by $\widehat{Hu}(n) = -\mathrm{i} \cdot \mathrm{sgn}(n)\widehat{u}(n)$, where we understand that $\mathrm{sgn}(0) = 0$. Since equation (1.1), as well as the truncated versions to be introduced below, preserves both reality and the mean value of u, we shall assume throughout this paper that u is real-valued and has mean zero. Under this restriction, (1.1) is a Hamiltonian PDE with conserved energy

$$E[u] = \int_{\mathbb{T}} \left(\tfrac{1}{2}|\partial_x^{1/2}u|^2 - \tfrac{1}{6}u^3\right). \tag{1.2}$$

Being completely integrable, it also has an infinite number of conserved quantities at the level of $H^{\sigma/2}$ for $0 \le \sigma \in \mathbb{Z}$, including the L^2 mass.

We briefly summarize the relevant previous study of (1.1). First, the classical energy method yields local well-posedness in $H^\sigma(\mathbb{T})$ for $\sigma > 3/2$ (see [17]). By conservation laws, this implies global well-posedness in (say) H^2. In [24], Tao introduced a gauge transform to prove the well-posedness result in H^1 for the Euclidean counterpart of (1.1). This approach was then adapted by Molinet–Ribaud [21] to prove the H^1 well-posedness in the periodic case. Then Molinet [19], [20] further improved this result to $H^{1/2}$ and then L^2. For the Euclidean version we now also have well-posedness in L^2 (see Burq–Planchon [8] and Ionescu–Kenig [16]).

Y. Deng: Department of Mathematics, Princeton University, Princeton, NJ 08544, USA; e-mail: yudeng@math.princeton.edu

Mathematics Subject Classification (2010): Primary 37L40; Secondary 35Q53, 37K05, 37K10

Starting from the pioneering work of Lebowitz–Rose–Speer [18] and Bourgain [2], there has been considerable interest in constructing Gibbs measures for Hamiltonian PDEs and proving their invariance. From the dynamical system point of view, this provides the natural invariant measure for the system (which is the first step in studying this system's long time behavior); from the PDE aspect this is also important since it tells us exactly how rough a space can be so that we still have *strong* solutions for *generic* initial data. In this regard, such results can be viewed as variations on the theme of classical low-regularity well-posedness; see [12] for a general discussion about the notion of well-posedness in probabilistic sense. We list here several important results in this field: Bourgain [2, 3, 4, 6, 7], Burq–Thomann–Tzvetkov [9], Burq–Tzvetkov [10, 11], Colliander–Oh [13], Nahmod–Oh–Rey-Bellet–Staffilani [22], Tzvetkov [26, 27].

The study of (1.1) along these lines was initiated in Tzvetkov [28] where the Gibbs measure was rigorously constructed (see [28] for details; this construction is also reviewed in Section 4 below). In order to to prove its invariance, one has to construct global flow on its support; since this measure is supported in spaces rougher than L^2 (namely $L^2(\mathbb{T})$ has measure zero), the well-posedness result of Molinet [20] will not apply. Nevertheless, in [28, Section 5], the author made several important observations regarding the behavior of the gauge transform and second Picard iteration for random data, which suggest that global well-posedness and measure invariance may still hold despite the low regularity.

In the current paper we will solve this problem by establishing the invariance of the Gibbs measure. To be precise, we will construct an almost-surely defined (and unique) global flow for (1.1) in some Besov-type space Z_1 rougher than L^2, and prove that the Gibbs measure is kept invariant by this flow.

Remark 1.1. Very recently, Tzvetkov–Visciglia [29, 30, 31] have constructed (and proved the invariance of) weighted Gaussian measures associated to the conserved quantities of (1.1) at the level of $H^{\sigma/2}$ for $\sigma \geq 4$, and Deng–Tzvetkov–Visciglia [15] proved the $\sigma \in \{2, 3\}$ case. The case $\sigma = 0$ still seems out of reach with current techniques.

1.1. Notation and preliminaries

Throughout this paper, the standard notations, such as $\lesssim$, $\gtrsim$ and $O(*)$, will always be used in terms of absolute values. The Japanese bracket $\langle x\rangle$ will be $(1+|x|^2)^{1/2}$ and $\mathbb{N}$ will denote the set of nonnegative integers; the characteristic function of a set E is denoted by $\mathbf{1}_E$ and if E is finite, its cardinality is denoted by $\#E$. We will use $\mathbb{P}_*$ to denote (spatial) frequency projections; for example $\mathbb{P}_+$ (or $\mathbb{P}_{\leq 0}$) will be the projection onto strictly positive (or nonpositive) frequencies, and $\mathbb{P}_{\gtrsim\lambda}$ will be the projection onto frequencies with absolute value $\gtrsim \lambda$. We may use the same (Roman or Greek) letter in different places, but its meaning will be clear from the context.

Define $\mathcal{V}$ to be the space of distributions on $\mathbb{T}$ that are real-valued and have mean zero; in other words, $f \in \mathcal{V}$ if and only if $\widehat{f}(-n) = \overline{\widehat{f}(n)}$ and $\widehat{f}(0) = 0$. Let $\mathcal{V}_N$ be the subspace of $\mathcal{V}$ containing functions of frequency not exceeding N (so that $\mathcal{V}_N$ is identified with $\mathbb{R}^{2N}$), and $\mathcal{V}_N^\perp$ be its orthogonal complement. Let Π_N and $\Pi_N^\perp$ be the projections to the corresponding spaces; we actually have $\Pi_N = \mathbb{P}_{\leq N}$ and $\Pi_N^\perp = \mathbb{P}_{>N}$.

We use a parameter $s > 0$ that will be chosen sufficiently small. The large constants C and small constants c may depend on s; any situation in which they are independent of s will be easily recognized. We choose a few other parameters, namely $(p, r, b, \tau, q, \kappa, \gamma, \epsilon)$, as follows:

$$p = \frac{2}{1-2s} + s^2, \quad r = 1/2 - 1/p, \quad b = 1/2 - s^{15/8}, \quad \tau = 8 - s^{13/8},$$
$$q = 1 + s^{3/2}, \quad \kappa = 1 - s^{5/4}, \quad \gamma = 2 - s^{2.5}, \quad \epsilon = s^{7/4}.$$

When s is small enough, we have the following hierarchy of smallness factors:

$$\begin{aligned} s^3 &\ll 2-\gamma \ll r - s = 1/2 - 1/p - s \ll 1/2 - b \\ &\ll \epsilon \ll 8 - \tau \ll q - 1 \ll 1 - \kappa \ll s \ll s^{1/2}. \end{aligned} \tag{1.3}$$

In (1.3) each $\ll$ symbol connects two numbers that actually differ in scale by a power of s. We will also use $0+$ to denote some small positive number (whether it depends on s will be clear from the context); the meanings of $0-$, and $a+$, $a-$ are then obvious. Finally, using these parameters, we can define the space Z_1 by

$$\|f\|_{Z_1} = \sup_{d\geq 0}\Big(\sum_{n\sim 2^d} 2^{rpd}|\widehat{f}(n)|^p\Big)^{1/p}. \tag{1.4}$$

Note that we are including $n = 0$ when $d = 0$.

In addition to (1.1), we will introduce finite-dimensional truncations of it. Fix a smooth, even cutoff function ψ on $\mathbb{R}$ which equals 1 on $[-1/2, 1/2]$ and vanishes outside $[-3/4, 3/4]$. Let $1 - \psi = \psi_0$. For a positive integer N, we define the multiplier S_N by

$$\widehat{S_N f}(n) = \psi(n/N)\widehat{f}(n). \tag{1.5}$$

We also allow $N = \infty$, in which case $S_\infty = 1$. The truncated equations are then

$$u_t + Hu_{xx} = S_N(S_N u \cdot S_N u_x). \tag{1.6}$$

Notice that (1.6) conserves the L^2 mass of u; also, if u is a solution of (1.6) whose spatial Fourier transform $\widehat{u}(n)$ is supported in $|n| \leq N$ for one time t, then this automatically holds for all time.

1.2. *The main results, and major difficulties*

With these preparations, we can now state our main results. The most precise and detailed versions are somewhat technical, and will be postponed to Section 13.

Theorem 1.2 (Local well-posedness). *For any $A > 0$, let $T = T(A) = C^{-1}e^{-CA}$; then for the metric space $\mathcal{BO}^T$ in Definition 12.1 containing $B^0([-T, T] \to Z_1)$, which denotes the space of* bounded *functions on $[-T, T]$ valued in Z_1, we have the following. For any f with $\|f\|_{Z_1} \leq A$, there exists a unique function $u \in \mathcal{BO}^T$ such that $u(0) = f$, and u satisfies* (1.1) *on $[-T, T]$ in the sense of distributions (we may define uu_x as a distribution for all $u \in \mathcal{BO}^T$; for details see Remark* 12.2*). Moreover, if we write $u = \Phi f$, then the map Φ, from the ball $\{f : \|f\|_{Z_1} \leq A\}$ to the metric space $\mathcal{BO}^T$, will be a Lipschitz extension of the classical solution map for regular data, and its image is bounded away from the zero element in $\mathcal{BO}^T$ by Ce^{CA}.*

Theorem 1.3 (Measure invariance). *Recall the Gibbs measure ν on $\mathcal{V}$ defined in* [28]*, which is absolutely continuous with respect to a Wiener measure ρ (see Section* 4.1 *for details). There exists a subset Σ of $\mathcal{V}$ with full ρ measure such that for each $f \in \Sigma$, equation* (1.1) *has a unique solution $u \in \bigcap_{T>0} \mathcal{BO}^T$ (in the sense described in Remark* 12.2*) with initial data f. If we denote $u = \Phi f = (\Phi_t f)_t$, then for each $t \in \mathbb{R}$ we get a map $f \mapsto \Phi_t f$ from Σ to itself. These maps form a one-parameter group, and each of them keeps invariant the Gibbs measure ν.*

Since we are solving (1.1) in Z_1, we would like to know that the solution u is continuous in t with values in Z_1; this is not true. The discontinuity, which already exhibits the subtlety of (1.1) below L^2, is due to a modulation factor needed to eliminate one logarithmically growing term (see Section 7.2), and can be characterized explicitly.

Theorem 1.4. (1) *Let $u \in \mathcal{BO}^T$ be the local solution described in Theorem* 1.2*. Let $u_k(t)$ denote the k-th Fourier coefficient at time t and define*

$$\Delta_n(t) = \int_0^t \frac{1}{2} \sum_{k=0}^{n} |u_k(t')|^2 \, dt' \tag{1.7}$$

for $n > 0$ and extend it to be odd for $n \leq 0$. Then $\Delta_n(t)$ grows at most logarithmically with n, and the function u^*, defined by

$$(u^*)_n(t) = e^{-\mathrm{i}\Delta_n(t)} u_n(t)$$

for all time, is continuous in t with values in Z_1.

(2) *Let $f \in \Sigma$ and let u be the global solution described in Theorem* 1.3*. Let the function $u^\#$, real-valued and having mean zero, be defined by*

$$(u^\#)_n(t) = e^{-\frac{\mathrm{i}t \log n}{8\pi}} u_n(t)$$

for all $n > 0$ and t. Then $u^\#$ is continuous in t with values in Z_1.

The first step in solving (1.1) (see [24] or [20]) is to use the gauge transform to obtain a more favorable nonlinearity; this already becomes problematic with infinite L^2 mass. In fact, when we use the gauge $w = \mathbb{P}_+(ue^{-\frac{\mathrm{i}}{2}\partial_x^{-1}u})$ as in [20], the evolution equation satisfied by w would be

$$(\partial_t - \mathrm{i}\partial_{xx})w = \frac{\mathrm{i}}{2}\partial_x\mathbb{P}_+\big(\partial_x^{-1}w \cdot \partial_x\mathbb{P}_-(\overline{w}\partial_x^{-1}w)\big) + \frac{\mathrm{i}}{4}\mathbb{P}_0(u^2)w + GT, \tag{1.8}$$

where GT represents good terms. Here one can recognize the term $\mathbb{P}_0(u^2)w$ that can be infinite for $u \in Z_1$. However, when we further analyze the cubic term above, we find another contribution, namely the "resonant" one, which is basically some constant multiple of $\mathbb{P}_0(|w|^2)w$. It then turns out that the coefficients match exactly to give a multiple of $\|w\|_{L^2}^2 - \|\mathbb{P}_+ u\|_{L^2}^2$. Since (at least heuristically)

$$w = \mathbb{P}_+(ue^{-\frac{i}{2}\partial_x^{-1}u}) = \mathbb{P}_+ u \cdot e^{-\frac{i}{2}\partial_x^{-1}u} + GT, \tag{1.9}$$

this expression will be finite even if u is only in Z_1.

The next obstacle to local theory is the failure of standard multilinear $X^{s,b}$ estimates, which play a crucial role in [20]. Recall from (1.8) that a typical nonlinearity of the transformed equation looks like

$$\partial_x \mathbb{P}_+\big(\partial_x^{-1} w \cdot \partial_x \mathbb{P}_-(\overline{w}\partial_x^{-1} w)\big). \tag{1.10}$$

If the frequency of $\partial_x^{-1} w$ appearing in $\overline{w}\partial_x^{-1} w$ is low, we may pretend this frequency is zero, obtaining a quadratic nonlinearity which is similar to the KdV equation. In fact, there is a similar failure of bilinear estimates in solving the KdV equation below $H^{-1/2}$, which is necessary in proving the invariance of white noise. This problem was solved in [23] by considering the second iteration, a strategy already used in [5]. We will use the same method, though the fact that our nonlinearity is only quadratic "to the first order" makes the argument a little more involved.

There is also a special cubic term, omitted in (1.8), which involves the function $z = \mathbb{P}_-(ue^{-\frac{i}{2}\partial_x^{-1}u})$. Recall that it is w, not z, that satisfies a good evolution equation; therefore z is not supposed to be bounded in any $X^{s,b}$ space where s is close to 0 and b close to $1/2$ (note z is basically w multiplied by a smoother function, but $X^{s,b}$ spaces are not closed under such multiplications). In [20], Molinet introduced the space $X^{-1,7/8}$ to accommodate z (he actually considered u, but the estimates for z will be the same). In our case, not only do we need (a slightly different version of) this space, but we also have to introduce an atomic space characterizing, roughly speaking, how z is "shifted" from w; see Section 2.2 for details.

Passing from local theory to global well-posedness and measure invariance is another challenge. The only known method is to produce finite-dimensional truncations such as (1.5), exploit the invariance of the (finite-dimensional) truncated Gibbs measures, and use a limiting procedure to pass to the original equation. This requires, among other things, uniform estimates for solutions to (1.6). The major difficulty here is that the gauge transform of [20] is now inadequate for eliminating all bad interactions. To see this, recall that when $w = \mathbb{P}_+(Mu)$ with some function M, then

$$\begin{aligned}(\partial_t - i\partial_{xx})w &= \mathbb{P}_+[-2iM \cdot \partial_{xx}\mathbb{P}_- u + u \cdot (\partial_t - i\partial_{xx})M] \\ &\quad + \mathbb{P}_+[M \cdot S(Su \cdot Su_x) - 2iM_x u_x],\end{aligned}$$

where we assume u satisfies (1.6) with $S = S_N$. The terms in the first bracket enjoy a smoothing effect and are (more or less) easier to bound, and those in the second bracket will be most troublesome. If $S = 1$, this second bracket can be made zero by choosing $M = e^{-\frac{i}{2}\partial_x^{-1}u}$; but this is impossible when $S = S_N$ with N finite but large. However, note that we only need to eliminate the "high-low" interactions where the factor Su contributes very low frequency and Su_x contributes high frequency, and this is indeed possible if we replace multiplication by M with some carefully chosen operator defined from a combination of S (which is a Fourier multiplier) and suitable multiplication operators. See Section 5.1 for details.

Finally, in order for the limiting procedure to work, we must compare a solution to (1.6) with a solution to (1.1). Since $\psi(n/N)$ equals 1 only for $|n| \leq N/2$, the difference will contain some term involving factors like $\mathbb{P}_{\gtrsim N}u$, which does not decay for large N due to the l^∞ nature of our Z_1 norm. Nevertheless, these bad terms eventually add up to zero, at least to first order, which is enough for our analysis. Note that the bad terms involve ψ factors which are unique to (1.6) and are not found in (1.1); this cancellation is really something of a miracle. See Section 6 for details.

1.3. Plan of this paper

In Sections 2 and 3 we will define the spacetime norms needed in the proof, and prove some linear estimates as well as auxiliary results. In Section 4 we provide the basic probabilistic arguments. We next introduce the gauge transform for (1.6) and derive the new equations; these will occupy Sections 5–7. From Section 8 to Section 12, we will prove our main a priori estimates. Combining these estimates with the standard probabilistic arguments, we will prove in Section 13 our main results, which are (local and almost sure global) well-posedness for (1.1), invariance of Gibbs measure, and modified continuity.

2. Spacetime norms

2.1. The easier norms

For a function u defined on $\mathbb{R} \times \mathbb{T}$, we define its spacetime Fourier transform $\widehat{u}_{n,\widetilde{\xi}}$ by

$$u(t,x) = \sum_n \int_{\mathbb{R}} \widehat{u}_{n,\widetilde{\xi}} e^{\mathrm{i}(nx+\widetilde{\xi}t)} \, d\widetilde{\xi},$$

and denote $\widetilde{u}_{n,\xi} = \widehat{u}_{n,\widetilde{\xi}} := \widehat{u}_{n,\xi-|n|n}$. Thus we have three ways to represent u: $u(t,x)$ as a function of t and x, $\widehat{u}_{n,\widetilde{\xi}}$ as a function of n and $\widetilde{\xi}$, and $\widetilde{u}_{n,\xi}$ as a function of n and ξ, where the ξ and $\widetilde{\xi}$ are always related by $\widetilde{\xi} = \xi - |n|n$. Since we will be dealing with more than one function, n and ξ may be replaced with other letters possibly with subscripts, say m_1 or β_2. To simplify the notation, when there is no confusion, we will omit the "hat" and "tilde" symbols above u; for example, if we talk about an expression involving $u_{m,\widetilde{\alpha}}$, it will actually mean $\widehat{u}_{m,\widetilde{\alpha}}$. The appearance of functions f defined on $\mathbb{T}$ will not be too frequent, but when they do appear, we will adopt the same convention and write for example f_n instead of $\widehat{f}(n)$.

We will need a number of norms in our proof. As a general convention, when we write a norm as l^2L^1, this will mean the $l_n^2 L_\xi^1$ norm for some $\widetilde{u}$ (which equals the $l_n^2 L_{\widetilde{\xi}}^1$ norm for $\widehat{u}$); the meaning of $L^1 l^2$ will thus be clear. The spacetime Lebesgue norms will be denoted by L^6L^6 etc. For example, in this notation system the expression $\|u\|_{l^\infty_{d\geq 0} l^p_{\sim 2^d} L^1}$ actually means

$$\sup_{d\geq 0}\Big(\sum_{n\sim 2^d}\|\widetilde{u}_{n,\xi}\|^p_{L^1_{\xi}}\Big)^{1/p}.$$

Next, observe that up to a constant,

$$\|u\|^6_{L^6L^6}=\sum_{n\in\mathbb{Z}}\int_{\mathbb{R}}\bigg|\sum_{n_1+n_2+n_3=n}\int_{\widetilde{\xi}_1+\widetilde{\xi}_2+\widetilde{\xi}_3=\widetilde{\xi}}\prod_{i=1}^{3}u_{n_i,\widetilde{\xi}_i}\bigg|^2\,d\widetilde{\xi}. \tag{2.1}$$

It follows that if $|u_{n,\xi}|\leq v_{n,\xi}$, then $\|u\|_{L^6L^6}\leq\|v\|_{L^6L^6}$. For any function u we define $\mathfrak{N}u$ by $(\mathfrak{N}u)_{n,\xi}=|u_{n,\xi}|$; then $\|\mathfrak{N}u\|_{L^6L^6}$ is a norm of u. Now we list the norms we will use:

$$\|u\|_{X_1}=\|\langle n\rangle^s\langle\xi\rangle^b u\|_{l^pL^2}, \tag{2.2}$$

$$\|u\|_{X_2}=\|\langle n\rangle^r u\|_{l^\infty_{d\geq 0}l^p_{\sim 2^d}L^1}, \tag{2.3}$$

$$\|u\|_{X_3}=\|\langle n\rangle^{-\epsilon}\mathfrak{N}u\|_{L^6L^6}, \tag{2.4}$$

$$\|u\|_{X_4}=\|\langle n\rangle^{-1}\langle\xi\rangle^{\kappa}u\|_{l^\gamma L^2}, \tag{2.5}$$

$$\|u\|_{X_5}=\|u\|_{l^\infty_{d\geq 0}L^q l^2_{\sim 2^d}}, \tag{2.6}$$

$$\|u\|_{X_6}=\|\langle n\rangle^r\langle\xi\rangle^{1/2+s^2}u\|_{l^2L^2}, \tag{2.7}$$

$$\|u\|_{X_7}=\|\langle n\rangle^r\langle\xi\rangle^{1/8}u\|_{l^\infty_{d\geq 0}l^p_{\sim 2^d}L^2}. \tag{2.8}$$

We also recall the norm Z_1 defined in Section 1.1, and rewrite it as

$$\|f\|_{Z_1}=\|\langle n\rangle^r f\|_{l^\infty_{d\geq 0}l^p_{\sim 2^d}}. \tag{2.9}$$

2.2. *Another norm*

We will need *another* spacetime norm, denoted by X_8, which is a little tricky to define.

Consider the space of functions u of $(n,\xi)\in\mathbb{Z}\times\mathbb{R}$, normed by

$$\|u\|_{\Phi}=\|u\|_{L^ql^2}. \tag{2.10}$$

The additive group $\mathbb{Z}$ acts on this space by

$$(\pi_{n_0}u)(n,\xi)=u(n+n_0,\xi+|n+n_0|(n+n_0)-|n|n). \tag{2.11}$$

If we write

$$\mathcal{S}:\mathbb{Z}\times\mathbb{R}\to\mathbb{Z}\times\mathbb{R},\qquad(n,\widetilde{\xi})\mapsto(n,\xi)=(n,\widetilde{\xi}+|n|n), \tag{2.12}$$

then we would have

$$\pi_{n_0}u=u\circ\mathcal{S}\circ T_{n_0}\circ\mathcal{S}^{-1}, \tag{2.13}$$

where $T_{n_0}:\mathbb{Z}\times\mathbb{R}\to\mathbb{Z}\times\mathbb{R}$ is the translation $(n,\widetilde{\xi})\mapsto(n+n_0,\widetilde{\xi})$. We then define the atomic $\mathcal{Y}$ norm by

$$\|u\|_{\mathcal{Y}}=\inf\Big\{\sum_i\langle n_i\rangle^{s^{1/2}}|\alpha_i|:u=\sum_i\alpha_iu_i,\ \|\pi_{n_i}u_i\|_{\Phi}\leq 1\Big\}. \tag{2.14}$$

The X_8 norm is then defined by

$$\|u\|_{X_8} = \sup_{d\geq 0} \|\mathbb{P}_{\sim 2^d} u\|_{\mathcal{Y}} := \sup_{d\geq 0} \|\mathbb{P}_{\sim 2^d} \widetilde{u}\|_{\mathcal{Y}}, \tag{2.15}$$

where the last equality is due to our convention.

Remark 2.1. In (2.14), the convergence takes place in a suitable weighted $L^1_{n,\xi}$ space. Therefore, if v is rapidly decaying in n and ξ (for example, $|v| \lesssim (|n| + |\xi| + 1)^{-100}$ will suffice), the sum $\sum_i \alpha_i (u_i, v)$ will converge absolutely to (u, v) provided that $\sum_i \langle n_i\rangle^{s^{1/2}} |\alpha_i|$ is finite, where (u, v) denotes (up to a constant) the standard pairing

$$(u, v) = \int_{\mathbb{R}\times\mathbb{T}} u(t, x)\overline{v(t, x)}\, dt\, dx = \sum_n \int_{\mathbb{R}} u_{n,\xi} \overline{v_{n,\xi}}\, d\xi. \tag{2.16}$$

2.3. The space in which we work

Define

$$\|u\|_{Y_1} = \|u\|_{X_1} + \|u\|_{X_2} + \|u\|_{X_4} + \|u\|_{X_5} + \|u\|_{X_7}, \tag{2.17}$$
$$\|u\|_{Y_2} = \|u\|_{X_2} + \|u\|_{X_3} + \|u\|_{X_4} + \|u\|_{X_8}. \tag{2.18}$$

Here, for each space $\mathcal{Z}$ (which can be Y_1, Y_2 or any other space) we define

$$\|u\|_{\mathcal{Z}^T} = \inf\{\|v\|_{\mathcal{Z}} : v|_{[-T,T]} = u|_{[-T,T]}\}. \tag{2.19}$$

This $[-T, T]$ may also be replaced by any interval I.

The main spacetime norms we shall use in the whole bootstrap argument are Y_1^T and Y_2^T, while other norms may be introduced whenever necessary.

3. Linear estimates, and more

Here we shall prove our main linear estimates, as well as some auxiliary results.

Proposition 3.1 (Strichartz estimates). *For any function u, we have*

$$\|u\|_{L^k L^k} \lesssim \|\langle n\rangle^{\sigma} \langle\xi\rangle^{\beta} u\|_{l^2 L^2} \tag{3.1}$$

provided that

$$(k, \sigma, \beta) \in \{(2, 0, 0), (4, 0, 3/8), (6, s^5, 1/2 + s^5), (\infty, 1/2 + s^5, 1/2 + s^5)\}. \tag{3.2}$$

Proof. When $(k, \sigma, \beta) = (2, 0, 0)$, the inequality (3.1) is simply Plancherel; when instead $(k, \sigma, \beta) = (\infty, 1/2 + s^5, 1/2 + s^5)$, this can also be easily proved by combining Hausdorff–Young and Hölder. When $(k, \sigma, \beta) = (4, 0, 3/8)$, the inequality reduces, after separating positive and negative frequencies and using time inversion, to the L^4 Strichartz estimate for the linear Schrödinger equation on $\mathbb{T}$, which is well-known: see for example [1, Proposition 2.6] or [25, Proposition 2.13].

When $(k, \sigma, \beta) = (6, s^5, 1/2 + s^5)$, (3.1) basically reduces to the L^6 Strichartz estimate proved in [1, Proposition 2.36]; for the reader's convenience we also include a proof here. By separating positive and negative frequencies and using time inversion, we only need to consider the case for the Schrödinger semigroup, thus *in this proof* our convention will change to $\xi = \widetilde{\xi} + n^2$. Now for any function u with the right hand side of (3.1) not exceeding 1, we may write $v_{n,\xi} = \langle n\rangle^{s^5}\langle\xi\rangle^{1/2+s^5}u_{n,\xi}$ using our (different) convention, and compute up to a constant that

$$(u^3)_{n,\widetilde{\xi}} = \sum_{n_1+n_2+n_3=n}\prod_{i=1}^{3}\langle n_i\rangle^{-s^5}(f_{n_1} * f_{n_2} * f_{n_3})_{\widetilde{\xi}+n_1^2+n_2^2+n_3^2}, \tag{3.3}$$

where

$$(f_n)_\xi = \langle\xi\rangle^{-1/2-s^5}v_{n,\xi}. \tag{3.4}$$

By our assumption we have $\|\langle\xi\rangle^{1/2+s^5}f_{n_i}\|_{L^2} \lesssim A_{n_i}$, where $\{A_n\}$ is some sequence satisfying $\|A\|_{l^2} \lesssim 1$. By (the Fourier version of) the product estimate for $H^\sigma(\mathbb{R})$ spaces, we deduce that

$$(f_{n_1} * f_{n_2} * f_{n_3})_\eta = \langle\eta\rangle^{-1/2-s^5}(g_{n_1n_2n_3})_\eta, \quad \|g_{n_1n_2n_3}\|_{L^2} \lesssim A_{n_1}A_{n_2}A_{n_3}. \tag{3.5}$$

Therefore we can estimate

$$|(u^3)_{n,\widetilde{\xi}}|^2 \lesssim \Big(\sum_{n_1+n_2+n_3=n}\prod_{i=1}^{3}\langle n_i\rangle^{-2s^5}\cdot\langle\widetilde{\xi}+n_1^2+n_2^2+n_3^2\rangle^{-1-2s^5}\Big) \times \Big(\sum_{n_1+n_2+n_3=n}|(g_{n_1n_2n_3})_{\widetilde{\xi}+n_1^2+n_2^2+n_3^2}|^2\Big).$$

Now to finish the proof it will suffice to show

$$\sum_{n_1+n_2+n_3=n}\prod_{i=1}^{3}\langle n_i\rangle^{-2s^5}\langle\widetilde{\xi}+n_1^2+n_2^2+n_3^2\rangle^{-1-2s^5} \leq C \tag{3.6}$$

when n and $\widetilde{\xi}$ are fixed. Now suppose the maximum (in absolute value) of n_i and $\Xi = \widetilde{\xi} + n_1^2+n_2^2+n_3^2$ is comparable to 2^d, and $\Xi \sim 2^{d'}$; then the summand is at most $2^{-d'-s^6(d+d')}$, so it will suffice to show that there are at most $2^{d'+s^7d}$ choices for (n_1, n_2, n_3). Since their can be at most $2^{d'}$ possibilities for $n_1^2+n_2^2+n_3^2$, we only need to show that there are at most 2^{s^7d} choices for (n_1, n_2, n_3) if we require $|n_i| \lesssim 2^d$, and fix $n_1+n_2+n_3 = n$ and $n_1^2+n_2^2+n_3^2$. But then $m_i = 3n_i - n$ will be integers for $i \in \{1, 2\}$, and $m_1^2+m_1m_2+m_2^2$ will be a fixed integer not exceeding $C2^{5d}$. The result then follows from the divisor estimate for the ring $\mathbb{Z}[e^{2\pi i/3}]$. □

By Proposition 3.1 and interpolation, we get a series of L^kL^k Strichartz estimates for all $2 \leq k \leq \infty$. It is these that we will actually use in the proof; we will not care too much about the exact numerology because there will be enough room whenever we use these estimates.

Proposition 3.2 (Relations between norms). *We have*

$$\|u\|_{X_3} \lesssim \|u\|_{X_1} + \|u\|_{X_4}, \qquad \|u\|_{X_8} \lesssim \|u\|_{X_5}, \tag{3.7}$$
$$\|u\|_{X_1} + \|u\|_{X_2} + \|u\|_{X_5} + \|u\|_{X_7} \lesssim \|u\|_{X_6}. \tag{3.8}$$

This in particular implies $\|u\|_{X_j} \lesssim \|u\|_{Y_1}$ *if* $1 \leq j \leq 8$ *and* $j \neq 6$.

Proof. By Proposition 3.1 and hierarchy (1.3) we know that

$$\|u\|_{X_3} \lesssim \|\langle n\rangle^{-\epsilon/2}\langle\xi\rangle^{1/2+s^5}u\|_{l^2L^2}. \tag{3.9}$$

Comparing this with the definition of X_1 and X_4, noticing that $\gamma < 2$ and by (1.3) and Hölder,

$$\|u\|_{X_1} \gtrsim \|\langle n\rangle^{-\epsilon/4}\langle\xi\rangle^{b}\|_{l^2L^2},$$

we will be able to prove the first inequality in (3.7) provided we can show

$$\langle n\rangle^{-\epsilon/2}\langle\xi\rangle^{1/2+s^5} \lesssim \langle n\rangle^{-1}\langle\xi\rangle^{\kappa} + \langle n\rangle^{-\epsilon/4}\langle\xi\rangle^{b}. \tag{3.10}$$

But this is clear since by (1.3), the left hand side is controlled by the first term on the right hand side if $\langle\xi\rangle \geq \langle n\rangle^{100}$, and by the second term if $\langle\xi\rangle < \langle n\rangle^{100}$. The second inequality in (3.7) is also easy, since we only need to prove $\|u\|_{\mathcal{Y}} \lesssim \|u\|_{L^q l^2}$, which is a direct consequence of the definition (2.14), if we choose to have only one term (with the corresponding $n_i = 0$) in the proposed atomic decomposition.

Now let us prove (3.8). The X_1 norm is controlled by the X_6 norm because $s < r$, $b < 1/2 + s^2$, and $2 < p$. For basically the same reason we can use Hölder to show $\|u\|_{X_2} + \|u\|_{X_7} \lesssim \|u\|_{X_6}$. Finally, to prove $\|u\|_{X_5} \lesssim \|u\|_{X_6}$, we only need to show that $\|g_\xi\|_{L^q} \lesssim \|\langle\xi\rangle^{1/2+s^2}g_\xi\|_{L^2}$, but this again follows from Hölder since $q > 1$. □

Next, we introduce the (cutoff) Duhamel operator $\mathcal{E}$ defined by

$$\mathcal{E}u(t,x) = \chi(t)\int_0^t \chi(t')\big(e^{-(t-t')H\partial_{xx}}u(t')\big)(x)\,dt', \tag{3.11}$$

where $\chi(t)$ is a cutoff function (compactly supported and equal to 1 in a neighborhood of 0). Here and below we shall use many such functions, but unless really necessary, we will not distinguish them and will denote them all by χ (for example, we write $\chi^2 = \chi$). We shall summarize the required linear estimates for $\mathcal{E}$ in Proposition 3.4 below, but before doing so, we need to introduce two more norms:

$$\|u\|_{X_9} = \|\langle n\rangle^{r}u\|_{l^\infty_{d\geq 0}L^{q'}l^p_{\sim 2^d}}, \tag{3.12}$$
$$\|u\|_{X_{10}} = \|\langle n\rangle^{r}\langle\xi\rangle^{-1/8}u\|_{l^\infty_{d\geq 0}L^{\tau}l^p_{\sim 2^d}}. \tag{3.13}$$

Lemma 3.3. *Suppose $v(t,x) = \mathcal{E}u(t,x)$. Then with constants c_j,*

$$v_{n,\xi} = c_1(\widehat{\chi} * (\eta^{-1}(\widehat{\chi} * u_{n,*})_\eta))_\xi + c_2\left(\int_{\mathbb{R}} \frac{(\widehat{\chi} * u_{n,*})_\eta}{\eta}\,d\eta\right) \cdot \widehat{\chi}_\xi. \tag{3.14}$$

Here the $1/\eta$ is to be understood as the principal value distribution. This operator obeys the following basic estimates, valid for all $\sigma, \beta \in \mathbb{R}$ and $1 \le h, k \le \infty$:

$$\|\langle n\rangle^\sigma \langle\xi\rangle^\beta \mathcal{E}u\|_{L^h l^k} \lesssim \|\langle n\rangle^\sigma \langle\xi\rangle^{\beta-1} u\|_{L^h l^k} + \|\langle n\rangle^\sigma \langle\xi\rangle^{-1} u\|_{l^k L^1}, \tag{3.15}$$
$$\|\langle n\rangle^\sigma \langle\xi\rangle^\beta \mathcal{E}u\|_{l^k L^h} \lesssim \|\langle n\rangle^\sigma \langle\xi\rangle^{\beta-1} u\|_{l^k L^h} + \|\langle n\rangle^\sigma \langle\xi\rangle^{-1} u\|_{l^k L^1}. \tag{3.16}$$

Note the reversed order of norms in the second term on the right hand side of (3.15). If moreover $\beta > 1 - 1/h$, we can remove the $l^k L^1$ norms. Finally, by commuting with $\mathbb{P}$ projections, we get similar estimates for norms like X_2 and X_5.

Proof. The computation (3.14) is basically done in [5]. In our case, noticing that multiplication by $\chi(t)$ corresponds to convolution with $\widehat{\chi}$ on the "tilde" side, we only need to express the Fourier transform of $\int_0^t u(t')\,dt'$ (which is exactly the Duhamel operator on the "tilde" side) in terms of $u(t)$. We compute

$$\int_0^t u(t')\,dt' = \frac{1}{2} u * \operatorname{sgn}(t) + \frac{1}{2}\int_{\mathbb{R}} u(t')\operatorname{sgn}(t')\,dt'. \tag{3.17}$$

On the Fourier side, these two terms give exactly the two terms in (3.14) after another convolution with $\widehat{\chi}$.

We will only prove (3.15), since the proof of (3.16) is basically the same; also notice that if $\beta > 1 - 1/h$, then

$$\|w\|_{l^k L^1} = \min\{\|w\|_{l^k L^1}, \|w\|_{L^1 l^k}\} \lesssim \min\{\|\langle\xi\rangle^\beta w\|_{l^k L^h}, \|\langle\xi\rangle^\beta w\|_{L^h l^k}\}$$

for $w_{n,\xi} = \langle n\rangle^\sigma \langle\xi\rangle^{-1} u_{n,\xi}$, by Hölder. Now to prove (3.15), we first consider the second term of (3.14). Due to its structure, we only need to prove for any function $z = z_\xi$ that

$$\left|\int_{\mathbb{R}} \frac{(z * \widehat{\chi})_\eta}{\eta}\,d\eta\right| \lesssim \|\langle\xi\rangle^{-1} z\|_{L^1}. \tag{3.18}$$

By considering $|\eta| \gtrsim 1$ and $|\eta| \lesssim 1$ separately and using the cancelation coming from the $1/\eta$ factor, we can control the left hand side by $\|\langle\eta\rangle^{-1}(z * \widehat{\chi})_\eta\|_{L^1}$ (which is easily bounded by the right hand side of (3.18)), plus another term bounded by $\|\langle\eta\rangle^{-1}\partial_\eta(z * \widehat{\chi})\|_{L^\infty}$. If we shift the derivative to $\widehat{\chi}$ to get rid of it, we can again bound this expression by the right hand side of (3.18).

Next, we consider the first term of (3.14). Again we consider the terms with $|\eta| \gtrsim 1$ and $|\eta| \lesssim 1$ separately (by introducing a smooth, even cutoff, ϕ_η say). The part where $|\eta| \gtrsim 1$ is easy, since convolution with $\widehat{\chi}_\xi$ is bounded on any weighted mixed norm Lebesgue space we have here, and $1/\eta$ is comparable to $\langle\eta\rangle^{-1}$ when restricted to the region $|\eta| \gtrsim 1$. Now for the region $|\eta| \lesssim 1$, we can actually prove for $y = y_\xi$ and arbitrary $K > 0$ that

$$\left|\left(\widehat{\chi} * \left(\frac{\phi_\eta}{\eta}(\widehat{\chi} * y)_\eta\right)\right)_\tau\right| \lesssim \langle\tau\rangle^{-K}\|\langle\xi\rangle^{-K} y\|_{L^1}, \tag{3.19}$$

which easily implies our inequality. To prove this, let $\widehat{\chi} * y = z$, and compute

$$\left(\widehat{\chi} * \left(\frac{\phi(\eta)}{\eta} z_\eta\right)\right)_\tau = \int_{|\eta|\lesssim 1} \widehat{\chi}_\tau \frac{\phi_\eta z_\eta - z_0}{\eta}\, d\eta + \int_{|\eta|\lesssim 1} \frac{\widehat{\chi}_{\tau-\eta} - \widehat{\chi}_\tau}{\eta} \phi_\eta z_\eta\, d\eta.$$

From this we can readily recognize a decay of $\langle\tau\rangle^{-K}$, and it will suffice to prove that $\sup_{|\eta|\lesssim 1} |z_\eta| \lesssim \|\langle\xi\rangle^{-K} y\|_{L^1}$, but this will be clear from the definition of z. □

Proposition 3.4. *We have*

$$\|\mathcal{E}u\|_{X_6} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_6}, \quad \|\mathcal{E}u\|_{X_4} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_4} \tag{3.20}$$

$$\|\mathcal{E}u\|_{X_1} + \|\mathcal{E}u\|_{X_2} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_1} + \|\langle\xi\rangle^{-1}u\|_{X_2} \lesssim \|u\|_{X_{10}}, \tag{3.21}$$

$$\|\mathcal{E}u\|_{X_7} \lesssim \|u\|_{X_{10}} \lesssim \|u\|_{X_9}, \quad \|\mathcal{E}u\|_{X_5} \lesssim \|u\|_{X_{10}}. \tag{3.22}$$

Moreover, suppose u is such that $u_{n,\xi}$ is supported in $\{(n,\xi) : n \sim 2^d,\ \xi \gtrsim 2^d\}$ for some d. Then

$$\|\mathcal{E}u\|_{X_5} + \|\mathcal{E}u\|_{X_7} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_1} + \|\langle\xi\rangle^{-1}u\|_{X_2}. \tag{3.23}$$

Finally, notice that all these estimates naturally imply the dual versions for the boundedness of $\mathcal{E}'$.

Proof. By checking the numerology, we see that (3.20) is a direct consequence of Lemma 3.3. To prove the first inequality in (3.21), we use Lemma 3.3 to conclude

$$\|\mathcal{E}u\|_{X_1} + \|\mathcal{E}u\|_{X_2} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_1} + \|\langle\xi\rangle^{-1}u\|_{X_2} + \|\langle n\rangle^s\langle\xi\rangle^{-1}u\|_{l^pL^1}, \tag{3.24}$$

and note that the last term can be controlled by $\|\langle\xi\rangle^{-1}u\|_{X_2}$ also. To prove that $\|\langle\xi\rangle^{-1}u\|_{X_2} \lesssim \|u\|_{X_{10}}$, one first commutes with $\mathbb{P}_{\sim 2^d}$, then controls the l^pL^1 norm by the L^1l^p norm, then uses Hölder (note the hierarchy (1.3)). To prove that $\|\langle\xi\rangle^{-1}u\|_{X_1} \lesssim \|u\|_{X_{10}}$, one first replaces the $\|\langle n\rangle^s * \|_{l^p}$ norm by the larger $\|\langle n\rangle^r * \|_{l_{d\geq 0} l^p_{n\sim 2^d}}$ norm, then commutes with $\mathbb{P}_{\sim 2^d}$, and controls the l^pL^2 norm by the L^2l^p norm and uses Hölder again. Along the same lines,

$$\|\mathcal{E}u\|_{X_7} \lesssim \|\langle\xi\rangle^{-1}u\|_{X_2} + \|\langle\xi\rangle^{-1}u\|_{X_7}, \tag{3.25}$$

$$\|\mathcal{E}u\|_{X_5} \lesssim \|\langle\xi\rangle^{-1}u\|_{l^\infty_{d\geq 0} l^2_{\sim 2^d} L^1} + \|\langle\xi\rangle^{-1}u\|_{X_5}, \tag{3.26}$$

where the first term on the right hand side of (3.26) is bounded by $\|\langle\xi\rangle^{-1}u\|_{X_2}$, and the second terms on both right hand sides are bounded by the X_{10} norm, by controlling the l^pL^2 norm by the L^2l^p norm and using Hölder. Also $\|u\|_{X_{10}} \lesssim \|u\|_{X_9}$ by Hölder. This proves (3.22).

Let us now prove (3.23). For the X_7 norm we use (3.25), and the support condition will easily allow us to control the second term on the right hand side of (3.25) by $\|\langle\xi\rangle^{-1}u\|_{X_1}$. For the X_5 norm, we only need to bound the second term on the right hand side of (3.26) by $\|\langle\xi\rangle^{-1}u\|_{X_1}$. Since we can restrict to $|n|\sim 2^d$ and $|\xi|\gtrsim 2^d$, we can bound this term by

$$\begin{aligned}\|\langle\xi\rangle^{-1}u\|_{L^q l^2} &\lesssim \|\langle\xi\rangle^{\sigma}u\|_{L^2 l^2} = \|\langle\xi\rangle^{\sigma}u\|_{l^2 L^2}\\ &\lesssim 2^{(\sigma-b+1)d}\|\langle\xi\rangle^{b-1}u\|_{l^2L^2} \lesssim 2^{(\sigma+\sigma'-b+1)d}\|\langle\xi\rangle^{-1}u\|_{X_1},\end{aligned}$$

where $\sigma' = 1/2-1/p-s>0$, $\sigma=-1/2-1/(2q')$ so that $\sigma+\sigma'-b+1<0$ by (1.3). □

Next we will prove two auxiliary results about our norms Y_j and Y_j^T, which are defined in Section 2.3. They will be used to validate our main bootstrap argument.

Proposition 3.5. *Suppose $j\in\{1,2\}$, and $u=u(t,x)\in Y_j$ is a function that vanishes at $t=0$. Then with a time cutoff χ (recall our convention about such functions) we have, uniformly in $T\lesssim 1$,*

$$\|\chi(T^{-1}t)u\|_{Y_j}\lesssim \|u\|_{Y_j}. \tag{3.27}$$

If u is smooth, then also

$$\lim_{T\to 0}\|\chi(T^{-1}t)u\|_{Y_j}=0. \tag{3.28}$$

Proof. We first assume $u\in Y_j$ and $u(0)=0$. We may also assume that u is supported in $|t|\lesssim 1$. Since on the "hat" or "tilde" side multiplication by $\chi(T^{-1}t)$ is just convolution with $T\widehat{\chi}_{T\xi}$, we need to prove the uniform boundedness of these operators on spaces involved in the definition of Y_j, as well as the corresponding limit result when u is smooth. The bound in X_3 is obtained by decomposing this convolution into translations (which preserve the X_3 norm) and integrating them using the boundedness of the L^1 norm of $T\widehat{\chi}_{T\xi}$. The bound in X_8 follows from the bound in $\mathcal{Y}$, which is valid because this convolution does not increase the Φ (or $L^q l^2$) norm, and commutes with the action described in Section 2.2; the bounds in X_2 and X_5 are shown in the same way. The remaining bounds will follow if we can bound this convolution in weighted norms $\|\langle\xi\rangle^{\sigma}y\|_{L^2}$, where $0\le\sigma<1$, for complex-valued functions y_ξ such that $\int_{\mathbb{R}} y_\xi\,d\xi=0$. Namely, we need to prove

$$\|\langle\eta\rangle^{\sigma}(y*T\widehat{\chi}_{T\xi})_\eta\|_{L^2}\lesssim \|\langle\xi\rangle^{\sigma}y_\xi\|_{L^2}. \tag{3.29}$$

Also, by Proposition 3.2 we can control the Y_1 and Y_2 norms by X_4 and X_6. Thus in order to prove (3.28), we only need to prove that the left hand side of (3.29) actually tends to zero when $T\to 0$, for any fixed Schwartz y with integral zero. By taking inverse Fourier transform, the problem can be reduced to proving

$$\|\chi(T^{-1}t)u\|_{H^\sigma}\lesssim \|u\|_{H^\sigma} \tag{3.30}$$

for $T\lesssim 1$, and the limit

$$\lim_{T\to 0}\|\chi(T^{-1}t)u(t)\|_{H^\sigma}=0, \tag{3.31}$$

for $u\in C_c^\infty$ such that $u(0)=0$. But these are proved, in a slightly different but equivalent setting, in [14, Lemma 2.8]. □

Proposition 3.6. *Suppose $u = u(t, x)$ is a smooth function defined on $\mathbb{R} \times \mathbb{T}$. Then for $j \in \{1, 2\}$, the function $T > 0 \mapsto \mathcal{M}(T) = \|u\|_{Y_j^T}$ satisfies $\mathcal{M}(T + 0) \leq C\mathcal{M}(T - 0)$ for all $0 < T \lesssim 1$, and also $\mathcal{M}(0+) \leq C\|u(0)\|_{Z_1}$.*

Proof. First we prove the estimate for $\mathcal{M}(0+)$. Let $u(0) = f$ and $v(t, x) = u(t, x) - e^{-tH\partial_{xx}} f(x)$. Since v is smooth and $v(0) = 0$, Proposition 3.5 shows that $\|v\|_{Y_j^T} \to 0$ when $T \to 0$. It then suffices to prove that for some cutoff $\chi(t)$, we have $\|\chi(t)e^{-tH\partial_{xx}} f\|_{Y_j} \lesssim \|f\|_{Z_1}$. Note that on the "tilde" side, the function $\chi(t)e^{-tH\partial_{xx}} f$ simply becomes $\widehat{\chi}_\xi f_n$; thus this inequality is basically trivial if we take into account that the Z_1 norm is stronger than the norm $\|\langle n\rangle^{-1} f\|_{L^\gamma}$, and the norm $\|f\|_{l^\infty_{d\geq 0} l^2_{\sim 2^d}}$.

Next, we shall prove that $\mathcal{M}(T + 0) \lesssim \mathcal{M}(T)$ for $0 < T \lesssim 1$. Namely, suppose u is a smooth function, and $0 < T \lesssim 1$ is such that $\|u\|_{Y_j^T} \leq 1$; we want to prove that $\|u\|_{Y_j^{T'}} \lesssim 1$ for some $T' > T$. Actually we only need to prove $\|u\|_{Y_j^{[-T,T']}} \lesssim 1$, since we can use the same argument to move the left point also. Now, due to the presence of X_2 norm in the definitions of both Y_j, our assumption implies $\|u(T)\|_{Z_1} \lesssim 1$, therefore by what we have just proved, $u_1 = e^{-(t-T)H\partial_{xx}} u(T)$ satisfies the estimate $\|u_1\|_{Y_j^{T'}} \lesssim 1$ for all $T < T' \lesssim 1$. Thus we only need to bound $\|u_2\|_{Y_j^{[-T,T']}}$ for some $T' > T$ and $u_2 = u - u_1$. Since $u_2(T) = 0$, by choosing δ small enough we can produce a function v coinciding with u_2 on $[T - 10\delta, T + 10\delta]$ such that $\|v\|_{Y_j} \lesssim 1$ by Proposition 3.5. Also since $\|u_2\|_{Y_j^T} \lesssim 1$, we may choose w coinciding with u_2 on $[-T, T]$ such that $\|w\|_{Y_j} \lesssim 1$. Note $v(T) = w(T) = 0$. Next, choose $\psi_1 \in C^\infty$ supported on $[-9, 10]$ that equals 1 on $[-1, 9]$. Define

$$u_3(t) = (1 - \psi_1(\delta^{-1}(t - T)))w(t) + \psi_1(\delta^{-1}(t - T))v(t). \tag{3.32}$$

Then we can verify that $u_3 = u_2$ on $[-T, T']$ with $T' = T + 9\delta$, and by Proposition 3.5 we have $\|u_3\|_{Y_j} \lesssim 1$, as desired.

Finally, let us prove that $\mathcal{M}(T) \lesssim \mathcal{M}(T - 0)$ for all $0 < T \lesssim 1$. Suppose $T_k \uparrow T$; we can find u^k coinciding with u on $[-T_k, T_k]$ such that $\|u^k\|_{Y_j} \leq 1$. Since $T \lesssim 1$, we may assume u^k are supported in $|t| \lesssim 1$. By the uniform boundedness in X_4 norm, and the fact that on the "tilde" side each u^k equals itself convolved with some $\widehat{\chi}_\xi$, we conclude that $(u^k)_{n,\xi}$ has second order ξ-derivatives bounded by (say) $\langle n\rangle^{10}$. We therefore extract a subsequence so that $\{u^k\}$, viewed as a sequence of maps from $\mathbb{R}_\xi$ to some weighted l^2_n space, converges uniformly in any $|\xi| \leq R$. In particular this implies the convergence as spacetime distributions; thus the limit, denoted by u^*, must coincide with u on $[-T, T]$. It therefore suffices to prove $\|u^*\|_{Y_j} \lesssim 1$. The bounds for the X_1, X_2, X_4 and X_7 norms immediately follow from distributional convergence; for X_3, note that the $|(u^k)_{n,\xi}|$ also converge uniformly to $|u_{n,\xi}|$ in any $|\xi| \leq R$ for any fixed n, thus $\mathfrak{N}u^k$ (recall Section 2.1 for definition) will converge to $\mathfrak{N}u^*$ as spacetime distributions, therefore the X_3 norm of u^* will also be bounded by $O(1)$.

It remains to bound the X_8 norm of u^*. By commuting with $\mathbb{P}_{\sim 2^d}$, we may assume that $\|u^k\|_{\mathcal{Y}} \leq 1$. For any bounded function $v = v_{n,\xi}$ with compact (n, ξ)-support, we have $(u^k, v) \to (u^*, v)$ with the standard pairing (u, v) as in (2.16). By the definition of the $\mathcal{Y}$ norm we can easily see that

$$|(u^k, v)| \le \|u^k\|_{\mathcal{Y}} \cdot \sup_{n_0 \in \mathbb{Z}} \langle n_0 \rangle^{-s^2} \|\pi_{n_0} v\|_{L^{q'} l^2} \le \sup_{n_0 \in \mathbb{Z}} \langle n_0 \rangle^{-s^2} \|\pi_{n_0} v\|_{L^{q'} l^2}. \tag{3.33}$$

Denote the right hand side by $\|v\|_{\mathcal{Z}}$. Then $|(u^*, v)| \le \|v\|_{\mathcal{Z}}$ for v with compact (n, ξ)-support. Now consider any v with $\|v\|_{\mathcal{Z}} \le 1$ (so in particular $v \in L^{q'} l^2$). We produce a sequence $v^R = v \cdot \mathbf{1}_{\{|v|+|n|+|\xi| \le R\}}$ so that $\|v^R\|_{\mathcal{Z}} \le 1$, and $v^R \to v$ in $L^{q'} l^2$, thus $(u^*, v^R) \to (u^*, v)$ (notice that $u^* \in X_4$ and is supported in some $|n| \sim 2^d$, thus we have $u^* \in L^q l^2$). This implies $|(u^*, v)| \le 1$ for all v such that $\|v\|_{\mathcal{Z}} \le 1$. Since a priori we have $u^* \in L^q l^2 \subset \mathcal{Y}$, and it is easily checked that $\mathcal{Y}$ is a Banach space, we may invoke the Hahn–Banach theorem to conclude $\|u^*\|_{\mathcal{Y}} \le 1$, provided that we can identify the dual space of $\mathcal{Y}$ with $\mathcal{Z}$. Now clearly each element in $\mathcal{Z}$ gives a linear functional on $\mathcal{Y}$ whose norm equals the $\mathcal{Z}$ norm; on the other hand, if we have a (bounded) linear functional on $\mathcal{Y}$, it must be bounded on $L^q l^2$, thus it is given by pairing with an element of $L^{q'} l^2$, and then by considering the action of $\mathbb{Z}$ on this function, we conclude that it is actually in $\mathcal{Z}$. □

4. Relevant probabilistic results

4.1. Review of the construction of Gibbs measure

In this section we briefly review the construction of the Gibbs measure ν as given in [28]. This measure is defined by adding a weight to some Wiener measure ρ, so we first describe the Wiener measure.

Consider a sequence $\{g_n\}_{n>0}$ of independent complex Gaussian random variables living on some ambient probability space $(\Omega, \mathcal{B}, \mathbb{P})$, which are normalized so that $\mathbb{E}(|g_n|^2) = 1$. We may also assume that $|g_n| = O(\langle n \rangle^{10})$ everywhere on Ω; this assumption is just in order to define the map $\mathbf{f}$ and is irrelevant otherwise. Letting $g_{-n} = \overline{g_n}$, we define the random series

$$\mathbf{f} : \Omega \ni \omega \mapsto \sum_{n \ne 0} \frac{g_n(\omega)}{2\sqrt{\pi |n|}} e^{\mathrm{i}nx} \in \mathcal{V} \tag{4.1}$$

as a map from Ω to $\mathcal{V}$ (recall that $\mathcal{V}$ is the subset of $\mathcal{D}'(\mathbb{T})$ containing real-valued distributions with vanishing mean). This then defines the Wiener measure ρ on $\mathcal{V}$ by $\rho(E) = \mathbb{P}(\mathbf{f}^{-1}(E))$. For each positive integer N, if we identify $\mathcal{V}$ with $\mathcal{V}_N \times \mathcal{V}_N^{\perp}$, then the measure $d\rho$ is identified with $d\rho_N \times d\rho_N^{\perp}$, with the latter two measures defined by

$$\rho_N(E) = \mathbb{P}((\Pi_N \mathbf{f})^{-1}(E)), \qquad \rho_N^{\perp}(E) = \mathbb{P}((\Pi_N^{\perp} \mathbf{f})^{-1}(E)).$$

Fix a compactly supported smooth cutoff ζ, $0 \le \zeta \le 1$, which equals 1 on some neighborhood of 0. Consider for each N the functions

$$\theta_N(f) = \zeta(\|\Pi_N f\|_{L^2}^2 - \alpha_N) e^{\frac{1}{3} \int_{\mathbb{T}} (S_N f)^3}, \tag{4.2}$$

$$\theta_N^{\sharp}(f) = \zeta(\|\Pi_N f\|_{L^2}^2 - \alpha_N) e^{\frac{1}{3} \int_{\mathbb{T}} (\Pi_N f)^3}, \tag{4.3}$$

where we recall $\Pi_N = \mathbb{P}_{\leq N}$ as in Section 2, S_N is as in (1.5), and

$$\alpha_N = \sum_{n=1}^{N} \frac{1}{n} = \mathbb{E}(\|\Pi_N \mathbf{f}\|_{L^2}^2).$$

Clearly θ_N and $\theta_N^\sharp$ only depend on $\Pi_N f$, thus they can also be understood as functions on $\mathcal{V}_N$. Define measures

$$d\nu_N = \theta_N \, d\rho, \qquad d\nu_N^\circ = \theta_N \, d\rho_N, \qquad d\nu_N^\sharp = \theta_N^\sharp \, d\rho, \qquad d\nu_N^\emptyset = \theta_N^\sharp \, d\rho_N.$$

Then we could identify $d\nu_N$ and $d\nu_N^\sharp$ with $d\nu_N^\circ \times d\rho_N^\perp$ and $d\nu_N^\emptyset \times d\rho_N^\perp$, respectively. Moreover, if we identify $\mathcal{V}_N$ with $\mathbb{R}^{2N}$ and thus denote the measure on $\mathcal{V}_N$ corresponding to the Lebesgue measure on $\mathbb{R}^{2N}$ by $\mathcal{L}_N$, then with some constant C_N,

$$d\nu_N^\circ = C_N \zeta(\|f\|_{L^2}^2 - \alpha_N) e^{-2E_N[f]} \, d\mathcal{L}_N, \tag{4.4}$$

$$d\nu_N^\emptyset = C_N \zeta(\|f\|_{L^2}^2 - \alpha_N) e^{-2E[f]} \, d\mathcal{L}_N, \tag{4.5}$$

with f here denoting some element of $\mathcal{V}_N$, the Hamiltonian E as in (1.2), and the truncated version E_N being

$$E_N[f] = \int_{\mathbb{T}} \big(\tfrac{1}{2}|\partial_x^{1/2} f|^2 - \tfrac{1}{6}(S_N f)^3\big). \tag{4.6}$$

The main result of [28] now reads as follows.

Proposition 4.1 ([28, Theorem 1]). *The sequence $\theta_N^\sharp$ converges in $L^r(d\rho)$ to some function θ for all $1 \leq r < \infty$, and if we define ν by $d\nu = \theta \, d\rho$, then $\nu_N^\sharp$ converges strongly to ν in the sense that the total variation of their difference tends to zero. This ν is defined to be the Gibbs measure for* (1.1).

Remark 4.2. Only weak convergence is claimed in [28], but an easy elaboration of the arguments there actually gives a much stronger convergence as stated in Proposition 4.1 above.

Remark 4.3. We note that the measure ν depends on the choice of ζ. In this regard we have the following easy observation: there exists a countable collection $\{\zeta^R\}_{R\in\mathbb{N}}$ with corresponding θ^R such that the union of $\mathcal{A}^R = \{f : \theta^R(f) > 0\}$ has full ρ measure. Note that $\mathcal{A}^R$ is the largest set on which ρ and ν^R are mutually absolutely continuous.

The finite-dimensional approximations we will actually use are ν_N instead of $\nu_N^\sharp$, thus we still need to prove the convergence of ν_N. However, the proof is essentially the same as the proof of Proposition 4.1, so we shall omit it here and only state the result.

Proposition 4.4. *The sequence θ_N converges in $L^r(d\rho)$ to the θ defined in Proposition* 4.1 *for all $1 \leq r < \infty$, and ν_N converges strongly to the ν defined in Proposition* 4.1 *in the sense that the total variation of their difference tends to zero.*

4.2. Compatibility with the Besov space

By elementary probabilistic arguments we can see that

$$\rho(f \in \mathcal{V} : \|f\|_{L^2} < \infty) = 0, \tag{4.7}$$
$$\rho(f \in \mathcal{V} : \|f\|_{H^{-\delta}} < \infty) = 1, \tag{4.8}$$

for all $\delta > 0$. Namely, the Wiener measure $d\rho$ (and hence the Gibbs measure $d\nu$) is compatible with $H^{-\delta}$ but not L^2, which is the essential difficulty in establishing the invariance result. In this section we show that this difficulty may be resolved by using the Besov space Z_1 defined in Section 1.1. First we prove a lemma.

Lemma 4.5. *Suppose that g_j $(1 \le j \le N)$ are independent normalized complex Gaussian random variables. Then*

$$\mathbb{P}\Big(\sum_{j=1}^{N} |g_j|^4 \ge \alpha N\Big) \le 4e^{-\frac{1}{120}\sqrt{\alpha N}}, \tag{4.9}$$

for all $\alpha > 1600$ and positive integer N.

Proof. Let $X = \sum_{j=1}^{N} |g_j|^4$. Since $\mathbb{E}(|g_j|^{4m}) = (2m)!$, we can estimate, for each integer $k \ge 1$, the k-th moment of X by

$$\begin{aligned}\mathbb{E}(X^k) &= \sum_{m_1+\cdots+m_N=k} \frac{k!}{m_1!\cdots m_N!} \times \mathbb{E}(|g_1|^{4m_1}\cdots|g_N|^{4m_N}) \\ &\le k! \sum_{m_1+\cdots+m_N=k} \prod_{j=1}^{N} \frac{(2m_j)!}{m_j!} \le k! 4^k \sum_{m_1+\cdots+m_N=k} \prod_{j=1}^{N} m_j!,\end{aligned}$$

since $\binom{2m}{m} \le 4^m$. From this, we see that (for $\epsilon > 0$)

$$\mathbb{E}(e^{\sqrt{\epsilon X}}) \le 2\mathbb{E}(\cosh\sqrt{\epsilon X}) \le 2 + 2\sum_{k\ge 1} \frac{\epsilon^k}{(2k)!}\mathbb{E}(X^k) \le 2 + \sum_{k\ge 1} \frac{(8\epsilon)^k}{k!} S_{N,k},$$

where

$$S_{N,k} = \sum_{m_1+\cdots+m_N=k} \prod_{j=1}^{N} m_j!, \tag{4.10}$$

which we shall now estimate. By identifying the nonzero terms in $(m_1, \ldots, m_N)$, we can rewrite $S_{N,k}$ as

$$S_{N,k} = \sum_{1\le r\le \min\{N,k\}} \binom{N}{r} S'_{k,r}, \tag{4.11}$$

where

$$S'_{k,r} = \sum_{m_1+\cdots+m_r=k,\ m_j\ge 1} \prod_{j=1}^{r} m_j!.$$

Clearly the number of choices of $(m_1, \ldots, m_r)$ is at most $\binom{k-1}{r-1} \leq 2^k$, and for each such choice we have

$$\prod_{j=1}^{r} m_j! \leq m_1 \cdots m_r \times \prod_{j=1}^{r}(m_j - 1)! \leq (k/r)^r \Big(\sum_{j=1}^{r}(m_j - 1)\Big)!$$
$$\leq e^{r\frac{k}{r}}(k-r)! \leq 3^k(k-r)!.$$

Therefore $S'_{k,r} \leq 6^k(k-r)!$. Next, notice that there are at most $k \leq 2^k$ choices of r, and $\binom{N}{r} \leq N^r/r!$, so we have

$$S_{N,k} \leq 12^k \max_{1\leq r\leq k} \frac{N^r(k-r)!}{r!}. \tag{4.12}$$

If the maximum in (4.12) is attained at $r = k$, it will be bounded by $N^k/k!$; otherwise it is attained at some $r < k$, which yields $N \leq (r+1)(k-r) \leq 2r(k-r)$. Therefore the maximum in this case is bounded by

$$\frac{N^r(k-r)!}{r!} \leq \frac{2^k r^r (k-r)^r (k-r)^{k-r}}{r!} \leq (6k)^k \leq 18^k k!.$$

Altogether we have $S_{N,k} \leq 216^k k! + (12N)^k/k!$, and hence

$$\mathbb{E}(e^{\sqrt{\epsilon X}}) \leq 2 + \sum_{k\geq 1}(1728\epsilon)^k + \sum_{k\geq 1}\frac{(384\epsilon N)^k}{(2k)!}, \tag{4.13}$$

which is clearly bounded by $4e^{20\sqrt{\epsilon N}}$ if we choose $\epsilon = \frac{1}{3456}$. Now if $\alpha > 1600$, we have

$$\mathbb{P}(X \geq \alpha N) \leq e^{-\sqrt{\epsilon\alpha N}}\mathbb{E}(e^{\sqrt{\epsilon X}}) \leq 4e^{-\frac{1}{120}\sqrt{\alpha N}}. \qquad \square$$

Now we can prove that the Wiener measure $d\rho$ is compatible with our Besov space Z_1:

Proposition 4.6. *With the measure ρ defined in Section 4.1, we have $\rho(Z_1) = 1$; more precisely,*

$$\rho(\{f \in \mathcal{V} : \|f\|_{Z_1} \leq K\}) \geq 1 - Ce^{-C^{-1}K^2} \tag{4.14}$$

for all $K > 0$.

Proof. We only need to prove (4.14). Setting C large, this inequality will be trivial when $K \leq 100$. When $K > 100$, we deduce from the definition that

$$\rho(\{f \in \mathcal{V} : \|f\|_{Z_1} > 100K\}) \leq \sum_{j\geq 0}\mathbb{P}\Big(\sum_{0<n\sim 2^j}|g_n|^p \geq K^p 2^j\Big). \tag{4.15}$$

By Hölder, $\sum_{0<n\sim 2^j}|g_n|^p \geq K^p 2^j$ implies $\sum_{0<n\sim 2^j}|g_n|^4 \geq K^4 2^j$. By Lemma 4.5, the latter has probability not exceeding $Ce^{-C^{-1}K^2 2^{j/2}}$ provided $K > 100$. Summing up over j, we see that

$$\rho(\{f \in \mathcal{V} : \|f\|_{Z_1} > K\}) \leq \sum_{j\geq 0} Ce^{-C^{-1}K^2 2^{j/2}} \leq Ce^{-C^{-1}K^2}. \qquad \square$$

5. The gauge transform I: Beating the derivative loss

From this section to Section 7, we will introduce the gauge transform for (1.6), and use it to derive the new equations. We fix a large positive integer N throughout, and drop the subscript N in S_N (we are allowing $N = \infty$, in which case the arguments should be modified slightly but no essential difference occurs). We also fix a smooth solution u to (1.6); note that smooth solutions are automatically global. When N is finite, we also assume that $\widehat{u}$ is supported in $|n| \leq N$ for all time.

The gauge transform we use is defined as a power series, thus in many occasions we will have to deal with summations over sequences of the form $(m_1, \ldots, m_\mu)$. To simplify the notation we will define, for such a sequence, the partial sums

$$m_{ij} = m_i + \cdots + m_j.$$

This notation will also be used for other sequences, say μ_i, which will always be nonnegative integers.

5.1. *The definition of w*

Let F be the unique mean-zero antiderivative of u, namely $F_n = \frac{1}{\mathrm{i}n} u_n$ for $n \neq 0$ and $F_0 = 0$. Define the operators $Q_0 : \phi \mapsto (Su) \cdot \phi$ and $P_0 : \phi \mapsto (SF) \cdot \phi$, as well as $Q = SQ_0S$ and $P = SP_0S$. Further, define the operator

$$M = e^{-\frac{\mathrm{i}P}{2}} = \sum_{\mu \geq 0} \frac{1}{\mu!} \left(-\frac{\mathrm{i}}{2}\right)^{\mu} P^{\mu}. \tag{5.1}$$

The function w will be defined by

$$w = \mathbb{P}_+(Mu). \tag{5.2}$$

We also define $v = Mu$, so that $w_n = v_n$ when $n > 0$, and $w_n = 0$ otherwise. The evolution equation satisfied by w can be computed as follows:

$$\begin{aligned}(\partial_t - \mathrm{i}\partial_{xx})w &= \mathbb{P}_+ M(\partial_t - \mathrm{i}\partial_{xx})u + \mathbb{P}_+[\partial_t, M]u - \mathrm{i}\mathbb{P}_+[\partial_{xx}, M]u \\ &= -2\mathrm{i}\mathbb{P}_+(M\mathbb{P}_- u_{xx}) + \mathbb{P}_+([\partial_t, M]u - \mathrm{i}[\partial_x, [\partial_x, M]]u) \\ &\quad + \mathbb{P}_+(MS(Su \cdot Su_x) - 2\mathrm{i}[\partial_x, M]u_x) \\ &= -2\mathrm{i}\mathbb{P}_+\partial_x(M\mathbb{P}_- u_x) & (5.3) \\ &\quad - 2\mathrm{i}\mathbb{P}_+\left([\partial_x, M] + \frac{\mathrm{i}}{2}MQ\right)u_x & (5.4) \\ &\quad + 2\mathrm{i}\mathbb{P}_+[\partial_x, M]\mathbb{P}_- u_x + \mathbb{P}_+([\partial_t, M] - \mathrm{i}[\partial_x, [\partial_x, M]])u. & (5.5)\end{aligned}$$

5.2. *The term in (5.3)*

By expanding M using (5.1), we can write

$$(5.3) = \sum_{\mu_1} \frac{(-1)^{\mu_1}}{2^{\mu_1}\mu_1!} \mathcal{K}^1_{\mu_1}, \tag{5.6}$$

where in Fourier space

$$(\mathcal{K}^1_{\mu_1})_{n_0} = 2\mathrm{i} \sum_{\mathbf{v}\in S^1_{n_0,\mu_1}} \Lambda^{1\mu_1}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_1}})u_{n_1}. \tag{5.7}$$

Here, the spatial frequency set is defined to be

$$S^1_{n_0,\mu_1} = \{\mathbf{v} = (m_1,\ldots,m_{\mu_1},n_1)\in\mathbb{Z}^{\mu_1+1} : m_i\neq 0,\ n_0>0,\ n_1<0,\ m_{1,\mu_1}+n_1=n_0\},$$

and the weight is

$$\Lambda^{1\mu_1}_{\mathbf{v}} = \prod_{i=1}^{\mu_1}\frac{1}{m_i}\psi\left(\frac{m_i}{N}\right)\times\prod_{i=2}^{\mu_1}\psi^2\left(\frac{m_{i,\mu_1}+n_1}{N}\right)n_0n_1\psi\left(\frac{n_0}{N}\right)\psi\left(\frac{n_1}{N}\right).$$

As the next step, we rewrite part of the weight as

$$\frac{1}{m_1\cdots m_{\mu_1}} = \frac{1}{n_0-n_1}\sum_{i=1}^{\mu_1}\frac{1}{m_1\cdots m_{i-1}m_{i+1}\cdots m_{\mu_1}}. \tag{5.8}$$

By renaming the variables, we obtain

$$(5.3) = \sum_{\mu_1\geq 1}\sum_{i=1}^{\mu_1}\frac{(-1)^{\mu_1}}{2^{\mu_1-1}\mu_1!}\mathcal{K}^1_{\mu_1 i}, \tag{5.9}$$

where in Fourier space

$$(\mathcal{K}^1_{\mu_1 i})_{n_0} = \mathrm{i}\sum_{\mathbf{v}\in S^{1.1}_{n_0,\mu_1}} \Lambda^{1\mu_1 i}_{\mathbf{v}}(u_{m_1}\cdots u_{m_{\mu_1-1}})u_{n_1}u_{n_2}. \tag{5.10}$$

The frequency set here is

$$S^{1.1}_{n_0,\mu_1} = \{\mathbf{v} = (m_1,\ldots,m_{\mu_1-1},n_1,n_2) : \mathbf{v}\in\mathbb{Z}^{\mu_1+1},\ m_j\neq 0,\ n_0>0,\ n_1<0,\ m_{1,\mu_1-1}+n_1+n_2=n_0\}, \tag{5.11}$$

and the weight is

$$\Lambda^{1\mu_1 i}_{\mathbf{v}} = \prod_{j=1}^{\mu_1-1}\frac{1}{m_j}\psi\left(\frac{m_j}{N}\right)\prod_{j=i+1}^{\mu_1}\psi^2\left(\frac{m_{j-1,\mu_1-1}+n_1}{N}\right) \times\prod_{j=2}^{i}\psi^2\left(\frac{m_{j,\mu_1-1}+n_1+n_2}{N}\right)\frac{n_0n_1}{|n_0|+|n_1|}\prod_{j=0}^{2}\psi\left(\frac{n_j}{N}\right). \tag{5.12}$$

5.3. The term in (5.4)

Since

$$[\partial_x, P] = Q, \tag{5.13}$$

we may compute

$$[\partial_x, M] = \sum_{\mu_1} \frac{1}{\mu_1!}\Big(-\frac{\mathrm{i}}{2}\Big)^{\mu_1}[\partial_x, P^{\mu_1}] = \sum_{\mu_1,\mu_2} \frac{1}{(\mu_{12}+1)!}\Big(-\frac{\mathrm{i}}{2}\Big)^{\mu_{12}+1} P^{\mu_1} Q P^{\mu_2}$$
$$= -\frac{\mathrm{i}}{2} M Q - \frac{\mathrm{i}}{2} \sum_{\mu_1,\mu_2} \frac{1}{(\mu_{12}+1)!}\Big(-\frac{\mathrm{i}}{2}\Big)^{\mu_{12}} P^{\mu_1}[Q, P^{\mu_2}]. \tag{5.14}$$

By expanding the commutator in (5.14), we can write

$$(5.4) = -\sum_{\mu_1,\mu_2} \frac{\mu_1+1}{(\mu_{12}+2)!}\Big(-\frac{\mathrm{i}}{2}\Big)^{\mu_{12}+1} \mathbb{P}_+ P^{\mu_1}[Q, P] P^{\mu_2} u_x. \tag{5.15}$$

Notice that

$$[Q, P] = S(Q_0 S^2 P_0 - P_0 S^2 Q_0) S, \tag{5.16}$$

we can thus write

$$(5.4) = \sum_{\mu_1,\mu_2} \frac{(-1)^{\mu_{12}}(\mu_1+1)}{2^{\mu_{12}+1}(\mu_{12}+2)!} \mathcal{K}^2_{\mu_1\mu_2}, \tag{5.17}$$

where in Fourier space

$$(\mathcal{K}^2_{\mu_1\mu_2})_{n_0} = \mathrm{i} \sum_{\mathbf{v}\in S^2_{n_0,\mu_1\mu_2}} \lambda^{2\mu_1\mu_2}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_{12}}}) u_{n_1} u_{n_2} u_{n_3}. \tag{5.18}$$

Here the frequency set is

$$S^2_{n_0,\mu_1\mu_2} = \{\mathbf{v} = (m_1, \dots, m_{\mu_{12}}, n_1, n_2, n_3) : \mathbf{v} \in (\mathbb{Z}^*)^{\mu_{12}+3},\ m_i \neq 0,\ n_1 n_2 n_3 \neq 0,$$
$$n_0 > 0,\ m_{1,\mu_{12}} + n_1 + n_2 + n_3 = n_0\}, \tag{5.19}$$

and the weight is

$$\lambda^{2\mu_1\mu_2}_{\mathbf{v}} = \frac{n_3}{n_2} \prod_{i=0}^{3} \psi\Big(\frac{n_i}{N}\Big) \prod_{i=1}^{\mu_{12}} \frac{1}{m_i} \psi\Big(\frac{m_i}{N}\Big) \prod_{i=1}^{\mu_2} \psi^2\Big(\frac{n_3 + m_{\mu_1+i,\mu_{12}}}{N}\Big)$$
$$\times \prod_{i=2}^{\mu_1+1} \psi^2\Big(\frac{n_1+n_2+n_3+m_{i,\mu_{12}}}{N}\Big)$$
$$\times \Big[\psi^2\Big(\frac{n_2+n_3+m_{\mu_1+1,\mu_{12}}}{N}\Big) - \psi^2\Big(\frac{n_1+n_3+m_{\mu_1+1,\mu_{12}}}{N}\Big)\Big].$$

Note that $S^2_{n_0,\mu_1\mu_2}$ is symmetric with respect to n_1 and n_3, so we can swap these two variables and rearrange terms to obtain

$$(\mathcal{K}^2_{\mu_1\mu_2})_{n_0} = \mathrm{i} \sum_{\mathbf{v}\in S^2_{n_0,\mu_1\mu_2}} \Lambda^{2\mu_1\mu_2}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_{12}}}) u_{n_1} u_{n_2} u_{n_3}, \tag{5.20}$$

where the frequency set $S^2_{n_0,\mu_1\mu_2}$ is as in (5.19), and the weight is

$$\Lambda_{\mathbf{v}}^{2\mu_1\mu_2} = \frac{1}{2n_2}\prod_{i=0}^{3}\psi\left(\frac{n_i}{N}\right)\prod_{i=1}^{\mu_{12}}\frac{1}{m_i}\psi\left(\frac{m_i}{N}\right)\prod_{i=2}^{\mu_1+1}\psi^2\left(\frac{n_1+n_2+n_3+m_{i,\mu_{12}}}{N}\right)$$
$$\times\left[n_3\psi^2\left(\frac{n_2+n_3+m_{\mu_1+1,\mu_{12}}}{N}\right)\prod_{i=1}^{\mu_2}\psi^2\left(\frac{n_3+m_{\mu_1+i,\mu_{12}}}{N}\right)\right.$$
$$-n_3\psi^2\left(\frac{n_1+n_3+m_{\mu_1+1,\mu_{12}}}{N}\right)\prod_{i=1}^{\mu_2}\psi^2\left(\frac{n_3+m_{\mu_1+i,\mu_{12}}}{N}\right)$$
$$+n_1\psi^2\left(\frac{n_1+n_2+m_{\mu_1+1,\mu_{12}}}{N}\right)\prod_{i=1}^{\mu_2}\psi^2\left(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\right)$$
$$\left.-n_1\psi^2\left(\frac{n_1+n_3+m_{\mu_1+1,\mu_{12}}}{N}\right)\prod_{i=1}^{\mu_2}\psi^2\left(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\right)\right]. \quad (5.21)$$

5.4. *The term in (5.5)*

Clearly we have

$$[\partial_t, M] = \sum_{\mu_1,\mu_2}\frac{1}{(\mu_{12}+1)!}\left(-\frac{\mathrm{i}}{2}\right)^{\mu_{12}+1}P^{\mu_1}[\partial_t, P]P^{\mu_2}, \quad (5.22)$$

where

$$[\partial_t, P] : \psi \mapsto S(SF_t\cdot S\psi); \quad (5.23)$$

also we may compute

$$[\partial_x,[\partial_x, M]] = \sum_{\mu_1,\mu_2}\frac{1}{(\mu_{12}+1)!}\left(-\frac{\mathrm{i}}{2}\right)^{\mu_{12}+1}[\partial_x, P^{\mu_1}QP^{\mu_2}]$$
$$= \sum_{\mu_1,\mu_2}\frac{1}{(\mu_{12}+1)!}\left(-\frac{\mathrm{i}}{2}\right)^{\mu_1+\mu_2+1}P^{\mu_1}[\partial_x, Q]P^{\mu_2}$$
$$+2\sum_{\mu_1,\mu_2,\mu_3}\frac{1}{(\mu_{13}+2)!}\left(-\frac{\mathrm{i}}{2}\right)^{\mu_{13}+2}P^{\mu_1}QP^{\mu_2}QP^{\mu_3}.$$

Using the fact that

$$[\partial_t, P] - \mathrm{i}[\partial_x, Q] : \psi \mapsto S(SG\cdot S\psi), \quad (5.24)$$

where

$$G = F_t - \mathrm{i}F_{xx} = -2\mathrm{i}\mathbb{P}_{-}u_x + \tfrac{1}{2}\big(S((Su)^2) - \mathbb{P}_0((Su)^2)\big), \quad (5.25)$$

we may write

$$(5.5) = \sum_{\mu_1,\mu_2}\frac{(-1)^{\mu_{12}}}{2^{\mu_{12}}(\mu_{12}+1)!}(\mathcal{K}^3_{\mu_1\mu_2}+\mathcal{K}^4_{\mu_1\mu_2}) + \sum_{\mu_1,\mu_2,\mu_3}\frac{(-1)^{\mu_{13}+1}}{2^{\mu_{13}+2}(\mu_{13}+2)!}\mathcal{K}^5_{\mu_1\mu_2\mu_3}. \quad (5.26)$$

Here, $\mathcal{K}^3_{\mu_1\mu_2}$ is defined in Fourier space as

$$(\mathcal{K}^3_{\mu_1\mu_2})_{n_0} = \mathrm{i} \sum_{\mathbf{v}\in S^3_{n_0,\mu_1\mu_2}} \Lambda^{3\mu_1\mu_2}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_{12}}}) u_{n_1} u_{n_2}, \tag{5.27}$$

with frequency set

$$S^3_{n_0,\mu_1\mu_2} = \{\mathbf{v} = (m_1,\ldots,m_{\mu_{12}},n_1,n_2) : \mathbf{v}\in\mathbb{Z}^{\mu_{12}+2},\ m_i\neq 0,\ n_1\neq 0,\ n_0>0,\\ n_2<0, m_{1,\mu_{12}}+n_1+n_2=n_0\} \tag{5.28}$$

and weight

$$\Lambda^{3\mu_1\mu_2}_{\mathbf{v}} = n_2 \prod_{i=1}^{\mu_{12}} \frac{1}{m_i}\psi\Big(\frac{m_i}{N}\Big) \prod_{i=0}^{2}\psi\Big(\frac{n_i}{N}\Big) \prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{n_1+n_2+m_{i,\mu_{12}}}{N}\Big)\\ \times\Big[\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_2+m_{\mu_1+i,\mu_{12}}}{N}\Big) - \prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\Big)\Big].$$

The term $\mathcal{K}^4_{\mu_1\mu_2}$ is defined as

$$(\mathcal{K}^4_{\mu_1\mu_2})_{n_0} = \frac{\mathrm{i}}{4}\mathbb{P}_0((Su)^2) \sum_{\mathbf{v}\in S^4_{n_0,\mu_1\mu_2}} \Lambda^{4\mu_1\mu_2}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_{12}}}) u_{n_1}, \tag{5.29}$$

with frequency set

$$S^4_{n_0,\mu_1\mu_2} = \{\mathbf{v} = (m_1,\ldots,m_{\mu_{12}},n_1) : \mathbf{v}\in\mathbb{Z}^{\mu_{12}+1},\ m_i\neq 0,\ n_1\neq 0,\ n_0>0,\\ m_{1,\mu_{12}}+n_1=n_0\} \tag{5.30}$$

and weight

$$\Lambda^{4\mu_1\mu_2}_{\mathbf{v}} = \psi\Big(\frac{n_0}{N}\Big)\psi\Big(\frac{n_1}{N}\Big)\prod_{i=1}^{\mu_{12}}\frac{1}{m_i}\psi\Big(\frac{m_i}{N}\Big)\\ \times\psi^2\Big(\frac{n_1+m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=2}^{\mu_{12}}\psi^2\Big(\frac{n_1+m_{i,\mu_{12}}}{N}\Big). \tag{5.31}$$

The term $\mathcal{K}^5_{\mu_1\mu_2\mu_3}$ is defined as

$$(\mathcal{K}^5_{\mu_1\mu_2\mu_3})_{n_0} = \mathrm{i} \sum_{\mathbf{v}\in S^5_{n_0,\mu_1\mu_2\mu_3}} \Lambda^{5\mu_1\mu_2\mu_3}_{\mathbf{v}} (u_{m_1}\cdots u_{m_{\mu_{13}}}) u_{n_1} u_{n_2} u_{n_3}, \tag{5.32}$$

with frequency set

$$S^5_{n_0,\mu_1\mu_2\mu_3} = \{\mathbf{v} = (m_1,\ldots,m_{\mu_{13}},n_1,n_2,n_3) : \mathbf{v}\in\mathbb{Z}^{\mu_{13}+3},\ m_i\neq 0,\ n_1n_2n_3\neq 0,\\ n_0>0,\ m_{1,\mu_{13}}+n_1+n_2+n_3=n_0\} \tag{5.33}$$

and weight

$$\Lambda_{\mathbf{v}}^{5\mu_1\mu_2\mu_3} = \prod_{i=2}^{\mu_1+1} \psi^2\Big(\frac{n_1+n_2+n_3+m_{i,\mu_{13}}}{N}\Big) \times \prod_{i=1}^{\mu_{13}} \frac{1}{m_i}\psi\Big(\frac{m_i}{N}\Big)\prod_{i=0}^{3}\psi\Big(\frac{n_i}{N}\Big)\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{n_3+m_{\mu_{12}+i,\mu_{13}}}{N}\Big) \times \Big[\psi^2\Big(\frac{n_1+n_2}{N}\Big)\prod_{i=2}^{\mu_2+1}\psi^2\Big(\frac{n_1+n_2+n_3+m_{\mu_1+i,\mu_{13}}}{N}\Big) + \psi^2\Big(\frac{n_1+n_2}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_3+m_{\mu_1+i,\mu_{13}}}{N}\Big) - 2\prod_{i=1}^{\mu_2+1}\psi^2\Big(\frac{n_2+n_3+m_{\mu_1+i,\mu_{13}}}{N}\Big)\Big]. \tag{5.34}$$

Next, we shall rewrite a part of the weight $\Lambda_{\mathbf{v}}^{3\mu_1\mu_2}$ as

$$\frac{1}{m_1\cdots m_{\mu_{12}}} = \frac{1}{m_1\cdots m_{12}}\psi_0\Big(\frac{n_1}{|n_0|+|n_2|}\Big) + \frac{1}{|n_0|+|n_2|-n_1}\psi\Big(\frac{n_1}{|n_0|+|n_2|}\Big)\sum_{i=1}^{\mu_{12}}\frac{1}{m_1\cdots m_{i-1}m_{i+1}\cdots m_{\mu_{12}}},$$

then rename the variables (separating the cases $i \le \mu_1$ and $i > \mu_1$) to obtain

$$(5.5) = \sum_{\mu_1,\mu_2}\frac{(-1)^{\mu_{12}}}{2^{\mu_{12}}(\mu_{12}+1)!}(\mathcal{K}^3_{\mu_1\mu_2 0}+\mathcal{K}^4_{\mu_1\mu_2}) + \sum_{\mu_1+\mu_2\ge 1}\sum_{i=1}^{\mu_{12}}\frac{(-1)^{\mu_{12}}}{2^{\mu_{12}}(\mu_{12}+1)!}\mathcal{K}^3_{\mu_1\mu_2 i} + \sum_{\mu_1,\mu_2,\mu_3}\frac{(-1)^{\mu_{13}+1}}{2^{\mu_{13}+2}(\mu_{13}+2)!}\mathcal{K}^5_{\mu_1\mu_2\mu_3}, \tag{5.35}$$

where in Fourier space

$$(\mathcal{K}^3_{\mu_1\mu_2 0})_{n_0} = \mathrm{i}\sum_{\mathbf{v}\in S^3_{n_0,\mu_1\mu_2}}\Lambda_{\mathbf{v}}^{3\mu_1\mu_2 0}(u_{m_1}\cdots u_{m_{\mu_{12}}})u_{n_1}u_{n_2}, \tag{5.36}$$

with frequency set $S^3_{n_0,\mu_1\mu_2}$ as in (5.28), and new weight

$$\Lambda_{\mathbf{v}}^{3\mu_1\mu_2 0} = \prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{n_1+n_2+m_{i,\mu_{12}}}{N}\Big)n_2\psi_0\Big(\frac{n_1}{|n_0|+|n_2|}\Big)\prod_{i=1}^{\mu_{12}}\frac{1}{m_i}\psi\Big(\frac{m_i}{N}\Big)\prod_{i=0}^{2}\psi\Big(\frac{n_i}{N}\Big) \times\Big[\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_2+m_{\mu_1+i,\mu_{12}}}{N}\Big)-\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\Big)\Big], \tag{5.37}$$

the other term will be

$$(\mathcal{K}^3_{\mu_1\mu_2 i})_{n_0} = \mathrm{i}\sum_{\mathbf{v}\in S^{3.1}_{n_0,\mu_1\mu_2}}\Lambda_{\mathbf{v}}^{3\mu_1\mu_2 i}(u_{m_1}\cdots u_{m_{\mu_{12}-1}})u_{n_1}u_{n_2}u_{n_3}, \tag{5.38}$$

where the new frequency set is

$$S^{3.1}_{n_0,\mu_1\mu_2} = \{\mathbf{v} = (m_1, \dots, m_{\mu_{12}-1}, n_1, n_2, n_3) : \mathbf{v} \in \mathbb{Z}^{\mu_{12}+2},\ m_i \neq 0,\ n_1 n_2 \neq 0,$$
$$n_0 > 0,\ n_3 < 0,\ m_{1,\mu_{12}-1} + n_1 + n_2 + n_3 = n_0\} \qquad (5.39)$$

and the new weight is

$$\Lambda^{3\mu_1\mu_2 i}_{\mathbf{v}} = \frac{n_3}{|n_0| + |n_3| - n_1} \psi\left(\frac{n_1}{|n_0| + |n_3|}\right) \prod_{j=1}^{\mu_{12}-1} \frac{1}{m_j} \psi\left(\frac{m_j}{N}\right)$$
$$\times \prod_{j=0}^{3} \psi\left(\frac{n_j}{N}\right) \prod_{j=2}^{i} \psi^2\left(\frac{n_1 + n_2 + n_3 + m_{j,\mu_{12}-1}}{N}\right) \prod_{j=i+1}^{\mu_1+1} \psi^2\left(\frac{n_1 + n_3 + m_{j-1,\mu_{12}-1}}{N}\right)$$
$$\times \left[\prod_{j=1}^{\mu_2} \psi^2\left(\frac{n_3 + m_{\mu_1+j-1,\mu_{12}-1}}{N}\right) - \prod_{j=1}^{\mu_2} \psi^2\left(\frac{n_1 + m_{\mu_1+j-1,\mu_{12}-1}}{N}\right)\right] \qquad (5.40)$$

for $1 \leq i \leq \mu_1$, and

$$\Lambda^{3\mu_1\mu_2 i}_{\mathbf{v}} = \frac{n_3}{|n_0| + |n_3| - n_1} \psi\left(\frac{n_1}{|n_0| + |n_3|}\right) \prod_{j=1}^{\mu_{12}-1} \frac{1}{m_j} \psi\left(\frac{m_j}{N}\right)$$
$$\times \prod_{j=0}^{3} \psi\left(\frac{n_j}{N}\right) \prod_{j=2}^{\mu_1+1} \psi^2\left(\frac{n_1 + n_2 + n_3 + m_{j,\mu_{12}-1}}{N}\right)$$
$$\times \left[\prod_{j=1}^{i-\mu_1} \psi^2\left(\frac{n_2 + n_3 + m_{\mu_1+j,\mu_{12}-1}}{N}\right) \prod_{j=i-\mu_1+1}^{\mu_2} \psi^2\left(\frac{n_3 + m_{\mu_1+j-1,\mu_{12}-1}}{N}\right)\right.$$
$$\left. - \prod_{j=1}^{i-\mu_1} \psi^2\left(\frac{n_1 + n_2 + m_{\mu_1+j,\mu_{12}-1}}{N}\right) \prod_{j=i-\mu_1+1}^{\mu_2} \psi^2\left(\frac{n_1 + m_{\mu_1+j-1,\mu_{12}-1}}{N}\right)\right] \qquad (5.41)$$

for $\mu_1 + 1 \leq i \leq \mu_{12}$.

5.5. *Summary*

Now we have obtained a first version of the equation satisfied by w, namely

$$(\partial_t - \mathrm{i}\partial_{xx})w = \sum_{\mu_1 \geq 1} \sum_{i=1}^{\mu_1} \frac{(-1)^{\mu_1}}{2^{\mu_1-1}\mu_1!} \mathcal{K}^1_{\mu_1 i} + \sum_{\mu_1,\mu_2} \frac{(-1)^{\mu_{12}}(\mu_1 + 1)}{2^{\mu_{12}+1}(\mu_{12} + 2)!} \mathcal{K}^2_{\mu_1\mu_2}$$
$$+ \sum_{\mu_1,\mu_2} \frac{(-1)^{\mu_{12}}}{2^{\mu_{12}}(\mu_{12} + 1)!} (\mathcal{K}^3_{\mu_1\mu_2 0} + \mathcal{K}^4_{\mu_1\mu_2}) + \sum_{\mu_1+\mu_2 \geq 1} \sum_{i=1}^{\mu_{12}} \frac{(-1)^{\mu_{12}}}{2^{\mu_{12}}(\mu_{12} + 1)!} \mathcal{K}^3_{\mu_1\mu_2 i}$$
$$+ \sum_{\mu_1,\mu_2,\mu_3} \frac{(-1)^{\mu_{13}+1}}{2^{\mu_{13}+2}(\mu_{13} + 2)!} \mathcal{K}^5_{\mu_1\mu_2\mu_3}, \qquad (5.42)$$

where:

- $\mathcal{K}^1_{\mu_1 i}$ is defined in (5.10)–(5.12);
- $\mathcal{K}^2_{\mu_1\mu_2}$ is defined in (5.19)–(5.21);

- $\mathcal{K}^3_{\mu_1\mu_2 0}$ is defined in (5.28), (5.36)–(5.37);
- $\mathcal{K}^3_{\mu_1\mu_2 i}$ is defined in (5.38)–(5.41);
- $\mathcal{K}^4_{\mu_1\mu_2}$ is defined in (5.29)–(5.31);
- $\mathcal{K}^5_{\mu_1\mu_2\mu_3}$ is defined in (5.32)–(5.34).

In the next section we will further examine the structure of these terms.

6. The gauge transform II: A miraculous cancellation

In this section we identify the bad resonant terms coming from each $\mathcal{K}^j$ term in (5.42). Our computation will show that these bad terms will eventually add up to zero, leaving only the better-behaved ones. Throughout this section we will use a variable k, and define $\theta = \psi(k/N)$, $\eta = \psi'(k/N)$.

6.1. The resonant terms in $\mathcal{K}^1$

In the expression (5.10), let $n_1 + n_2 = 0$. Noticing that $n_1 < 0$, we get a sum

$$-\mathrm{i}n_0 \sum_{k>0} |u_k|^2 \sum_{m_1+\cdots+m_{\mu_1-1}=n_0} \Delta \cdot \frac{u_{m_1}\cdots u_{m_{\mu_1-1}}}{m_1\cdots m_{\mu_1-1}}, \tag{6.1}$$

where we always assume $m_i \neq 0$, and the factor

$$\Delta = \prod_{j=1}^{\mu_1-1} \psi\left(\frac{m_j}{N}\right) \prod_{j=i+1}^{\mu_1} \psi^2\left(\frac{k - m_{j-1,\mu_1-1}}{N}\right) \times \prod_{j=2}^{i} \psi^2\left(\frac{m_{j,\mu_1-1}}{N}\right) \frac{k}{|n_0|+|k|} \psi\left(\frac{n_0}{N}\right) \psi^2\left(\frac{k}{N}\right).$$

We then replace each variable in this expression, except k, by zero (strictly speaking, we should replace n by m_{1,μ_1-1} and cancel each m_j in the numerator before this process, but the results will be the same and no estimate is affected), and get a term which reads

$$-\mathrm{i}n_0 \sum_{k>0} |u_k|^2 \sum_{m_1+\cdots+m_{\mu_1-1}=n_0} \theta^{2(\mu_1-i+1)} \prod_{i=1}^{\mu_1-1} \frac{u_{m_i}}{m_i}. \tag{6.2}$$

Noting that the summation over the m_i's gives exactly $((\mathrm{i}F)^{\mu_1-1})_{n_0}$, we can sum over μ_1 and i to get

$$\begin{aligned}(\mathcal{R}^1)_{n_0} &= -\mathrm{i} \sum_{k>0} |u_k|^2 \sum_{\mu_1\geq 1} \frac{(-1)^{\mu_1}}{2^{\mu_1-1}\mu_1!} n_0((\mathrm{i}F)^{\mu_1-1})_{n_0} \sum_{i=1}^{\mu_1} \theta^{2(\mu_1-i+1)} \\ &= -\sum_{k>0} |u_k|^2 \sum_{\mu_1\geq 1} \frac{(-1)^{\mu_1}}{2^{\mu_1-1}\mu_1!} (\partial_x(\mathrm{i}F)^{\mu_1-1})_{n_0} \sum_{i=1}^{\mu_1} \theta^{2(\mu_1-i+1)} \\ &= \sum_{k>0}\sum_{\mu\geq 0} \frac{\mathrm{i}|u_k|^2}{(\mu+2)!} \left(u\left(-\frac{\mathrm{i}F}{2}\right)^{\mu}\right)_{n_0} \cdot \mathcal{C}^1,\end{aligned} \tag{6.3}$$

where

$$\mathcal{C}^1 = -\tfrac{1}{2}(\mu+1)(\theta^2+\theta^4+\cdots+\theta^{2\mu+4}), \tag{6.4}$$

and we have dropped the dependence of $\mathcal{C}^1$ on k and μ for simplicity.

6.2. *The resonant terms in* $\mathcal{K}^2$

In the expression (5.20), let $n_2+n_3=0$ to obtain a term

$$\frac{\mathrm{i}}{2}\sum_{k\neq 0}|u_k|^2 \sum_{n_1+m_1+\cdots+m_{\mu_{12}}=n_0} \Delta\cdot\frac{u_{m_1}\cdots u_{m_{\mu_{12}}}}{m_1\cdots m_{\mu_{12}}}u_{n_1}, \tag{6.5}$$

where

$$\begin{aligned}\Delta = {} & \psi^2\Big(\frac{k}{N}\Big)\prod_{i=0}^{1}\psi\Big(\frac{n_i}{N}\Big)\prod_{i=1}^{\mu_{12}}\psi\Big(\frac{m_i}{N}\Big)\prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{n_1+m_{i,\mu_{12}}}{N}\Big)\\ & \times\Big[-\psi^2\Big(\frac{m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ & \quad+\psi^2\Big(\frac{k+n_1+m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ & \quad-\frac{n_1}{k}\psi^2\Big(\frac{n_1+m_{\mu_1+1,\mu_{12}}-k}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ & \quad+\frac{n_1}{k}\psi^2\Big(\frac{n_1+m_{\mu_1+1,\mu_{12}}+k}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{12}}}{N}\Big)\Big].\end{aligned} \tag{6.6}$$

We then discard the last two summands in the bracket, and in what remains replace each variable except k by zero to get

$$\frac{\mathrm{i}}{2}\sum_{k\neq 0}|u_k|^2 \sum_{n_1+m_1+\cdots+m_{\mu_{12}}=n_0}(\theta^{2\mu_2+4}-\theta^{2\mu_2+2})\cdot\frac{u_{m_1}\cdots u_{m_{\mu_{12}}}}{m_1\cdots m_{\mu_{12}}}u_{n_1}. \tag{6.7}$$

Since the summation over m_i and n_1 gives exactly $(u\cdot(iF)^{\mu_{12}})_{n_0}$, we can then sum over μ_1 and μ_2 to obtain an expression which involves a sum over all $k\neq 0$. We may include a factor of 2 and restrict to $k>0$ (since θ is even in k), and then take into account the symmetry with respect to n_1 and n_3 (namely, we are considering also the term where $n_1+n_2=0$) to include another factor of 2, and the final expression will be

$$\begin{aligned}(\mathcal{R}^{2.1})_{n_0} &= 2\mathrm{i}\sum_{k>0}|u_k|^2\sum_{\mu\geq 0}\frac{(-1)^\mu}{2^{\mu+1}(\mu+2)!}(u(\mathrm{i}F)^\mu)_{n_0}\sum_{\mu_2=0}^{\mu}(\mu-\mu_2+1)(\theta^{2\mu_2+4}-\theta^{2\mu_2+2})\\ &= \sum_{k>0}\sum_{\mu\geq 0}\frac{\mathrm{i}|u_k|^2}{(\mu+2)!}\Big(u\Big(-\frac{\mathrm{i}F}{2}\Big)^\mu\Big)_{n_0}\cdot\mathcal{C}^2,\end{aligned} \tag{6.8}$$

where

$$\mathcal{C}^2 = -(\mu+1)\theta^2+(\theta^4+\cdots+\theta^{2\mu+4}). \tag{6.9}$$

The other possibility is when $n_1 + n_3 = 0$. In this case we rename n_2 as n_1 and get

$$\frac{\mathrm{i}}{2}\sum_{k\neq 0}|u_k|^2 \sum_{n_1+m_1+\cdots+m_{\mu_{12}}=n_0} \Delta\cdot\frac{u_{m_1}\cdots u_{m_{\mu_{12}}}u_{n_1}}{m_1\cdots m_{\mu_1-1}n_1}, \tag{6.10}$$

where

$$\begin{aligned}\Delta = {}&\psi^2\Big(\frac{k}{N}\Big)\prod_{i=0}^{1}\psi\Big(\frac{n_i}{N}\Big)\prod_{i=1}^{\mu_{12}}\psi\Big(\frac{m_i}{N}\Big)\prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{n_1+m_{i,\mu_{12}}}{N}\Big)\\ &\times k\bigg[\psi^2\Big(\frac{k+n_1+m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ &\quad-\psi^2\Big(\frac{k-n_1-m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k-m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ &\quad-\psi^2\Big(\frac{m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{12}}}{N}\Big)\\ &\quad+\psi^2\Big(\frac{m_{\mu_1+1,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k-m_{\mu_1+i,\mu_{12}}}{N}\Big)\bigg].\end{aligned} \tag{6.11}$$

Next, we examine the terms in the bracket, which basically can be written, for some σ_j which are linear combinations of n_1 and m_i, as $\prod_j \psi^2((k+\sigma_j)/N)-\prod_j \psi^2((k-\sigma_j)/N)$. We then replace this expression by $4\theta^{2\mu-1}\eta\sum_j \sigma_j/N$, where μ is the number of factors. If we plug into (6.11) this and the expression for each σ_j, cancel each n_1 or m_i factor with the corresponding denominator in (6.10), and finally replace each variable other than k by zero, we will get a term which, up to a rearrangement of variables, reads

$$\begin{aligned}2\mathrm{i}\sum_{k\neq 0}\frac{k}{N}|u_k|^2 &\sum_{n_1+m_1+\cdots+m_{\mu_{12}}=n_0}\frac{u_{m_1}\cdots u_{m_{\mu_{12}}}}{m_1\cdots m_{\mu_{12}}}u_{n_1}\\ &\times\Big(\frac{(\mu_2+1)(\mu_2+2)}{2}\theta^{2\mu_2+3}\eta-\frac{\mu_2(\mu_2+1)}{2}\theta^{2\mu_2+1}\eta\Big).\end{aligned} \tag{6.12}$$

We may restrict to $k > 0$ since η is odd, and then sum over μ_1 and μ_2 to obtain

$$\begin{aligned}(\mathcal{R}^{2.2})_{n_0} &= 4\mathrm{i}\sum_{k>0}\frac{k\eta}{N}|u_k|^2\sum_{\mu\geq 0}\sum_{\mu_2=0}^{\mu}\frac{(-1)^\mu}{2^{\mu+1}(\mu+2)!}(u(\mathrm{i}F)^\mu)_{n_0}(\mu-\mu_2+1)\\ &\qquad\times\Big(\frac{(\mu_2+1)(\mu_2+2)}{2}\theta^{2\mu_2+3}-\frac{\mu_2(\mu_2+1)}{2}\theta^{2\mu_2+1}\Big)\\ &= \sum_{k>0}\sum_{\mu\geq 0}\frac{\mathrm{i}k\eta|u_k|^2}{N(\mu+2)!}\Big(u\Big(-\frac{\mathrm{i}F}{2}\Big)^\mu\Big)_{n_0}\cdot\mathcal{D}^2,\end{aligned} \tag{6.13}$$

where

$$\mathcal{D}^2 = 2\theta^3+\cdots+\mu(\mu+1)\theta^{2\mu+1}+(\mu+1)(\mu+2)\theta^{2\mu+3}. \tag{6.14}$$

6.3. The resonant terms in $\mathcal{K}^3$

In the expression (5.36), let $n_1 + n_2 = 0$. Noting that $n_2 < 0$, we obtain a term

$$\mathrm{i}\sum_{k>0}|u_k|^2 \sum_{m_1+\cdots+m_{\mu_{12}}=n_0} \Delta \cdot \frac{u_{m_1}\cdots u_{m_{\mu_{12}}}}{m_1\cdots m_{\mu_{12}}}, \tag{6.15}$$

where

$$\Delta = -k\psi_0\Big(\frac{k}{|n_0|+|k|}\Big)\psi^2\Big(\frac{k}{N}\Big)\prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{m_{i,\mu_{12}}}{N}\Big)\prod_{i=1}^{\mu_{12}}\psi\Big(\frac{m_i}{N}\Big)$$
$$\times\psi\Big(\frac{n_0}{N}\Big)\Big[\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k-m_{\mu_1+i,\mu_{12}}}{N}\Big)-\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{12}}}{N}\Big)\Big]. \tag{6.16}$$

Then we replace the term in the bracket by $-4\theta^{2\mu_2-1}\eta\sum_i(m_{\mu_1+i,\mu_{12}}/N)$, cancel the corresponding m_j factor in the denominator, and replace all the variables except k by zero to obtain, after a rearrangement of variables, the sum

$$4\mathrm{i}\sum_{k>0}\frac{k}{N}|u_k|^2\frac{\mu_2(\mu_2+1)}{2}\theta^{2\mu_2+1}\eta\sum_{n_1+m_1+\cdots+m_{\mu_{12}-1}=n_0}\frac{u_{m_1}\cdots u_{m_{\mu_{12}-1}}}{m_1\cdots m_{\mu_{12}-1}}u_{n_1}. \tag{6.17}$$

Then we sum over μ_1 and μ_2 to obtain

$$(\mathcal{R}^{3.1})_{n_0} = 4\mathrm{i}\sum_{k>0}\frac{k\eta}{N}|u_k|^2\sum_{\mu\ge 0}\frac{(-1)^{\mu+1}}{2^{\mu+1}(\mu+2)!}(u(\mathrm{i}F)^\mu)_{n_0}\Big(\sum_{\mu_2=0}^{\mu}\frac{\mu_2(\mu_2+1)}{2}\theta^{2\mu_2+1}\Big)$$
$$=\sum_{k>0}\sum_{\mu\ge 0}\frac{\mathrm{i}k\eta|u_k|^2}{N(\mu+2)!}\Big(u\Big(-\frac{\mathrm{i}F}{2}\Big)^\mu\Big)_{n_0}\cdot\mathcal{D}^3, \tag{6.18}$$

where

$$\mathcal{D}^3 = -\big(2\theta^3+\cdots+\mu(\mu+1)\theta^{2\mu+1}+(\mu+1)(\mu+2)\theta^{2\mu+3}\big). \tag{6.19}$$

Next, in the expression (5.38), let $n_2 + n_3 = 0$; noting that $n_3 < 0$, we get a term

$$\mathrm{i}\sum_{k>0}|u_k|^2\sum_{n_1+m_1+\cdots+m_{\mu_{12}-1}=n_0}\Delta\cdot\frac{u_{m_1}\cdots u_{m_{\mu_{12}-1}}}{m_1\cdots m_{\mu_{12}-1}}u_{n_1}, \tag{6.20}$$

where

$$\Delta = \frac{-k}{|k|+|n_0|-n_1}\psi\Big(\frac{n_1}{|k|+|n_0|}\Big)\psi^2\Big(\frac{k}{N}\Big)\prod_{j=1}^{\mu_{12}-1}\psi\Big(\frac{m_j}{N}\Big)\prod_{j=0}^{1}\psi\Big(\frac{n_j}{N}\Big)$$
$$\times\prod_{j=2}^{i}\psi^2\Big(\frac{n_1+m_{j,\mu_{12}-1}}{N}\Big)\prod_{j=i+1}^{\mu_1+1}\psi^2\Big(\frac{k-n_1-m_{j-1,\mu_{12}-1}}{N}\Big)$$
$$\times\Big[\prod_{j=1}^{\mu_2}\psi^2\Big(\frac{k-m_{\mu_1+j-1,\mu_{12}-1}}{N}\Big)-\prod_{j=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+j-1,\mu_{12}-1}}{N}\Big)\Big]$$

for $1 \le i \le \mu_1$, and

$$\Delta = \frac{-k}{|k|+|n_0|-n_1}\psi\left(\frac{n_1}{|k|+|n_0|}\right)\psi^2\left(\frac{k}{N}\right)\prod_{j=1}^{\mu_{12}-1}\psi\left(\frac{m_j}{N}\right)\prod_{j=2}^{\mu_1+1}\psi^2\left(\frac{n_1+m_{i,\mu_{12}-1}}{N}\right)$$
$$\times\prod_{j=0}^{1}\psi\left(\frac{n_j}{N}\right)\left[\prod_{j=1}^{i-\mu_1}\psi^2\left(\frac{m_{\mu_1+j,\mu_{12}-1}}{N}\right)\prod_{j=i-\mu_1+1}^{\mu_2}\psi^2\left(\frac{k-m_{\mu_1+j-1,\mu_{12}-1}}{N}\right)\right.$$
$$\left.-\prod_{j=1}^{i-\mu_1}\psi^2\left(\frac{k+n_1+m_{\mu_1+j,\mu_{12}-1}}{N}\right)\prod_{j=i-\mu_1+1}^{\mu_2}\psi^2\left(\frac{n_1+m_{\mu_1+j-1,\mu_{12}-1}}{N}\right)\right]$$

for $\mu_1+1\leq i\leq\mu_1+\mu_2$. Then we replace every variable other than k by zero, and sum over i to obtain

$$-\mathrm{i}\sum_{k>0}|u_k|^2\sum_{n_1+m_1+\cdots+m_{\mu_{12}-1}=n_0}\frac{u_{m_1}\cdots u_{m_{\mu_{12}-1}}}{m_1\cdots m_{\mu_{12}-1}}u_{n_1}$$
$$\times\Big(\sum_{i=1}^{\mu_1}(\theta^{2\mu_{12}-2i+4}-\theta^{2\mu_1-2i+4})+\sum_{i=\mu_1+1}^{\mu_{12}}(\theta^{2\mu_{12}-2i+2}-\theta^{2i-2\mu_1+2})\Big).$$

We then sum over μ_1 and μ_2 to get

$$(\mathcal{R}^{3.2})_{n_0}=-\mathrm{i}\sum_{k>0}|u_k|^2\sum_{\mu\geq0}\frac{(-1)^{\mu+1}}{2^{\mu+1}(\mu+2)!}(u(\mathrm{i}F)^\mu)_{n_0}$$
$$\times\Big(\sum_{\mu_2=0}^{\mu+1}\sum_{i=1}^{\mu+1-\mu_2}(\theta^{2\mu-2i+6}-\theta^{2\mu-2i-2\mu_2+6})+\sum_{\mu_2=0}^{\mu+1}\sum_{i=\mu+2-\mu_2}^{\mu+1}(\theta^{2\mu-2i+4}-\theta^{2i+2\mu_2-2\mu})\Big)$$
$$=\sum_{k>0}\sum_{\mu\geq0}\frac{\mathrm{i}|u_k|^2}{(\mu+2)!}\left(u\left(-\frac{\mathrm{i}F}{2}\right)^\mu\right)_{n_0}\cdot\mathcal{C}^3,\qquad(6.21)$$

where

$$\mathcal{C}^3=\tfrac{1}{2}\big((\mu+1)\theta^2+(-\mu-1)\theta^4+(-\mu+1)\theta^6+\cdots+(\mu-1)\theta^{2\mu+4}\big).\qquad(6.22)$$

6.4. The resonant terms in $\mathcal{K}^4$

The whole term $\mathcal{K}^4$ should be viewed as resonant. Here we simply expand $\mathbb{P}_0((Su)^2)=2\sum_{k>0}\theta^2|u_k|^2$, and replace every variable in (5.31) by zero (after extracting the $\prod_i m_i^{-1}$ factor, as before) to obtain

$$(\mathcal{R}^4)_{n_0}=\frac{\mathrm{i}}{2}\sum_{k>0}\theta^2|u_k|^2\sum_{\mu\geq0}\frac{(-1)^\mu}{2^\mu(\mu+1)!}(u(\mathrm{i}F)^\mu)_{n_0}\times\sum_{\mu_2=0}^{\mu}1$$
$$=\sum_{k>0}\sum_{\mu\geq0}\frac{\mathrm{i}|u_k|^2}{(\mu+2)!}\left(u\left(-\frac{\mathrm{i}F}{2}\right)^\mu\right)_{n_0}\cdot\mathcal{C}^4,\qquad(6.23)$$

where

$$\mathcal{C}^4=\tfrac{1}{2}(\mu+1)(\mu+2)\theta^2.\qquad(6.24)$$

6.5. *The resonant terms in* $\mathcal{K}^5$

In the expression (5.32), consider the contribution where $n_1 + n_2 = 0$, $n_2 + n_3 = 0$, or where $n_1 + n_3 = 0$. For each of these cases, we perform the same operation as in the above sections, and collect all the resulting terms (and rearrange the variables) to obtain

$$\mathrm{i}\sum_{k\neq 0}|u_k|^2 \sum_{m_1+\cdots+m_{\mu_{13}}+n_1=n_0} \Delta \cdot \frac{u_{m_1}\cdots u_{m_{\mu_{13}}}}{m_1\cdots m_{\mu_{13}}} u_{n_1}, \tag{6.25}$$

where

$$\begin{aligned}
\Delta = {}& \psi^2\Big(\frac{k}{N}\Big)\prod_{i=2}^{\mu_1+1}\psi^2\Big(\frac{n_1+m_{i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_{13}}\psi\Big(\frac{m_i}{N}\Big)\prod_{i=0}^{1}\psi\Big(\frac{n_i}{N}\Big)\\
&\times\Big[2\psi^2\Big(\frac{k-n_1}{N}\Big)\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{k+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=2}^{\mu_2+1}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad+2\psi^2\Big(\frac{k-n_1}{N}\Big)\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{k+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{k+m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad-2\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{k+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_2+1}\psi^2\Big(\frac{m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad+\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{n_1+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=2}^{\mu_2+1}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad+\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{n_1+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n_1+m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad-2\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{n_1+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_2+1}\psi^2\Big(\frac{k+n_1+m_{\mu_1+i,\mu_{13}}}{N}\Big)\\
&\quad-2\prod_{i=1}^{\mu_3}\psi^2\Big(\frac{k+m_{\mu_{12}+i,\mu_{13}}}{N}\Big)\prod_{i=1}^{\mu_2+1}\psi^2\Big(\frac{k+n_1+m_{\mu_1+i,\mu_{13}}}{N}\Big)\Big].
\end{aligned}$$

Then we replace each variable other than k by zero, obtaining

$$\mathrm{i}\sum_{k\neq 0}|u_k|^2 \sum_{m_1+\cdots+m_{\mu_{13}}+n_1=n_0} \frac{u_{m_1}\cdots u_{m_{\mu_{13}}}}{m_1\cdots m_{\mu_{13}}} u_{n_1} 2\theta^2(\theta^{2\mu_3+2}-\theta^{2\mu_2+2}+1-\theta^{2\mu_3}). \tag{6.26}$$

Again we restrict to $k > 0$ and sum over μ_1, μ_2, μ_3 to get

$$\begin{aligned}
(\mathcal{R}^5)_{n_0} &= 4\mathrm{i}\sum_{k>0}|u_k|^2\sum_{\mu\geq 0}\frac{(-1)^{\mu+1}}{2^{\mu+2}(\mu+2)!}(u(iF)^\mu)_{n_0}\\
&\qquad\times\sum_{\mu_1+\mu_2+\mu_3=\mu}(\theta^{2\mu_3+4}-\theta^{2\mu_2+4}+\theta^2-\theta^{2\mu_3+2})\\
&= \sum_{k>0}\sum_{\mu\geq 0}\frac{\mathrm{i}|u_k|^2}{(\mu+2)!}\Big(u\Big(-\frac{\mathrm{i}F}{2}\Big)^\mu\Big)_{n_0}\cdot\mathcal{C}^5,
\end{aligned}$$

where

$$\mathcal{C}^5 = -\tfrac{1}{2}\mu(\mu+1)\theta^2 + \mu\theta^4 + (\mu-1)\theta^6 + \cdots + 2\theta^{2\mu} + \theta^{2\mu+2}. \tag{6.27}$$

6.6. When put together...

Now we can directly verify from the above computations that

$$\mathcal{C}^1 + \mathcal{C}^2 + \mathcal{C}^3 + \mathcal{C}^4 + \mathcal{C}^5 = 0, \tag{6.28}$$

$$\mathcal{D}^2 + \mathcal{D}^3 = 0, \tag{6.29}$$

which then implies

$$\mathcal{R}^1 + \mathcal{R}^{2.1} + \mathcal{R}^{2.2} + \mathcal{R}^{3.1} + \mathcal{R}^{3.2} + \mathcal{R}^4 + \mathcal{R}^5 = 0. \tag{6.30}$$

6.7. What remains?

Here we analyze what remains after we subtract from each $\mathcal{K}^j$ term the resonant contribution, and deduce a second version of the equation satisfied by w. To simplify the argument, we need to introduce a few more notions.

Definition 6.1. We say a function $f : \mathbb{Z} \to \mathbb{R}$ is *slowly varying of type* 1, or $f \in SV_1$, if $|f(n)| \leq C$ and

$$|f(n+1) - f(n)| \leq C\langle n\rangle^{-1} \tag{6.31}$$

for some constant C. We say f is *slowly varying of type* 2, or $f \in SV_2$, if we have

$$|f(n+1) - f(n)| \leq C\langle n\rangle^{-1}(|f(n)| + |f(n+1)|) \tag{6.32}$$

for some constant C. For a function $f : \mathbb{Z}^\mu \to \mathbb{R}$, we say it is slowly varying of type 1 or 2 if it satisfies the above inequalities for each single variable when the other variables are fixed, with uniformly bounded constants.

Proposition 6.2. *The following functions are in SV_1:*

(1) $\phi(f_1, \ldots, f_k)$*, where* $f_j \in SV_1$ *and* $\phi : \mathbb{R}^k \to \mathbb{R}$ *is Lipschitz;*
(2) $\phi(f_1, \ldots, f_k)$*, where* $f_j \in SV_2$*,* ϕ *is smooth and is constant outside some compact set.*

The following functions are in SV_2:

(3) *any monomial (say* n_1^2 *or* $n_2 n_3$*), or characteristic function of any set generated by* $\{n_j > 0\}$ *and* $\{n_j < 0\}$*;*
(4) *products and reciprocals of functions in* SV_2 *(with* $1/f$ *defined to be* 1 *at points where* $f = 0$*);* $\max(f, g)$, $\min(f, g)$ *and* $f + g$ *for nonnegative* $f, g \in SV_2$*;*
(5) $|f|$, $\langle f\rangle$ *and* $(\max(f, 0))^\lambda$*, where* $f \in SV_2$ *and* $\lambda > 0$.

Proof. Omitted. □

Proposition 6.3. *We have*

$$(\partial_t - \mathrm{i}\partial_{xx})w = \mathcal{H} = \sum_{\mu} C_\mu \mathcal{H}_\mu, \tag{6.33}$$

where $|C_\mu| \leq C^\mu/\mu!$ *with some absolute constant* C*, and* $\mathcal{H}_\mu = \mathcal{H}_\mu^2 + \mathcal{H}_\mu^3 + \mathcal{H}_\mu^4$*. The* $\mathcal{H}^j$ *terms can be written as*

$$(\mathcal{H}_\mu^2)_{n_0} = \mathrm{i} \sum_{n_1+n_2+m_1+\cdots+m_\mu=n_0} \min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\} \cdot \Theta_\mu^2 \prod_{l=1}^{2} u_{n_l} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i}, \tag{6.34}$$

$$(\mathcal{H}_\mu^j)_{n_0} = \mathrm{i} \sum_{n_1+\cdots+n_j+m_1+\cdots+m_\mu=n_0} \Theta_\mu^j \prod_{l=1}^{j} u_{n_l} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i}, \qquad j \in \{3,4\}, \tag{6.35}$$

for positive n_0*. For each* (μ, j)*, the function*

$$\Theta_\mu^j = \Theta_\mu^j(n_0, n_1, \ldots, n_j, m_1, \ldots, m_\mu), \qquad j \in \{2,3,4\},$$

is a linear combination of products $\mathbf{1}_E \cdot \Theta$*, where* E *is some set generated by the sets* $\{n_h + n_l = 0\}$*,* $1 \leq h < l \leq j$*, and* Θ *is* slowly varying of type 1 (*later we may slightly abuse the notation and use the term "*Θ *factor" or "*Θ^j *factor" to refer to both the* Θ_μ^j *and the* Θ *here*)*; note in particular they are* real-valued*. Moreover:*

(i) *When* $j = 2$*,* Θ *is nonzero only when*

$$\max_i \langle m_i\rangle \ll (\mu+1)^{-2} \min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\}. \tag{6.36}$$

(ii) *When* $j = 3$*, if* E *is contained in* $\{n_1 + n_2 = 0\}$ but not $\{|n_1| = |n_2| = |n_3|\}$*, we must have*

$$|\Theta| \lesssim \min\left\{1, \frac{\langle n_0\rangle + \langle n_3\rangle}{\langle n_1\rangle}\right\}. \tag{6.37}$$

The same holds for other permutations of $(1,2,3)$*.*

(iii) *When* $j = 4$*, we have*

$$|\Theta| \lesssim \left(\max_{0\leq l\leq 4} \langle n_l\rangle\right)^{-1}. \tag{6.38}$$

Proof. The estimate on the coefficients C_μ, whose choice will be clear from the expressions we have, is elementary based on the factorial decay we have, and the simple observation that

$$\frac{(\mu_1 + \cdots + \mu_k)!}{\mu_1! \cdots \mu_k!} \leq k^{\mu_1+\cdots+\mu_k}, \tag{6.39}$$

where in practice we always have (say) $k \leq 30$. Next we shall examine the terms left after the subtraction of resonant ones, and define the Θ factors. We will first prove the boundedness of Θ and properties (i)–(iii), and then show that $\Theta \in SV_1$.

Before proceeding, let us make one useful observation. If we have a term (temporarily called a *term of type R* for convenience) of type (6.34) in which the Θ factor is bounded and is accompanied by some $\mathbf{1}_E$ with $E \subset \{n_1 + n_2 \neq 0\}$, then we can use a smooth cutoff (similar to ψ_1 or ψ_2) to separate the part where (6.36) holds, and the part where $\langle m_i\rangle \gtrsim (\mu+1)^{-2}\min\{\langle n_0\rangle,\langle n_1\rangle,\langle n_2\rangle\}$ for some $1\leq i\leq \mu$; in the former case we have $\mathcal{H}_\mu^2$, and in the latter case we promote m_i and rename it n_3 to obtain $\mathcal{H}_{\mu-1}^3$ (since here the Θ^3 factor is bounded, $n_1+n_2\neq 0$, and if $n_l+n_3=0$ for $l\in\{1,2\}$, then the Θ^3 factor will have an n_3 in the denominator, and at most $\langle n_0\rangle$ in the numerator). Also we may assume that all the n_l and m_i variables are nonzero and $\lesssim N$.

The first contribution we need to consider is when none of the equalities we proposed in obtaining the $\mathcal{R}^j$ terms hold; these include the contribution from each $\mathcal{K}^j$ which we discuss separately.

For the part in $\mathcal{K}_{\mu_1 i}^1$ we have $n_1+n_2\neq 0$. If $\langle n_2\rangle \gtrsim \langle n_0\rangle+\langle n_1\rangle$, then we have a term of type R and obtain either $\mathcal{H}_{\mu_1-1}^2$ or $\mathcal{H}_{\mu_1-2}^3$. Now if $\langle n_2\rangle \ll \langle n_0\rangle+\langle n_1\rangle$, then $\langle m_j\rangle \gtrsim \langle n_0\rangle+\langle n_1\rangle$ for at least one j, so we can promote that m_j and rename it n_3 to obtain $\mathcal{H}_{\mu_1-2}^3$, due to a similar argument as above and the restriction $n_1+n_2\neq 0$.

For the part of $\mathcal{K}_{\mu_1\mu_2}^2$, no $n_h+n_l=0$ happens. In the expression (5.21), first assume $\langle n_3\rangle$ (or, by symmetry, $\langle n_1\rangle$) is $\lesssim \min\{\langle n_0\rangle,\langle n_1\rangle,\langle n_3\rangle\}$. Then the first two terms in the bracket on the right hand side of (5.21) contribute at most $O(\langle n_3\rangle)$, so for these terms we may relegate n_2 (rename it by some m_i) to obtain a contribution of type R. For the last two terms in the bracket, the contribution is at most $N^{-1}\langle n_1\rangle(\langle n_2\rangle+\langle n_3\rangle)$, which is a sum of two terms. One of them is at most $\langle n_3\rangle$ and can be treated as above; the other can be canceled by the n_2^{-1} factor and we get $\mathcal{H}_{\mu_{12}}^3$ (since we have pre-assumed that no n_h+n_l can be zero).

Next suppose (say) $\langle n_0\rangle \ll \langle n_3\rangle \ll \langle n_1\rangle$. In this case the first two terms in the bracket on the right hand side of (5.21) contribute at most $\langle n_3\rangle$, and at least one of $\langle m_j\rangle$ or $\langle n_2\rangle$ must be $\gtrsim \langle n_1\rangle$ here, so we get $\mathcal{H}_{\mu_{12}-1}^3$ after making appropriate promotion or relegations; the last two terms contribute at most $N^{-1}\langle n_1\rangle(\langle n_2\rangle+\langle n_3\rangle)$, which is bounded either by $\langle n_3\rangle$ (which can be treated the same way as above), or $N^{-1}\langle n_1\rangle\langle n_2\rangle$ (which is canceled by the n_2^{-1} to obtain $\mathcal{H}_{\mu_{12}-1}^3$).

The only remaining possibility is $\langle n_0\rangle \ll \langle n_1\rangle \sim \langle n_3\rangle$. We may write

$$\tau_1(n)=\prod_{i=1}^{\mu_2}\psi^2\Big(\frac{n+m_{\mu_1+i,\mu_{12}}}{N}\Big), \tag{6.40}$$

$$\tau_2(n)=\psi^2\Big(\frac{n+n_2+m_{\mu_1+1,\mu_{12}}}{N}\Big)\tau_1(n), \tag{6.41}$$

so the net contribution in the bracket will be

$$\big(n_3\tau_2(n_3)+n_1\tau_2(n_1)\big)-\psi^2\big(n_3\tau_1(n_3)+n_1\tau_1(n_1)\big)$$

with some factor ψ. Since we can write

$$n_3\tau_j(n_3)+n_1\tau_j(n_1)=(n_1+n_3)\tau_j(n_3)+n_1(\tau_j(n_1)-\tau_j(n_3)), \tag{6.42}$$

and n_1+n_3 is a linear combination of n_0, n_2 and m_i, the first term on the right hand side of (6.42) will be bounded either by $\langle n_0\rangle$ (in which case we have a term of type R), or by $\langle n_2\rangle$ (in which case we obtain $\mathcal{H}^3_{\mu_{12}-1}$), or by some $\langle m_j\rangle$ (in which case we relegate n_2 and promote m_j to obtain $\mathcal{H}^3_{\mu_{12}-1}$ under the restriction $\langle n_1\rangle\sim\langle n_3\rangle$). The contribution of the second term will be bounded by $N^{-1}\langle n_1\rangle$ times either $\langle n_0\rangle$ (in which case we have a term of type R), $\langle n_2\rangle$ (in which case we have a part of $\mathcal{H}^3_{\mu_{12}-1}$), or some $\langle m_j\rangle$ (in which case we relegate n_2 and promote m_j to get $\mathcal{H}^3_{\mu_{12}-1}$).

For the part of $\mathcal{K}^3_{\mu_1\mu_2 0}$ we have $n_1+n_2\neq 0$. By the assumptions about this term, if $\langle n_0\rangle\gtrsim\langle n_2\rangle$, we will have a term of type R. Now assuming $\langle n_0\rangle\ll\langle n_2\rangle$, we can extract from the bracket in (5.37) a factor of n_0/N or m_i/N. If we have an n_0/N factor then the net Θ factor will be $\lesssim\langle n_0\rangle$ and we again have a term of type R; if we have an m_i/N factor then we may cancel it with the $1/m_i$ factor, promote this m_i and rename it n_3, to obtain $\mathcal{H}^3_{\mu_{12}-2}$. Notice that in this case the Θ factor is bounded by $\langle n_2\rangle/N\lesssim 1$, $n_1+n_2\neq 0$, and if $n_2+n_3=0$, we must have $\langle n_1\rangle\gtrsim\langle n_2\rangle$.

For the part of $\mathcal{K}^3_{\mu_1\mu_2 i}$ we have $n_2+n_3\neq 0$. We claim that this part is $\mathcal{H}^3_{\mu_{12}-1}$. In fact, this will be the case if both n_1+n_3 and n_1+n_2 are nonzero since the Θ factor is bounded; if $n_1+n_3=0$, then from the assumptions about the $\mathcal{K}^3_{\mu_1\mu_2 i}$ term we have $\langle n_0\rangle\gtrsim\langle n_3\rangle$, so we also have $\mathcal{H}^3_{\mu_{12}-1}$; if $n_1+n_2=0$, then either $\langle n_0\rangle$ or $\langle n_3\rangle$ must be $\gtrsim\langle n_1\rangle$, so we still have $\mathcal{H}^3_{\mu_{12}-1}$.

For the part of $\mathcal{K}^5_{\mu_1\mu_2\mu_3}$, no $n_h+n_l=0$ happens. In this case the Θ factor is clearly bounded, thus we obtain $\mathcal{H}^3_{\mu_{13}}$.

Next, we have the "error term" which is some resonant contribution in $\mathcal{K}^j$ (for example, the contribution in $\mathcal{K}^1_{\mu_1 i}$ where $n_1+n_2=0$) minus the corresponding $\mathcal{R}^j$. In this term we may specify some k (for example, in the term corresponding to $\mathcal{K}^1_{\mu_1 i}$ we will have $n_1=-k$ and $n_2=k$). From the computations made before, we can see that the corresponding terms may be written in an appropriate form so that the Θ factor is bounded even without subtracting $\mathcal{R}^j$. Note that here we may need to promote some m_i so that we can include m_i^{-1} in Θ to cancel certain factors (for example when dealing with $\mathcal{K}^1_{\mu_1 i}$). Therefore, *before subtracting the* $\mathcal{R}^j$ *terms*, the resonant contributions can be written in the form of (6.35), with $j=3$, the Θ factor bounded, and (say) $n_1=-k$, $n_2=k$. In particular, if $\langle n_0\rangle+\langle n_3\rangle\gtrsim\langle k\rangle$, we will obtain $\mathcal{H}^3$ and subtraction of $\mathcal{R}^j$ will not affect this. Now we assume $\langle n_0\rangle+\langle n_3\rangle\ll\langle k\rangle$.

After the subtraction of the $\mathcal{R}^j$ factors, the Θ will remain bounded; moreover, it can be checked case-by-case that in the remaining term, we gain an additional factor of

$$\min\left\{1,\frac{1}{\langle k\rangle}\left(\langle n_0\rangle+\langle n_3\rangle+\sum_{i=1}^{\mu}\langle m_i\rangle\right)\right\},\tag{6.43}$$

if $n_1=-k$ and $n_2=k$. For example, say we are replacing $\prod_j\psi^2((k+\sigma_j)/N)-\prod_j\psi^2((k-\sigma_j)/N)$ by $4\theta^{2\mu-1}\eta\sum_j\sigma_j/N$; then the error term we introduce is at most $O(N^{-2}\langle\sigma_j\rangle^2)$, which is then at most $O(N^{-2}\langle n_l\rangle^2)$ or $O(N^{-2}\langle m_i\rangle^2)$ for some i and $l\in\{0,3\}$. Since this contribution can be canceled by other factors to produce a bounded Θ even if we replace the power of 2 by 1 (which will be the case if we do not subtract the $\mathcal{R}^j$), we will have in the error term an additional factor as in (6.43). The other factors are treated in the same way, provided that in some cases we replace the N in the denominator by something larger than $\langle k\rangle$. This guarantees that either we obtain $\mathcal{H}^3$, or we may promote some m_i to obtain $\mathcal{H}^4$.

Next, notice that in obtaining $\mathcal{R}^{2.1}$, we have discarded the last two terms in the bracket on the right hand side of (6.6). However, they add up to produce a factor of at most $N^{-1}\langle n_1\rangle$, thus they can be included in $\mathcal{H}^3$. Finally, there are terms where at least *two* of the proposed equalities hold (these terms appear due to the inclusion-exclusion principle), for example we have the term where $n_1 + n_2 = n_2 + n_3 = 0$ in $\mathcal{K}^5_{\mu_1\mu_2\mu_3}$; but by the discussion above, the corresponding Θ factor will be bounded, thus they can also be included in $\mathcal{H}^3$.

Now we only need to show $\Theta \in SV_1$. This will follow from Proposition 6.2, since it can be checked that all the Θ factors are formed using rules (1) through (5) in that proposition, with rule (2) used at least once (in particular, all the cutoff factors we introduce will be in SV_1). □

7. The gauge transform III: The final substitution

Starting from equations (6.33)–(6.35), we need to make further substitutions before we can state and prove the main estimates. Here we introduce one more notation, namely when we write g^ω for a function g, where $\omega \in \{-1, 1\}$, this will mean g if $\omega = 1$, and $\overline{g}$ if $\omega = -1$. Also in the following, we will use the letter υ to represent a function that can be either u or v.

7.1. *From u to w*

Recalling that $v = Mu$ and $w = \mathbb{P}_+ v$, we have

$$u = \sum_\mu \frac{\mathrm{i}^\mu}{2^\mu \mu!} P^\mu v,$$

which then implies, for $n > 0$,

$$u_n = \sum_\mu \frac{1}{2^\mu \mu!} \sum_{n_1 + m_1 + \cdots + m_\mu = n} \Psi_\mu \cdot v_{n_1} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i}, \tag{7.1}$$

where

$$\Psi_\mu = \Psi_\mu(n, n_1, m_1, \ldots, m_\mu)$$

is a product of ψ factors. When $n < 0$, since $u_n = \overline{u_{-n}}$, we have instead

$$u_n = \sum_\mu \frac{(-1)^\mu}{2^\mu \mu!} \sum_{n_1 + m_1 + \cdots + m_\mu = n} \Psi_\mu \cdot (\overline{v})_{n_1} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i}, \tag{7.2}$$

where we note $(\overline{v})_n = \overline{v_{-n}}$. By replacing each u_{n_l} in (6.34) and (6.35) with one of the above expressions, we can prove

Proposition 7.1. *We have*

$$(\partial_t - \mathrm{i}\partial_{xx})w = \mathcal{J} = \sum_{\mu} C_\mu \mathcal{J}_\mu, \tag{7.3}$$

where $|C_\mu| \lesssim C^\mu/\mu!$ *and the nonlinearity is written as*

$$\mathcal{J}_\mu = \sum_{j\in\{2,3,3.5,4,4.5\}} \sum_{\omega\in\{-1,1\}^{\lfloor j\rfloor}} \mathcal{J}_\mu^{\omega j} \tag{7.4}$$

with

$$(\mathcal{J}_\mu^{\omega j})_{n_0} = \mathrm{i} \sum_{n_1+\cdots+n_j+m_1+\cdots+m_\mu=n_0} \phi_\mu^j \prod_{l=1}^{j} (w^{\omega_l})_{n_l} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i} \tag{7.5}$$

for $j \in \{2, 3\}$, *and*

$$(\mathcal{J}_\mu^{\omega j})_{n_0} = \mathrm{i} \sum_{n_1+\cdots+n_{\lfloor j\rfloor}+m_1+\cdots+m_\mu=n_0} \phi_\mu^j \prod_{l=1}^{\lfloor j\rfloor} (\upsilon^{\omega_l})_{n_l} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i} \tag{7.6}$$

for $j \in \{3.5, 4, 4.5\}$. *Here the real-valued weights*

$$\phi_\mu^j = \phi_\mu^j(n_0, n_1, \ldots, n_{\lfloor j\rfloor}, m_1, \ldots, m_\mu), \tag{7.7}$$

where $j \in \{2, 3, 3.5, 4, 4.5\}$, *satisfy the following.*

(i) *When* $j = 2$, *we have*

$$|\phi_\mu^2| \lesssim \min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\};$$

also ϕ_μ^2 *is nonzero only when*

$$\min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\} \gg (\mu+1)^2 \max_i \langle m_i\rangle.$$

(ii) *When* $j = 3$, *we have* $|\phi_\mu^3| \lesssim 1$; *also when* $n_1 + n_2 = 0$, *and neither* n_1 *nor* n_2 *is related to* n_3 *by* m (*here and after, we say two* n *variables are "related by* m*" if their sum or difference belongs to some fixed, finite set of linear combinations of the* m *variables*), *we have*

$$|\phi_\mu^3| \lesssim \min\left\{1, \frac{\langle n_0\rangle + \langle n_3\rangle}{\langle n_1\rangle}\right\},$$

and the estimate also holds for other permutations of $(1, 2, 3)$. *Also, when all three of* (n_1, n_2, n_3) *are related by* m, *we are allowed to have* $(\upsilon^{\omega_l})_{n_l}$ *instead of* $(w^{\upsilon_l})_{n_l}$ *in* (7.5) *for* $j = 3$.

(iii) *When* $j = 3.5$, *we have*

$$|\phi_\mu^{3.5}| \lesssim \frac{\min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2+n_3\rangle\}}{\max\{\langle n_2\rangle, \langle n_3\rangle\}}. \tag{7.8}$$

Moreover, we can replace the υ *in* $(\upsilon^{\omega_1})_{n_1}$ *in* (7.6) *for* $j = 3.5$ *by* w; *also, if*

$$\left(\max_{0\le l\le 3} \langle n_l\rangle\right)^{1/2} \ll \min_{0\le l\le 3} \langle n_l\rangle,$$

then n_2 *and* n_3 *must have opposite sign.*

(iv) *When* $j = 4$, *we have*

$$|\phi_\mu^4| \lesssim \Big(\max_{0\le l\le 4} \langle n_l\rangle^{1/20} + \min_{1\le l\le 4}\langle n_l\rangle\Big)^{-1}.$$

(v) *When* $j = 4.5$, *we have* $n_1 + n_2 \neq 0$, *and*

$$|\phi_\mu^{4.5}| \lesssim (\langle n_3\rangle + \langle n_4\rangle)^{-1},$$

and this factor is nonzero only if

$$\max\{\langle n_3\rangle, \langle n_4\rangle\} + \max_i \langle m_i\rangle \ll (\max\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\})^{1/10};$$

also, whenever

$$\langle n_l\rangle \gtrsim (\max\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\})^{1/10}$$

for some $l \in \{1, 2\}$, *we can replace the* υ *in* $(\upsilon^{\omega_l})_{n_l}$ *in* (7.6) *for* $j = 4.5$ *by* w.

(vi) *When* $j = 3$, *suppose* $n_0 = n_1 = n, -n_2 = n_3 = k$, $\langle k\rangle \ll (\mu+3)^{-11}\langle n\rangle$, *and (if necessary) restrict to the set* $\{k > 0\}$ *or* $\{k < 0\}$, *then the* ϕ_μ^3 *factor will be a function of* n, k *and other variables. This function can then be divided into two parts, with the first part satisfying*

$$|\phi_\mu^3(n, k, m_1, \ldots, m_\mu)| \lesssim \frac{\min\{\langle n\rangle, \langle k\rangle\}}{\max\{\langle n\rangle, \langle k\rangle\}}, \tag{7.9}$$

and the second part satisfying

$$|\phi_\mu^3(\mathbf{w}) - \phi_\mu^3(\mathbf{w}')| \lesssim \langle k\rangle^{-1}, \tag{7.10}$$

where $\mathbf{w} = (n, k, m_1, \ldots, m_\mu)$, *and* $\mathbf{w}'$ *differs from* $\mathbf{w}$ *by* 1 *in exactly one component; if this component is* n *or* k, *then the right hand side of* (7.10) *should be replaced by* $\langle n\rangle^{-1}$.

Proof. We will first prove (i) through (v) as well as (7.9); the proof of (7.10) will be left to the end. Since each $\mathcal{H}^4$ term is also a $\mathcal{J}^4$ term, we only need to consider the expressions (6.34) and (6.35) with $j \in \{2, 3\}$. We replace each u_{n_l}, where $1 \le l \le j$, by either (7.1) or (7.2), depending on whether n_l is positive or negative, to obtain

$$\mathcal{H}_{\mu_0}^j = \sum_{\omega;\mu_1,\ldots,\mu_j} C_{\mu_0}\frac{\omega_1^{\mu_1}\cdots\omega_j^{\mu_j}}{2^{\mu_{1j}}\mu_1!\cdots\mu_j!}\mathcal{H}_{\mu_0\cdots\mu_j}^{\omega j} \tag{7.11}$$

for all μ_0 and $j \in \{2, 3\}$, where $\omega = (\omega_1, \ldots, \omega_j) \in \{-1, 1\}^j$, and

$$(\mathcal{H}_{\mu_0\cdots\mu_j}^{\omega j})_{n_0} = \sum_{\mathbf{w}\in V_{n_0,\mu_0\cdots\mu_j}^{\omega j}} \Theta_{\mathbf{w}}^{\mu_0\cdots\mu_j j}\prod_{l=1}^j (\upsilon^{\omega_l})_{n_l'}\prod_{l=0}^j\prod_{i=1}^{\mu_l}\frac{u_{(m^l)_i}}{(m^l)_i}. \tag{7.12}$$

Here the frequency set is

$$V_{n_0,\mu_0\cdots\mu_j}^{\omega j} = \{\mathbf{w} = ((n_l, n_l')_{1\le l\le j}, ((m^l)_i)_{1\le i\le\mu_l;0\le l\le j}) : n_l = n_l' + (m^l)_{1\mu_l}, \ \omega_l n_l > 0,$$
$$n_1 + \cdots + n_j + (m^0)_{1\mu_0} = n_0\}. \tag{7.13}$$

Note that the free variables are n_l' and $(m^l)_i$, and they satisfy a constraint

$$\sum_{l=1}^{j} n_l' + \sum_{l=0}^{j} \sum_{i=1}^{\mu_l} (m^l)_i = n_0$$

as well as several inequalities. Also the weight is

$$\Theta_{\mathbf{w}}^{\mu_0\mu_1\mu_2 2} = \Theta_{\mu_0}^2(n_0, n_1, n_2, (m^0)_1, \dots, (m^0)_{\mu_0}) \times \min_{0\le l\le 2} \langle n_l\rangle \cdot \prod_{l=1}^{j} \Psi_{\mu_l}(n_l, n_l', (m^l)_1, \dots, (m^l)_{\mu_l}), \tag{7.14}$$

$$\Theta_{\mathbf{w}}^{\mu_0\cdots\mu_3 3} = \Theta_{\mu_0}^3(n_0, \dots, n_3, (m^0)_1, \dots, (m^0)_{\mu_0}) \times \prod_{l=1}^{j} \Psi_{\mu_l}(n_l, n_l', (m^l)_1, \dots, (m^l)_{\mu_l}). \tag{7.15}$$

Our argument will be an enumerative examination of all the possible terms, and this can be greatly simplified with the following lemma, which we will assume for now, and prove after the proof of this main proposition.

Lemma 7.2. *We say a term has type A if it has the form* (7.12)*, with some factor Θ' in place of $\Theta_{\mathbf{w}}^{\mu_0\cdots\mu_j j}$, which is bounded by*

$$|\Theta'| \lesssim \min_{0\le j\le 2} \langle n_j\rangle \cdot \min\left\{1, \frac{\langle (m^l)_i\rangle}{\langle n_l\rangle + \langle n_l'\rangle}\right\}, \qquad j=2, \tag{7.16}$$

$$|\Theta'| \lesssim \min\left\{1, \frac{\langle (m^l)_i\rangle}{\langle n_l\rangle + \langle n_l'\rangle}\right\}, \qquad j=3, \tag{7.17}$$

for some $l \ge 1$ and $1 \le i \le \mu_l$. Moreover assume that (1) *either there is some $h \ne l$ such that $n_l' + n_h' = 0$, or no $n_j' + n_k' = 0$ regardless of whether j or k is equal to l;* (2) *either $(m^l)_i n_l' < 0$, or the v in $(v^{\omega_l})_{n_l'}$ is replaced by w. Then this term will be $\mathcal{J}^b$ for some $b \in \{3, 3.5, 4, 4.5\}$.*

We now start to analyze the sum (7.12). Note that the $\Theta_{\mu_0}^2$ in (7.14) and the $\Theta_{\mu_0}^3$ in (7.15) are fixed linear combinations of products $\mathbf{1}_E \cdot \Theta$ (recall Proposition 6.3), so we only need to consider one product of this type.

First, we collect the terms in (7.12) where $n_h' + n_l' = 0$ for some $1 \le h \ne l \le j$. We fix such a pair (h, l) and fix a $k > 0$ (the case $k = 0$ being trivial) so that $n_h' = k$ and $n_l' = -k$, then we fix ω and all the μ's except for μ_h and μ_l, and fix all the variables except for $(m^h)_i$ and $(m^l)_i$. There are then two possibilities.

(1) If $(\omega_h, \omega_l) \ne (1, -1)$, say $\omega_l = 1$, then from (7.13) we have $(m^l)_{1\mu_l} - k > 0$, which implies $\langle k\rangle \lesssim \langle (m^i)_i\rangle$ for some i, and we may assume that $(m^l)_i$ has opposite sign to k. Therefore we get a term of type A and reduce to Lemma 7.2.

(2) If $(\omega_h, \omega_l) = (1, -1)$, then in particular we may replace the v in $(v^{\omega_h})_{n_h'}$ and $(v^{\omega_l})_{n_l'}$ by w in (7.12). Now we make the restriction that $\langle (m^h)_i\rangle \ll (\mu+1)^{-2}\langle k\rangle$ for all $1 \le i \le \mu_h$ and the same for l, where μ is the sum of all μ_j, including μ_h and μ_l. It is

important to notice that this restriction depends only on $\mu_h+\mu_l$; also, the remaining part is of type A and can be treated using Lemma 7.2.

Next, assume $\mu_h+\mu_l>0$; we will replace the Ψ_{μ_h} factor in (7.14) and (7.15) by $\psi^{2\mu_h}(n_h'/N)$ and the same for l; thus the modified version of $\Psi_{\mu_h}\Psi_{\mu_l}$ will depend only on $\mu_h+\mu_l$. Also we may replace the n_h and n_l appearing in Θ factors of $\Theta_{\mu_0}^{j}$ functions, as well as $\min_{0\leq j\leq 2}\langle n_j\rangle$, by n_h' $(=k)$ and n_l' $(=-k)$; note that we are *not* doing this for the $\mathbf{1}_E$ factor. Now, since the Θ factors and the Ψ factors are in SV_1, $\min_{0\leq j\leq 2}\langle n_j\rangle$ is in SV_2, and we already have $\langle n_h\rangle\sim\langle n_h'\rangle$ and the same for l, we can easily show that the error introduced in this way will be of type A.

Now, apart from the $\mathbf{1}_E$ factors, we have replaced $\Theta_{\mathbf{w}}^{\mu_0\cdots\mu_j j}$ with some Θ' independent of the $(m^h)_i$ and the $(m^l)_i$ variables. Regarding the $\mathbf{1}_E$ factor, let us consider the case $E=\{n_{l'}+n_{h'}=0\}$. If $\{l',h'\}=\{l,h\}$, this factor will again be independent of the chosen m variables (since it only depends on $(m^h)_{1\mu_h}+(m^l)_{1\mu_l}$ which is fixed). Therefore, up to an error term which only involves the summation where $j=3$, the weight $\Theta_{\mathbf{w}}^{\mu_0\cdots\mu_3 3}$ factor is bounded, and all three of (n_1',n_2',n_3') are related by m (thus it will be $\mathcal{J}^3$), we may assume that Θ' is completely independent of the $(m^h)_i$ and $(m^l)_i$ variables.

Next, we will fix μ_h and μ_l, so that we are summing over $(m^h)_i$ and $(m^l)_i$, the restriction being

$$(m^h)_{1\mu_h}+(m^l)_{1\mu_l}=\mathrm{cst},\qquad \max\{\langle (m^h)_i\rangle,\langle (m^l)_{i'}\rangle\}\ll(\mu+1)^{-2}\langle k\rangle,\tag{7.18}$$

where the constant depends on the other fixed variables, and the summand will be

$$\prod_{i=1}^{\mu_h}\frac{u_{(m^h)_i}}{(m^h)_i}\prod_{i=1}^{\mu_l}\frac{u_{(m^l)_i}}{(m^l)_i}.\tag{7.19}$$

Note that when each $(m^h)_i$ and $(m^l)_{i'}$ is small, the restriction

$$(m^h)_{1\mu_h}>-k,\qquad (m^l)_{1\mu_l}<k,\tag{7.20}$$

which comes from (7.13), *will be void*. Now we can see that this sum actually depends *only on* $\mu_h+\mu_l$, thus when we sum over μ_h fixing $\mu_h+\mu_l$, we will get zero, since

$$\sum_{\mu_1+\mu_2=\mu>0}\frac{(-1)^{\mu_2}}{\mu_1!\mu_2!}=0.\tag{7.21}$$

Therefore all the terms in this case can be treated using Lemma 7.2.

We still need to consider when $\mu_h=\mu_l=0$. In this case we have $n_h=k$ and $n_l=-k$. Note in particular we must have $j=3$ due to the restriction (i) in Proposition 6.3; we may assume $h=1$ and $l=2$, so the $\Theta_{\mathbf{w}}^{\mu_0\cdots\mu_3 3}$ factor is bounded by

$$\min\left\{1,\frac{\langle n_0\rangle+\langle n_3\rangle}{\langle k\rangle}\right\}\tag{7.22}$$

provided $n_3\neq\pm k$, which we may assume since otherwise all three of (n_1',n_2',n_3') will be related by m and we will get $\mathcal{J}^3$. Now if $\langle (m^3)_i\rangle\ll(\mu_3+1)^{-2}\langle n_3\rangle$ for all i, then

$\langle n_3\rangle \sim \langle n_3'\rangle$ and the v in $(v^{\omega_3})_{n_3'}$ may be replaced by w, thus we will have $\mathcal{J}^3$; otherwise we may promote some $(m^3)_i$ and rename it n_4, and it can be easily checked that this part will be $\mathcal{J}^4$.

Now we collect the terms where no two n_l' add to zero. Among these, we will first take out the part where $\langle (m^l)_i\rangle \gtrsim (\mu_{0l}+1)^{-2}\langle n_i\rangle$ for at least one $1 \leq l \leq j$ and at least one $1 \leq i \leq \mu_l$, since this again will be of type A. In what remains, we will have $\langle n_l'\rangle \sim \langle n_l\rangle$, and that $\omega_l n_l' > 0$, and we may replace the v in $(v^{\omega_l})_{n_l'}$ by w. Now when $j = 3$, we already obtain a part of $\mathcal{J}^3$. Finally, when $j = 2$ we separate the cases where

$$\max_{l,i}\langle (m^l)_i\rangle \ll (\mu_{0j}+1)^{-2}\min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\} \tag{7.23}$$

or otherwise, again by inserting smooth cutoffs. If (7.23) holds we get a part of $\mathcal{J}^2$; if (7.23) fails, we can promote some $(m^l)_i$ and call it n_3 so that the new Θ factor is bounded, and then replace u_{n_3} by (7.1) or (7.2), introducing the n_3' and $(m^3)_i$ variables. Now, if $\langle (m^3)_i\rangle \ll \langle n_3\rangle$ for all i, so that $\langle n_3'\rangle \sim \langle n_3\rangle$ and the v in $(v^{\omega_3})_{n_3'}$ may be replaced by w, we get $\mathcal{J}^3$ due to the same argument as in the proof of Proposition 6.3; otherwise we could promote some $(m^3)_i$ to be n_4. We then obtain $\mathcal{J}^4$ if one of n_3, n_3', n_4 or the remaining m variables is $\gtrsim (\max\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\})^{1/10}$, and obtain a part of $\mathcal{J}^{4.5}$ otherwise.

Finally, to prove part (vi), first notice that in (7.10) we may assume each $\langle m_i\rangle \ll \langle k\rangle$ also, since otherwise we will have $\mathcal{J}^4$ or $\mathcal{J}^{4.5}$. It can then be checked that *in this particular case*, every term in $\mathcal{J}^3$ will involve no characteristic functions other than $\mathbf{1}_{\{m_i\neq 0\}}$ (which can always be ignored) or $\mathbf{1}_{\{k>0\}}$ or $\mathbf{1}_{\{k<0\}}$ (which is taken care of when we restrict to positive or negative k); in particular any $\mathbf{1}_E$ factor introduced before will be constant here. What remains in the ϕ^3 factor are simple linear fractions and cutoff functions, and for them (7.10), which is a stronger property than being in SV_1, can be directly verified. □

Proof of Lemma 7.2. Fix the l and the i in (7.17), and first suppose $j = 2$. We may assume $l = 2$ and promote the $(m^2)_i$ by calling it n_3. Then the new Θ factor will be bounded by

$$\frac{\min\{\langle n_0\rangle, \langle n_1\rangle, \langle n_2\rangle\}}{\max\{\langle n_2\rangle, \langle n_3\rangle, \langle n_2'\rangle\}}.$$

Notice that

$$\langle n_1\rangle \lesssim \langle n_1'\rangle + \sum_i \langle (m^1)_i\rangle, \qquad \langle n_2\rangle \lesssim \langle n_2' + n_3\rangle + \sum_i \langle (m^2)_i\rangle,$$

thus the Θ factor will be bounded either by

$$\frac{\min\{\langle n_0\rangle, \langle n_1'\rangle, \langle n_2'+n_3\rangle\}}{\max\{\langle n_2'\rangle, \langle n_3\rangle\}} \tag{7.24}$$

or by some

$$\min\left\{1, \frac{\langle (m^l)_i\rangle}{\max\{\langle n_2\rangle, \langle n_3\rangle, \langle n_2'\rangle\}}\right\}, \qquad l\in\{1,2\},\ 1\leq i\leq \mu_l.$$

In the former case the bound (7.8) is already verified, and we will have a part of $\mathcal{J}^{3.5}$ if $\omega_1 n_1' > 0$ and n_2' has different sign with n_3. If $\omega_1 n_1' \leq 0$, we find for some $1 \leq i \leq \mu_1$ that $\langle n_1 \rangle + \langle n_1' \rangle \lesssim \langle (m^1)_i \rangle$ and we are reduced to the latter case above. Now in the latter case, we promote $(m^l)_i$ and call it n_4, so that we get an expression of the form

$$\sum_{n_1+\cdots+n_4+m_1+\cdots+m_\mu=n_0} \Phi \cdot \prod_{l=1}^{4} (\upsilon^{\omega_l})_{n_l} \prod_{i=1}^{\mu} \frac{u_{m_i}}{m_i}, \tag{7.25}$$

where we may assume $\omega_l(n_l+\lambda_l) > 0$, where λ_l is some linear combination of the m variables, and the υ in $(\upsilon^{\omega_l})_{n_l}$ can be replaced by v for $l \in \{1, 2\}$; also the Φ factor is bounded by $(\langle n_2 \rangle + \langle n_3 \rangle + \langle n_4 \rangle)^{-1}$. Now if one of n_l $(2 \leq l \leq 4)$ or m_i is $\gtrsim \max\{\langle n_0 \rangle, \langle n_1 \rangle\}^{1/20}$ we will obtain a part of $\mathcal{J}^4$; otherwise we must have $\omega_1 n_1 > 0$ and thus we are in $\mathcal{J}^{4.5}$.

Now, if we have some term similar to $\mathcal{J}^{3.5}$ (i.e. with Θ factor bounded by (7.24)), with $\omega_1 n_1 > 0$ but n_2 and n_3 have the same sign (note the n_j here was n_j' before we renamed it), then from the definition of type A terms, we can replace the υ in $(\upsilon^{\omega_j})_{n_j'}$ by w for $j \in \{1, 2\}$. Next, we replace u_{n_3} by (7.1) or (7.2) according to the sign of n_3, introducing the n_3' and $(m^3)_i$ variables. Under the assumption $\langle n_3 \rangle \gg \max_{0 \leq l \leq 3} \langle n_l \rangle^{1/2}$, we may assume $\langle (m^3)_i \rangle \ll \langle n_3 \rangle^{1/4}$, otherwise we will have $\mathcal{J}^4$. In particular we have $(w^{\omega_3})_{n_3'}$ and the weight will be bounded by $\min\{\langle n_0 \rangle, \langle n_1 \rangle\}/\max\{\langle n_2 \rangle, \langle n_3' \rangle\}$. Since n_2 will have the same sign as n_3', we *cannot* have $n_0 = n_1$ or $n_2 + n_3' = 0$; then we can check that this term will be $\mathcal{J}^3$, and that it satisfies (7.9).

Now assume $j = 3$. We may assume $l = 3$, and by a similar argument we will obtain an expression of form (7.25), but with Φ factor bounded only by $(\langle n_3 \rangle + \langle n_4 \rangle)^{-1}$. If we can assume that some other n_j (say n_2) is related to $n_3 + n_4$ by m, then we can reduce to the case just studied. Otherwise we must have $n_1 + n_2 \neq 0$. Now we may assume that n_3, n_4 and all the m variables are $\ll (\max\{\langle n_0 \rangle, \langle n_1 \rangle, \langle n_2 \rangle\})^{1/20}$ or we are in $\mathcal{J}^4$; also, if for some $l \in \{1, 2\}$ we have $\langle n_l \rangle \gtrsim (\max\{\langle n_0 \rangle, \langle n_1 \rangle, \langle n_2 \rangle\})^{1/10}$ (we make this restriction by inserting a smooth cutoff), then $\omega_l n_l > 0$ and we can replace υ in $(\upsilon^{\omega_l})_{n_l}$ by w. Therefore we will have $\mathcal{J}^{4.5}$. □

7.2. *From w to w^**

We still need to remove from the right hand side of (7.3) the part that cannot be controlled directly, by means of a substitution which will be described in the following proposition.

Proposition 7.3. *We can define w^*, for each fixed time, by*

$$(w^*)_n = e^{-i\Delta_n} w_n, \tag{7.26}$$

where the Δ factors are

$$\Delta_n(t) = \int_0^t \delta_n(t')\, dt', \tag{7.27}$$

and the δ factors are

$$\delta_n = \left[\frac{1}{2}\psi^4\left(\frac{n}{N}\right) + \frac{2n}{N}\psi^3\left(\frac{n}{N}\right)\psi'\left(\frac{n}{N}\right)\right] \sum_{k=0}^{n} |w_k|^2 \tag{7.28}$$

for $n > 0$; notice we may replace the w_k by $(w^*)_k$ in this expression. We then extend δ_n and Δ_n to $n \leq 0$ so that they are odd in n, and define u^* and v^* by

$$(\upsilon^*)_n = e^{-\mathrm{i}\Delta_n}\upsilon_n, \qquad \upsilon \in \{u, v\}.$$

With these definitions, we have

$$(\partial_t - \mathrm{i}\partial_{xx})w^* = \mathcal{N}^2(w^*, w^*) + \sum_{j\in\{3,2.5,4,4.5\}} \mathcal{N}^j, \tag{7.29}$$

where $\mathcal{N}^j = \sum_\mu \sum_{\omega\in\{-1,1\}^{\lfloor j\rfloor}} C_\mu^{\omega j}\mathcal{N}_\mu^{\omega j}$ for each j with $|C_\mu^{\omega j}| \leq C^\mu/\mu!$. The nonlinearities are

$$(\mathcal{N}_\mu^{\omega 2}(f,g))_{n_0} = \sum_{n_1+n_2+m_1+\cdots+m_\mu=n_0} \Phi_\mu^2 \cdot e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})}(f^{\omega_1})_{n_1}(g^{\omega_2})_{n_2}\prod_{i=1}^{\mu}\frac{u_{m_i}}{m_i},$$

$$(\mathcal{N}_\mu^{\omega 3}(f,g))_{n_0} = \sum_{n_1+n_2+n_3+m_1+\cdots+m_\mu=n_0} \Phi_\mu^3 \cdot e^{-\mathrm{i}\Delta_{n_0}}\prod_{l=1}^{3}(w^{\omega_l})_{n_l}\prod_{i=1}^{\mu}\frac{u_{m_i}}{m_i},$$

$$(\mathcal{N}_\mu^{\omega j})_{n_0} = \sum_{n_1+\cdots+n_{\lfloor j\rfloor}+m_1+\cdots+m_\mu=n_0} \Phi_\mu^j \cdot e^{-\mathrm{i}\Delta_{n_0}}\prod_{l=1}^{\lfloor j\rfloor}(\upsilon^{\omega_l})_{n_l}\prod_{i=1}^{\mu}\frac{u_{m_i}}{m_i}$$

for $n_0 > 0$ and $j \in \{3.5, 4, 4.5\}$. Here $\Phi_\mu^j = \Phi_\mu^j(n_1, \ldots, n_{\lfloor j\rfloor}, m_1, \ldots, m_\mu)$, and these factors (and the corresponding terms they come from) satisfy the requirements in parts (i), (iii), (iv), (v) *of Proposition* 7.1 *for $j = 2, 3.5, 4, 4.5$, respectively.*

Finally, when $j = 3$ and we only consider the case where $\Phi_\mu^3 \neq 0$, we have one of the following: (a) *either three of the four variables $(-n_0, n_1, n_2, n_3)$ are related by m (in which case we are allowed to have υ instead of w), or no two of them add up to zero, and $|\Phi_\mu^3| \lesssim 1$;* (b) *up to some permutation, $n_1 + n_2 = 0 \neq n_0 - n_3$ and $|\Phi_\mu^3| \lesssim \min\{1, (\langle n_0\rangle + \langle n_3\rangle)/\langle n_1\rangle\}$;* (c) *up to some permutation, $n_0 = n_1$, and either $n_2 + n_3 \neq 0$ and $|\Phi_\mu^3| \lesssim 1$, or $n_2 = -n_3$ and $|\Phi_\mu^3| \lesssim \min\{\langle n_0\rangle, \langle n_2\rangle\}/\max\{\langle n_0\rangle, \langle n_2\rangle\}$.*

Proof. In the nonlinearity in Proposition 7.1, we will call any contribution a "manageable error" if it can be included in $\mathcal{J}^j$ with $j \in \{2, 3.5, 4, 4.5\}$, or it belongs to $\mathcal{J}^3$ with $n_0 = n_1 = n$ and $-n_2 = n_3 = k$ as in part (vi) of Proposition 7.1, with the ϕ factor satisfying (7.9). Now we collect the terms in $\mathcal{J}^3$ where $n_0 = n_1$, $n_2 + n_3 = 0$, $\langle n_2\rangle \ll (\mu + 3)^{-12}\langle n_1\rangle$, as well as the terms corresponding to other permutations. The sum of these terms can be written as

$$\mathrm{i}\cdot w_{n_0}\sum_{0<k\ll n_0}|w_k|^2\sum_{\mu\geq 0}C_\mu\sum_{m_1+\cdots+m_\mu=0}\Gamma\cdot\prod_{i=1}^{\mu}\frac{u_{m_i}}{m_i}, \tag{7.30}$$

where Γ is a certain function of n_0, k and m_i satisfying (7.10). Now, up to manageable errors, we can assume that Γ depends only on n_0 (effectively we replace k and m_i by zero in the expression of Γ); then the right hand side of (7.30) will be *exactly* $\mathrm{i}\delta_{n_0}w_{n_0}$, where

$$\delta_{n_0} = (\mathcal{L}_{n_0}(F))_0 \cdot \sum_{0<k\ll n_0}|w_k|^2, \tag{7.31}$$

where $(\cdot)_0$ means the projection onto zero frequency, and $\mathcal{L}_{n_0}$ is some holomorphic function depending on n_0, and F is the mean-zero antiderivative of u as before. Now we can define the δ_n and Δ_n factors accordingly and make the substitution, thus getting rid of the term in (7.30). The terms $\mathcal{J}^j$ with $j \neq 3$ are transformed into $\mathcal{N}^j$ without any change; for the remaining terms in $\mathcal{J}^3$, we can see by an easy enumeration that the coefficient Φ_μ^3 will meet our requirements.

Now, to compute the function $\mathcal{L}_{n_0}$ in (7.31), we need to track every part of the $\mathcal{J}^3$ term as enumerated in Sections 5 and 6, and in the proof of Proposition 7.1 (also we disregard those satisfying (7.9) since they will be manageable errors). For each of them we perform the reduction above and obtain a corresponding contribution to $\mathcal{L}_{n_0}$. It turns out that the computations here are very similar to those in Section 6 (although we will have slightly different terms), and each single contribution to $\mathcal{L}_{n_0}$ will be of the form

$$\exp\left(\frac{\mathrm{i}\theta^2 F}{2}\right) \cdot \sum_{\mu} \frac{1}{(\mu+2)!}\left(-\frac{\mathrm{i}F}{2}\right)^{\mu} \cdot \left(\mathcal{E} + \frac{n_0\eta}{N}\mathcal{F}\right)$$

similar to those appearing in Section 6, where $\theta = \psi(n_0/N)$, $\eta = \psi'(n_0/N)$, $\mathcal{E}$ and $\mathcal{F}$ are simple expressions involving a power of θ, similar to the $\mathcal{C}$ and $\mathcal{D}$ expressions appearing in Section 6.

Note that the above computations are completely algebraic; in the end we add up all contributions to find that $\mathcal{L}_{n_0}(F)$ is actually a constant (depending on n_0), namely

$$\mathcal{L}_{n_0}(F) = \frac{1}{2}\theta^4 + \frac{2n_0}{N}\theta^3\eta.$$

Finally in (7.31), the sum corresponding to $k \sim n$ is a manageable error, thus we get (7.28). □

Remark 7.4. In fact, all we need for the estimates below is that δ_n grows (in some appropriate sense) at most logarithmically in n, and that it is real-valued. The first property does not require any computation and follows immediately from the boundedness of the coefficients involved in $\mathcal{J}^3$, while the second property is also heuristically clear due to conservation of mass (and Gibbs measure).

8. The a priori estimate I: The general setting

In this section we state our main estimate that works for a single solution. Its proof will occupy Sections 9–11. There will be another version concerning the difference of two solutions, which will be stated and proved in Section 12.

8.1. The bootstrap

Let us fix a smooth solution u, defined on $\mathbb{R} \times \mathbb{T}$, to the equation (1.6), with the parameter $1 \ll N \leq \infty$. In what follows we will assume $N < \infty$, since the case $N = \infty$ will follow from a similar (and simpler) argument. The main estimate can then be stated as follows.

Proposition 8.1. *There exists an absolute constant C such that the following holds. Suppose $\|u(0)\|_{Z_1} \leq A$ for some large A. Then within a short time $T = C^{-1}e^{-CA}$, for the functions v and w defined in Section 5, and the functions u^*, v^* and w^* defined in Section 7, we have*

$$\|w^*\|_{Y_1^T} + \|v^*\|_{Y_2^T} + \|u^*\|_{Y_2^T} \leq Ce^{CA}, \tag{8.1}$$

$$\|\langle \partial_x \rangle^{-s^3} u\|_{(X_2 \cap X_3 \cap X_4)^T} \leq CA. \tag{8.2}$$

Here the space $X_2 \cap X_3 \cap X_4$ is normed by $\|\cdot\|_{X_2} + \|\cdot\|_{X_3}\| + \|\cdot\|_{X_4}$, for which we can easily show that Proposition 3.6 still holds.

Remark 8.2. The constant C will depend on the constants in the inequalities in earlier sections, such as Propositions 3.6 and 7.3. To make this clear, we will now use C_0 to denote any (large) constant that can be bounded by the constants appearing in those inequalities.

In the proof of Proposition 8.1 we will use a bootstrap argument. The starting point is

Proposition 8.3. *The estimates (8.1) and (8.2) are true, with C replaced by C_0, when $T > 0$ is sufficiently small.*

Proof. Noting that $u^*(0) = u(0)$ and the same holds for v^* and w^*, and also $w(0) = \mathbb{P}_+ v(0)$, by invoking Proposition 3.6, we only need to prove that $\|u(0)\|_{Z_1} \leq C_0 A$ and $\|v(0)\|_{Z_1} \leq C_0 e^{C_0 A}$. The first inequality follows from our assumption, so we only need to prove that $\|Mu\|_{Z_1} \lesssim C_0 e^{C_0\|u\|_{Z_1}}$. By the definition of M, we only need to prove that

$$\|P^\mu u\|_{Z_1} \leq C_0^\mu \|u\|_{Z_1}^{\mu+1} \tag{8.3}$$

for all μ. Now we clearly have

$$|(P^\mu u)_{n_0}| \lesssim \sum_{n_1} |u_{n_1}| \cdot |z_{n_0 - n_1}|, \tag{8.4}$$

where

$$z_m = \sum_{m_1 + \cdots + m_\mu = m} \prod_{i=1}^{\mu} \frac{|u_{m_i}|}{\langle m_i \rangle}. \tag{8.5}$$

Since when $m = m_{1\mu}$ we have $\langle m \rangle \leq C_0^\mu \langle m_1 \rangle \cdots \langle m_\mu \rangle$, we conclude that

$$\sum_m \langle m \rangle^{1/4} |z_m| \lesssim C_0^\mu \prod_{i=1}^{\mu} \sum_{m_i} \frac{|u_{m_i}|}{\langle m_i \rangle^{3/4}} \lesssim (C_0 \|u\|_{Z_1})^\mu, \tag{8.6}$$

where the last inequality follows from

$$\begin{aligned}\sum_m \frac{|u_m|}{\langle m \rangle^{3/4}} &\lesssim \sum_d 2^{-3d/4} \sum_{m \sim 2^d} |u_m| \lesssim \sum_d 2^{(-3/4+1-1/p-r)d} \|\langle m \rangle^r u_m\|_{l^p_{m\sim 2^d}} \\ &\lesssim \sum_d 2^{-d/4} \|u\|_{Z_1} \lesssim \|u\|_{Z_1}.\end{aligned}$$

Now using (8.6), we will be able to prove (8.3) once we can prove

$$\|(u_{n+m})_{n\in\mathbb{Z}}\|_{Z_1} \lesssim \langle m\rangle^{1/4}\|u\|_{Z_1}. \tag{8.7}$$

To prove this, by definition we need to control $\|\langle n\rangle^r u_{n+m}\|_{l^p_{n\sim 2^d}}$ for each d. If $m \ll 2^d$ this is easy, since we then have $n+m \sim 2^d$ and also $\langle n\rangle^r \lesssim \langle m\rangle^{1/4}\langle n+m\rangle^r$. Now if $m \sim 2^{d'} \gtrsim 2^d$, we can use $\langle n\rangle^r \lesssim \langle m\rangle^{1/8}\langle n+m\rangle^r$ and

$$\|\langle n+m\rangle^r u_{n+m}\|_{l^p_{n\sim 2^d}} \lesssim \|\langle n\rangle^r u_n\|_{l^p_{n\lesssim 2^{d'}}} \lesssim (d'+1)\|u\|_{Z_1} \lesssim \langle m\rangle^{1/8}\|u\|_{Z_1} \tag{8.8}$$

to complete the proof. □

Starting from Proposition 8.3 and with the help of Proposition 3.6, it is easily seen that we only need to prove

Proposition 8.4. *Suppose C_j is large enough depending on C_{j-1} for $1 \leq j \leq 2$, and $0 < T \leq C_2^{-1}e^{-C_2 A}$. If the inequalities*

$$\|w^*\|_{Y_1^T} + \|v^*\|_{Y_2^T} + \|u^*\|_{Y_2^T} \leq C_1 e^{C_1 A}, \tag{8.9}$$

$$\|\langle\partial_x\rangle^{-s^3} u\|_{(X_2\cap X_3\cap X_4)^T} \leq C_1 A \tag{8.10}$$

hold, then they must hold with C_1 replaced by C_0.

The rest of this section, as well as Sections 9 and 10, is devoted to the proof of the estimate for w^* in Proposition 8.4; in Section 11 we consider the other three functions. During the whole proof, the inequalities (8.9) and (8.10) will be assumed.

8.2. *The extensions*

By the definition of Y_j^T norms, we have globally defined functions u'', v'', w'' and u''' which agree with u^*, v^*, w^* and u respectively on $[-T, T]$, and satisfy the inequalities (8.9) and (8.10) with the superscript T in the norms removed. By inserting a time cutoff $\chi(t)$, we may assume that they are all supported in $|t| \leq 1$. We then define the factors δ_n and Δ_n for all time as in (7.27) and (7.28), with w^* and u replaced by w'' and u''' respectively. We may also define functions u' by $(u')_n = e^{\mathrm{i}\Delta_n}(u'')_n$; the functions v' and w' are defined similarly.

Now we could interpret the bilinear form $\mathcal{N}^2$ and terms $\mathcal{N}^j$ on the right hand side of (7.29), by replacing each u with u''', each w with w', each υ with υ' (note υ is either u or v), each δ_n and Δ_n with what we defined above. If we then choose some $0 < \mathcal{T} \leq T$ and define the function z by $z(t) = w''(t)$ for $t \in [-\mathcal{T}, \mathcal{T}]$ and $(\partial_t - \mathrm{i}\partial_{xx})z(t) = 0$ on both $(-\infty, -\mathcal{T}]$ and $[\mathcal{T}, +\infty)$, then we can check that this function z satisfies the equation

$$(\partial_t - \mathrm{i}\partial_{xx})z = \mathbf{1}_{[-\mathcal{T},\mathcal{T}]}(t)\mathcal{N}^2(z,z) + \mathbf{1}_{[-\mathcal{T},\mathcal{T}]}(t)\sum_{j\in\{3,3.5,4,4.5\}}\mathcal{N}^j, \tag{8.11}$$

with initial data $z(0) = w(0)$. Using the time cutoff $\chi(t)$, we can define $y(t) = \chi(t)z(t)$. From (8.11) we conclude that

$$y = \chi(t)e^{\mathrm{i}t\partial_{xx}}w(0) + \mathcal{E}(\mathbf{1}_{[-\mathcal{T},\mathcal{T}]}\cdot\mathcal{N}^2(y,y)) + \sum_{j\in\{3,3.5,4,4.5\}}\mathcal{E}(\mathbf{1}_{[-\mathcal{T},\mathcal{T}]}\cdot\mathcal{N}^j). \quad (8.12)$$

Since w'' is smooth on $[-T,T]$, we conclude that $\mathcal{T}\mapsto y$ is a continuous map from $(0,T]$ to Y_1; also it is clear that when $\mathcal{T}$ is sufficiently small we have $\|y\|_{Y_1}\le C_0e^{C_0A}$. Thus in order to prove the estimate for w^*, we only need to prove

Proposition 8.5. *Suppose $y\in Y_1$ is a function satisfying (8.12) with $0<\mathcal{T}\le C_2^{-1}e^{-C_2A}$, and $\|y\|_{Y_1}\le C_1e^{C_1A}$. Then $\|y\|_{Y_1}\le C_0e^{C_0A}$.*

In what follows, we will use T instead of $\mathcal{T}$ for simplicity; note that $T\le C_2^{-1}e^{-C_2A}$. Before proceeding, let us prove a few results concerning the exponential factors $e^{\pm\mathrm{i}\Delta_n(t)}$. The first lemma is a general feature.

Lemma 8.6. *Suppose $h_j = h_j(t)$, $j\in\{0,1\}$, are two functions of t, and define $J_j(t) = \chi(t)e^{\mathrm{i}H_j(t)}$, where $H_j(t) = \int_0^t h_j(t')\,dt'$. Then*

$$\|\langle\xi\rangle(J_1-J_0)^\wedge(\xi)\|_{L^k}\lesssim\|(h_1-h_0)^\wedge\|_{L^1}(1+\|\widehat{h_1}\|_{L^1}+\|\widehat{h_0}\|_{L^1})^2 \quad (8.13)$$

for all $1\le k\le\infty$.

Proof. Recall from Section 3 that $\chi=\chi(t)$ is some time cutoff that may vary from place to place. Thanks to this factor, we only need to prove (8.13) for $k=1$. Next, noticing that

$$J_1-J_0 = \mathrm{i}\chi\cdot(H_1-H_0)\int_0^1 e^{\mathrm{i}(\theta H_1+(1-\theta)H_0)}\,d\theta, \quad (8.14)$$

we only need to prove (8.13) for a fixed θ. Let $h = h_1-h_2$ and $h_\theta=\theta h_1+(1-\theta)h_0$, let H, H_θ be defined accordingly, and define $\chi\cdot He^{\mathrm{i}H_\theta}=\Phi$. Then

$$\partial_x\Phi = (\chi'\cdot H+\chi\cdot h+\mathrm{i}\chi\cdot Hh_\theta)e^{\mathrm{i}H_\theta}, \quad (8.15)$$

which implies

$$\begin{aligned}\|\langle\xi\rangle\widehat{\Phi}(\xi)\|_{L^1}&\lesssim\|\widehat{\Phi}\|_{L^1}+\|\widehat{\partial_x\Phi}\|_{L^1}\\&\lesssim\|(\chi\cdot e^{\mathrm{i}H_\theta})^\wedge\|_{L^1}\cdot(\|\widehat{\chi h}\|_{L^1}+\|\widehat{\chi H}\|_{L^1}+\|\widehat{\chi H}\|_{L^1}\|\widehat{\chi h_\theta}\|_{L^1})\\&\lesssim\|\chi\cdot e^{\mathrm{i}H_\theta}\|_{H^1}\cdot(\|\widehat{h}\|_{L^1}+\|\chi H\|_{H^1}+\|\chi H\|_{H^1}(\|\widehat{h_1}\|_{L^1}+\|\widehat{h_0}\|_{L^1}))\\&\lesssim\|\widehat{h}\|_{L^1}(1+\|\widehat{h_0}\|_{L^1}+\|\widehat{h_1}\|_{L^1})^2,\end{aligned}$$

where H^1 is the standard Sobolev norm. □

Proposition 8.7. *We have*

$$\|\widehat{\delta_n}\|_{L^1}\le C_0C_1e^{C_0C_1A}\log(2+|n|), \quad (8.16)$$

$$\|(\delta_{n+1}-\delta_n)^\wedge\|_{L^1}\le C_0C_1e^{C_0C_1A}\langle n\rangle^{-1/2}. \quad (8.17)$$

Proof. Recall from Proposition 7.3 that

$$\delta_n = C(n) \cdot \sum_{k=0}^{n} |(w'')_k|^2, \tag{8.18}$$

where clearly $|C(n)| \lesssim 1$ and $|C(n+1) - C(n)| \lesssim \langle n\rangle^{-1}$. Now, using the fact that $\|\widehat{fg}\|_{L^1} \leq \|\widehat{f}\|_{L^1}\|\widehat{g}\|_{L^1}$, we obtain

$$\begin{aligned}\|\widehat{\delta_n}\|_{L^1} &\lesssim \sum_{k\lesssim\langle n\rangle} \|\widehat{(w'')_k}\|_{L^1}\|\widehat{(\overline{w''})_{-k}}\|_{L^1} \\ &\lesssim \|w''\|_{l^2_{k\lesssim\langle n\rangle}L^1}\|\overline{w''}\|_{l^2_{k\lesssim\langle n\rangle}L^1} \lesssim C_0C_1e^{C_0C_1A}\log(2+|n|).\end{aligned} \tag{8.19}$$

Here we have used the fact that

$$\|w''\|_{l^2_{k\lesssim\langle n\rangle}L^1} \lesssim \log(2+|n|)\cdot\|w''\|_{l^\infty_{d\geq 0}l^2_{k\sim 2^d}L^1} \lesssim \log(2+|n|)\|w''\|_{X_2}, \tag{8.20}$$

and the same estimate for $\overline{w''}$.

The estimate for the difference is proved in the same way, by using the inequality $|C(n+1)-C(n)| \lesssim \langle n\rangle^{-1}$. In fact, we get a power $\langle n\rangle^{-1}\log(2+|n|)$, which is better than $\langle n\rangle^{-1/2}$. □

Remark 8.8. Note that all our norms are invariant under complex conjugation. Occasionally we will make restrictions such as $n_l > 0$ which breaks this symmetry, but such information is only used in controlling the weights and the nonresonance factors, thus in terms of norm estimates for a single function, we will basically view w and $\overline{w}$ as the same function.

Proposition 8.9. *For any function h, let h' be defined by $(h')_n = \chi(t)e^{\pm \mathrm{i}\Delta_n}h_n$ for each fixed time. Then*

$$\|\langle\partial_x\rangle^{-s^3}h'\|_{X_j} \leq O_{C_1}(1)e^{C_0C_1A}\|h\|_{X_j} \tag{8.21}$$

for $1 \leq j \leq 7$.

Proof. Apart from X_3, all the other norms we are considering are (some Besov versions of) $\|\langle n\rangle^\sigma\langle\xi\rangle^\beta u\|_{l^kL^h}$ or $\|\langle n\rangle^\sigma\langle\xi\rangle^\beta u\|_{L^hl^k}$ with $\beta < 1$, and in the latter case we have $\sigma = \beta = 0$. Since the map $h \mapsto h'$ commutes with $\mathbb{P}$ projections, we only need to consider these kinds of norms. Notice that on the $\widetilde{u}$ side, this map is just a convolution with the Fourier transform of $\chi(t)e^{\pm\mathrm{i}\Delta_n(t)}$ for each $\widetilde{u}_n$. Thus to prove the result for the l^kL^h norm, we only need to prove that convolution with this function is bounded with respect to the weighted norm $\|\langle\xi\rangle^\beta\cdot\|_{L^h}$ by $O_{C_1}(1)e^{C_0C_1A}\langle n\rangle^{s^3}$. An elementary argument show that this bound does not exceed the norm

$$\|\langle\xi\rangle(\chi(t)e^{\pm\mathrm{i}\Delta_n(t)})^\wedge(\xi)\|_{L^1},$$

which is bounded by $C_0C_1^3e^{C_0C_1A}(\log(2+|n|))^3$, thanks to Lemma 8.6 and Proposition 8.7.

Now let us consider the $L^h l^k$ norm and the X_3 norm. Let $I_n(\xi)$ be the Fourier transform of $\chi(t)e^{\pm i\Delta_n(t)}$. Then we conclude that

$$\|\langle n\rangle^{-s^3}(h')_{n,\xi}\|_{L^h l^k} \lesssim \int_{\mathbb{R}} \|\langle n\rangle^{-s^3} h_{n,\xi-\eta} I_n(\eta)\|_{L^h l^k}\, d\eta$$
$$\lesssim \int_{\mathbb{R}} \sup_n \langle n\rangle^{-s^3}|I_n(\eta)| \cdot \|h\|_{L^h l^k}\, d\eta; \tag{8.22}$$

note the same argument also works for X_3. Therefore we need to bound the expression

$$\int_{\mathbb{R}} \sup_n \langle n\rangle^{-s^3}|I_n(\xi)|\, d\xi$$

by $O_{C_1}(1)e^{C_0C_1A}$. By performing a dyadic summation in n, we only need to bound

$$\int_{\mathbb{R}} \max_{n\sim 2^d} |I_n(\xi)|\, d\xi \tag{8.23}$$

by $O_{C_1}(1)e^{C_0C_1A}(d+2)^{O(1)}$. Now suppose $|\xi| \lesssim 2^{10d}$. Then we simply use Proposition 8.7 as well as the L^∞ estimate of Lemma 8.6 to bound this contribution by $O_{C_1}(1)e^{C_0C_1A}$ times $(d+2)^{O(1)} \int_{|\xi|\lesssim 2^{20d}} \langle\xi\rangle^{-1}\, d\xi = (d+2)^{O(1)}$. If $|\xi| \gg 2^{10d}$, we may replace the "maximum" in this expression by summation (during which we lose a power 2^d), then use the L^1 estimate of Lemma 8.6 and the largeness of ξ to gain a power 2^{10d}. Thus in any case we obtain the desired estimate. ◻

9. The a priori estimate II: Quadratic and cubic terms

We now begin the proof of Proposition 8.5, the starting point being (8.12). The linear term is clearly bounded in Y_1 by $C_0e^{C_0A}$, so we only need to bound the $\mathcal{N}^j$ terms. There will be a large number of cases, and they are ordered according to the difficulty level. In this section we will be able to treat every term except $\mathcal{N}^{3.5}$.

Proposition 9.1. *For each $j \in \{2, 3, 3.5, 4, 4.5\}$, define*

$$\mathcal{M}^j = \mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^j), \tag{9.1}$$

where we may write $\mathcal{N}^2$ or $\mathcal{N}^2(y, y)$ depending on the context. Then

$$\|\mathcal{M}^2\|_{X_4} \le O_{C_1}(1)e^{C_0C_1A}T^{0+}, \tag{9.2}$$
$$\sum_{j\in\{3,3.5,4,4.5\}} \|\langle n\rangle^{-1/20}\langle\xi\rangle^{\kappa}(\mathcal{M}^j)_{n,\xi}\|_{l^2L^2} \le O_{C_1}(1)e^{C_0C_1A}T^{0+}. \tag{9.3}$$

Remark 9.2. Since

$$\|\langle n\rangle^{-1}\langle\xi\rangle^{\kappa}u\|_{l^\gamma L^2} \le C_0\|\langle n\rangle^{-1/20}\langle\xi\rangle^{\kappa}u\|_{l^2L^2} \tag{9.4}$$

by Hölder, the inequalities (9.2) and (9.3) will imply $\|y\|_{X_4} \le C_0e^{C_0A}$, due to the restriction $T \le C_2^{-1}e^{-C_2A}$.

Proof of Proposition 9.1. In this proof, as well as the following ones, we will use the $\lesssim$ and $\gtrsim$ symbols, with the convention that all the implicit constants are $\leq O_{C_1}(1)e^{C_0C_1A}$. Note that in the estimate for any possible multilinear term, the total number of appearances of all functions other than u''' is bounded by 10, thus as long as we only use the norm $\|\langle\partial_x\rangle^{-s^3}u'''\|_{X_2\cap X_3\cap X_4}$ (which is bounded by C_1A) for the function u''', the implicit constants will be bounded by

$$(O_{C_1}(1)e^{C_0C_1A})^{C_0}\sum_{\mu}\frac{C_0^{\mu}}{\mu!}(C_0C_1A)^{\mu}\leq O_{C_1}(1)e^{C_0C_1A} \tag{9.5}$$

and are thus under control. We also need to be careful with the *sharp* cutoff $\mathbf{1}_{[-T,T]}$. Denote by $\phi_\xi=(e^{\mathrm{i}T\xi}-e^{-\mathrm{i}T\xi})/(\mathrm{i}\xi)$ the Fourier transform of $\mathbf{1}_{[-T,T]}$; note that $|\phi_\xi|\lesssim\min(T,1/(\langle\xi\rangle))$, and $\|\phi\|_{L^{1+}(\{|\xi|\geq K\})}\lesssim T^{0+}\langle K\rangle^{0-}$.

First let us prove $\|\mathcal{M}^2\|_{X_4}\lesssim T^{0+}$. As above, we may fix $\mu\geq 0$ and $\omega\in\{-1,1\}^2$ (though we will not add any sub- or superscript for simplicity). Choose a function g such that $\|g\|_{X_4'}\leq 1$ and define $f=\mathcal{E}'g$. Also define f' by $(f')_n=e^{\mathrm{i}\Delta_n}f_n$ and y' similarly; these notations will be standard throughout the proof. Since f has compact time support, we may insert $\chi(t)$ in the definition of f', so that we can use the arguments in the proof of Proposition 8.9. The same comment applies for later discussions.

From the bound $\|g\|_{X_4'}\leq 1$ we see by Proposition 3.4 that $\|\langle n_0\rangle\langle\alpha_0\rangle^{1-\kappa}f_{n_0,\alpha_0}\|_{l^{\gamma'}L^2}\lesssim 1$, which then implies, thanks to (Hölder and) an argument similar to the proof of Proposition 8.9, that

$$\|\langle n_0\rangle^{1-O(s^{2.5})}\langle\alpha_0\rangle^{1-\kappa}(f')_{n_0,\alpha_0}\|_{l^2L^2}\lesssim 1. \tag{9.6}$$

Using Plancherel, we now only need to bound the expression

$$\begin{aligned}
\mathcal{S}&=\sum_{n_0}\int_{\mathbb{R}}\overline{f_{n_0,\alpha_0}}\cdot(\mathbf{1}_{[-T,T]}\mathcal{N}^2)_{n_0,\alpha_0}\,d\alpha_0\\
&=\sum_{n_0=n_1+n_2+m_1+\cdots+m_\mu}\int_{\mathbb{R}}\Phi^2\cdot\overline{f_{n_0,\alpha_0}}\\
&\qquad\times\bigg(\mathbf{1}_{[-T,T]}e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})}\prod_{l=1}^{2}(y^{\omega_l})_{n_l}\prod_{i=1}^{\mu}\frac{(u''')_{m_i}}{m_i}\bigg)^{\wedge}(\alpha_0-|n_0|n_0)\,d\alpha_0\\
&=\sum_{n_0=n_1+n_2+m_1+\cdots+m_\mu}\int_{\mathbb{R}}\Phi^2\cdot\overline{(f')_{n_0,\alpha_0}}\\
&\qquad\qquad\times\bigg(\mathbf{1}_{[-T,T]}\prod_{l=1}^{2}((y')^{\omega_l})_{n_l}\prod_{i=1}^{\mu}\frac{(u''')_{m_i}}{m_i}\bigg)^{\wedge}(\alpha_0-|n_0|n_0)\,d\alpha_0\\
&=\sum_{n_0=n_1+n_2+\cdots+m_\mu}\int_{(T)}\Phi^2\cdot\overline{(f')_{n_0,\alpha_0}}\prod_{l=1}^{2}((y')^{\omega_l})_{n_l,\alpha_l}\cdot\phi_{\alpha_3}\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}.
\end{aligned}$$

Here (T) indicates integration over the set

$$\{(\alpha_0,\ldots,\alpha_3,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_{13}+\beta_{1\mu}+\Xi\},$$

which is a hyperplane in $\mathbb{R}^{\mu+4}$ (recall the notation $\alpha_{13} = \alpha_1 + \alpha_2 + \alpha_3$), with respect to the standard measure

$$\prod_{l=1}^{3} d\alpha_l \cdot \prod_{i=1}^{\mu} d\beta_i,$$

with the non-resonance (NR) factor

$$\Xi = |n_0|n_0 - \sum_{l=1}^{2} |n_l|n_l - \sum_{i=1}^{\mu} |m_i|m_i. \tag{9.7}$$

Note that we are using the convention that $u_{n,\alpha}$ stands for $\widetilde{u}_{n,\alpha}$; also we may always restrict to $n_0 > 0$.

Noticing that the m variables are all $\ll \min_{0\leq l\leq 2}\langle n_l\rangle$ (again here we may have harmless polynomial factors in μ), we can check from (9.7) that

$$|\Xi| \sim \min_{0\leq l\leq 2}\langle n_l\rangle \cdot \max_{0\leq l\leq 2}\langle n_l\rangle. \tag{9.8}$$

We will first take the summation-integration over the set where $\sum_{l=0}^{2}\langle n_l\rangle \sim 2^d$, and then sum over d. In this case, at least one of the α and β variables must be $\gtrsim 2^d$. Now, with a loss of $2^{O(s^{2.5})d}$, we can replace the $1 - O(s^{2.5})$ exponent in (9.6) by 1. Also noticing that $|\Phi^2| \lesssim \langle n_0\rangle$, we may further (upon taking absolute values) remove this Φ factor and the $\langle n\rangle$ factor in (9.6) simultaneously.

With these reductions, we then proceed to the estimate of $\mathcal{S}$. First assume $\langle\alpha_0\rangle \gtrsim 2^d$; thus we gain from the bound (9.6) a power $2^{(1-\kappa)d}$, while after exploiting this, we still have (for the function f'^* obtained from extracting from f' the $\langle\alpha_0\rangle$ factor) $\|f'^*\|_{l^2L^2} \lesssim 1$. In the same way, we can use the X_1 and X_4 bounds for y to deduce some bound for y' (see Proposition 8.9), and strengthen the bound to $\|\langle n_l\rangle^{s^2}\langle\alpha_l\rangle^{1/2+s^2}(y')_{n_l,\alpha_l}\|_{l^2L^2} \lesssim 1$ at a price of at most $2^{O(1/2-b)d}$.

We then fix all the m and β variables to get a sub-summation-integration that is bounded by (with C being irrelevant constants)

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} &\lesssim \sum_{n_0=n_1+n_2+C}\int_{\widetilde{\alpha_0}=\widetilde{\alpha_1}+\widetilde{\alpha_2}+\widetilde{\alpha_3}+C} |(f')_{n_0,\widetilde{\alpha_0}}|\,|\phi_{0,\widetilde{\alpha_3}}| \cdot \prod_{l=1}^{2}|((y')^{\omega_l})_{n_l,\widetilde{\alpha_l}}| \\ &\lesssim \big\| |\widehat{f'}| * |\widehat{(y')^{\omega_1}}| * |\widehat{(y')^{\omega_2}}| * |\widehat{\phi}| \big\|_{l^\infty L^\infty} \\ &\lesssim \|f'\|_{l^2L^2}\|\mathfrak{N}(y')^{\omega_1}\|_{L^{6+}L^{6+}}\|\mathfrak{N}(y')^{\omega_2}\|_{L^3L^3}\|\widehat{\phi}\|_{l^{1+}L^{1+}},\end{aligned} \tag{9.9}$$

where $\widetilde{\alpha_l} = \alpha_l - |n_l|n_l$, and ϕ is viewed as a function of (t,x) that is supported at $n=0$ (so that $\widetilde{\alpha_3} = \alpha_3$); also recall the $\mathfrak{N}$ notation defined in Section 2.1. The right hand side will be bounded by T^{0+} by our (reduced) assumptions and Strichartz estimates, provided we choose $6+$ to be $6 + cs^2$ with some small c, and choose $1+$ accordingly.

Now we sum over m_i and integrate β_i, exploiting the bound $\|\langle m_i\rangle^{-1}u'''\|_{l^1L^1} \leq C_1A$, to bound the whole summation-integration for a single d; taking into account the gains and losses from the reductions made before and exploiting (1.3), we conclude that the part of $\mathcal{S}$ considered above is bounded by $T^{0+}2^{(0-)d}$, which allows us to sum over d.

Next, assume that $\langle\alpha_1\rangle \gtrsim 2^d$ (the case for α_2 will follow by symmetry). In this case we do not gain from the bound (9.6), so that we still have $\|\langle\xi\rangle^{1-\kappa} f'\|_{l^2L^2} \lesssim 1$, which, via Strichartz, allows us to control $\|\mathfrak{N}f'\|_{L^{2+}L^{2+}}$, where $2+$ is some $2+c(1-\kappa)$. Instead, we gain from the bound

$$\|\langle n_1\rangle^{s^2}\langle\alpha_1\rangle^{1/2+s^2}(y')_{n_1,\alpha_1}\|_{l^2L^2} \lesssim 1$$

as above (with a loss of $2^{O(1/2-b)d}$) and change the exponent $\langle\alpha_1\rangle^{1/2+s^2}$ to $\langle\alpha_1\rangle^{1/2-c(1-\kappa)}$ to gain the power $2^{c(1-\kappa)d}$, and the reduced bound will allow us to control $\mathfrak{N}y'$ (in the form of $\mathfrak{N}(y')^{\omega_1}$) in $L^{6-}L^{6-}$ with $6-$ here being $6-c(1-\kappa)$. Choosing the constants c appropriately, we can then proceed as in (9.9), with the f' factor estimated in $L^{2+}L^{2+}$, two y' factors estimated in $L^{6-}L^{6-}$ and L^3L^3 respectively and the ϕ factor in $l^{1+}L^{1+}$, to get the desired bound. In the same spirit, if $\langle\alpha_3\rangle \gtrsim 2^d$, we will use the $L^{2+}L^{2+}$ bound for $\mathfrak{N}f'$ (with $2+$ being $2+c(1-\kappa)$), $L^{6+}L^{6+}$ and L^3L^3 bound for $\mathfrak{N}y'$ (with $6+$ being $6+cs^2$) and $l^{1+}L^{1+}$ bound for ϕ (with $1+$ being $1+c(1-\kappa)$; note that we gain a power $2^{c(1-\kappa)d}$ here due to the largeness of α_3) to conclude. Again we gain at least $2^{c(1-\kappa)d}$ and lose at most $2^{O(1/2-b)d}$ so we have enough room for summation in d.

Next, assume that $\langle\beta_i\rangle \gtrsim 2^d$ for some i. If for this i we also have $\langle m_i\rangle \gtrsim 2^{d/30}$, then we would bound $|m_i|^{-1} \lesssim 2^{-d/90}\langle m_i\rangle^{-2/3}$ to gain a power of 2^{cd} and proceed as above, since we still have

$$\|\langle m_i\rangle^{-2/3}(u''')_{m_i,\beta_i}\|_{l^1L^1} \lesssim \|\langle\partial_x\rangle^{-s^3}u'''\|_{X_2} \leq C_0C_1A, \tag{9.10}$$

which allows us to sum over m_i and integrate over β_i. If instead $\langle m_i\rangle \lesssim 2^{d/30}$, we could use the X_4 bound of $\langle\partial_x\rangle^{-s^3}u'''$ and Proposition 8.9 to bound

$$\|\langle m_i\rangle^{-3/2}\langle\beta_i\rangle^{9/10}(y')_{m_i,\beta_i}\|_{l^2L^2} \lesssim 1,$$

and exploit the largeness of β_i to gain a power $2^{d/20}$ and reduce the above bound to

$$\|\langle m_i\rangle^{2}\langle\beta_i\rangle^{3/5}(y')_{m_i,\beta_i}\|_{l^2L^2} \lesssim 1,$$

which would imply $\|y'\|_{l^1L^1} \lesssim 1$ so that we can still apply the argument above, sum over m_i and integrate over β_i. This concludes the proof of (9.2).

Now let us prove (9.3). Let g, f and f' be as before, but with the new bound

$$\|\langle n_0\rangle^{1/30}\langle\alpha_0\rangle^{1-\kappa}f'\|_{l^2L^2} \lesssim 1.$$

Note that the estimate for f' is again easily deduced from the estimate for g and the same type of argument as in the proof of Propositions 3.3 and 8.9. To bound $\mathcal{M}^3$ and $\mathcal{M}^{3.5}$, we need to bound

$$\mathcal{S} = \sum_{n_0=n_1+n_2+n_3+\cdots+n_\mu}\int_{(T)}\Phi^j\cdot\overline{(f')_{n_0,\alpha_0}} \times ((w')^{\omega_1})_{n_1,\alpha_1}\prod_{l=2}^{3}(z^l)_{n_l,\alpha_l}\cdot\phi_{\alpha_4}\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}, \tag{9.11}$$

with (T) indicating integration over the set

$$\{(\alpha_0,\ldots,\alpha_4,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_{14}+\beta_{1\mu}+\Xi\},\tag{9.12}$$

with respect to the standard measure, with the NR factor

$$\Xi=|n_0|n_0-\sum_{l=1}^{3}|n_l|n_l-\sum_{i=1}^{\mu}|m_i|m_i.\tag{9.13}$$

Here z^l or $\overline{z^l}$ equals u', v' or w' for each l. Again we assume $\sum_{l=0}^{3}\langle n_l\rangle\sim 2^d$. By losing at most $2^{O(\epsilon)d}$, we may assume that w' satisfies the same bound as y' before, and $\mathfrak{N}z^l$ is bounded in X_4 and L^6L^6. Also note that $|\Phi^j|\lesssim 1$ in any situation.

If $\langle n_0\rangle+\langle\alpha_0\rangle\gtrsim 2^{d/90}$, we may gain a power $2^{c(1-\kappa)d}$ (note our loss is at most $2^{O(\epsilon)d}$) from the bound of f', and reduce this bound to $\|f'\|_{l^2L^2}\lesssim 1$. Then we can first fix the m_i and β_i variables and obtain the $\mathcal{S}_{\mathrm{sub}}$, estimate in the same was as in (9.9), then sum over m_i and integrate over β_i. The only difference with (9.9) is that now $\mathcal{S}_{\mathrm{sub}}$ contains five functions instead of four; however, here we may estimate the f' factor in L^2L^2, the $\mathfrak{N}w'$ factor in $L^{6+}L^{6+}$ with 6+ being $6+cs^2$, the $\mathfrak{N}z^l$ factors in L^6L^6 and the ϕ factor in $l^{1+}L^{1+}$ so that we can still close the argument.

If $\langle\alpha_1\rangle\gtrsim 2^{d/90}$, we may perform the same reduction as in the estimate of $\|\mathcal{M}^2\|_{X_4}$ before, gain a power of $2^{c(1-\kappa)d}$ and use Strichartz and the reduced bound to control $\|\mathfrak{N}w'\|_{L^{6-}L^{6-}}$, where 6− is $6-c(1-\kappa)$. We may now control $\mathfrak{N}f'$ in $L^{2+}L^{2+}$ with the 2+ being $2+c(1-\kappa)$, then control $\mathfrak{N}z^l$ in L^6L^6 and ϕ in some $l^{1+}L^{1+}$. The exponents will match if we choose the constants c appropriately.

If $\langle\alpha_2\rangle\gtrsim 2^{d/4}$ (the α_3 case being identical), we have two possibilities. If $j=3$ then z^2 is also taken from $\{w',\overline{w'}\}$ so that we are in the same situation as above. If $j=3.5$ then either $\langle n_2\rangle\gtrsim 2^{d/89}$ and we gain a power 2^{cd} from the Φ factor thanks to (7.8) and the assumption that $\langle n_0\rangle\ll 2^{d/90}$, or $\langle n_2\rangle\lesssim 2^{d/89}$ and we can exploit the X_4 bound of z^l, gain a power 2^{cd}, and use the reduced estimate to bound $\|\mathfrak{N}z^2\|_{L^6L^6}$ (again, as we already did in the X_4 estimate before). In any case we gain a power $2^{c(1-\kappa)d}$, lose at most $2^{O(\epsilon)d}$, and can control the reduced $\mathcal{S}_{\mathrm{sub}}$ expression.

If $\langle\alpha_4\rangle\gtrsim 2^{d/90}$, we can again control $\mathfrak{N}f'$ in $L^{2+}L^{2+}$ with the 2+ being $2+c(1-\kappa)$, then control the $\mathfrak{N}w'$ in $L^{6+}L^{6+}$ (with 6+ being $6+cs^2$), $\mathfrak{N}z^l$ factors in L^6L^6 and ϕ in $l^{1+}L^{1+}$ with 1+ being $1+c(1-\kappa)$, with c chosen appropriately. Note that since α_4 is large, we will gain a power $2^{c(1-\kappa)d}$ from the $l^{1+}L^{1+}$ bound of ϕ. Moreover, if $\langle m_i\rangle\gtrsim 2^{d/90}$ for some i, we can repeat the argument made before to gain a (small) 2^{cd} power from this factor alone while keeping the ability to sum over m_i and integrate over β_i, and reduce to the above cases. Similarly, if $\langle\beta_i\rangle\gtrsim 2^{d/10}$, we can also gain this 2^{cd} power by using the bound for $\|\langle\partial_x\rangle^{-s^3}u'''\|_{X_4}$.

Finally, if none of the above holds, we must have $\langle n_0\rangle\ll 2^{d/90}$ and $|\Xi|\ll 2^{d/4}$. We may also assume $\langle m_i\rangle\ll 2^{d/90}$ or we are reduced to one of the cases above. Thus from (9.13) we deduce

$$\big||n_1|n_1+|n_2|n_2+|n_3|n_3\big|\ll 2^{d/4}.\tag{9.14}$$

Note that we may assume $j=3$, since when $j=3.5$, one of $\langle n_2\rangle$ and $\langle n_3\rangle$ must be $\gtrsim 2^d$ and we gain a power 2^{cd} from the weight Φ so that we can proceed as above. Now if the minimum of $\langle n_l\rangle$ for $1\leq l\leq 3$ is at least $\gtrsim 2^{d/9}$, then we will be in the same situation as in (9.7) and the expression in (9.14) has to be $\gtrsim 2^d$. Therefore we may further assume $\langle n_3\rangle \ll 2^{d/9}$, and it will be clear that the NR factor can be small only if $n_1+n_2=0$. However, in this case we gain from the factor Φ a positive power 2^{cd}, due to parts (b) and (c) in the requirements for $\mathcal{N}^3$ in Proposition 7.3. This allows us to complete the estimate in the same way as above.

Notice that in estimating $\mathcal{M}^3$ above, we have ignored the term where three of $(-n_0,n_1,n_2,n_3)$ are related by m and we are allowed to have υ instead of w (in the discussion here, they will be υ' and w' respectively). To handle this term, simply fix the m and β variables and bound the Φ factor by 1 (we may assume $\langle m_i\rangle \ll 2^{d/90}$ or we gain a power 2^{cd} and can proceed as above). We can bound the resulting $\mathcal{S}_{\text{sub}}$ (note that we are restricting to $n_l=c_l\pm n_0\sim 2^d$)

$$\mathcal{S}_{\text{sub}} \lesssim \sum_{n_0}\int_{\alpha_0=\alpha_1+\cdots+\alpha_4+c_4(n_0)} |(f')_{n_0,\alpha_0}|\cdot\prod_{l=1}^{3}|(z^l)_{c_l\pm n_0,\alpha_l}|\cdot|\phi_{\alpha_4}|$$
$$\lesssim T^{0+}\sum_{n_0}\|\widehat{(f')_{n_0}}\|_{L^2}\prod_{l=1}^{3}\|\widehat{(z^l)_{c_l\pm n_0}}\|_{L^1}\lesssim T^{0+}\|f'\|_{l^4_{\sim 2^d}L^2}\prod_{l=1}^{3}\|z^l\|_{l^4_{\sim 2^d}L^1}\lesssim T^{0+}2^{-cd},$$

where c_j are constants (or functions of n_0). Thus this term is also acceptable.

Now let us bound $\mathcal{M}^4$ and $\mathcal{M}^{4.5}$. The quantity we need to control is now

$$\mathcal{S}=\sum_{n_0=n_1+\cdots+n_4+\cdots+m_\mu}\int_{(T)}\Phi^j\cdot\overline{(f')_{n_0,\alpha_0}}\prod_{l=1}^{4}(z^l)_{n_l,\alpha_l}\cdot\phi_{\alpha_5}\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}, \tag{9.15}$$

with (T) indicating integration over the set

$$\{(\alpha_0,\ldots,\alpha_5,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_{15}+\beta_{1\mu}+\Xi\}, \tag{9.16}$$

with respect to the standard measure, with the NR factor

$$\Xi=|n_0|n_0-\sum_{l=1}^{4}|n_l|n_l-\sum_{i=1}^{\mu}|m_i|m_i. \tag{9.17}$$

Here z^l or $\overline{z^l}$ equals u', υ' or w' for each l. We assume the maximum of the n variables is $\sim 2^d$, and with a loss of at most $2^{O(\epsilon)d}$, we may assume that $\mathfrak{N}w'$ satisfies the same estimate as before, and $\mathfrak{N}\upsilon'$ is bounded in X_4 and L^6L^6 (again, it is the modified versions of w' and υ' that satisfy the estimates).

If $j=4$ we may assume (up to a permutation) that $|\Phi|\lesssim 2^{-d/90}\langle n_3\rangle^{-2/3}$. Due to the presence of the $\langle n_3\rangle^{-2/3}$ factor, we may also fix n_3 and α_3 when we fix the m_i and β_i variables. Once these variables are fixed, we only need to control the resulting $\mathcal{S}_{\text{sub}}$. But since we gain a power $2^{d/90}$ from the Φ factor, the estimate for $\mathcal{S}_{\text{sub}}$ will be easy; we simply bound $\mathfrak{N}f'$ in $L^{2+}L^{2+}$, bound $\mathfrak{N}z^l$ in L^6L^6 for $l\in\{1,2,4\}$, and bound ϕ in $l^{1+}L^{1+}$. This proves (9.3) for $j=4$.

If $j = 4.5$, then all the m_i variables, as well as n_3 and n_4, are $\ll 2^{d/10}$. This in particular implies that either $\langle n_0\rangle \gtrsim 2^{d/10}$ so that we gain from the $\langle n_0\rangle^{1/30}$ factor in the bound for f', or the NR factor has $|\Xi| \gtrsim 2^d$ (note that $n_1 + n_2 \neq 0$) so we can gain from one of the α or β variables. Note that $|\Phi| \lesssim (\langle n_3\rangle + \langle n_4\rangle)^{-1}$, thus when we fix the m_i and β_i variables, we always have the choice of also fixing (n_3, α_3) or (n_4, α_4). This means that though the $\mathcal{S}_{\text{sub}}$ expression seems to involve six factors, in practice we will always use only five of them. The rest will be basically the same as before. If we gain at least $2^{c(1-\kappa)d}$ from n_0, α_0, α_3 (or similarly α_4) or some β_i, then we will fix (and then sum and integrate over) (m_j, β_j) and (n_3, α_3) to produce $\mathcal{S}_{\text{sub}}$, then control $\mathfrak{N}f'$ in $L^{2+}L^{2+}$, $\mathfrak{N}z^l$ $(l \in \{1, 2, 4\})$ in L^6L^6, ϕ in $l^{1+}L^{1+}$ with $2+$ being $2 + c(1-\kappa)$ and $1+$ defined accordingly. If we gain from α_l for $l \in \{1, 2\}$ (say $l = 1$), we will again fix (m_j, β_j) and (n_3, α_3). To estimate $\mathcal{S}_{\text{sub}}$, we control $\mathfrak{N}f'$ in $L^{2+}L^{2+}$ with $2+$ being $2 + c(1-\kappa)$, $\mathfrak{N}z^2$ and $\mathfrak{N}z^4$ in L^6L^6, ϕ in $l^{1+}L^{1+}$, and $\mathfrak{N}z^1$ in $L^{6-}L^{6-}$ with $6-$ being $6 - c(1-\kappa)$. Here, if $\langle n_1\rangle \gtrsim 2^{d/10}$, z^1 will be either w' or $\overline{w'}$ so that we can get the $L^{6-}L^{6-}$ bound from the same arguments as before; otherwise $\langle n_1\rangle \lesssim 2^{d/10}$, and we can use the X_4 bound for $\mathfrak{N}z^1$ to deduce the $L^{6-}L^{6-}$ bound with a gain of 2^{cd}. This concludes the proof of Proposition 9.1. ◻

Proposition 9.3. *We have*

$$\sum_{j\in\{3,4,4.5\}}\sum_{j'\in\{1,2,5,7\}} \|\mathcal{M}^j\|_{X_{j'}} \lesssim T^{0+}. \tag{9.18}$$

Proof. Note that $\mathcal{M}_j$ involves a sum over the n_l and m_i variables. We shall first prove the bound for the terms where $j = 4$, or $j = 3$ and $n_0 \notin \{n_1, n_2, n_3\}$, or $j = 4.5$ and $n_0 \notin \{n_1, n_2\}$. By (3.8), we only need to bound this part of $\mathcal{M}^j$ in X_6.

Let the functions g and f, f' be as usual, with $\|g\|_{X_6'} \leq 1$. This would imply

$$\|\langle n_0\rangle^{-s-O(s^3)}\langle\alpha_0\rangle^{1/2-O(s^2)}(f')_{n_0,\alpha_0}\|_{l^2L^2} \lesssim 1$$

(we have done this kind of reduction many times before). What we need to control is the same quantity $\mathcal{S}$ with $j \in \{3, 4, 4.5\}$ as in (9.11) and (9.15), and we assume the maximal n_l variable is $\sim 2^d$ as usual. With a loss of $2^{O(\epsilon)d}$, we may assume that w' and v' satisfy the good bounds appearing in the proof of Proposition 9.1. Using Strichartz, we can deduce from the bound for f' as above the $L^2L^2 \cap L^4L^4$ bound for $\mathfrak{N}f'$ with a loss of $2^{O(s)d}$.

Now we will be able to bound the $\mathcal{S}$ expression in (9.15) easily. In fact, if we gain from anything except α_0, we can repeat the argument in the proof of (9.3), but with the $c(1-\kappa)$ involved in various $2+$ or $6-$ replaced by c (since we now have the $L^{2+c}L^{2+c}$ control for $\mathfrak{N}f'$), and check that in these cases we always gain a power 2^{cd}, which will be enough to cover the loss $2^{O(s)d}$. If we gain from α_0, this gain will be 2^{cd}, with a loss of at most $2^{O(s)d}$, and we can bound the reduced $\mathfrak{N}f'$ factor in $L^{2+c}L^{2+c}$, so this contribution will be acceptable. On the other hand, if we do not gain anything from any of the variables or weights, it must be the case that $j = 4.5$ and $|\Xi| \ll 2^{d/4}$. Since all the m variables as well as n_3 and n_4 are assumed to be $\ll 2^{d/10}$, we then deduce that

$$\big||n_0|n_0 - |n_1|n_1 - |n_2|n_2\big| \ll 2^{d/4}.$$

By repeating the argument in the proof of Proposition 9.1, we see that this can happen only if $n_1+n_2=0$, or $n_0=n_1$, or $n_0=n_2$; but all these possibilities contradict our assumptions.

Next, assume $j=3$. Recall that $\sum_{l=0}^{3}\langle n_l\rangle\sim 2^d$, and that the $\mathcal{S}$ we need to estimate is bounded by

$$|\mathcal{S}|\lesssim\sum_{n_0=n_1+n_2+n_3+m_1+\cdots+m_\mu}\int_{(T)}\overline{|(f')_{n_0,\alpha_0}|}\prod_{l=1}^{3}|((w')^{\omega_l})_{n_l,\alpha_l}|\cdot|\phi_{\alpha_4}|\prod_{i=1}^{\mu}\left|\frac{(u''')_{m_i,\beta_i}}{m_i}\right|.$$

First assume that some of the α or β variables is at least $2^{d/90}$. Then, by the same argument we made before (notice that the n_l variables for $1\leq l\leq 3$ correspond to the function w' or $\overline{w'}$, which, up to a loss of $2^{O(s)d}$, satisfies the estimate $\|\langle n_l\rangle^{s^2}\langle\alpha_l\rangle^{1/2+s^2}w'\|_{l^2L^2}\lesssim 1$), we can gain a power 2^{cd} from the corresponding factor, then fix the m_i and β_i variables (and sum and integrate over them afterwards), produce the $\mathcal{S}_{\mathrm{sub}}$ term, and estimate it by controlling $\mathfrak{N}f'$ in $L^{2+}L^{2+}$, $\mathcal{N}(w')^{\omega_l}$ in $L^{6-}L^{6-}$ with $2+$ and $6-$ being $2+c$ and $6-c$ respectively, and finally control ϕ in $l^{1+}L^{1+}$.

Now let us assume that all α and β variables are $\ll 2^{d/90}$; we may assume that all m_i variables are $\ll 2^{d/90}$ also. Thus, the variables $(-n_0,n_1,n_2,n_3)$ will satisfy the conditions in the following lemma.

Lemma 9.4. *Suppose four numbers $n_0,\ldots,n_3$ satisfy*

$$n_0+n_1+n_2+n_3=K_1,\qquad |n_0|n_0+|n_1|n_1+|n_2|n_2+|n_3|n_3=K_2,$$

where K_j are constants such that

$$|K_1|+|K_2|\ll 2^{d/40},\qquad \max_{0\leq l\leq 3}\langle n_l\rangle\sim 2^d.$$

Then one of the following must hold:

(i) *Up to some permutation, $n_0+n_1=n_2+n_3=0$. In particular, this can happen only if $K_1=K_2=0$.*

(ii) *Up to some permutation, $n_0+n_1=0$, $\langle n_0\rangle\sim 2^d$, and $\langle n_2\rangle+\langle n_3\rangle\ll 2^{d/40}$. Note that it is possible that (say) $n_1+n_2=0$ and n_0,n_3 are small.*

(iii) *No two of n_l add to zero.* Under this restriction *we must have $\langle n_l\rangle\gtrsim 2^{0.9d}$ for each l; moreover, if we fix K_1, K_2 and any single n_l, there will be at most $\lesssim 2^{s^3d}$ choices for the quadruple (n_0,n_1,n_2,n_3).*

Proof of Lemma 9.4. Suppose some $\langle n_l\rangle$ is $\ll 2^{0.9d}$ (say for $l=0$); then one of $\langle n_l\rangle$ for $1\leq l\leq 3$ must also be $\ll 2^{0.9d}$, since otherwise we would have

$$\big||n_1|n_1+|n_2|n_2+|n_3|n_3\big|\gtrsim\max_{1\leq l\leq 3}\langle n_l\rangle\cdot\min_{1\leq l\leq 3}\langle n_l\rangle\gtrsim 2^{1.9d}$$

while $|n_0|^2\lesssim 2^{1.8d}$, which is impossible. Now assume that $\langle n_1\rangle\ll 2^{0.9d}$; then in particular $\langle n_2+n_3\rangle\ll 2^d$, thus $n_2n_3<0$ as well as $|n_2-n_3|\sim 2^d$. Suppose $n_0+n_1=k$ and $n_2+n_3=l$; then $|k+l|\leq c2^{d/40}$ and

$$2^d|l|\leq 2^{0.9d}\langle k\rangle+2^{d/40}.$$

Now if $l \neq 0$, this inequality cannot hold, since it would require $|k| \gg |l|$, which implies $\langle k\rangle \lesssim 2^{d/40}$, so that the right hand side will be at most $2^{(0.9+1/40)d}$ and the left hand side is at least 2^d (note that here we may assume that 2^d is larger than some constant which is polynomial in μ, since the summation for small values of 2^d will be trivial). Therefore we must have $n_2 + n_3 = 0$. If also $n_0 + n_1 = 0$, we will be in case (i); otherwise $k \neq 0$, so that we always have $\big||n_0|n_0 + |n_1|n_1\big| \gtrsim |n_0| + |n_1|$, which then implies that $\langle n_0\rangle + \langle n_1\rangle \ll 2^{d/40}$ and we will be in case (ii).

Now assume that $\langle n_l\rangle \gtrsim 2^{0.9d}$ for each l. By the discussion above, we cannot have any $n_h + n_l = 0$ (unless we are in case (i)), so we will be in case (iii). Finally, suppose we fix K_1, K_2 and n_0. The requirement $n_h + n_l \neq 0$ implies that each $\langle n_l\rangle$ is $\gtrsim 2^{0.9d}$, so without loss of generality we may assume $n_0 > 0 > n_1$. Now n_2 and n_3 cannot have the same sign since $|K_1| \lesssim 2^{d/40}$, thus we may assume $n_2 > 0$ and $n_3 < 0$. Therefore we will have

$$n_0 + n_1 + n_2 + n_3 = K_1, \quad n_0^2 - n_1^2 + n_2^2 - n_3^2 = K_2,$$

which implies

$$(n_2 + n_1)(n_2 + n_3) = \tfrac{1}{2}(K_1^2 - 2K_1 n_0 + K_2).$$

By our assumptions, the right hand side is a nonzero constant whose absolute value does not exceed 2^{2d}. The result now follows from the standard divisor estimate, since knowing $n_2 + n_1$ and $n_2 + n_3$ will allow us to recover the whole quadruple. □

Proceeding to the estimate of the $\mathcal{M}^3$ term, we can see that the only possibility is case (iii) in Lemma 9.4 (since we have required $n_0 \notin \{n_1, n_2, n_3\}$; also if $n_1 + n_2 = 0$ and n_0, n_3 are small, we will gain a power 2^{cd} from the weight Φ so we can argue as above to close the estimate). In this case we will use a completely different argument.

Recall that up to a loss of $2^{O(s^3)d}$ we may assume that with small c,

$$\|\langle n_0\rangle^{-s}\langle\alpha_0\rangle^{1/2-c} f'\|_{l^2L^2} \lesssim 1; \tag{9.19}$$

also by invoking the X_1 norm of w we obtain the estimate

$$\|\langle n_l\rangle^{s}\langle\alpha_l\rangle^{1/2-c} w'\|_{l^pL^2} \lesssim 1 \tag{9.20}$$

with a loss of at most $2^{O(s^3)d}$. Since now $2^{0.9d} \lesssim n_l \lesssim 2^d$, we may remove the $\langle n_0\rangle^{-s}$ and $\langle n_l\rangle^s$ factors in (9.19) and (9.20), and gain at least $2^{1.7sd}$. Therefore, by fixing m_i and β_i first, we will be able to get the desired result if we can prove that

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} &= \sum_{n_0 = n_1 + n_2 + n_3 + K_1} \int_{(T)} \prod_{l=0}^{3} |(A^l)_{n_l,\alpha_l}| \cdot \min\{T, 1/\langle\alpha_4\rangle\} \\ &\lesssim 2^{O(s^3)d} T^{0+} \prod_{l=0}^{3} \|\langle\alpha_l\rangle^{1/2-c} A^l\|_{l^{2+c}L^2},\end{aligned} \tag{9.21}$$

provided c is a small absolute constant, where the (T) integral is taken over the set

$$\Big\{(\alpha_0, \ldots, \alpha_4) : \alpha_0 = \alpha_{14} + |n_0|n_0 - \sum_{l=1}^{3} |n_l|n_l + K_2\Big\}, \tag{9.22}$$

and we restrict to the region where no two of $(-n_0, n_1.n_2, n_3)$ add to zero, $\sum_l \langle n_l\rangle \sim 2^d$, the NR factor satisfies $\big||n_0|n_0 - \sum_{l=1}^3 |n_l|n_l\big| \ll 2^{d/40}$ and $|K_1| + |K_2| \ll 2^{d/40}$.

We will use an interpolation argument to prove (9.21); in fact, it will suffice to prove the estimate when we replace the parameter set $(1/2 - c, 2 + c)$ with $(2/5, 2)$ or $(3, 4)$. When we have $(2/5, 2)$ we will be able to control $\mathfrak{N}A^l$ in $L^{4+}L^{4+}$ for each l, so that we can control the α_4 factor in $l^{1+}L^{1+}$, and invoke the argument used many times before to conclude. When we have $(3, 4)$, assuming the norm of each A^l is one, we will get

$$\|\langle\widetilde{\alpha_l} + |n_l|n_l\rangle((A^l)_{n_l})^\wedge(\widetilde{\alpha_l})\|_{L^1} \lesssim \|\langle\widetilde{\alpha_l} + |n_l|n_l\rangle^3((A^l)_{n_l})^\wedge(\widetilde{\alpha_l})\|_{L^2} =: A^l_{n_l},$$

with

$$\|A^l_{n_l}\|_{l^4} \lesssim \|\langle\widetilde{\alpha_l} + |n_l|n_l\rangle^3 A^l\|_{l^4\widetilde{L}^2} \lesssim 1. \tag{9.23}$$

Therefore when we fix $(n_0, \ldots, n_3)$ and integrate over $(\alpha_0, \ldots, \alpha_4)$, we get

$$\begin{aligned}\mathcal{S}'_{\text{sub}} \lesssim T^{0+}\int_{\mathbb{R}^4}\Big\langle\widetilde{\alpha_0} - \sum_{l=1}^3\widetilde{\alpha_l} - K_2\Big\rangle^{-1+s^3}\prod_{l=0}^3\langle\widetilde{\alpha_l} + |n_l|n_l\rangle^{-1}&\\ \times\prod_{l=0}^3\langle\widetilde{\alpha_l} + |n_l|n_l\rangle|((A^l)_{n_l})^\wedge(\widetilde{\alpha_l})|\cdot\prod_{l=1}^4 d\widetilde{\alpha_l}&\\ \lesssim T^{0+}\Big\langle|n_0|n_0 - \sum_{l=1}^3|n_l|n_l + K_2\Big\rangle^{-1+s^3}\prod_{l=0}^3 A^l_{n_l}.&\end{aligned}$$

We then sum this over (n_l); by Hölder, we only need to bound the sum

$$\sum_{(n_0,\ldots,n_3)}\Big\langle|n_0|n_0 - \sum_{l=1}^3|n_l|n_l + K_2\Big\rangle^{-1+s^4}(A^0_{n_0})^4.$$

If we fix $|n_0|n_0 - \sum_{l=1}^3 |n_l|n_l = K_3$ with $|K_3| \ll 2^{d/40}$, the corresponding sum will be $\lesssim 2^{O(s^3)d}$, since each n_0 appears at most this many times due to Lemma 9.4; also the sum over K_3 will contribute at most $\sum_{|K_3|\lesssim 2^{d/40}}\langle K_3 - K_2\rangle^{-1+s^3} = 2^{O(s^3)d}$. This completes the proof for $\mathcal{M}^3$.

What remains is when $j \in \{3, 4.5\}$ and (say) $n_0 = n_1$. Note that the case when three of n_l are related by m will be treated at the end of the proof. In both cases we will use the expressions (9.11) and (9.15), but with f' replaced by f, and $(w')^{\omega_1}$ replaced by $(w'')^{\omega_1}$ (if $j = 3$), z^1 replaced by y^1 (if $j = 4.5$). This is easily justified by definition and the fact that $n_0 = n_1$. We will assume $\sum_{l\geq 2}\langle n_l\rangle \sim 2^{d'}$, then fix d and d'. Here we will use a new bound for f. Recall from Proposition 3.4 that $\|g\|_{X'_j} \lesssim 1$ for some $j \in \{1, 2, 5, 7\}$ implies $\|f\|_{X'_9} \lesssim 1$, or equivalently

$$\|f\|_{L^q l^{p'}_{\sim 2^d}} \lesssim 2^{rd}T_d, \tag{9.24}$$

where T_d is such that

$$\sum_{d\geq 0} T_d \lesssim 1. \tag{9.25}$$

In the easier case $j = 4.5$, we will be able to fix m_i and β_i, then estimate $\mathcal{S}_{\mathrm{sub}}$ by (note we have all the restrictions made above, say $\langle n_0\rangle \sim 2^d$)

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} &\lesssim \sum_{n_0,\, n_2+n_3+n_4=c_1} (\langle n_3\rangle + \langle n_4\rangle)^{-1} \int_{(T)} |f_{n_0,\alpha_0}|\,|(y^1)_{n_0,\alpha_1}| \prod_{l=2}^{4} |(z^l)_{n_l,\alpha_l}| \cdot |\phi_{\alpha_5}| \\ &\lesssim T^{0+} \sum_{n_0,n_2+n_3+n_4=c_1} (\langle n_3\rangle + \langle n_4\rangle)^{-1} \|\widehat{f_{n_0}}\|_{L^q} \|\widehat{(y^1)_{n_0}}\|_{L^1} \prod_{l=2}^{4} \|\widehat{(z^l)_{n_l}}\|_{L^1} \\ &\lesssim T^{0+}2^{rd}T_d \cdot 2^{-rd} \cdot \|z^2\|_{l^3L^1} \prod_{l=3}^{4} \|\langle n_l\rangle^{-1/2} z^l\|_{l^{6/5}L^1} \lesssim T^{0+}2^{-csd'}T_d, \end{aligned} \tag{9.26}$$

using (9.24) for f, the X_2 bound for y^1, and slightly weaker bounds for z^l that follow from Proposition 8.9. Here c_j are constants, and the (T) integral is taken over the set

$$\Big\{(\alpha_0, \ldots, \alpha_5) : \alpha_0 = \alpha_{15} - \sum_{l=2}^{4} |n_l| n_l + c_2\Big\}. \tag{9.27}$$

The reason we can gain $2^{csd'}$ is that in (9.26) we can restrict some n_l, where $2 \leq l \leq 4$, to be $\sim 2^{d'}$ before using the corresponding control for z^l (for example, when $n_2 \sim 2^d$ we will have $\|z^2\|_{l^3_{n_2\sim 2^{d'}}L^1} \lesssim 2^{-csd'}$). If we then sum over m_i, integrate over β_i, and sum over d, d', we will get the desired estimate.

In the harder case $j = 3$, we will assume $\langle m_i\rangle + \langle\beta_i\rangle \ll 2^{d'/90}$. In fact, if this does not hold, we will gain a power $2^{cd'}$ from this term and estimate the $\mathcal{S}_{\mathrm{sub}}$ as above, except that we estimate $(w')^{\omega_2}$ and $(w')^{\omega_3}$ in l^2L^1 with a loss of $2^{O(s)d'}$ to conclude (note in particular we estimate f and $(w'')^{\omega_1}$ exactly as above, so we do not gain or lose any power of 2^d). In the same way, we may assume $\langle\alpha_4\rangle \ll 2^{d'/90}$ in (9.11). Now if $n_2 + n_3 = 0$, we must have $|\Phi| \lesssim 2^{-|d-d'|}$. Also we may replace z^2 and z^3 in (9.11) with y^2 and y^3 (in the same way we replace f' and $(w')^{\omega_1}$ with f and $(w'')^{\omega_1}$; note that we have not made any restrictions on α_2 and α_3). Then we may fix m_i and β_i (here the m variables satisfy some linear relation which we ignore) and bound

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} &\lesssim 2^{-|d-d'|} \sum_{n_0,n_2} \int_{\alpha_0=\alpha_1+\cdots+\alpha_4+c_2} |f_{n_0,\alpha_0}| \\ &\qquad\qquad \times |((w'')^{\omega_1})_{n_0,\alpha_1}| \cdot |(y^2)_{n_2,\alpha_2}| \cdot |(y^3)_{-n_2,\alpha_3}| \cdot |\phi_{\alpha_4}| \\ &\lesssim T^{0+}2^{-|d-d'|} \sum_{n_0,n_2} \|\widehat{f_{n_0}}\|_{L^q} \|(((w'')^{\omega_1})_{n_0})^\wedge\|_{L^1} \|\widehat{(y^2)_{n_2}}\|_{L^1} \|\widehat{(y^3)_{-n_2}}\|_{L^1} \\ &\lesssim T^{0+}2^{-|d-d'|} \|f\|_{l^{p'}_{\sim 2^d}L^q} \|w''\|_{l^p_{\sim 2^d}L^1} \|y^2\|_{l^2_{\sim 2^{d'}}L^1} \|y^3\|_{l^2_{\sim 2^{d'}}L^1} \lesssim T^{0+}2^{-|d-d'|}T_d,\end{aligned}$$

where the c_j are constants, and we are restricting to $n_0 \sim 2^d$, $n_2 \sim 2^{d'}$. Then we may sum and integrate over (m_i, β_i), and sum over d, d' to bound this part by T^{0+}.

Now assume $j = 3$, $n_2 + n_3 \neq 0$, and all the restrictions made before hold. In particular we have $n_2 \sim n_3 \sim 2^{d'}$ and $|\Xi'| \gtrsim 2^{d'}$ where $\Xi' = |n_2|n_2 + |n_3|n_3$ (again we may assume $2^{d'}$ is large, otherwise we proceed as before). Fixing m_i and β_i, we then need to bound

$$\mathcal{S}_{\text{sub}} = \sum_{n_0,n_2} \int_{(T)} |f_{n_0,\alpha_0}| \cdot |((w'')^{\omega_1})_{n_0,\alpha_1}| \cdot |(z^2)_{n_2,\alpha_2}| \cdot |(z^3)_{c_1-n_2,\alpha_3}| \cdot |\phi_{\alpha_4}|,$$

where $c_1 - n_2 = n_3$, the (T) integral is over the set

$$\{(\alpha_0, \ldots, \alpha_4) : \alpha_0 = \alpha_{14} - \Xi' + c_2\}, \tag{9.28}$$

and the $c_j \ll 2^{d'/10}$ are constants. Also each z^l *here* is either w' or $\overline{w'}$. Now, by Proposition 3.4, we can show that $\|g\|_{X'_j} \leq 1$ for some $j \in \{1, 2, 5, 7\}$ implies

$$\|f\|_{X'_{10}} = \|\langle n_0\rangle^{-r}\langle\alpha_0\rangle^{1/8} f\|_{l^1_{d\geq 0} L^{\tau'} l^{p'}_{n_0\sim 2^d}} \lesssim 1. \tag{9.29}$$

For the w'' we will use the X_7 bound, and for z^l we will simply use the X_1 bound. Now, since at least one α_l must be $\gtrsim 2^{d'}$, we will gain some $2^{cd'}$ from the $\langle\alpha_l\rangle$ weight in one of the above bounds. If $l = 0$, we can then estimate f in $l^{p'}L^{\tau'}$ by $2^{rd}T_d$, w'' in l^pL^1 by 2^{-rd} (recall we are restricting to $n_0 \sim 2^d$ and $n_2 \sim 2^{d'}$), and $z^{2,3}$ in l^2L^1 with a loss of $2^{O(s)d'}$, so that we can use Hölder to conclude. If $l = 1$, we simply replace the l^pL^1 bound by the l^pL^2 bound and argue as in the case $l = 0$. If $l \in \{2, 3\}$, we may replace the l^2L^1 bound for z^l by the l^2L^2 bound and argue as in the case $l = 0$. If $l = 4$ we simply gain from the ϕ factor. This completes the proof for the $n_0 = n_1$ case.

Finally, assume that $j = 3$, and three of n_l are related by m. We may assume that $\langle m_i\rangle \ll 2^{d/90}$, so that $n_l \sim 2^d$ for each l. Then we fix m_i and β_i, so that n_l are uniquely determined by n_0. The corresponding $\mathcal{S}_{\text{sub}}$ will be bounded by

$$T^{0+}\|f'\|_{l^4_{\sim 2^d}L^q}\|z^1\|_{l^4_{\sim 2^d}L^1}\|z^2\|_{l^4_{\sim 2^d}L^1}\|z^3\|_{l^4_{\sim 2^d}L^1} \lesssim T^{0+}2^{-sd},$$

due to a similar computation as in the proof of Proposition 9.1. □

Now we start to consider the $\mathcal{M}^2$ term. Fixing the functions g, f, f' and the scale d as before, we need to bound the expression

$$\mathcal{S} = \sum_{n_0=n_1+n_2+m_1+\cdots+m_\mu} \int_{(T)} \Phi^2 \cdot \overline{f_{n_0,\alpha_0}} \prod_{l=1}^{2} (y^{\omega_l})_{n_l,\alpha_l} \cdot \phi_{\alpha_3} \times (\chi e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})})^\wedge(\alpha_4) \prod_{i=1}^{\mu} \frac{(u''')_{m_i,\beta_i}}{m_i}. \tag{9.30}$$

Here the (T) integration is over the set

$$\{(\alpha_0, \ldots, \alpha_4, \beta_1, \ldots, \beta_\mu) : \alpha_0 = \alpha_{14} + \beta_{1\mu} + \Xi\}$$

with the NR factor

$$\Xi = |n_0|n_0 - \sum_{l=1}^{2} |n_l|n_l - \sum_{i=1}^{\mu} |m_i|m_i. \tag{9.31}$$

Note that we may insert χ since f has compact time support. Suppose the minimum of $\langle n_l \rangle$ is $\sim 2^h$ and also fix h. Then $\langle m_i \rangle \ll 2^h$, so that $|\Xi| \gtrsim 2^{d+h}$; also $h \leq d + O(1)$ and $|\Phi^2| \lesssim 2^h$. Note that one of α or β variables must be $\gtrsim 2^{d+h}$. We first treat the easy cases, which we collect in the following proposition.

Proposition 9.5. *Let $\mathcal{S}$ be defined in (9.30), where all the restrictions made above are assumed. If $h < 0.9d$, or*

$$\langle \alpha_0 \rangle + \langle \alpha_3 \rangle + \langle \alpha_4 \rangle + \sum_{i=1}^{\mu} (\langle m_i \rangle + \langle \beta_i \rangle) \gtrsim 2^{d/90}, \tag{9.32}$$

then the corresponding contribution will be bounded by $T^{0+} 2^{(0-)d}$.

Proof. First assume $h \geq 0.9d$. If $\beta_i \gtrsim 2^{d+h}$ for some i, we may use the X_4 bound for $\langle \partial_x \rangle^{-s^3} u'''$ to gain a power $2^{0.99(d+h)}$ and then estimate this $(u''')_{m_i,\beta_i}$ factor in $L^2 L^2$. Next we may bound

$$|(\chi e^{\mathrm{i}(\Delta_{n_1} + \Delta_{n_2} - \Delta_{n_0})})^\wedge(\alpha_4)| \lesssim 2^{s^3 d} \langle \alpha_4 \rangle^{-1} \tag{9.33}$$

by Lemma 8.6 and Proposition 8.7, and estimate the right hand side (again, viewed as a function of space-time supported at $n = 0$) as well as the ϕ_{α_3} factor in $l^{1+} L^{1+}$. We then fix (m_j, β_j) for $j \neq i$ to produce an $\mathcal{S}_{\text{sub}}$ involving $(u''')_{m_i,\beta_i}$, which we estimate by controlling $\mathfrak{N} f$ in $L^{6-} L^{6-}$, $\mathfrak{N} y^{\omega_l}$ in $L^{6+} L^{6+}$ (using the X_1 bound for y and the norm for f deduced from the $X_{6'}$ bound for g; here the $6-$ and $6+$ are $6 + O(s)$). In this process we lose at most $2^{O(s)d}$, but the gain $2^{0.99(d+h)}$ (even after canceling the 2^h loss coming from the Φ^2 weight) will allow us to cancel the Φ^2 factor and still gain 2^{cd}.

Next, suppose $\langle \alpha_3 \rangle \gtrsim 2^{d+h}$. By using (9.33) and losing a harmless $2^{O(s)d}$ factor, the argument for α_4 can be done in the same way. Let c_j be constants (or functions of n_l); recalling we are restricting to $\sum_l \langle n_l \rangle \sim 2^d$, we may fix m_i and β_i, and bound

$$\begin{aligned}
\mathcal{S}_{\text{sub}} &\lesssim T^{0+} \sum_{n_0 = n_1 + n_2 + c_1} \int_{\alpha_0 = \alpha_1 + \cdots + \alpha_4 + c_2(n_0, \ldots, n_2)} 2^h \\
&\qquad\qquad \times |f_{n_0,\alpha_0}| \cdot \prod_{l=1}^{2} |(y^{\omega_l})_{n_l,\alpha_1}| \cdot \frac{\mathbf{1}_{\{\alpha_3 \gtrsim 2^{d+h}\}}}{\langle \alpha_3 \rangle^{0.9} \langle \alpha_4 \rangle} \\
&\lesssim T^{0+} \sum_{n_0 = n_1 + n_2 + c_1} 2^{-0.62d} \|\widehat{f_{n_0}}\|_{L^q} \|\widehat{(y^{\omega_1})_{n_1}}\|_{L^1} \|\widehat{(y^{\omega_2})_{n_2}}\|_{L^1} \\
&\lesssim T^{0+} 2^{-cd} \|\langle n_0 \rangle^{-0.2} f\|_{l^{3/2} L^q} \|\langle n_1 \rangle^{-0.2} y\|_{l^{3/2} L^1} \|\langle n_2 \rangle^{-0.2} y\|_{l^{3/2} L^1} \lesssim T^{0+} 2^{-cd}.
\end{aligned}$$

Thus this term is also acceptable.

Next, assume that $\langle \alpha_1 \rangle \gtrsim 2^{d+h}$ (the α_2 case is proved in the same way), and that one of $\alpha_0, \alpha_3, \alpha_4, m_i$ or β_i is $\gtrsim 2^{d/90}$. We then use (9.33) to bound the exponential factor and fix (m_i, β_i). To estimate the resulting $\mathcal{S}_{\text{sub}}$, we use the $\langle \alpha_1 \rangle^b$ factor in the X_1 bound for y to cancel the Φ^2 factor which is at most 2^h and bound the resulting y^{ω_1} factor in $L^2 L^2$, then bound $\mathfrak{N} f$ and $\mathfrak{N} y^{\omega_2}$ in $L^{4+} L^{4+}$, and bound the factors involving α_3 and α_4 in $l^{1+} L^{1+}$, where $4+$ is some $4 + c$. In this process we may lose $2^{O(s)d}$, but since another α_l or (m_i, β_i) is $\gtrsim 2^{d/90}$, we will be able to gain 2^{cd} from this factor (since the $L^{4+} L^{4+}$ Strichartz estimate allows for some room), and we will find this term acceptable.

The only remaining case is when $\langle\alpha_0\rangle \gtrsim 2^{d+h}$. By basically the same argument as above, we may assume that the other α_l and (m_i, β_i) are all $\ll 2^{d/90}$. Also recall that two of n_l $(0 \le l \le 2)$ are $\sim 2^d$ and the third is $\sim 2^h$. Now we may use the bound (9.33), then fix (α_3, α_4) and all (m_i, β_i) to produce

$$|\mathcal{S}_{\text{sub}}| \lesssim 2^{(b-s)h-(1-b)d} \sum_{n_0=n_1+n_2+c_1} \int_{\alpha_0=\alpha_1+\alpha_2+\Xi'+c_2} A_{n_0,\alpha_0} B_{n_1,\alpha_1} C_{n_2,\alpha_2}, \tag{9.34}$$

where the $c_j \ll 2^{d/10}$ are constants, the factor Ξ' is

$$\Xi' = |n_0|n_0 - |n_1|n_1 - |n_2|n_2, \tag{9.35}$$

and the relevant functions are defined by

$$A_{n_0,\alpha_0} = \langle n_0\rangle^{-s}\langle\alpha_0\rangle^{1-b}|f_{n_0,\alpha_0}|,$$
$$B_{n_1,\alpha_1} = \langle n_1\rangle^{s}|(y^{\omega_1})_{n_1,\alpha_1}|, \quad C_{n_2,\alpha_2} = \langle n_2\rangle^{s}|(y^{\omega_2})_{n_2,\alpha_2}|.$$

Also note that when we sum over m_i, and integrate over β_i and (α_3, α_4), we will gain T^{0+} and lose at most $2^{O(s^3)d}$.

Now we estimate $\mathcal{S}_{\text{sub}}$. If $\|g\|_{X'_m} \le 1$ for some $m \in \{1, 2\}$, by using Proposition 3.4, we may assume $\|\langle\alpha_0\rangle f\|_{X'_j} \lesssim 1$ for some $j \in \{1, 2\}$ (this relies on the fact that $\mathcal{E}$ can be written as the sum of two linear operators that are bounded from each W_j to X_1 separately, where $\|u\|_{W_j} = \|\langle\xi\rangle^{-1}u\|_{X_j}$). If $\|g\|_{X'_m} \lesssim 1$ for some $m \in \{5, 7\}$, since we may insert a $\mathbf{1}_E$ factor to $\overline{f_{n_0,\alpha_0}}$ with $E = \{n_0 \sim 2^{d'}, \alpha_0 \gtrsim 2^{d'}\}$ with $d' \in \{d, h\}$, we can use (3.23) and again assume $\|\langle\alpha_0\rangle f\|_{X'_j} \lesssim 1$ for some $j \in \{1, 2\}$. Next, notice that $|\alpha_0 - \Xi'| \lesssim 2^{d/10}$, so α_0 is also restricted to some set of measure $O(2^{1.1d})$ for each fixed n_0. Since α_0 is restricted to be $\gtrsim 2^{1.9d}$ and $n_0 \lesssim 2^d$, we will have

$$\begin{aligned}\|\langle\alpha_0\rangle f\|_{X'_1} &\lesssim 2^{O(s)d}\|\langle\alpha_0\rangle^{0.6} f\|_{l^2L^2} \lesssim 2^{(O(s)+0.55)d}\|\langle\alpha_0\rangle^{0.6} f\|_{l^2L^\infty}\\ &\lesssim 2^{(0.6-0.4\times 1.9)d}\|\langle\alpha_0\rangle f\|_{l^2L^\infty} \lesssim 2^{-cd}\|\langle\alpha_0\rangle f\|_{X'_2},\end{aligned}$$

thus we may furthermore assume $j = 1$.

Now, using this bound for f and the X_1 bound for y, we deduce that

$$\|A\|_{l^{p'}L^2} + \|\langle\alpha_1\rangle^b B\|_{l^pL^2} + \|\langle\alpha_2\rangle^b C\|_{l^pL^2} \lesssim 2^{O(s^2)d}.$$

Let us define

$$\mathbf{B}_{n_1} = \|\langle\widetilde{\alpha_1} + |n_1|n_1\rangle^b \widehat{B_{n_1}}(\widetilde{\alpha_1})\|_{L^2}$$

and $\mathbf{C}_{n_3}$ similarly, so that $\|\mathbf{B}\|_{l^p} + \|\mathbf{C}\|_{l^p} \lesssim 2^{O(s^2)d}$. Then we will have the estimate

$$\|\langle\alpha_0 - \Xi' - c_2\rangle^{2b-1/2}(\widehat{B_{n_1}} * \widehat{C_{n_2}})(\alpha_0 - |n_0|n_0 - c_2)\|_{L^2_{\alpha_0}} \lesssim \mathbf{B}_{n_1}\mathbf{C}_{n_2},$$

which, after taking Fourier transform, follows from the standard one-dimensional inequality $\|fg\|_{H^{2b-1/2}} \lesssim \|f\|_{H^b}\|g\|_{H^b}$. Now we will be able to control $\mathcal{S}_{\text{sub}}$ by

$$\mathcal{S}_{\text{sub}} \lesssim 2^{\lambda}\Big(\sum_{n_0}\Big\|\sum_{n_1+n_2=n_0-c_1}(\widehat{B_{n_1}} * \widehat{C_{n_2}})(\alpha_0 - |n_0|n_0 - c_2)\Big\|^p_{L^2_{\alpha_0}}\Big)^{1/p},$$

where $\lambda = (b-s)h + (b-1+O(s^2))d$, and the square of the inner L^2 norm is

$$\begin{aligned}
\mathcal{J}^2 &= \int_{\mathbb{R}} \Big| \sum_{n_1+n_2=n_0-c_1} (\widehat{B_{n_1}} * \widehat{C_{n_2}})(\alpha_0 - |n_0|n_0 - c_2) \Big|^2 \, d\alpha_0 \\
&\lesssim \int_{\mathbb{R}} \Big(\sum_{n_1+n_2=n_0-c_1} \langle \alpha_0 - \Xi' - c_2 \rangle^{1-4b} \Big) \, d\alpha_0 \\
&\quad \times \Big(\sum_{n_1+n_2=n_0-c_1} \langle \alpha_0 - \Xi' - c_2 \rangle^{4b-1} |(\widehat{B_{n_1}} * \widehat{C_{n_2}})(\alpha_0 - |n_0|n_0 - c_2)|^2 \Big) \\
&\lesssim \sup_{\alpha_0} \Big(\sum_{n_1+n_2=n_0-c_1} \langle \alpha_0 - \Xi' - c_2 \rangle^{1-4b} \Big) \\
&\quad \times \sum_{n_1+n_2=n_0-c_1} \int_{\mathbb{R}} \langle \alpha_0 - \Xi' - c_2 \rangle^{4b-1} |(\widehat{B_{n_1}} * \widehat{C_{n_2}})(\alpha_0 - |n_0|n_0 - c_2)|^2 \, d\alpha_0 \\
&\lesssim \sup_{\alpha_0} \Big(\sum_{n_1+n_2=n_0-c_1} \langle \alpha_0 - \Xi' - c_2 \rangle^{1-4b} \Big) \cdot \sum_{n_1+n_2=n_0-c_1} \mathbf{B}_{n_1}^2 \mathbf{C}_{n_2}^2 .
\end{aligned}$$

Next we claim that for fixed n_0 and α_0 we have

$$\sum_{n_1+n_2=n_0-c_1} \langle \alpha_0 - \Xi' - c_2 \rangle^{-3/4} \lesssim 1. \tag{9.36}$$

In fact, if $n_1 n_2 < 0$, then $\alpha_0 - \Xi' - c_2$ is a linear expression in n_1 with leading coefficient $k = \pm(n_0 - c_1)/2 \gtrsim 2^{0.9d}$ (we assume d is large enough), so any two summands in (9.36) differ by at least k, while there are $\lesssim 2^d$ summands. The sum is thus bounded by

$$1 + \sum_{h=1}^{2^d} (kh)^{-3/4} \lesssim 1 + k^{-3/4} 2^{d/4} \lesssim 1. \tag{9.37}$$

If $n_1 n_2 > 0$, then $\alpha_0 - \Xi' - c_2$ equals $\pm\frac{1}{2}(n_1 - n_2)^2$ plus a constant, so similarly we only need to prove

$$\sum_{k\in\mathbb{Z}} \langle \alpha - k^2 \rangle^{-3/4} \lesssim 1$$

for each α, but this is again easily proved by separating the cases $\langle k \rangle^2 \lesssim \langle \alpha \rangle$ and otherwise, and applying elementary inequalities.

Now we are able to bound

$$\begin{aligned}
\mathcal{S}_{\mathrm{sub}} &\lesssim 2^{\lambda} \Big(\sum_{n_0} \Big(\sum_{n_1+n_2=n_0-c_1} \mathbf{B}_{n_1}^2 \mathbf{C}_{n_2}^2 \Big)^{p/2} \Big)^{1/p} \\
&\lesssim 2^{\lambda + (1/2 - 1/p)d} \Big(\sum_{n_0} \sum_{n_1+n_2=n_0-c_1} \mathbf{B}_{n_1}^p \mathbf{C}_{n_2}^p \Big)^{1/p} \lesssim 2^{\lambda + (1/2-1/p)d} \|\mathbf{B}\|_{l^p} \|\mathbf{C}\|_{l^p},
\end{aligned}$$

where we notice that

$$\begin{aligned}
\lambda + (1/2 - 1/p)d &= (b-s)h + \big(b - 1/2 - 1/p + O(s^2)\big)d \\
&\lesssim (2b-1)d + \big(1/2 - s - 1/p + O(s^2)\big)d,
\end{aligned}$$

and this is $\leq -c(1/2-b)d$ by (1.3). We may then sum and integrate over the previously fixed variables to get a desirable estimate for $\mathcal{S}$.

Finally, suppose $h < 0.9d$. Since at least one α_l or β_i will be $\gtrsim 2^{d+h}$, we may repeat the arguments above; using the inequality $2^{b(d+h)} \gtrsim 2^{cd+h}$ that holds for $h < 0.9d$, we will be able to gain an additional power of 2^{cd} after canceling the Φ^2 weight, which will allow us to close the estimate as above. This completes the proof. □

What remains to be bounded, denoted by $\mathcal{S}^E$, is actually the same summation-integration as $\mathcal{S}$, but restricted to the region $h \geq 0.9d$ and with the additional factor $\mathbf{1}_E$, where

$$E = \{\langle\alpha_0\rangle \vee \langle\alpha_3\rangle \vee \langle\alpha_4\rangle \vee \langle m_i\rangle \vee \langle\beta_i\rangle \ll 2^{d/90},\ \forall i\},$$

with $a \vee b$ meaning $\max\{a,b\}$. Now let $E_l = \{\langle\alpha_l\rangle \ll 2^{d/90}\}$ for $l \in \{1,2\}$. We have

$$\mathbf{1}_E = \mathbf{1}_{E\cap E_1} + \mathbf{1}_{E\cap E_2} + \mathbf{1}_{E-(E_1\cup E_2)}. \tag{9.38}$$

By symmetry, we need to bound $\mathcal{S}^{E\cap E_1}$ and $\mathcal{S}^{E-(E_1\cup E_2)}$ (whose meaning is obvious). In the latter case, we may assume that $\alpha_1 \gtrsim 2^{d+h}$, and also $\alpha_2 \gtrsim 2^{d/90}$, so we can estimate this part in the same was as in the proof of Proposition 9.5.

It remains to bound $\mathcal{S}^{E\cap E_1}$. Let $E\cap E_1 = F$. Using (3.14) and (8.12) we may compute

$$\begin{aligned}(y^{\omega_2})_{n_2,\alpha_2} &= (\chi(t)e^{-H\partial_{xx}}w^{\omega_2}(0))_{n_2,\alpha_2} + \big(\mathcal{E}(\mathbf{1}_{[-T,T]}\cdot\mathcal{N}^2(y,y))^{\omega_2}\big)_{n_2,\alpha_2}\\ &\quad+\sum_{j\in\{3,3.5,4,4.5\}}(\mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^j)^{\omega_2})_{n_2,\alpha_2}\\ &=\sum_{j\in\{0,3,3.5,4,4.5\}}((\mathcal{M}^j)^{\omega_2})_{n_2,\alpha_2} + (\mathcal{L}^1)_{n_2,\alpha_2} + (\mathcal{L}^2)_{n_2,\alpha_2}.\end{aligned} \tag{9.39}$$

Here we denote $\mathcal{M}^0 = \chi(t)e^{\mathrm{i}\partial_{xx}}w(0)$, and

$$\begin{aligned}(\mathcal{L}^1)_{n_2,\alpha_2} &= c_1\int_{\mathbb{R}^2}\frac{\widehat{\chi}(\alpha_2-\gamma_2)\widehat{\chi}(\gamma_2-\gamma_1)}{\gamma_2}\mathcal{I}_{n_2,\gamma_1}\,d\gamma_1\,d\gamma_2,\\ (\mathcal{L}^2)_{n_2,\alpha_2} &= c_2\widehat{\chi}(\alpha_2)\cdot\int_{\mathbb{R}^2}\frac{\widehat{\chi}(\gamma_2-\gamma_1)}{\gamma_2}\mathcal{I}_{n_2,\gamma_1}\,d\gamma_1\,d\gamma_2,\end{aligned}$$

where $\mathcal{I} = (\mathbf{1}_{[-T,T]}\mathcal{N}^2(y,y))^{\omega_2}$. Interpreting the singular integral as a principal value, we may compute that

$$\begin{aligned}&\left|\int_{\mathbb{R}}\frac{\widehat{\chi}(\gamma_2-\gamma_1)}{\gamma_2}\,d\gamma_2\right| \lesssim \frac{1}{\langle\gamma_1\rangle},\\ &\left|\int_{\mathbb{R}}\frac{\widehat{\chi}(\alpha_2-\gamma_2)\widehat{\chi}(\gamma_2-\gamma_1)}{\gamma_2}\,d\gamma_2\right| \lesssim \frac{1}{(\langle\alpha_2\rangle+\langle\gamma_1\rangle)\langle\alpha_2-\gamma_1\rangle^{1/s}},\\ &\left|\nabla_{\alpha_2,\gamma_1}\int_{\mathbb{R}}\frac{\widehat{\chi}(\alpha_2-\gamma_2)\widehat{\chi}(\gamma_2-\gamma_1)}{\gamma_2}\,d\gamma_2\right| \lesssim \frac{1}{(\langle\alpha_2\rangle+\langle\gamma_1\rangle)^2\langle\alpha_2-\gamma_1\rangle^{1/s}},\end{aligned}$$

where the third inequality can be proved by integrating by parts in γ_2. Now, to treat the first three terms in (9.39), we may use Proposition 9.1, the easy observation that

$$\|\langle n_2\rangle^{-1/20}\langle\alpha_2\rangle^{\kappa}(\mathcal{M}^0)_{n_2,\alpha_2}\|_{l^2L^2} \lesssim \|\langle n\rangle^{-1/20}(w(0))_n\|_{l^2} \lesssim 1,$$

together with the following

Proposition 9.6. *If we consider the sum* (9.30) *with the factor* $\mathbf{1}_F$*, and* y^{ω_2} *replaced by some function* ζ *satisfying*

$$\|\langle n_2\rangle^{-1/20}\langle\alpha_2\rangle^{\kappa}\zeta_{n_2,\alpha_2}\|_{l^2L^2}\lesssim 1, \tag{9.40}$$

then this contribution can be bounded by T^{0+}.

Proof. Since in F we will have $\langle\alpha_2\rangle\gtrsim 2^{d+h}$, we can gain a power $2^{0.999(d+h)}$ from the $\langle\alpha_2\rangle^{\kappa}$ factor in the bound for ζ. After exploiting this, we may estimate ζ in L^2L^2 with a loss $2^{(1/20+O(s))d}$. Then we fix (m_i,β_i) as usual, and use the inequality (9.33) to bound the factor involving α_4. To bound the resulting $\mathcal{S}_{\mathrm{sub}}$ term, we estimate ζ in L^2L^2, $\mathfrak{N}f$ and $\mathfrak{N}y$ in $L^{4+}L^{4+}$ (where 4+ equals $4+c$) with a loss of $2^{O(s)d}$, the α_3 and α_4 factors in $l^{1+}L^{1+}$. Note that here we will gain a power T^{0+}, and the total power of 2^d we may lose is at most $2^{(1.1+O(s))d}$, which is smaller than the gain $2^{0.999(d+h)}$. Then we sum over m_i and integrate over β_i to conclude. □

Next consider the contribution of $\mathcal{L}^2$. Since we are in F (thus $\alpha_2\gtrsim 2^d$), the gain from $\widehat{\chi}(\alpha_2)$ will overwhelm any possible loss in terms of 2^d. Therefore we may even fix all the n, m and β variables and estimate the integral in α variables and γ_1 only; but we can easily estimate this integral by controlling all the factors except $\langle\gamma_1\rangle^{-1}|\mathcal{I}_{n_2,\gamma_1}|$ in L^{1+} (since the expression now has a convolution structure in the α variables), and estimate the $\langle\gamma_1\rangle^{-1}|\mathcal{I}_{n_2,\gamma_1}|$ factor in L^1. This last estimate is due to (the proof of) Proposition 9.1, which implies

$$\|\langle\gamma_1\rangle^{-1}\mathcal{I}_{n_2,\gamma_1}\|_{L^1}\lesssim\|\langle\gamma_1\rangle^{\kappa-1}\mathcal{I}_{n_2,\gamma_1}\|_{L^2}\lesssim 2^{O(1)d}.$$

It then remains to bound the $\mathcal{L}^1$ contribution. After integrating over γ_2, we may rename the variable $\alpha_2-\gamma_1$ as γ_2, and reduce to estimating (up to a constant)

$$\mathcal{S}^F=\sum_{n_0=n_1+n_2+m_1+\cdots+m_\mu}\int_{(T)}\mathbf{1}_F\cdot\Phi^2\cdot\overline{f_{n_0,\alpha_0}}\cdot(y^{\omega_1})_{n_1,\alpha_1}\cdot\phi_{\alpha_3}$$
$$\times(\chi e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})})^\wedge(\alpha_4)\cdot\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}\cdot\eta(\gamma_1,\gamma_2)\cdot\mathcal{I}_{n_2,\gamma_1},$$

where η is some function bounded by

$$|\eta(\gamma_1,\gamma_2)|\lesssim\frac{1}{\langle\gamma_1\rangle\langle\gamma_2\rangle^{1/s}},\qquad|\partial_{\gamma_1}\eta(\gamma_1,\gamma_2)|\lesssim\frac{1}{\langle\gamma_1\rangle^2\langle\gamma_2\rangle^{1/s}},$$

and the (T) integral is taken over the set

$$\{(\alpha_0,\alpha_1,\alpha_3,\alpha_4,\beta_1,\ldots,\beta_\mu,\gamma_1,\gamma_2):\alpha_0=\alpha_1+\alpha_{34}+\beta_{1\mu}+\gamma_{12}+\Xi\}$$

with the NR factor as in (9.31). Clearly we may also assume $\langle\gamma_2\rangle\ll 2^{d/90}$ and add this restriction into F (or we simply gain a large power of 2^d and proceed as above); after doing this we will have $F\subset\{\langle\gamma_1\rangle\gtrsim 2^{d+h}\}$.

Next, note that $\overline{\mathcal{N}^2(y,y)}=\overline{\mathcal{N}}^2(\overline{y},\overline{y})$, where $\overline{\mathcal{N}^2}$ is another bilinear form that differs from $\mathcal{N}^2$ only in the Φ^2 weights; moreover, the Φ weight for $\overline{\mathcal{N}^2}$ will satisfy all the bounds we have for the Φ weight for $\mathcal{N}^2$. Thus we only need to bound the above expression with $\mathcal{I}_{n_2,\gamma_1}$ replaced by $(\mathbf{1}_{[-T,T]}\mathcal{N}^2(y^{\omega_2}, y^{\omega_2}))_{n_2,\gamma_1}$. Clearly we may also fix the parameters μ' and ω' in $\mathcal{N}_{\mu'}^{\omega'2}$ and reduce to estimating

$$\mathcal{S}' = \sum_{n_0=n_1+n_5+n_6+m_1+\cdots+m_\nu} \int_{(T)} \Phi^2(\Phi^2)' \cdot \overline{f_{n_0,\alpha_0}} \times \prod_{l\in\{1,5,6\}} (y^{\omega_l})_{n_l,\alpha_l} \cdot \phi_{\alpha_3}\phi_{\alpha_7} \cdot \prod_{i=1}^{\nu} \frac{(u''')_{m_i,\beta_i}}{m_i} \times \int_{\alpha_4+\alpha_8=\alpha_9} \mathbf{1}_F \eta(\gamma_1, \gamma_2) \cdot (\chi e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})})^\wedge(\alpha_4)(\chi e^{\mathrm{i}(\Delta_{n_5}+\Delta_{n_6}-\Delta_{n_2})})^\wedge(\alpha_8),$$

where $\nu = \mu + \mu'$ and the (T) integral is taken over the set

$\{(\alpha_0, \alpha_1, \alpha_3, \alpha_5, \alpha_6, \alpha_7, \alpha_9, \beta_1, \ldots, \beta_\nu, \gamma_2) : \alpha_0 = \alpha_1 + \alpha_3 + \alpha_{57} + \alpha_9 + \beta_{1\nu} + \gamma_2 + \Xi'\}$

with the new NR factor

$$\Xi' = |n_0|n_0 - |n_1|n_1 - |n_5|n_5 - |n_6|n_6 - \sum_{i=1}^{\nu} |m_i|m_i.$$

The Φ^2 and $(\Phi^2)'$ are functions of the n and m variables that are bounded by $\min_{l\in\{0,1,2\}}\langle n_l\rangle$ and $\min_{l\in\{2,5,6\}}\langle n_l\rangle$ respectively. The other implicit variables are $n_2 = n_5 + n_6 + m_{\mu+1,\nu}$ and

$$\gamma_1 = \alpha_0 - \alpha_1 - \alpha_{34} - \beta_{1\mu} - \gamma_2 - \Xi = \alpha_{58} + \beta_{\mu+1,\nu} + (\Xi' - \Xi),$$

where Ξ is the same as in (9.31). Also recall from the definition of $\mathcal{N}^2$ that $n_0 \neq n_1$ and $n_5 + n_6 \neq 0$.

Next, let $\max\{\langle n_2\rangle, \langle n_5\rangle, \langle n_6\rangle\} \sim 2^{d'}$ so that $d' \geq h \geq 0.9d$, and fix d' also. In the expression for $\mathcal{S}'$, we may assume

$$\langle m_i\rangle + \langle \beta_i\rangle + \langle \alpha_j\rangle \ll 2^{d'/70} \tag{9.41}$$

for all $\mu + 1 \leq i \leq \nu$ and $j \in \{5, 6, 7, 9\}$ (note we already have this for $1 \leq i \leq \mu$ and $j \in \{0, 1, 3\}$ due to the factor $\mathbf{1}_F$). In fact, if any one of these does not hold, we may bound $|\eta| \lesssim 2^{-(d+h)}\langle\gamma_2\rangle^{-10}$ and $|\Phi^2(\Phi^2)'| \lesssim 2^{d+h}$ (so that the weight is canceled by the part of the η factor), then use (9.33) to bound the α_4 and α_8 factors by $\langle\alpha_4\rangle^{-1}$ and $\langle\alpha_8\rangle^{-1}$ respectively with a loss $2^{O(s^3)d'}$. Then we fix (m_i, β_i) to produce $\mathcal{S}'_{\text{sub}}$, and estimate it by bounding the γ_2 and α_l factors for $l \in \{3, 4, 7, 8\}$ in $l^{1+}L^{1+}$, and bounding the $\mathfrak{N}f$ and $\mathfrak{N}y$ factors in $L^{4+}L^{4+}$ (with $4+$ being $4 + c$). Note that in the whole process we lose at most $2^{O(s)d'}$; but by our assumptions at least one (m_i, β_i) or α_l must be $\gtrsim 2^{d'/70}$, so we will be able to gain some $2^{cd'}$ power from the corresponding factor (again using the room available for $L^{4+}L^{4+}$ Strichartz estimate) to complete the estimate.

Now we may assume all the variables mentioned above are small. This in particular implies that $|\Xi'| \ll 2^{d'/70}$. By Lemma 9.4 (combined with the restrictions made above, such as $h \geq 0.9d$), we can conclude that either (i) $n_0 = n_5$ and $n_1 + n_6 = 0$ (or with 5 and 6 switched); or (ii) no two of $(-n_0, n_1, n_5, n_6)$ add to zero, and $\langle n_l\rangle \gtrsim 2^{0.9d'}$ for $l \in \{0, 1, 5, 6\}$. In case (ii), we have in particular

$$\langle n_1\rangle^s\langle n_5\rangle^s\langle n_6\rangle^s \gtrsim 2^{1.5sd'}\langle n_0\rangle^r, \tag{9.42}$$

thus we may gain a power $2^{csd'}$ from the $\langle n_l\rangle$ weights (after canceling $\Phi^2(\Phi^2)'$ by the η factor) if we use the X_2 bound for y and the bound for f deduced from the X_6' bound for g. Then we simply bound the α_4 and α_8 factors using (9.33) with $2^{O(s^2)d}$ loss, take absolute value of everything, then fix m_i and β_i to produce a term $\mathcal{S}_{\text{sub}}$ that has basically the same form as the left hand side of (9.21), with possibly some additional loss of $2^{O(s^2)d}$, and with the $\min\{T, \langle\alpha_4\rangle^{-1}\}$ factor in (9.21) replaced by $T^{0+}\langle\alpha_4\rangle^{-1+s^4}$, which is due to the estimate

$$\int_{\alpha_3+\alpha_4+\alpha_7+\alpha_8+\gamma_2=\alpha_{10}} \min\{T, \langle\alpha_3\rangle^{-1}\}\cdot\langle\gamma_2\rangle^{-10}\prod_{l\in\{4,7,8\}}\langle\alpha_l\rangle^{-1}\lesssim T^{0+}\langle\alpha_{10}\rangle^{-1+s^4}.$$

We can then repeat the proof of (9.21) to conclude (notice that every variable is now $\lesssim 2^{O(1)d'}$).

Now we consider case (i), so that $d'=d$. We will first replace the $\eta(\gamma_1,\gamma_2)$ factor appearing in the expression of $\mathcal{S}'$ by $\eta(\gamma_1',\gamma_2)$, where $\gamma_1'=\gamma_1-\alpha_8$. Note that γ_1' depends on α_4 only through $\alpha_9=\alpha_4+\alpha_8$. When we estimate the difference caused by this substitution, since we still have the restriction $\mathbf{1}_F$, we will have $\gamma_1\sim 2^{d+h}$, so we will gain a power $2^{2(d+h)-d'/70}$, which is more than enough to cancel $\Phi^2(\Phi^2)'$, thus this part will be acceptable. We also note that the assumption (9.41) allows us to insert another characteristic function which depends on α_4 only through α_9; the presence of this function (as well as the part of $\mathbf{1}_F$ independent of α_4) will allow us to deduce $\langle\gamma_1'\rangle\sim 2^{d+h}$. Therefore, if we *remove* the part in $\mathbf{1}_F$ depending on α_4, the error we create will be a summation-integration of the type $\mathcal{S}'$, but restricted to some set on which we have $|\eta|\lesssim 2^{-(d+h)}\langle\gamma_2\rangle^{-10}$ (note that here we already have $\eta(\gamma_1',\gamma_2)$ instead of $\eta(\gamma_1,\gamma_2)$), *as well as* $\langle\alpha_4\rangle\gtrsim 2^{d'/90}$. Then we will be able to take absolute values, cancel $\Phi^2(\Phi^2)'$ by the η factor, and gain a power $2^{cd'}$ from the assumption about α_4, and proceed exactly as above.

After we have made the above substitutions, the integral with respect to α_4 (or α_8) will be *exactly*

$$\int_{\mathbb{R}}(\chi e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}-\Delta_{n_0})})^{\wedge}(\alpha_4)(\chi e^{\mathrm{i}(\Delta_{n_0}-\Delta_{n_1}-\Delta_{n_2})})^{\wedge}(\alpha_9-\alpha_4)\,d\alpha_4=\widehat{\chi^2}(\alpha_9).$$

Then we will get rid of this integration, then take absolute values, fix (m_i,β_i) (again we ignore the restriction that the m_i must add to zero) to obtain an expression

$$\begin{aligned}\mathcal{S}_{\text{sub}}&\lesssim\sum_{n_0,n_1}\int_{(T)}2^{2h}|f_{n_0,\alpha_0}|\cdot\prod_{l\in\{1,5,6\}}|(y^{\omega_l})_{n_l,\alpha_l}|\\&\qquad\times\prod_{l\in\{3,7\}}\min\{T,1/\langle\alpha_l\rangle\}\cdot 2^{-d-h}\langle\alpha_9\rangle^{-10}\langle\gamma_2\rangle^{-10}\\&\lesssim 2^{-|d-h|}T^{0+}\sum_{n_0,n_1}\|\widehat{f_{n_0}}\|_{L^q}\prod_{l\in\{1,5,6\}}\|\widehat{(y^{\omega_l})_{n_l}}\|_{L^1}.\end{aligned}$$

where the c_j are constants, $n_5=n_0$, $n_6=-n_1$, the summation is restricted to the set

$$\{(n_0,n_1):\max\{\langle n_0\rangle,\langle n_1\rangle\sim 2^d,\ \min\{\langle n_0\rangle,\langle n_1\rangle,\langle n_0-n_1\rangle\}\sim 2^h\},$$

and the (T) integration is taken over the set

$$\{(\alpha_0,\alpha_1,\alpha_3,\alpha_5,\alpha_6,\alpha_7,\alpha_9,\gamma_2):\alpha_0=\alpha_1+\alpha_3+\alpha_{57}+\alpha_9+\gamma_2+c_2\};$$

note that the restriction we make here is enough to guarantee that $|\eta|\lesssim 2^{-(d+h)}\langle\gamma_2\rangle^{-10}$.

Now, if we restrict to $n_0\sim 2^{d''}$ and $n_1\sim 2^{d'''}$, then up to an additive constant they are between h and d, and the restricted $\mathcal{S}_{\mathrm{sub}}$ is bounded by $2^{-|d-h|}T_{d''}$ due to (9.24). We may sum over d and h for fixed d'' and d''' to obtain a bound $2^{-|d''-d'''|}T_{d''}$, then sum over d'' and d''' to conclude.

10. The a priori estimate III: A special term

In this section we prove the following proposition, with which we will be able to close the proof of Proposition 8.5.

Proposition 10.1. *We have*

$$\sum_{j\in\{1,2,5,7\}}\|\mathcal{M}^{3.5}\|_{X_j}\lesssim T^{0+}. \tag{10.1}$$

Proof. Define the functions g, f and f', and fix the scale 2^d as usual. Note in particular that $\|g\|_{X_6'}\le 1$, so that

$$\|\langle n_0\rangle^{-s}\langle\alpha_0\rangle^{1/2-O(s^2)}f'\|_{l^2L^2}\lesssim 1. \tag{10.2}$$

Now, according to a computation similar to those made before (for example, in the proof of Propositions 9.1 and 9.3), we can write the expression $\mathcal{S}$ we need to bound in two ways:

$$\mathcal{S}=\sum_{n_0=n_1+n_2+n_3+m_1+\cdots+m_\mu}\int_{(T)}\Phi^{3.5}\cdot\overline{(f')_{n_0,\alpha_0}} \times((w')^{\omega_1})_{n_1,\alpha_1}\prod_{l=2}^{3}(z^l)_{n_l,\alpha_l}\cdot\phi_{\alpha_4}\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}, \tag{10.3}$$

$$\mathcal{S}=\sum_{n_0=n_1+n_2+n_3+m_1+\cdots+m_\mu}\int_{(T)}\Phi^{3.5}\cdot\overline{f_{n_0,\alpha_0}}((w'')^{\omega_1})_{n_1,\alpha_1} \times\prod_{l=2}^{3}(y^l)_{n_l,\alpha_l}\cdot\phi_{\alpha_4}(\chi e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}+\Delta_{n_3}-\Delta_{n_0})})^\wedge(\alpha_5)\prod_{i=1}^{\mu}\frac{(u''')_{m_i,\beta_i}}{m_i}. \tag{10.4}$$

Here the (T) integration in (10.3) is over the set

$$\{(\alpha_0,\ldots,\alpha_4,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_{14}+\beta_{1\mu}+\Xi\}, \tag{10.5}$$

while the (T) integration in (10.4) is over the set

$$\{(\alpha_0,\ldots,\alpha_5,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_{15}+\beta_{1\mu}+\Xi\}, \tag{10.6}$$

both with the NR factor

$$\Xi = |n_0|n_0 - \sum_{l=1}^{3} |n_l|n_l - \sum_{i=1}^{\mu} |m_i|m_i. \tag{10.7}$$

Also each z^l or $\overline{z^l}$ equals u', v' or w', and y^l or $\overline{y^l}$ equals u'', v'' or w''.

First we treat the case when

$$\min_{0\leq l\leq 3} \langle n_l\rangle \gtrsim 2^{2d/3}. \tag{10.8}$$

In this situation, n_2 and n_3 must have opposite sign (note that here we are again assuming 2^d is large enough). By symmetry, we may assume $n_2 > 0$ and $n_3 < 0$; also note that $n_0 > 0$.

Next, we may assume that $\langle m_i\rangle + \langle\beta_i\rangle + \langle\alpha_4\rangle \ll 2^{d/90}$ for all i, since otherwise we will be able to gain a power 2^{cd} from the corresponding factor alone, and estimate the expression (10.3) by controlling $\mathfrak{N}f'$ in $L^{2+}L^{2+}$, $\mathfrak{N}w'$ in $L^{6+}L^{6+}$, $\mathfrak{N}z^l$ in L^6L^6 and ϕ in $l^{1+}L^{1+}$ with a loss of at most $2^{(s+O(\epsilon))d}$. Notice that the loss from the $\langle n_0\rangle^{-s}$ factor in (10.2) is at most 2^{sd}, while the loss from other places is at most $2^{O(\epsilon)d}$. In the same way, we will also be done if $\langle m_i\rangle \gtrsim 2^{1.2sd}$ for some i, or when $|\Xi| \gtrsim 2^{(1+1.01s)d}$. In fact, in the former case we invoke the X_3 norm for $\langle\partial_x\rangle^{-s^3}u'''$ to gain a power of $2^{(1+c)sd}$ to cancel the 2^{sd} loss, then fix m_j and β_j for $j \neq i$ to produce $\mathcal{S}_{\mathrm{sub}}$, which is estimated by controlling $\mathfrak{N}f'$ in L^4L^4, $\mathfrak{N}w'$, $\mathfrak{N}z^l$ and $\mathfrak{N}u'''$ in L^6L^6, ϕ in some $l^{1+}L^{1+}$ with a loss of at most $2^{O(\epsilon)d}$. In the latter case at least one α_l must be $\gtrsim 2^{(1+1.01s)d}$. If $l \in \{0, 1\}$, we could gain 2^{cd} from the corresponding factor and proceed as above (since the 2^{cd} gain will overwhelm any loss). If $l \in \{2, 3\}$ (say $l = 2$), we invoke the X_4 norm for z^2; noticing that $1-\kappa = s^{5/4}$, we will gain at least $2^{1.001sd}$ from z^2 and estimate the reduced function in l^2L^2. This will cancel the 2^{sd} loss from f' and we can fix all m_i and β_i, then bound $\mathcal{S}_{\mathrm{sub}}$ by controlling $\mathfrak{N}f'$ in $L^{6-}L^{6-}$, $\mathfrak{N}w'$ in $L^{6+}L^{6+}$ (where $6-$ and $6+$ differ from 6 by cs^2 with appropriately chosen c), $\mathfrak{N}z^2$ in L^2L^2, $\mathfrak{N}z^3$ in L^6L^6, ϕ in $l^{1+}L^{1+}$ with a loss of at most $2^{O(\epsilon)d}$.

Now, we have $\langle m_i\rangle \ll 2^{1.2sd}$ and $|\Xi| \ll 2^{(1+1.01s)d}$. Since $n_0, n_2 > 0 > n_3$ and $\langle n_1\rangle \gtrsim 2^{2d/3}$, we can easily see that $n_1 > 0$, which implies

$$|n_0^2 - n_1^2 - n_2^2 + n_3^2| \ll 2^{(1+1.01s)d}. \tag{10.9}$$

Note that $|n_2 + n_3| \gg 2^{d/2}$ (otherwise we gain 2^{cd} from the weight and everything will again be easy; also this will imply $n_0 \neq n_1$), we write $n_2 + n_3 = k$ and $n_0 - n_1 = l$ so that $l - k = O(2^{1.2sd})$. We deduce from (10.9) and elementary algebra that

$$|k| \cdot |n_0 + n_1 - n_2 + n_3| \lesssim 2^{(1+1.2s)d}, \tag{10.10}$$

which implies that $\max\{\langle n_2\rangle, \langle n_3\rangle\} \sim 2^d$. Since we will be done if we gain $2^{(1+c)sd}$ from the weight, we may then assume $|n_2 + n_3| \gtrsim 2^{(1-1.01s)d}$.

Next, we claim that we may assume $|n_0 - n_2| + |n_1 + n_3| \ll 2^{1.9sd}$. In fact, the difference between $n_0 - n_2$ and $n_1 + n_3$ is already $O(2^{1.2sd})$, so if one of them is $\gtrsim 2^{1.9sd}$, the factor $n_0+n_1-n_2+n_3$ in (10.10) will be at least $2^{1.9sd}$ also. This would force k to be $\ll 2^{(1-0.7s)d}$. Noting that $\max\{\langle n_2\rangle,\langle n_3\rangle\}\sim 2^d$, we will gain $2^{0.7sd}$ from the weight $\Phi^{3.5}$. Therefore, we will still be able to close the estimate if we can gain more than $2^{0.3sd}$ elsewhere, for example, when $|\Xi|\gtrsim 2^{(1+0.31s)d}$ or when $\langle m_i\rangle\gtrsim 2^{0.4sd}$ for some i. If we assume further that $|\Xi|\ll 2^{(1+0.31s)d}$ and $\langle m_i\rangle\ll 2^{0.4sd}$, then (10.10) will hold with the right hand side replaced by $2^{(1+0.4s)d}$. This would then force $|k|\lesssim 2^{(1-1.5s)d}$, which is impossible since we have already had $|k|\gtrsim 2^{(1-1.01s)d}$.

Note that all the restrictions made above concern only the n_l, m_i, β_i and α_4 variables, so we still have the freedom of choosing (10.3) or (10.4). After making these restrictions, we will now choose (10.4) and analyze the exponential factor first. Noting that

$$\|(\delta_{n_1}+\delta_{n_2}+\delta_{n_3}-\delta_{n_0})^\wedge\|_{L^1}\lesssim 2^{-d/4} \tag{10.11}$$

by Proposition 8.7, we deduce from Lemma 8.6 that

$$\|\langle\alpha_5\rangle J_{(n)}(\alpha_5)\|_{L^\mu}\lesssim 2^{-d/8} \tag{10.12}$$

for all $1\leq\mu\leq\infty$ with

$$J_{(n)}(\alpha_5)=\big(\chi(t)\cdot(e^{\mathrm{i}(\Delta_{n_1}+\Delta_{n_2}+\Delta_{n_3}-\Delta_{n_0})}-1)\big)^\wedge(\alpha_5). \tag{10.13}$$

By a similar argument to the proof of Proposition 8.9, we deduce that (where, of course, the supremum is taken over (n) such that $\sum_l\langle n_l\rangle\sim 2^d$)

$$\int_{\mathbb{R}}\sup_{n_0,\dots,n_3}|J_{(n)}(\alpha_5)|\,d\alpha_5\lesssim 2^{-d/9}. \tag{10.14}$$

Therefore, if we replace in (10.4) the exponential factor by $J_{(n)}$, we will be able to first fix α_5 and then integrate over it, and gain a power 2^{cd} from this process. Once α_5 is fixed and the $J_{(n)}$ factor is removed with a 2^{cd} gain, we will be in the same situation as considered before. We can then fix m_i and β_i to produce $\mathcal{S}_{\text{sub}}$, and estimate it by controlling $\mathfrak{N}f$ in $L^{2+}L^{2+}$, $\mathfrak{N}w''$ and $\mathfrak{N}y^l$ in L^6L^6, ϕ in $l^{1+}L^{1+}$ with a loss $2^{O(s)d}$.

Now we may replace the exponential factor in (10.4) by $\widehat{\chi}(\alpha_5)$. We can actually get rid of this factor since f and f' is supposed to have compact t support. Therefore, we are reduced to estimating

$$\mathcal{S}=\sum_{n_0=n_1+n_2+n_3+m_1+\cdots+m_\mu}\int_{(T)}|\Phi^{3.5}|\cdot|\overline{f_{n_0,\alpha_0}}| \times|((w'')^{\omega_1})_{n_1,\alpha_1}|\prod_{l=2}^{3}|(y^l)_{n_l,\alpha_l}|\cdot|\phi_{\alpha_4}|\prod_{i=1}^{\mu}\left|\frac{(u''')_{m_i,\beta_i}}{m_i}\right|, \tag{10.15}$$

where the integral (T) is taken over the set (10.5). Starting from this point we will no longer use the equivalence of (10.3) and (10.4), so we will assume here that each $\langle\alpha_l\rangle$ is $\ll 2^{(1+1.01s)d}$, since otherwise we may proceed as above (note that the bounds for f, w'' and y^l are better than those for f', w' and z^l). For the same reason, we may assume $\langle\alpha_0\rangle+\langle\alpha_1\rangle\ll 2^{d/9000}$ (otherwise we may gain 2^{cd}, then control $\mathfrak{N}f'$ in $L^{2+}L^{2+}$, $\mathfrak{N}w''$

in $L^{6-}L^{6-}$, $\mathfrak{N}y^l$ in L^6L^6 and ϕ in $l^{1+}L^{1+}$ with $2+$ and $6-$ being $2+c$ and $6-c$ respectively).

To estimate (10.15), we recall the bound (9.24) in the proof of Proposition 9.3. Suppose that $n_0 \sim n_2 \sim 2^{d'}$ and $n_1 \sim n_3 \sim 2^{d''}$ (note that $|n_0 - n_2|$ and $|n_1 + n_3|$ are small). Then we have $\max\{d', d''\} = d + O(1)$, as well as

$$|\Phi^{3.5}| \lesssim 2^{-|d'-d''|}. \tag{10.16}$$

Here, instead of fixing d, we will fix all of d, d', d'', then sum over d' and d''. By (10.16), we may assume

$$\min\{d', d''\} \geq (1 - 1.01s)d,$$

so in particular $d' \sim d'' \sim d$. Once we fix $\langle n_3 \rangle \sim 2^{d''}$, we can invoke the X_8 norm of y^3 to write y^3 (now restricted to frequency $\sim 2^{d''}$) as a sum

$$y^3 = \sum_j \gamma_j \pi_{k_j} y^{(j)}, \quad \sum_j \langle k_j \rangle^{s^{1/2}} |\gamma_j| \lesssim 1, \tag{10.17}$$

such that $\|y^{(j)}\|_{L^q l^2} \lesssim 1$ for each j. See Section 2.2. We only need to consider a single j; namely we need to bound $\mathcal{S}$ provided $y^3 = \langle k \rangle^{-s^{1/2}} \pi_k y''$, where y'' is some function satisfying $\|y''\|_{L^q l^2} \lesssim 1$. Next, if $\langle k \rangle \gtrsim 2^{d/90}$, we will gain a power $2^{cs^{1/2}d}$ from the coefficient in y^3; we then fix m_i and β_i. To estimate the resulting $\mathcal{S}_{\text{sub}}$, we can control $\mathfrak{N}f$ in $L^{6+}L^{6+}$, $\mathfrak{N}w''$ and $\mathfrak{N}y^2$ in L^6L^6 and ϕ in $l^{1+}L^{1+}$ with a loss of at most $2^{O(s)d}$ (where $6+$ is $6+cs$ and $1+$ is defined accordingly; also note that we have assumed $\langle \alpha_0 \rangle \lesssim 2^{2d}$, as well as the bound for f deduced from the X_6' bound for g). We can then close the estimate if we can control y'' (and hence $\pi_k y''$) in l^2L^2. This can be achieved by inserting a $\chi(t)$ factor to every term in (10.17), which, while doing nothing to the equality and the $L^q l^2$ norms, allows us to control the $L^2 l^2$ norm by the $L^q l^2$ norm. Thus here we also get the desired estimate.

We now assume $\langle k \rangle \ll 2^{d/90}$. We will fix k and each (m_i, β_i) to obtain some constants $K_1 \ll 2^{1.2sd}$ and $K_2 \ll 2^{d/90}$, and produce

$$2^{-|d'-d''|}\mathcal{S}_{\text{sub}} = 2^{-|d'-d''|} \sum_{n_0 = n_1 + n_2 + n_3 + K_1} \int_{(T)} |\overline{f_{n_0,\alpha_0}}| \times |((w'')^{\omega_1})_{n_1,\alpha_1}| \cdot |(y^2)_{n_2,\alpha_2}| \cdot |(\pi_k y'')_{n_3,\alpha_3}| \cdot |\phi_{\alpha_4}|,$$

where the (T) integral is taken over the set

$$\Big\{(\alpha_0, \ldots, \alpha_4) : \alpha_0 = \alpha_{14} + |n_0|n_0 - \sum_{l=1}^{3} |n_l|n_l + K_2\Big\}, \tag{10.18}$$

with all the restrictions made above taking effect. Now if $n_0 - n_2 \in \{0, K_1 - k\}$, we can bound

$$\begin{aligned}\mathcal{S}_{\text{sub}} &\lesssim T^{0+} \sum_{n_0 \sim 2^{d'}} \sum_{n_1 \sim 2^{d''}} \|\widehat{f_{n_0}}\|_{L^q} \|(((w'')^{\omega_1})_{n_1})^\wedge\|_{L^1} \|((y^2)_{n_0+c_0})^\wedge\|_{L^1} \|((y'')_{n_1+c_1})^\wedge\|_{L^q} \\ &\lesssim T^{0+} T_{d'} \|\langle n_2 \rangle^r y^2\|_{l^p_{n_2 \sim 2^{d'}} L^1} \cdot \|(w'')^{\omega_1}\|_{l^2_{n_1 \sim 2^{d''}} L^1} \|y''\|_{l^2_{n_1 \sim 2^{d''}} L^q} \lesssim T^{0+} T_{d'},\end{aligned}$$

using the bound (9.24) for f, the X_2 bound for w'' and y^2, and the $L^q l^2$ bound for y'', where the c_j are constants, small compared to $2^{d'}$ and $2^{d''}$, such that $n_j \sim n_0 + c_j$ for $j \in \{0, 1\}$. If we then sum and integrate over m_i and β_i, then multiply by $2^{-|d'-d''|}$ and sum over d' and d'', we will get a quantity bounded by T^{0+}.

Assume $n_0 - n_2 \notin \{0, K_1 - k\}$. Let $\lambda = n_0 - n_2$. We can rewrite the expression for $\mathcal{S}_{\text{sub}}$ as

$$\mathcal{S}_{\text{sub}} \lesssim T^{0+} 2^{-0.999sd+O(s)|d'-d''|} \sum_{n_0,n_1,\lambda} \int_{\mathbb{R}^4} A_{n_0,\alpha_0} B_{n_1,\alpha_1} C_{n_0-\lambda,\alpha_2} \times D_{\lambda-n_1+c_1,\alpha_3'} \langle \alpha_0 - \alpha_1 - \alpha_2 - \alpha_3' - \Xi' + c_2 \rangle^{-1-s^2} \prod_{l=0}^{2} d\alpha_l \cdot d\alpha_3',$$

where the $c_j \ll 2^{d/90}$ are constants, and the summation-integration is restricted to the subset where all the restrictions made above are satisfied by $(n_0, \dots, n_3, \alpha_0, \dots, \alpha_4)$ which is defined in terms of our new variables (as well as the intermediate variable n_3') by

$$n_2 = n_0 - \lambda, \quad n_3 = n_3' - k, \quad \alpha_3 = \alpha_3' - |n_3'|n_3' + |n_3|n_3,$$
$$n_3' = \lambda - n_1 - K_1 + k, \quad \alpha_4 = \alpha_0 - \alpha_{13} - |n_0|n_0 + \sum_{l=1}^{3} |n_l|n_l - K_2.$$

We can check from the assumptions made above that no two of $(-n_0, n_1, n_2, n_3')$ add to zero. Moreover, Ξ' is defined by

$$\Xi' = \Xi'(n_0, n_1, \lambda) = |n_0|n_0 - |n_1|n_1 - |n_2|n_2 - |n_3'|n_3',$$

and the relevant functions are defined by

$$A_{n_0,\alpha_0} = \langle n_0 \rangle^{-r} |f_{n_0,\alpha_0}|, \qquad B_{n_1,\alpha_1} = \langle n_1 \rangle^{r} |((w'')^{\omega_1})_{n_1,\alpha_1}|,$$
$$C_{n_2,\alpha_2} = \langle n_2 \rangle^{r} |(y^2)_{n_2,\alpha_2}|, \qquad D_{n_3',\alpha_3'} = |(y'')_{n_3',\alpha_3'}|.$$

When restricted to appropriate subsets (for example, we must have $n_0 \sim n_2 \sim 2^{d'}$ and $n_1 \sim n_3 \sim 2^{d''}$), these functions will satisfy

$$\|A\|_{L^1 l^{p'}} + \|B\|_{L^1 l^p} + \|C\|_{l^p L^1} + \|D\|_{L^1 l^2} \lesssim 2^{0.0002sd}. \tag{10.19}$$

In fact, due to the restrictions we made, we can bound all the variables by $2^{O(1)d}$; so when we replace the L^q norm by the L^1 norm we lose (by Hölder) at most $2^{O(q-1)d}$. Thus the bound for A follows from (9.24), and the bound for D follows from our assumption about y''. The bound for C follows from the X_2 bound for y^2, while for B we simply estimate (note that $\langle \alpha_1 \rangle \ll 2^{d/9000}$)

$$\|B\|_{L^1 l^p} \lesssim 2^{O(1-q)d} \|\langle n_1 \rangle^r \langle \alpha_1 \rangle^b (w'')^{\omega_1}\|_{L^2 l^p} \lesssim 2^{O(1-q)d} 2^{sd/8000} \|\langle n_1 \rangle^s \langle \alpha_1 \rangle^b (w'')^{\omega_1}\|_{L^p l^p} \lesssim 2^{0.0002sd}, \tag{10.20}$$

using the X_1 bound for w''. Note that by inserting $\chi(t)$ to w'', we may control the $l^p L^p$ norm by the $l^p L^2$ norm.

Now we need to estimate $\mathcal{S}_{\text{sub}}$ under the assumption of (10.19). First replace the bounds in (10.19) by 1, so that we only need to bound the summation-integration part of $\mathcal{S}_{\text{sub}}$ by $2^{0.998sd}$. Fix α_0 and α_1 which are $\ll 2^{0.02d}$ (then integrate over them), we may assume A and B are functions of n_0 and n_1 only, and are bounded in $l^{p'}$ and l^p respectively. We then bound (with the $c_j \ll 2^{d/90}$ being constants)

$$\begin{aligned}\mathcal{S}'_{\text{sub}} &= \int_{\mathbb{R}^2} d\alpha_2\, d\alpha'_3 \cdot \sum_{n_0,n_1,\lambda} A_{n_0} B_{n_1} C_{n_0-\lambda,\alpha_2} D_{\lambda-n_1+c_1,\alpha'_3} \langle \alpha_2+\alpha'_3+\Xi'+c_3\rangle^{-1-s^2}\\ &\lesssim \sum_{\rho\in\mathbb{Z}} \langle\rho\rangle^{-1-s^2} \int_{\mathbb{R}^2} d\alpha_2\, d\alpha'_3 \sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} A_{n_0} B_{n_1} C_{n_0-\lambda,\alpha_2} D_{\lambda-n_1+c_1,\alpha'_3}\\ &\lesssim \sup_{\rho} \int_{\mathbb{R}^2} d\alpha_2\, d\alpha'_3 \cdot \Big(\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} A_{n_0}^2 C_{n_0-\lambda,\alpha_2}^2\Big)^{1/2}\\ &\qquad \times \Big(\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} B_{n_1}^4\Big)^{1/4}\Big(\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} D^4_{\lambda-n_1+c_1,\alpha'_3}\Big)^{1/4},\end{aligned}\tag{10.21}$$

where we write $\alpha_2+\alpha'_3+\Xi'+c_3=\Xi''$ for simplicity. Now for any positive function E_{n_1} of n_1, when ρ and α_2, α'_3 are fixed, we may bound

$$\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} E_{n_1} \lesssim \sum_{n_1} E_{n_1} \sum_{(n_0,\lambda):\, \Xi'=c''} 1 \lesssim 2^{O(s^4)d}\|E\|_{l^1}$$

(where c' and c'' are constants depending on α_2, α'_3 and ρ), thanks to part (iii) of Lemma 9.4 (or actually, an argument similar to the proof of that part). The same inequality holds if we replace n_1 by $\lambda - n_1 + c_1$ (which equals n'_3 plus a constant). Therefore we can bound the second factor in (10.21) by $2^{O(s^4)d}\|B\|_{l^4} \lesssim 2^{O(s^4)d}$, and the third factor by $2^{O(s^4)d}\|D_{\cdot,\alpha'_3}\|_{l^4}$. Ignoring the $2^{O(s^4)d}$ factors, we thus bound (10.21) by

$$\begin{aligned}\mathcal{S}'_{\text{sub}} &\lesssim \sup_{\rho} \int_{\mathbb{R}^2} d\alpha_2\, d\alpha'_3 \cdot \|D_{\cdot,\alpha'_3}\|_{l^4} \cdot \Big(\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} A_{n_0}^2 C_{n_0-\lambda,\alpha_2}^2\Big)^{1/2}\\ &\lesssim \sup_{\rho,\alpha'_3} \int_{\mathbb{R}} \Big(\sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} A_{n_0}^2 C_{n_0-\lambda,\alpha_2}^2\Big)^{1/2} d\alpha_2\\ &\lesssim \sup_{\rho,\alpha'_3} \int_{\mathbb{R}} d\alpha_2 \cdot \sum_{(n_0,n_1,\lambda):\, \lfloor\Xi''\rfloor=\rho} A_{n_0} C_{n_0-\lambda,\alpha_2}.\end{aligned}$$

Now we fix ρ and α'_3. Noticing $n_0-\lambda=n_2$, and that

$$A_{n_0} \le F_{n_2} := \Big(\sum_{|m-n_2|\lesssim 2^{1.9sd}} A_m^{p'}\Big)^{1/p'}\tag{10.22}$$

because $|n_0-n_2| \lesssim 2^{1.9sd}$, we proceed to estimate

$$\begin{aligned}\mathcal{S}''_{\text{sub}} &\lesssim \int_{\mathbb{R}} d\alpha_2 \cdot \sum_{(n_1,n_2,\lambda):\, \lfloor\Xi''\rfloor=\rho} F_{n_2} C_{n_2,\alpha_2} \lesssim 2^{O(s^4)d} \sum_{n_2} F_{n_2} \int_{\mathbb{R}} C_{n_2,\alpha_2}\, d\alpha_2\\ &\lesssim 2^{O(s^4)d}\|F\|_{l^{p'}}\|C\|_{l^p L^1}.\end{aligned}$$

Here we have again used the divisor estimate as above. Finally, notice that

$$\|F\|_{l^{p'}}^{p'} = \sum_{n_2} \sum_{|m-n_2|\lesssim 2^{1.9sd}} A_m^{p'} \lesssim 2^{1.9sd} \|A\|_{l^{p'}}^{p'},$$

so we deduce that $\mathcal{S}''_{\text{sub}} \lesssim 2^{1.91sd/p'} \lesssim 2^{0.998sd}$, as desired.

It remains to consider the case where $\langle n_l \rangle \ll 2^{2d/3}$ for some l. Note that if $\langle m_i \rangle + \langle \beta_i \rangle + \langle \alpha_4 \rangle \gtrsim 2^{d/90}$ for some i, or the weight $|\Phi^{3.5}|$ is $\lesssim 2^{-cd}$ or the NR factor (as defined in (10.7)) satisfies $|\Xi| \gtrsim 2^{(1+c)d}$, we will be done using the same arguments as before. This in particular includes the cases when (i) three of the n_l are $\gtrsim 2^{3d/4}$ and the remaining one is $\ll 2^{2d/3}$; (ii) at least two of the n_l are $\ll 2^{3d/4}$, and $\langle n_2 \rangle + \langle n_3 \rangle \gtrsim 2^{4d/5}$; (iii) both n_2 and n_3 are $\ll 2^{4d/5}$, and $n_0 n_1 < 0$.

Now we assume that $n_0 n_1 > 0$ and $\langle n_2 \rangle + \langle n_3 \rangle \ll 2^{4d/5}$. Let $n_0 - n_1 = k$ and $n_2 + n_3 = l$, so that $|k - l| \ll 2^{d/90}$. If $l \lesssim 2^{d/80}$, we must have $\langle n_2 \rangle + \langle n_3 \rangle \ll 2^{d/70}$ (or we gain from the Φ factor). These two variables being small means that we will be able to repeat the argument made before and gain 2^{cd} even if $|\Xi|$ is bounded below by $2^{0.99d}$ instead of $2^{(1+c)d}$. But when $|\Xi| \ll 2^{0.99d}$, it is clear that we must have $k = 0$. If $l \gg 2^{d/80}$, we will have $k \sim l$, so that

$$\big||n_0|n_0 - |n_1|n_1\big| \gtrsim 2^d |k| \gg 2^{4d/5}|l| \gg \big||n_2|n_2 + |n_3|n_3\big|,$$

which implies $|\Xi| \gtrsim 2^{81d/80}$, contradicting our assumptions. Thus in any case we deduce that $n_0 = n_1 \sim 2^d$. Now we may use the expression (10.3) for $\mathcal{S}$, but with f' and w' replaced with f and w'' respectively (see the proof of Proposition 9.3; note that we have made no restrictions for α_0 or α_1).

Next, suppose $\langle n_2 \rangle + \langle n_3 \rangle \sim 2^{d'}$; we may assume $d' < d/10$, otherwise we will gain a power 2^{cd} from the weight $\Phi^{3.5}$ (note that $n_2 + n_3$ equals a linear combination of the m variables since $n_0 = n_1$). We will fix d and d' (then sum over them). If $\langle m_i \rangle \ll 2^{d'/2}$ for all i, then we gain a power $2^{cd'}$ from the weight $\Phi^{3.5}$; otherwise we have $\langle m_i \rangle \gtrsim 2^{d'/2}$ for some i, so we may extract a power $2^{cd'}$ from the $1/m_i$ factor (without affecting summability in m_i). In any case, we will be able to fix m_i and β_i and sum over them later, and the $\mathcal{S}_{\text{sub}}$ term can be bounded by

$$\begin{aligned}\mathcal{S}_{\text{sub}} &\lesssim 2^{-cd'} \sum_{n_0,n_2} \int_{(T)} |f_{n_0,\alpha_0}| \cdot |((w'')^{\omega_1})_{n_0,\alpha_1}| \cdot |(z^2)_{n_2,\alpha_2}| \cdot |(z^3)_{c_1-n_2,\alpha_3}| \cdot \min\left\{T, \frac{1}{\langle \alpha_4 \rangle}\right\} \\ &\lesssim 2^{-cd'} T^{0+} \sum_{n_0,n_2} \|\widehat{f_{n_0}}\|_{L^q} \|(((w'')^{\omega_1})_{n_0})^\wedge\|_{L^1} \prod_{l=2}^{3} \|\langle n_l \rangle^{-c} \widehat{(z^l)_{n_l}}\|_{L^1} \\ &\lesssim 2^{-cd'} T^{0+} \cdot 2^{rd} T_d \cdot 2^{-rd} \lesssim 2^{-cd'} T^{0+} T_d.\end{aligned}$$

using the bound (9.24) for f and the X_2 bound for w'', where the c_j are constants, $n_3 = c_1 - n_2$, and the (T) integral is over the set

$$\{(\alpha_0, \ldots, \alpha_4) : \alpha_0 = \alpha_{14} - |n_2|n_2 - |c_1 - n_2|(c_1 - n_2) + c_2\}.$$

Now we can (sum over m_i and integrate over β_i and then) sum over d and d' to conclude that $\mathcal{S}$ is bounded by T^{0+}. This proves Proposition 10.1. □

11. The a priori estimate IV: The remaining estimates

In this section we will construct appropriate extensions of u^*, v^* and u so that the improved versions of (8.9) and (8.10) hold. Note that we have already constructed a function, denoted by $w^{(4)}$, that coincides with w^* on $[-T, T]$, and satisfies $\|w^{(4)}\|_{Y_1} \leq C_0 e^{C_0 A}$. We will fix this function in later discussions. In particular, we may (starting from this point) redefine the δ_n and Δ_n factors as in (7.27) and (7.28) by replacing w^* with $w^{(4)}$ (instead of w'') and u with u'''.

11.1. The extension of u

Fix a scale K so that $K = C_{1.5}e^{C_{1.5}A}$ where $C_{1.5}$ is large enough depending on C_1, and the C_2 defined before is large enough depending on $C_{1.5}$. In order to construct a function $u^{(5)}$ that coincides with u on $[-T, T]$ and satisfies

$$\|\langle\partial_x\rangle^{-s^3}u^{(5)}\|_{X_2} + \|\langle\partial_x\rangle^{-s^3}u^{(5)}\|_{X_3} + \|\langle\partial_x\rangle^{-s^3}u^{(5)}\|_{X_4} \leq C_0 A, \tag{11.1}$$

we only need to construct $\mathbb{P}_{>K}u^{(5)}$ and $\mathbb{P}_{\leq K}u^{(5)}$ separately.

To construct $\mathbb{P}_{>K}u^{(5)}$, simply note that u'' coincides with u^* on $[-T, T]$, and we have $\|u''\|_{Y_2} \leq C_1 e^{C_1 A}$; thus if we define $(u^{(5)})_n = e^{i\Delta_n}(u'')_n$, where Δ_n is redefined as above, then $\mathbb{P}_{>K}u^{(5)}$ will equal $\mathbb{P}_{>K}u$ on $[-T, T]$, and we have

$$\|\langle\partial_x\rangle^{-s^4}u^{(5)}\|_{X_j} \lesssim O_{C_1}(1)e^{C_0 C_1 A} \tag{11.2}$$

for $j \in \{2, 3, 4\}$, thanks to Proposition 8.9. Here note that the s^3 exponent in that proposition can actually be replaced by s^4 (which is clear from the proof), and the current (δ_n, Δ_n) also satisfies Proposition 8.7 (in the same way as the (δ_n, Δ_n) defined in Section 8 does). Since we are restricting to high frequencies, the inequality (11.2) will easily imply

$$\|\langle\partial_x\rangle^{-s^3}\mathbb{P}_{>K}u^{(5)}\|_{X_j} \leq A$$

for $j \in \{2, 3, 4\}$, which is what we need for $\mathbb{P}_{>K}u^{(5)}$.

Now let us construct $\mathbb{P}_{\leq K}u^{(5)}$. Recalling that the function u satisfies the equation (1.6), and the Y_2 norm of $\chi(t)e^{-tH\partial_{xx}}u(0)$ is clearly bounded by $C_0 A$, we only need to prove

$$\left\| \int_0^t e^{-(t-t')H\partial_{xx}}\mathbb{P}_{\neq 0}((S_N u(t'))^2)\,dt' \right\|_{(X^{-1/s,\kappa})^T} \lesssim T^{0+}, \tag{11.3}$$

with the implicit constants bounded by $O_{C_{1.5}}(1)e^{C_0 C_{1.5}A}$, where $X^{\sigma,\beta}$ is the standard space normed by $\|\langle n\rangle^\sigma\langle\xi\rangle^\beta\cdot\|_{l^2L^2}$. Define the function $u^{(7)}$ by (7.1) and (7.2), with the u appearing on the right hand side replaced by u''', and v replaced by $\mathbb{P}_{\leq 0}v''' + w'''$ with v''' defined by $(v''')_n = e^{i\Delta_n}(v'')_n$ and w''' similarly, so that $u^{(7)}$ coincides with u on $[-T, T]$ (note that the Δ_n here is different from the Δ_n defined in Section 8; later we will further modify the definition of Δ_n, and this will be clearly stated at that time). We claim that

$$\left\|\mathcal{E}\big(\mathbf{1}_{[-T,T]}\mathbb{P}_{\neq 0}((S_N u^{(7)})^2)\big)\right\|_{X^{-1/s,\kappa}} \lesssim T^{0+}. \tag{11.4}$$

This implies (11.3), since the two functions on the left hand side of (11.3) and (11.4) coincide on $[-T, T]$.

Let $\mathcal{N} = \mathbb{P}_{\neq 0}((S_N u^{(7)})^2)$, we will have

$$\mathcal{N}_{n_0} = \sum_{(\omega_1,\omega_2)\in\{-1,1\}^2} \sum_{\mu_1,\mu_2} \frac{\omega_1^{\mu_1}\omega_2^{\mu_2}}{2^{\mu_{12}}\mu_1!\mu_2!} \times \sum_{n_1+n_2+m_1+\cdots+m_{\mu_{12}}=n_0} \Psi \cdot \prod_{l=1}^{2}(z^{\omega_l})_{n_l} \prod_{i=1}^{\mu_{12}} \frac{(u''')_{m_i}}{m_i}, \qquad (11.5)$$

where $z = \mathbb{P}_{\leq 0}v''' + w'''$, Ψ is the product of some ψ factors and two characteristic functions $\mathbf{1}_{E_1}\mathbf{1}_{E_2}$, where

$$E_1 = \{\omega_1(n_1 + m_{1\mu_1}) > 0\}, \qquad E_2 = \{\omega_2(n_2 + m_{\mu_1+1,\mu_{12}}) > 0\}.$$

Now, by the same argument as in the proof of Proposition 7.1 (note that $n_0 \neq 0$), we can rewrite the right hand side of (11.5) as a sum of the same form, but either with Ψ bounded by 1 and $n_1 + n_2 \neq 0$, or with Ψ bounded by $\frac{\langle m_i\rangle+\langle n_0\rangle}{\langle n_1\rangle+\langle m_i\rangle+\langle n_0\rangle}$ for some i.

To prove (11.4), we will use the function g and f as in the previous sections, and fix the scale d as before; we are then reduced to estimating (with $\mu = \mu_{12}$)

$$\mathcal{S} = \sum_{n_0=n_1+n_2+m_1+\cdots+m_\mu} \int_{(T)} \overline{f_{n_0,\alpha_0}}\,\phi_{\alpha_3} \prod_{l=1}^{2}(z^{\omega_l})_{n_l,\alpha_l} \cdot \prod_{i=1}^{\mu} \frac{(u''')_{m_i,\beta_i}}{m_i},$$

where ϕ is the Fourier transform of $\mathbf{1}_{[-T,T]}$ and the (T) integration is taken over the set

$$\{(\alpha_0,\ldots,\alpha_3,\beta_1,\ldots,\beta_\mu) : \alpha_0 = \alpha_{13} + \beta_{1\mu} + \Xi\},$$

with the NR factor

$$\Xi = |n_0|n_0 - |n_1|n_1 - |n_2|n_2 - \sum_{i=1}^{\mu}|m_i|m_i.$$

We may assume that $\langle n_0\rangle$ and $\langle m_i\rangle$ are all $\ll 2^{d/90}$; otherwise, since we can gain some small power of $\langle m_i\rangle$ and any large power of $\langle n_0\rangle$ (because of the $-1/s$ index), we will be able to gain some power 2^{cd}. Then we simply fix (m_i, β_i) to produce $\mathcal{S}_{\text{sub}}$, then bound f in L^2L^2, $\Re z$ in L^6L^6 and ϕ_{α_3} in $l^{1+}L^{1+}$ with $2^{O(s)d}$ loss to conclude. Now, since n_0 and all m_i are small, we have either $n_1 + n_2 \neq 0$ (which implies $|\Xi| \gtrsim 2^d$) or $|\Psi| \lesssim 2^{-cd}$ (so we can proceed as above). In this case at least one of the α or β variables must be $\gtrsim 2^d$; since we will also have $\omega_l n_l > 0$ and hence $z = w'''$ which is bounded in Y_1 by $C_1 e^{C_1 A}$, we will always gain a power of at least $2^{c(1-\kappa)d}$ from the corresponding factor, then proceed as before to estimate $\mathcal{S}_{\text{sub}}$ and then $\mathcal{S}$, with a loss of at most $2^{O(\epsilon)d}$. Finally, noting that we always gain a power T^{0+} which overwhelms any loss $O_{C_{1.5}}(1)e^{O_{C_{1.5}}(1)A}$, we have already proved (11.4).

Next, noting that $(u^*)_n = e^{-\mathrm{i}\Delta_n}u_n$ on the interval $[-T, T]$, we have

$$(\partial_t + H\partial_{xx})(u^*)_n = e^{-\mathrm{i}\Delta_n}(\partial_t + H\partial_{xx})u_n - \mathrm{i}e^{-\mathrm{i}\Delta_n}(\delta_n u_n).$$

The first term on the right hand side can be bounded in $X^{-2/s,\kappa-1}$ using Proposition 8.9 and what we proved above, while the second term is easily bounded in the stronger space $X^{-10,0}$, by $O_{C_{1.5}}(1)e^{O_{C_{1.5}}(1)A}$. Therefore by the same argument, we can construct an extension of $\mathbb{P}_{\leq K}u^*$ that satisfies (8.9).

11.2. The extensions of u^* and v^*

Now, in order to construct appropriate extensions of $\mathbb{P}_{>K}u^*$ and v^*, we need the following

Proposition 11.1. *Let δ_n and Δ_n be redefined using (7.27) and (7.28). This time with w^* replaced by $w^{(4)}$ and u replaced by $u^{(5)}$. Then the new factors will satisfy Proposition 8.7 with the constants being $C_0e^{C_0A}$ instead of $O_{C_1}(1)e^{C_0C_1A}$.*

Now suppose h, k and h', k' are four functions, supported in $|t|\lesssim 1$, that are related by $(h')_n = e^{\mathrm{i}\Delta_n}h_n$ and $(k')_n = e^{\mathrm{i}\Delta_n}k_n$. Assume that

$$(h')_{n_0} = \sum_{\mu} C_\mu \sum_{n_0=n_1+m_1+\cdots+m_\mu} \Psi\cdot(k')_{n_1}\prod_{i=1}^{\mu}\frac{(u^{(5)})_{m_i}}{m_i} \tag{11.6}$$

with Ψ bounded. Then $\|h\|_{Y_2}\lesssim C_0e^{C_0A}\|k\|_{Y_2}$. Moreover, if Ψ is nonzero only when $\langle m_i\rangle\gtrsim K$ for some i (again, the constant here may involve polynomial factors of μ), then $\|h\|_{Y_2}\lesssim K^{0-}\|k\|_{Y_2}$.

Proof. The estimates of δ_n and Δ_n are proved in the same way as in Proposition 8.7; notice that all the relevant norms bounded by $O_{C_1}(1)e^{C_0C_1A}$ there are now bounded by $C_0e^{C_0A}$ in this updated version, thanks to the construction of $w^{(4)}$ in previous sections and the construction of $u^{(5)}$ above.

Now we need to bound $\|h\|_{X_j}$ for $j\in\{2,3,4,8\}$. By fixing and then summing over μ, we may assume that

$$|h_{n_0,\alpha_0}|\leq C_0\sum_{n_0=n_1+m_1+\cdots+m_\mu}\int_{(T)}|k_{n_1,\alpha_1}|\cdot|(\chi e^{\mathrm{i}(\Delta_{n_1}-\Delta_{n_0})})^\wedge(\alpha_2)|\prod_{i=1}^{\mu}\left|\frac{(u^{(5)})_{m_i,\beta_i}}{m_i}\right|,$$

where the integration is taken over the set

$$\{(\alpha_1,\alpha_2,\beta_1,\ldots,\beta_\mu):\alpha_0=\alpha_1+\alpha_2+\beta_{1\mu}+\Xi\},$$

and the NR factor is

$$\Xi = |n_0|n_0-|n_1|n_1-\sum_{i=1}^{\mu}|m_i|m_i.$$

Throughout the proof we will only use the $X_{j'}$ norm for $\langle\partial_x\rangle^{-s^3}u^{(5)}$ for $j'\in\{2,3,4\}$, and it is important to notice that these norms are bounded by C_0A instead of C_1A.

First assume $j=4$. We introduce the function g with $\|g\|_{X_4'}\lesssim 1$, so that we only need to estimate $\mathcal{S}:=(g,h)$. This is a summation-integration we have seen many times before; to analyze it, we notice that either $\langle\Xi\rangle$, or one of $\langle\alpha_l\rangle$ (where $l\in\{1,2\}$) or $\langle\beta_i\rangle$, must be $\gtrsim\langle\alpha_0\rangle$.

Suppose $\langle\alpha_0\rangle\lesssim\langle\Xi\rangle$. Let the maximum of $\langle n_0\rangle$, $\langle n_1\rangle$ and all $\langle m_i\rangle$ be $\sim 2^d$ (and we fix d); then $\langle\alpha_0\rangle\lesssim 2^{2d}$. If among the variables n_0 and m_i, at least two are $\gtrsim 2^{(1-s^2)d}$, then we will gain a net power $2^{c(1-\kappa)d}$ from the weights in the X_4' bound for g, or from the $|m_i|^{-1}$ weights appearing in $\mathcal{S}$. Then we will be able to bound the $(\chi e^{\mathrm{i}(\Delta_{n_1}-\Delta_{n_0})})^\wedge(\alpha_2)$ factor using some inequality similar to (9.33), fix the irrelevant (m_j,β_j) variables to produce $\mathcal{S}_{\mathrm{sub}}$, then estimate it by bounding $\mathfrak{N}g$ in $L^{2+}L^{2+}$, $\mathfrak{N}k$ and the two $\mathfrak{N}u^{(5)}$ factors in L^6L^6 and the $(\chi e^{\mathrm{i}(\Delta_{n_1}-\Delta_{n_0})})^\wedge(\alpha_2)$ factor in $l^{1+}L^{1+}$, where $2+$ is some $2+cs^2$, with a further loss of at most $2^{O(\epsilon)d}$. We then sum over the (m_j, β_j) variables and sum over d to conclude the estimate for $\mathcal{S}$. If instead only one of them can be $\gtrsim 2^{(1-s^2)d}$ (again, assume d is large enough), then this variable and n_1 must both be $\sim 2^d$. Let the maximum of all the remaining variables be $\sim 2^{d'}$ where $d' \leq (1-s^2)d$ is also fixed; then we will have $|\alpha_0| \lesssim 2^{d+d'}$. Since we will be able to gain a power $2^{c(1-\kappa)(d+d')}$ from the weights, we can proceed in the same way as above.

Next, suppose $\langle\alpha_0\rangle \lesssim \langle\alpha_2\rangle$. By invoking (8.17) we may get an estimate better than (9.33) for the α_2 factor, namely

$$\|\langle\alpha_2\rangle(\chi e^{\mathrm{i}(\Delta_{n_1}-\Delta_{n_0})})^\wedge(\alpha_2)\|_{L^\sigma} \lesssim C_0 e^{C_0 A}\sum_{i=1}^{\mu}\langle m_i\rangle^{s^5} \tag{11.7}$$

for all $1 \leq \sigma \leq \infty$; the $\lesssim$ here allows for a polynomial factor in μ. Therefore, by losing a tiny power of some m_i, we may cancel the α_0 weight in the X_4' bound for g and still bound the α_2 factor in L^2, then fix (m_i, β_i) and produce $\mathcal{S}_{\mathrm{sub}}$, and estimate it by

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} &\lesssim \sum_{n_0}\langle n_0\rangle^{-1}\|\langle n_0\rangle\langle\alpha_0\rangle^{-\kappa}g_{n_0,\alpha_0}\|_{L^2_{\alpha_0}}\|k_{n_0+c_1,\alpha_1}\|_{L^1_{\alpha_1}}\\ &\lesssim \|\langle n_0\rangle^{-1}\langle n_0\rangle\langle\alpha_0\rangle^{-\kappa}g_{n_0,\alpha_0}\|_{l^{3/2}L^2}\cdot\|k_{n_1,\alpha_1}\|_{l^3L^1} \lesssim 1,\end{aligned}$$

where the c_j are constants. If instead $\langle\alpha_0\rangle \lesssim \langle\alpha_1\rangle$, we can invoke the α_1 weight in the X_4 norm for k to cancel the α_0 weight, then notice that $\langle n_1\rangle \lesssim \langle n_0\rangle + \langle m_i\rangle$ for some i, then bound the α_2 factor in L^1 and fix all the other (m_j, β_j) to produce $\mathcal{S}_{\mathrm{sub}}$. If $\langle n_1\rangle \lesssim \langle m_i\rangle$ we will estimate

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} \lesssim &\sum_{n_0=n_1+m_i+c_1}\frac{\langle n_1\rangle}{\langle n_0\rangle\langle m_i\rangle}\|\langle n_0\rangle\langle\alpha_0\rangle^{-\kappa}g_{n_0,\alpha_0}\|_{L^2_{\alpha_0}}\\ &\qquad\times\|\langle n_1\rangle^{-1}\langle\alpha_1\rangle^{\kappa}k_{n_1,\alpha_1}\|_{L^2_{\alpha_1}}\|(u^{(5)})_{m_i,\beta_i}\|_{L^1_{\beta_i}}\\ \lesssim &\|\langle\alpha_0\rangle^{-\kappa}g\|_{l^1L^2}\|\langle n_1\rangle^{-1}\langle\alpha_1\rangle^{\kappa}k\|_{l^\gamma L^2}\cdot\|u^{(5)}\|_{l^{\gamma'}L^1} \lesssim 1,\end{aligned}$$

where the c_j are constants; note that $\|u^{(5)}\|_{l^{\gamma'}L^1}$ can be controlled by the X_2 norm of $\langle\partial_x\rangle^{-4s^3}u^{(5)}$ due to (1.3). If $\langle n_1\rangle \lesssim \langle n_0\rangle$ we will instead estimate the g factor above in $l^{\gamma'}L^2$, the k factor in $l^\gamma L^2$, and the $u^{(5)}$ factor with weight $\langle m_i\rangle^{-1}$ in l^1L^1. Finally, if $\langle\alpha_0\rangle \lesssim \langle\beta_i\rangle$ for some i, we will cancel the $\langle\alpha_0\rangle$ weight by the $\langle\beta_i\rangle$ weight, then fix (m_j, β_j) and again get $\mathcal{S}_{\mathrm{sub}}$, which we estimate by

$$\begin{aligned}\mathcal{S}_{\mathrm{sub}} \lesssim &\sum_{n_0=n_1+m_i+c_1}\langle c_1\rangle^{-s}\langle n_0\rangle^{-s}\langle n_1\rangle^{-c(2-\gamma)}\|\langle n_0\rangle^{s}\langle\alpha_0\rangle^{-\kappa}g_{n_0,\alpha_0}\|_{L^2_{\alpha_0}}\\ &\qquad\times\|\langle n_1\rangle^{c(2-\gamma)}k_{n_1,\alpha_1}\|_{L^1_{\alpha_1}}\|\langle m_i\rangle^{-1}\langle\beta_i\rangle^{\kappa}(u^{(5)})_{m_i,\beta_i}\|_{L^2_{\beta_i}}\\ \lesssim &\|\langle n_0\rangle^{s}\langle\alpha_0\rangle^{-\kappa}g\|_{l^1L^2}\|\langle n_1\rangle^{c(2-\gamma)}k\|_{l^{\gamma'}L^1}\cdot\|\langle\partial_x\rangle^{-4s^3}u^{(5)}\|_{X_4} \lesssim 1,\end{aligned}$$

where the c_j are constants, and again note that we can gain any small power of c_1, since $\pm c_1$ is the sum of all m_j where $j \neq i$.

Next, let us assume $j \in \{2, 3, 8\}$. In this case we only use the l^1L^1 norm of $m_i^{-1}(u^{(5)})_{m_i,\beta_i}$, so we will be free to lose any power $\langle m_i\rangle^c$ for small c. Therefore we may fix each (m_i, β_i), invoke (11.7) to fix α_2 also (by an argument similar to the proof of Proposition 8.9), then reduce to bounding $\|z\|_{X_j}$ in terms of $\|k\|_{X_j}$, provided

$$|z_{n_0,\alpha_0}| \leq |k_{n_0+c_1,\alpha_0+|n_0+c_1|(n_0+c_1)-|n_0|n_0+c_2}|.$$

But since the bound we get is allowed to grow like $\langle c_1\rangle^{s^{1/3}}$ (note that $-c_1$ is the sum of all m_i, and we are allowed to lose $\langle m_i\rangle^c$ for small c), this will be easy if we examine X_2, X_3 and $\mathcal{Y}$ separately (in particular, we will use the definition of the $\mathcal{Y}$ norm). The only thing we need to address is the $\langle n\rangle$ weights in the definition of X_2 and X_3, and the step of taking supremum when obtaining the X_8 norm from the $\mathcal{Y}$ norm; however, by a standard argument we can show that through these we will lose at most $\langle c_1\rangle^{O(s)}$ power, which is acceptable.

Finally, we may check that throughout the above proof, we only need to use the X_j' norms of $\langle\partial_x\rangle^{-2s^3}u^{(5)}$ instead of $\langle\partial_x\rangle^{-s^3}u^{(5)}$; thus we will gain a power K^{0+} if we make the restriction $m_i \gtrsim K$ for some i. □

To see how Proposition 11.1 allows us to construct extensions of $\mathbb{P}_{>K}u^*$ and v^*, we first note that u^* is real-valued, so we only need to construct an extension of $\mathbb{P}_{>+K}u^*$ (which is an abbreviation of $\mathbb{P}_+\mathbb{P}_{>K}u^*$). Now, in Proposition 11.1 we may choose k to be an arbitrary extension of v^* and h to be some extension of u^* (and choose h' and k' accordingly) so that (11.6) holds with appropriate coefficients (cf. (7.1) and (7.2)).

Exploiting the freedom in the choice of k, we will set $\mathbb{P}_+k = w^{(4)}$ and $\mathbb{P}_{\leq 0}k = \mathbb{P}_{\leq 0}v''$. The part coming from $\mathbb{P}_+k$ is bounded in Y_2 (before or after the $\mathbb{P}_{>+K}$ projection) by $C_0e^{C_0A}$ due to Proposition 11.1, since we already have $\|w^{(4)}\|_{Y_2} \lesssim \|w^{(4)}\|_{Y_1} \leq C_0e^{C_0A}$. As for the part coming from $\mathbb{P}_{\leq 0}k$, we must have $n_0 > K$ and $n_1 \leq 0$ in (11.6), so the Ψ factor will be nonzero only when $\langle m_i\rangle \gtrsim (\mu+2)^{-2}K$ for some i, thus we may again use Proposition 11.1 to bound this part in Y_2 by $O_{C_1}(1)e^{C_0C_1A}K^{0-} \leq 1$, since we have $\|v''\|_{Y_2} \leq C_1e^{C_1A}$. This completes the construction for the extension of u^*.

Now, to construct the extension of v^*, simply set the k in Proposition 11.1 to be $u^{(4)}$ (which is the extension of u^* we just constructed) and h to be some extension of v^* so that (11.6) holds with appropriate coefficients. Then this extension will do the job, since we already have $\|u^{(4)}\|_{Y_2} \leq C_0e^{C_0A}$. This finally completes the proof of Proposition 8.1.

12. The a priori estimate V: Controlling the difference

The main purpose of this section is to provide necessary estimates for differences of two solutions to (1.6). First we need to introduce some notation, including the definition of the metric space $\mathcal{BO}^T$, which will be used also in Section 13.

12.1. Preparations

Definition 12.1. Suppose $\mathcal{Q} = (u'', v'', w'', u''')$ and $\mathcal{Q}' = (u^{\dagger\dagger}, v^{\dagger\dagger}, w^{\dagger\dagger}, u^{\dagger\dagger\dagger})$ are two quadruples of functions defined on $\mathbb{R}\times\mathbb{T}$. We define their distance by

$$\mathfrak{D}_\sigma(\mathcal{Q},\mathcal{Q}') = \|\langle\partial_x\rangle^{-\sigma}(w''-w^{\dagger\dagger})\|_{Y_1} + \|\langle\partial_x\rangle^{-\sigma}(v''-v^{\dagger\dagger})\|_{Y_2} + \|\langle\partial_x\rangle^{-\sigma}(u''-u^{\dagger\dagger})\|_{Y_2} + \|\langle\partial_x\rangle^{-s^3-\sigma}(u'''-u^{\dagger\dagger\dagger})\|_{X_2\cap X_3\cap X_4},$$

for $\sigma \in \{0, s^5\}$. In particular, if $\sigma = \mathcal{Q}' = 0$, we define the triple norm

$$|||\mathcal{Q}||| := \mathfrak{D}_0(\mathcal{Q}, 0) = \|w''\|_{Y_1} + \|v''\|_{Y_2} + \|u''\|_{Y_2} + \|\langle\partial_x\rangle^{-s^3} u'''\|_{X_2\cap X_3\cap X_4}.$$

Next, suppose u and u^- are functions defined on $I \times \mathbb{T}$ for some interval I. We will define the functions (u^*, v^*, w^*) corresponding to u and some M, and (u^+, v^+, w^+) corresponding to u^- and some N (note the definition depends on the choice of the origin in $\Delta_n(t) = \int^t \delta_n(t')\,dt'$, but this will not affect the triple norm $|||\cdot|||$; this does affect estimates for differences, but we need them only when $I = [-T, T]$ or its translation, in which case the choice of origin is canonical), as in Sections 5 and 7, and then define

$$\mathfrak{D}_\sigma^{I,MN}(u, u^-) = \inf_{\mathcal{Q},\mathcal{Q}'} \mathfrak{D}_\sigma(\mathcal{Q}, \mathcal{Q}'), \tag{12.1}$$

where the infimum is taken over all quadruples $\mathcal{Q}$ and $\mathcal{Q}'$ that extends (u^*, v^*, w^*, u) and (u^+, v^+, w^+, u^-) from $I \times \mathbb{T}$ to $\mathbb{R} \times \mathbb{T}$, respectively. We will also define $|||u|||_I^M = \mathfrak{D}_0^{I,MM}(u, 0) = \inf_{\mathcal{Q}} |||\mathcal{Q}|||$; these notations can be written in a more familiar way as

$$\begin{aligned}\mathfrak{D}_\sigma^{I,MN}(u, u^-) &= \|\langle\partial_x\rangle^{-\sigma}(w^* - w^+)\|_{Y_1^I} + \|\langle\partial_x\rangle^{-\sigma}(v^* - v^+)\|_{Y_2^I}\\ &\quad + \|\langle\partial_x\rangle^{-\sigma}(u^* - u^+)\|_{Y_2^I} + \|\langle\partial_x\rangle^{-s^3-\sigma}(u - u^-)\|_{(X_2\cap X_3\cap X_4)^I},\\ |||u|||_I^M &= \|w^*\|_{Y_1^I} + \|v^*\|_{Y_2^I} + \|u^*\|_{Y_2^I} + \|\langle\partial_x\rangle^{-s^3} u\|_{(X_2\cap X_3\cap X_4)^I}.\end{aligned}$$

Also, if $M = N = \infty$ we will omit it. Now we can define the metric space

$$\mathcal{BO}^I = \{u : |||u|||_I = |||u|||_I^\infty < \infty\}, \tag{12.2}$$

with the distance function given by $\mathfrak{D}_0^I$ (we will also use $\mathfrak{D}_{s^5}^I$, which is also well-defined on $\mathcal{BO}^I$). Finally, when $I = [-T, T]$, we may use T in place of I in sub- or superscripts, so this contains the definition of $\mathcal{BO}^T$.

Remark 12.2. If $u \in \mathcal{BO}^T$, we may define $uu_x = \frac{1}{2}\partial_x(\mathbb{P}_{\neq 0}u^2)$ as a distribution on $[-T, T]$ through an argument similar to the one in Section 11. More precisely, we may uniquely define the function

$$h(t) = \int_0^t e^{-(t-t')H\partial_{xx}}\big(u(t')\partial_x u(t')\big)\,dt' \tag{12.3}$$

as an element of $(X^{-1/s,\kappa})^T$.

In particular, we may define $u \in \mathcal{BO}^T$ to be a solution to (1.1) on $[-T, T]$, if u satisfies the integral version of (1.1) with the evolution term defined as in (12.3). Clearly this definition is independent of the choice of origin, and $[-T, T]$ may be replaced by any interval I.

Moreover, since the arguments in Section 11 allow for some room, the map sending u to h in (12.3) is continuous with respect to the *weak* distance function $\mathfrak{D}_{s^5}^T$ (or $\mathfrak{D}_{s^5}$ if we consider the map sending the quadruple $\mathcal{Q}$ to h). This fact will be important in the proof of Theorem 13.1.

Proposition 12.3. *Let B_t^0 be the space of* bounded *functions of t into some Banach space. Suppose u and u^- are two functions defined on $I\times\mathbb{T}$, and choose corresponding extensions $\mathcal{Q}=(u'',v'',w'',u''')$ and $\mathcal{Q}'=(u^{\dagger\dagger},v^{\dagger\dagger},w^{\dagger\dagger},u^{\dagger\dagger\dagger})$ corresponding to M and N, where $M\geq N$.*

$$\|u\|_{B_t^0(I\to Z_1)}\lesssim\|\mathcal{Q}\|,\qquad \|u\|_{B_t^0(I\to Z_1)}\lesssim\|u\|_I^M. \tag{12.4}$$

Concerning differences, we only have the weaker estimates

$$\|\langle\partial_x\rangle^{-s^5}(u-u^-)\|_{C_t^0(I\to Z_1)}\lesssim O_{\|\mathcal{Q}\|,\|\mathcal{Q}'\|}(1)\cdot\big(\mathfrak{D}_{s^5}(\mathcal{Q},\mathcal{Q}')+N^{0-}\big), \tag{12.5}$$

$$\|\langle\partial_x\rangle^{-\theta}(u-u^-)\|_{C_t^0(I\to Z_1)}\lesssim O_{\theta,\|\mathcal{Q}\|,\|\mathcal{Q}'\|}(1)\cdot\big(\mathfrak{D}_0(\mathcal{Q},\mathcal{Q}')+N^{0-}\big), \tag{12.6}$$

for all $\theta>0$, where the constant may also depend on the upper bound of the length of I.

Proof. We may assume $I=[-T,T]$ with $T\lesssim 1$. The inequalities in (12.4) follow directly from the definition (and the fact that $u(t)$ and $u''(t)$ have the same Z_1 norm for $t\in[-T,T]$); the proofs of (12.5) and (12.6) are similar, so we only prove (12.5). Assume $\|\mathcal{Q}\|+\|\mathcal{Q}'\|\lesssim 1$ and $\mathfrak{D}_{s^5}(\mathcal{Q},\mathcal{Q}')\leq\varepsilon$, we will define Δ_n and Δ_n^- corresponding to $\mathcal{Q}$ and $\mathcal{Q}'$ as in (7.27) and (7.28) using functions (w'',u''') and $(w^{\dagger\dagger},u^{\dagger\dagger\dagger})$ respectively, then set u' and $u^\dagger$ to be extensions of u and u^-, defined by $(u')_n=\chi(t)e^{\mathrm{i}\Delta_n}(u'')_n$ and similarly for $u^\dagger$. Since

$$\|\langle\partial_x\rangle^{-s^5}(u''-u^{\dagger\dagger})(t)\|_{Z_1}\lesssim\mathfrak{D}_{s^5}(\mathcal{Q},\mathcal{Q}')\lesssim\varepsilon, \tag{12.7}$$

we only need to estimate the function z defined by $z_n=(u'')_n(e^{\mathrm{i}\Delta_n}-e^{\mathrm{i}\Delta_n^-})$. Due to the bound $\|\mathcal{Q}\|\lesssim 1$ which implies the bound for the Z_1 norm of each $u''(t)$, we only need to prove

$$|\chi(t)(e^{\mathrm{i}\Delta_n}-e^{\mathrm{i}\Delta_n^-})(t)|\lesssim(\varepsilon+N^{0-})\langle n\rangle^{s^5} \tag{12.8}$$

for each n and t. Using the arguments in Lemma 8.6, it suffices to prove the bound for $\delta_n-\delta_n^-$, but if we use (7.28), this will be clear from the strong bounds on w'' and $w^{\dagger\dagger}$, and the weak bound on their difference. □

12.2. Statement and proof

Now suppose u is a smooth function solving 1.6) on $[-T,T]$. The arguments in Sections 8–10 actually give us a way to update a given quadruple $\mathcal{Q}=(u'',v'',w'',u''')$ extending (u^*,v^*,w^*,u) to a new quadruple $\mathcal{Q}'=(u^{(4)},v^{(4)},w^{(4)},u^{(5)})$, which remains to be an extension, and satisfies better bounds. We define $\mathfrak{I}$ to be the map from the set of extensions to itself, that sends $\mathcal{Q}$ to $\mathcal{Q}'$. Using the arguments from Sections 8–10, we can prove

Proposition 12.4. *Let C_1 be large enough, C_2 large enough depending on C_1, and $0<T\leq C_2^{-1}e^{C_2A}$. Suppose u is a smooth function solving (1.6) on $[-T,T]$, and $\mathcal{Q}$ is an extension satisfying*

$$\|\mathcal{Q}\|\leq C_1e^{C_1A},\qquad \|\langle\partial_x\rangle^{-s^3}u'''\|_{X_2\cap X_3\cap X_4}\leq C_1A. \tag{12.9}$$

Then the same estimate will hold if we replace $\mathcal{Q}$ by $\mathfrak{I}\mathcal{Q}$.

Now we can state the main proposition in this section, namely

Proposition 12.5. *Let C_1, C_2 and T be as in Proposition 12.4. Suppose u and u^- are two smooth functions solving (1.6) with truncations S_N and S_M respectively, where $1 \ll N \leq M \leq \infty$, $\mathcal{Q}$ and $\mathcal{Q}'$ are two quadruples corresponding to u and u^- respectively, such that (12.9) holds, and that*

$$\mathfrak{D}_{s^5}(\mathcal{Q}, \mathcal{Q}') \leq B \tag{12.10}$$

for some $B > 0$. Then

$$\mathfrak{D}_{s^5}(\mathfrak{I}\mathcal{Q}, \mathfrak{I}\mathcal{Q}') \leq B/2 + O_{C_1}(1)e^{C_0 C_1 A}\big(\|\langle\partial_x\rangle^{-s^5}(u(0) - u^-(0))\|_{Z_1} + N^{0-}\big), \tag{12.11}$$

where C_0 is any constant appearing in previous sections. In particular,

$$\mathfrak{D}_{s^5}^{T,NM}(u, u^-) \leq O_{C_2,A}(1)\big(\|\langle\partial_x\rangle^{-s^5}(u(0) - u^-(0))\|_{Z_1} + N^{0-}\big), \tag{12.12}$$

provided $\|u(0)\|_{Z_1} + \|u^-(0)\|_{Z_1} \leq A$ for some large A. Moreover, if $M = N$, we may replace the $\mathfrak{D}_{s^5}$ distance by the $\mathfrak{D}_0$ distance and remove the N^{0-} term on the right hand side of (12.12).

Proof. When we take differences in the case $M = N$, the right hand side will involve only factors like $u - u^-$ and not the ones like $\mathbb{P}_{\geq N}u$, thus we will not have an N^{0-} term on the right hand side. Also, it is easy to see from the proof below that removing the $\langle n\rangle^{-s^5}$ weight will only make arguments easier. Thus we will focus on (12.12) now. By an iteration using Proposition 12.4, we only need to prove (12.11) assuming (12.9) and (12.10).

Recall the functions δ_n, δ_n^-, Δ_n, Δ_n^- and y, y^- that come from the two quadruples $\mathcal{Q}$ and $\mathcal{Q}'$ in the same way as in Section 8.2. The two functions y and y^- will satisfy two equations with the form of (8.12) separately. Clearly we may also assume all relevant functions are supported in $|t| \lesssim 1$. To bound the first part of $\mathfrak{D}_{s^5}(\mathfrak{I}\mathcal{Q}, \mathfrak{I}\mathcal{Q}')$ requires proving

$$\|\langle\partial_x\rangle^{-s^5}(y - y^-)\|_{Y_1} \leq B/10 + O_{C_1}(1)e^{C_0 C_1 A}(\theta + N^{0-}), \tag{12.13}$$

where we denote $\|\langle\partial_x\rangle^{-s^5}(u(0) - u^-(0))\|_{Z_1} = \theta$ for simplicity.

By another bootstrap argument, we may assume (12.13) holds with right hand side multiplied by $O_{C_1}(1)$. Recall the equations

$$y = \chi(t)e^{it\partial_{xx}}w(0) + \mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^2(y, y)) + \sum_{j\in\{3,3.5,4,4.5\}} \mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^j), \tag{12.14}$$

$$y^- = \chi(t)e^{it\partial_{xx}}w^-(0) + \mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^{2-}(y^-, y^-)) + \sum_{j\in\{3,3.5,4,4.5\}} \mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^{j-}), \tag{12.15}$$

where $\mathcal{N}^j$ and $\mathcal{N}^{j-}$ are suitable nonlinearities; to bound $y - y^-$, we will first bound

$$\sum_{j\in\{0,3,3.5,4,4.5\}} \|\langle\partial_x\rangle^{-s^5}(\mathcal{M}^j - \mathcal{M}^{j-})\|_{Y_1},$$

where the definitions of $\mathcal{M}^j$ and $\mathcal{M}^{j-}$ are clear (the term $j = 0$ corresponds to the linear term which can be bounded by $\theta + N^{0-}$, so we will omit this below).

Here it is important to note that *all the bounds in the previous sections are proved directly using multilinear estimates*, thus they will automatically imply the corresponding estimates for differences. In fact, when we try to estimate $\mathcal{M}^j - \mathcal{M}^{j-}$ by introducing some (g, f) and forming an $\mathcal{S}$ expression, there are a few possibilities:

(1) Suppose we take the difference $y - y^-$, or (for example) some $v'' - v^{\dagger\dagger}$ directly. Then one of the y or v'' factors appearing in the previous sections will be replaced by this difference. Note that if we estimate this difference in the weakened norm $\|\langle\partial_x\rangle^{-s^5}\cdot\|_{Y_j}$ (we use the $X_2 \cap X_3 \cap X_4$ norm for $u''' - u^{\dagger\dagger\dagger}$, but the proof will be the same), we will get a bound $O_{C_1}(1)e^{C_0C_1A}(B + \theta + N^{0-})$ which is what we need; the loss coming from using this weaker norm can be recovered from the fact that we only need to estimate the weaker norm of $\mathcal{M}^j - \mathcal{M}^{j-}$. To be precise, for each multilinear estimate we proved in the previous sections, suppose the term we bound in the weaker norm (i.e. the norm involving $\langle\partial_x\rangle^{-s^5}$) corresponds to the variable n_l; then one of the following must hold: (i) we can gain a power $2^{(0+)d}$ in the estimate, where $0+$ is at least $cs^{2.5}$, and we also have $\langle n_l\rangle \lesssim 2^d$; in this case it will suffice to use this weaker norm in all the discussions before, so this part will be acceptable; (ii) we have $\langle n_0\rangle \gtrsim \langle n_l\rangle$ (for example, when $n_0 = n_l$ and the other variables are small compared to them). In this case, since we only need to estimate the output $y - y^-$ in the weaker norm, we will gain a power $\langle n_0\rangle^{s^5}$ compared to the proof in the previous sections, which is enough to cancel the loss $\langle n_l\rangle^{s^5}$, thus this part is also acceptable; (iii) we have $\langle n_0\rangle \sim 2^d$ and $\langle n_l\rangle \sim 2^{d'}$, and the expression $\mathcal{S}$ involves the factor $2^{-|d-d'|}$ (this appears, for example, in various "resonant" cases in Section 9 and Proposition 10.1, and is characterized by the need to use (9.24)). In this case we lose at most $2^{s^5|d-d'|}$ from the additional weights compared to the proof in the previous sections, which can be canceled by the $2^{-|d-d'|}$ factor, so it will still be acceptable. To conclude, we can estimate this part of $y - y^-$ in the weaker norm as

$$T^{0+}O_{C_1}(1)e^{C_0C_1A}(B + \theta + N^{0-}),$$

by repeating the arguments in the previous sections, with minor modifications illustrated above.

(2) Suppose we take the difference of the Φ weights. The difference will satisfy the same bounds as the weights themselves; moreover it is nonzero only when some m or n variable is $\gtrsim N$. Therefore we may replace one of the y or v'' factors appearing in the previous sections by $\mathbb{P}_{\geq N}y$ or $\mathbb{P}_{\geq N}v''$. We then proceed as in case (1), estimating this particular factor in the weakened norm to gain a power N^{0-}, and bound the whole expression in the same way as in case (1).

(3) Suppose we take (for example) the difference $v' - v^{\dagger}$, where $(v')_n = e^{\mathrm{i}\Delta_n}(v'')_n$ and $(v^{\dagger})_n = e^{\mathrm{i}\Delta_n^-}(v^{\dagger\dagger})_n$; alternatively, suppose we take the difference

$$e^{\mathrm{i}(\pm\Delta_{n_0}\pm\Delta_{n_1}\pm\cdots)} - e^{\mathrm{i}(\pm\Delta_{n_0}^-\pm\Delta_{n_1}^-\pm\cdots)}.$$

It turns out that whenever we need to estimate these factors, we will always gain (from these factors themselves, or from elsewhere) some power $2^{(0+)d}$ where $0+$ is at least

$cs^{2.5}$, and 2^d controls every relevant variable (for typical examples, see the estimate of $J_{(n)}(\alpha_5)$ as defined in (10.13) in the proof of Proposition 10.1, as well as the last part of Section 9). Here we may use Proposition 8.6 to reduce the estimation of the difference of these exponential factors to the estimation of the differences $\delta_n - \delta_n^-$ themselves. Since we can bound functions like $w'' - w^{\dagger\dagger}$ in the weaker norm by $O_{C_1}(1)e^{C_0C_1A}(B+\theta+N^{0-})$, we will be able to obtain estimates similar to the ones in Proposition 8.7, but with the coefficient $C_0C_1e^{C_0C_1A}$ on the right hand side replaced by $O_{C_1}(1)e^{C_0C_1A}(B+\theta+N^{0-})$, with a loss of at most $\langle n\rangle^{O(s^5)}$ which is dwarfed by the power we gain. Finally, we may use the T^{0+} gain coming from the evolution to cancel the $O_{C_1}(1)e^{C_0C_1A}$ factor, thus this part is also acceptable.

Next we need to control the difference of the $\mathcal{M}^2$ terms. We will follow the proof in Section 9, and the part of the proof where no second iteration is needed can be completed in the same way as above. As for the remaining part, what we do in Section 9 is basically rewriting

$$\mathcal{N}^6(y,y) = \sum_{j\in\{0,3,3.5,4,4.5\}} \mathcal{N}^6(y,\mathcal{M}^j) + \mathcal{N}^6(y,\mathcal{E}(\mathbf{1}_{[-T,T]}\mathcal{N}^2(y,y)))$$

where $\mathcal{N}^6$ is the part of $\mathcal{N}^2$ under consideration; we may also rewrite $\mathcal{N}^{6-}(y^-,y^-)$ in the same way. When we take the difference, we may control the first term on the right hand side using the bound for $\mathcal{M}^j - \mathcal{M}^{j-}$ as in Proposition 9.1 (actually we have a slightly weaker version, but this will suffice); as for the second term, since it is bounded in Section 9 via multilinear estimates, we can again treat the difference in the same way as above. This completes the proof for the bound of $w^* - w^+$.

Next, recall that the other parts of $\mathfrak{I}\mathcal{Q}$ and $\mathfrak{I}\mathcal{Q}'$ such as $u^{(5)}$ and $u^{[5]}$, $u^{(4)}$, $u^{[4]}$, $v^{(4)}$ and $v^{[4]}$ are constructed in the same way as in Section 11, where the scale K is taken to be $K = C_{1.5}e^{C_{1.5}A}$ with $C_{1.5}$ large enough depending on C_1, but small compared to C_2. Note that we may redefine Δ_n and Δ_n^- when necessary. Now to prove (12.10), we need to bound the differences such as $u^{(4)} - u^{[4]}$ in the weaker norm by $O_{C_1}(1)e^{C_0C_1A}(K^{0-}B + \theta + N^{0-})$. But this can again be achieved by combining the argument above with the proof in Section 11, if we notice two things:

(1) In the proof of Proposition 11.1, we can always gain some power $\langle m_i\rangle^{cs^{2.5}}$ for each m_i, so we will be able to cover the loss coming from using only the weaker norm if we take the difference of the exponential factors (cf. (11.7)), or if we take $u^{(5)} - u^{[5]}$. For the same reason, if we lose a power $\langle n_1\rangle^{s^5}$ we will be able to recover it from the gain $\langle n_0\rangle^{s^5}$.

(2) From the above we already know that the weaker norm of $w^{(4)} - w^{[4]}$ can be bounded by $O_{C_1}(1)e^{C_0C_1A}(T^{0+}B+\theta+N^{0-})$. We may then prove the same bound (possibly with some $O_{C_1}(1)$ factors) for $\mathbb{P}_{>K}(u^{(5)}-u^{[5]})$, $\mathbb{P}_{\leq K}(u^{(5)}-u^{[5]})$, $\mathbb{P}_{\leq K}(u^{(4)}-u^{[4]})$, $\mathbb{P}_{>K}(u^{(4)}-u^{[4]})$ and $v^{(4)}-v^{[4]}$ *in that order*, in the same way as in Section 11 (note T^{-1} is assumed to be larger than any power of K).

Therefore we will be able to bound all the differences and thus complete the proof of Proposition 12.5. □

13. Proof of the main results

With Propositions 8.1 and 12.5, it is now easy to prove our main results. Since the argument in this section will be more or less standard, we may present only the most important steps.

13.1. Local well-posedness and stability

Theorem 13.1 (Precise version of Theorem 1.2). *There exists a constant C such that, when we choose any $A > 0$ and $0 < T \leq C^{-1}e^{-CA}$, the following hold:*

(1) Existence: *For any $f \in \mathcal{V}$ with $\|f\|_{Z_1} \leq A$, there exists some $u \in \mathcal{BO}^T$ such that $|||u|||_T \leq Ce^{CA}$ and u satisfies equation (1.1), in the sense described in Remark 12.2, with initial data $u(0) = f$.*

(2) Continuity: *Let the solution described in part* (1) *be $u = \Phi f = (\Phi_t f)_t$. Suppose $\|f\|_{Z_1} \leq A$ and $\|g\|_{Z_1} \leq A$. Then for each $\varepsilon > 0$, we have*

$$\sup_{|t|\leq T} \|\langle\partial_x\rangle^{-s^5}(\Phi_t f - \Phi_t g)\|_{Z_1} + \mathfrak{D}^T_{s^5}(\Phi f, \Phi g) \leq O_{C,A}(1)\|\langle\partial_x\rangle^{-s^5}(f-g)\|_{Z_1},$$

$$\sup_{|t|\leq T} \|\langle\partial_x\rangle^{-\varepsilon}(\Phi_t f - \Phi_t g)\|_{Z_1} + \mathfrak{D}^T_0(\Phi f, \Phi g) \leq O_{\varepsilon,C,A}(1)\|f-g\|_{Z_1}.$$

(3) Short-time stability: *Let $u = \Phi f$ as in part* (2)*, and let Φ^N be the solution flow of* (1.6) *and $u^N = \Phi^N \Pi_N f$. Then*

$$\lim_{N\to\infty}\Big(\mathfrak{D}^{T,N\infty}_{s^5}(u^N, u) + \sup_{|t|\leq T}\|\langle\partial_x\rangle^{-s^5}(u^N(t) - u(t))\|_{Z_1}\Big) = 0.$$

(4) Uniqueness: *For any* other *time T', suppose u and u^- are two elements of $\mathcal{BO}^{T'}$ with the same initial data, and they both solve* (1.1)*. Then $u = u^-$ (on $[-T', T']$).*

(5) Long-time existence: *Consider any $f \in Z_1$, and define u^N as in* (3)*. Suppose that for some* other *time T' and some subsequence $\{N_k\}$,*

$$\sup_k |||u^{N_k}|||^{N_k}_{T'} < \infty. \tag{13.1}$$

Then there exists a solution $u \in \mathcal{BO}^{T'}$ to (1.1) *with initial data f.*

Proof. Suppose $f \in Z_1$ and $\|f\|_{Z_1} \leq A$, and let $0 < T \leq C_2^{-1}e^{-C_2 A}$ with constants as in Propositions 12.4 and 12.5. Consider u^N as defined in (3); using Proposition 8.1, we may choose for each N some quadruple $\mathcal{Q}_N$ corresponding to u^N that satisfies (12.9). We define

$$\mathcal{Q}^N = \mathfrak{I}^N \mathcal{Q}_N. \tag{13.2}$$

It is clear from Propositions 12.4 and 12.5 that

$$|||\mathcal{Q}^N||| \leq C_1 e^{C_1 A}, \tag{13.3}$$

$$\lim_{N,M\to\infty} \mathfrak{D}_{s^5}(\mathcal{Q}^M, \mathcal{Q}^N) = 0. \tag{13.4}$$

By a simple completeness argument we can then find some $\mathcal{Q}$ so that $\mathfrak{D}_{s^5}(\mathcal{Q}^N, \mathcal{Q}) \to 0$ (in particular $\mathcal{Q}$ will have initial data f), and by an argument similar to the proof of Proposition 3.6 we deduce that $\|\!|\mathcal{Q}|\!\| \leq C_1 e^{C_1 A}$. By using Remark 12.2, we can now pass to the limit and show that the quadruple $\mathcal{Q}$ gives a solution $u \in \mathcal{BO}^T$ of (1.1) on the interval $[-T, T]$. This proves existence.

Parts (2) and (3) will follow from basically the same argument. In fact, for each (f, g), we may construct $\mathcal{Q}^N$ and $\mathcal{Q}^{N-}$ corresponding to $\Phi^N \Pi_N f$ and $\Phi^N \Pi_N g$ as above, so that they have uniformly bounded triple norm, and moreover

$$\mathfrak{D}_{s^5}(\mathcal{Q}^N, \mathcal{Q}^{N-}) \lesssim \|\langle \partial_x \rangle^{-s^5}(f - g)\|_{Z_1} + N^{0-}.$$

Using Proposition 12.3 and passing to the limit, we obtain the result in (2). The result in (3) follows from comparing $\mathcal{Q}^N$ with $\mathcal{Q}$ and using Proposition 12.3 also.

As for part (5), we will deduce it merely from the condition that $\|\!|u^{N_k}|\!\|_{T'}^{N_k} \leq A$ and

$$\|\langle \partial_x \rangle^{-s^5}(u^{N_k} - u)(0)\|_{Z_1} \to 0, \tag{13.5}$$

which is clearly satisfied in our setting. Choose some τ small enough depending on A; then $\|u(0)\|_{Z_1} \leq C_0 A$ implies we can solve (1.1) on $[-\tau, \tau]$, and from (Proposition 12.5 and) what we just proved, we also have

$$\|\langle \partial_x \rangle^{-s^5}(u^{N_k} - u)(\pm\tau)\|_{Z_1} \to 0, \tag{13.6}$$

and therefore

$$\|u(\pm\tau)\|_{Z_1} \leq \limsup_{N\to\infty} \|u^{N_k}(\pm\tau)\|_{Z_1} \leq C_0 A. \tag{13.7}$$

This information will allow us to restart from time $\pm\tau$, and thus obtain a solution to (1.1) on $[-2\tau, 2\tau]$. Repeating this, we will finally get a solution on $[-T', T']$, which we can prove to be in $\mathcal{BO}^{T'}$ using partitions of unity. This proves (conditional) global existence.

Finally, we need to prove uniqueness. Let u and u^- be two solutions to (1.1) that both belong to $\mathcal{BO}^{T'}$ and have the same initial data. Let their strong norms be bounded by A, and choose τ small enough depending on A. To prove that $u = u^-$ on $[-\tau, \tau]$, we need to prove the following claim: if for quadruples $\mathcal{Q}$ and $\mathcal{Q}'$ corresponding to u and u^- respectively, we have

$$\|\!|\mathcal{Q}|\!\| + \|\!|\mathcal{Q}'|\!\| \leq A, \qquad \mathfrak{D}_{s^5}(\mathcal{Q}, \mathcal{Q}') \leq K, \tag{13.8}$$

then with $\mathcal{Q}$ replaced by $\mathfrak{I}\mathcal{Q}$ and $\mathcal{Q}'$ by $\mathfrak{I}\mathcal{Q}'$, the inequalities will hold with A unchanged and K replaced by $K/2$. Thus we need to repeat the whole argument from Section 8 to Section 12 *without* the smoothness assumption. Fortunately, since we have chosen $\tau \leq \tau(A)$, we do not need the bootstrap argument (which requires a priori smoothness) in bounding the evolution term; however, we do need this in Section 8 when we try to obtain a first bound for $\|y\|_{Y_1}$.

This difficulty can be overcome as follows: first, we may check for every part of Sections 8, 9 and 10 that in order to bound y in Y_1 using the evolution equation (8.12), it will suffice to bound y in some weaker space Y_1^w defined by (cf. Section 2.3)

$$\|u\|_{Y_1^w} = \|u\|_{X_1^w} + \|u\|_{X_2^w} + \|u\|_{X_4^w} + \|u\|_{X_5^w} + \|u\|_{X_7^w}. \tag{13.9}$$

Here to obtain the X_j^w norm, we weaken the X_j by decreasing the powers b in (2.2), κ in (2.5) and $1/8$ in (2.8) by s^5, and increasing the indices 1 in (2.3) and q in (2.6) by s^5. Notice that any power of n and any l^p norm remain unchanged. Therefore, we only need to show that the linear map L defining y from w'' (see Section 8.2) is bounded from Y_1 to Y_1^w, since this combined with the proof from Sections 8 to 10 will give us a stronger bound of y in Y_1 and close the estimate (note that after the end of Section 10, no arguments will depend on smoothness, and we will be able to finish just as Sections 11 and 12).

Now, suppose $\|u\|_{Y_1} \leq 1$; we can easily show that $\|Lu\|_{X_2^w} + \|Lu\|_{X_5^w} \lesssim 1$ using the decomposition

$$Lu = u \cdot \mathbf{1}_{[-T,T]}(t) + \chi(t)\mathbf{1}_{[T,\infty)}(t)e^{-(t-T)H\partial_{xx}}u(T) \\ + \chi(t)\mathbf{1}_{(-\infty,-T]}(t)e^{-(t+T)H\partial_{xx}}u(-T). \tag{13.10}$$

In fact, the last two terms in (13.10) are bounded in X_2^w and X_5^w because $u(\pm T)$ is bounded in Z_1, and the Fourier transform of $\chi(t)\mathbf{1}_{[T,\infty)}(t)$ is in L^k for $k > 1$; the first term is bounded because convolution with the Fourier transform of $\mathbf{1}_{[-T,T]}$ (which decays like $\langle\xi\rangle^{-1}$ uniformly for $T \lesssim 1$) is bounded from L_ξ^k to $L_\xi^{k'}$ for all $k' > k$. Now to bound Lu in X_j^w for $j \in \{1, 4, 7\}$, we only need to bound the operator

$$\widetilde{L} : f(t) \mapsto \mathbf{1}_{[-T,T]}(t)f(t) + \chi(t)\mathbf{1}_{[T,\infty)}f(T) + \chi(t)\mathbf{1}_{(-\infty,-T]}f(-T) \tag{13.11}$$

from H_t^h to $H_t^{h-\theta}$ for any $\theta > 0$. By direct computations we can bound $\widetilde{L}$ from H^1 to itself, thus (by interpolation) it suffices to bound $\widetilde{L}$ from $H^{1/2+\theta}$ to $H^{1/2-\theta}$. But this result is well-known for the first part of $\widetilde{L}$, and trivial (given the decay of the Fourier transform of $\chi(t)\mathbf{1}_{[T,\infty)}(t)$) for the last two parts. □

13.2. The Hamiltonian structure and global well-posedness

In this section we will denote any constant by C, since they no longer make any difference. We fix some large time T, and recall the energy functional

$$E_N[f] = \int_{\mathbb{T}} \left(\tfrac{1}{2}|\partial_x^{1/2} f|^2 - \tfrac{1}{6}(S_N f)^3\right) \tag{13.12}$$

defined in Section 4.1. If we introduce the symplectic form

$$\omega(u, v) = \int_{\mathbb{T}} u \cdot (\partial_x^{-1} v)$$

in the (finite-dimensional) space $\mathcal{V}_N$, then a simple computation shows that the Hamiltonian equation with respect to the symplectic form ω and the functional E_N is (up to a sign depending on the convention) the truncated equation (1.6). By Liouville's Theorem, the solution flow $\{\Phi_t^N\}_{t\in\mathbb{R}}$ will preserve the measure $\mathcal{L}_N$ which corresponds to the Lebesgue measure on $\mathbb{R}^{2N}$ (see Section 4.1). Since this flow also preserves the L^2 norm as well as the Hamiltonian E_N, we see that

$$\nu_N^\circ(E) = \nu_N^\circ(\Phi_t^N(E)) \tag{13.13}$$

for all time t and all Borel sets $E \subset \mathcal{V}_N$.

Next, for any $f \in \mathcal{V}$, consider the functions $u^N(t) = \Phi_t^N \Pi_N f$, which are the solutions to (1.6) with initial data $u^N(0) = \Pi_N f$. Thus $f \mapsto u^N$ is a map from $\mathcal{V}$ to $\mathcal{BO}^T$ depending on N, therefore we may denote $\|u^N\|_T^N = J_N(f)$.

Choose a large positive integer M, a parameter A depending on M, and define

$$\Omega_{N,A} = \{g \in \mathcal{V}_N : \|g\|_{Z_1} > A\}.$$

Then

$$\nu_N^\circ(\Omega_{N,A}) = \nu_N(\Pi_N^{-1}(\Omega_{N,A})) \leq \nu_N(\{f \in \mathcal{V} : \|f\|_{Z_1} > A\}) \leq Ce^{-C^{-1}A^2}, \quad (13.14)$$

where the last inequality follows from Proposition 4.6, Cauchy–Schwarz, and the fact that $\|\theta_N\|_{L^2(d\rho)} = O(1)$ (which is part of Proposition 4.4). Therefore if we introduce

$$\Omega_{N,M,A} = \bigcup_{j=-M}^{M} (\Phi_{jT/M}^N)^{-1}(\Omega_{N,A}),$$

we will have

$$\nu_N^\circ(\Omega_{N,M,A}) \leq CMe^{-C^{-1}A^2}. \quad (13.15)$$

If we choose $A = A(M) = C'\sqrt{\log M}$ with some sufficiently large C', then the inequality (13.15) will imply $\nu_N^\circ(\Omega_{N,M,A}) \leq CM^{-3}$. Now if $g \notin \Omega_{N,M,A}$, we must have

$$\Phi_{jT/M}^N(g) \notin \Omega_{N,A(M)}$$

for all $|j| \leq M$. By Proposition 8.1, this implies

$$\max_{|j|\leq M} \|(\Phi_t^N g)_t\|_{[(j-1)T/M,(j+1)T/M]}^N \leq Ce^{CC'\sqrt{\log M}}, \quad (13.16)$$

provided $T/M \leq C^{-1}e^{-CA(M)}$, which is clearly true when M is large enough depending on T. Using partitions of unity, we easily see that (13.16) implies

$$\|(\Phi_t^N g)_t\|_T^N \leq CM^C,$$

again when M is large enough depending on T. Thus we have proved

$$\nu_N(\{f \in \mathcal{V} : J_N(f) > CM^C\}) \leq CM^{-3} \quad (13.17)$$

for all $M > M(T)$, and hence (recall Section 4.1 for the definition of θ_N)

$$\sup_N \int_{\mathcal{V}} \log(J_N(f) + 2)\theta_N(f)\, d\rho(f) < \infty. \quad (13.18)$$

Since $\theta_N(f)$ converges to $\theta(f)$ almost surely after passing to a subsequence, we may use Fatou's Lemma to conclude that except for a set with zero ρ measure, for each f with $\theta(f) > 0$, there exists a sequence $N_k \uparrow \infty$ so that $J_{N_k}(f) \leq C$ for some C. By part (5) of Theorem 13.1, this would imply the existence of a solution $u \in \mathcal{BO}^T$ to (1.1) on $[-T, T]$ with initial data f. Finally, by Remark 4.3 we may choose a sequence of Gibbs measures $\{\theta^R\}$ so that for almost every $f \in \mathcal{V}$ we have at least one $\theta^R(f) > 0$; then we take another countable intersection with respect to T, to arrive at

Proposition 13.2. *For almost every $f \in \mathcal{V}$ with respect to the Wiener measure ρ, there exists a unique global solution u to (1.1) with initial data f such that $u \in \mathcal{BO}^T$ for each $T > 0$.*

13.3. The global flow and invariance of Gibbs measure

In this section we will restate and prove Theorem 1.3.

Theorem 13.3 (Restatement of Theorem 1.3). *Let the Wiener measure ρ be defined as in Section 4.1. There exists a subset $\Sigma \subset \mathcal{V}$ such that $\rho(\mathcal{V} - \Sigma) = 0$ and the following holds: For any $f \in \Sigma$ there exists a unique global solution u to (1.1) with initial data f such that $u \in \mathcal{BO}^T$ for all $T > 0$. Moreover, if $u = \Phi f = (\Phi_t f)_t$, then these Φ_t form a measurable transformation group from Σ to itself. Finally, suppose the Gibbs measure ν is defined as in Section 4.1 (using some cutoff function ζ). Then each Φ_t keeps ν invariant.*

Proof. We define Σ to be the set of all $f \in \mathcal{V}$ such that there exists a solution u to (1.1) with initial data f that belongs to $\mathcal{BO}^T$ for all $T > 0$. We first show that Σ is measurable; in fact, we have $\Sigma = \bigcap_T \bigcup_A \Sigma_{AT}$, where Σ_{AT} is the set of f such that a solution u exists in $\mathcal{BO}^T$ with $\|u(t)\|_{Z_1} \leq A$ for all $|t| \leq T$. Now, divide $[-T, T]$ into M equal intervals where M is large enough depending on A. Then by local theory, the solution map Φ_t is well-defined (and measurable) on each subinterval. Now we can (iteratively) see that Σ_{AT} is a finite intersection of sets, each being the pre-image of the previous one under a measurable map, so Σ_{AT} is measurable.

Proposition 13.2 guarantees that $\rho(\mathcal{V} - \Sigma) = 0$; also the map Φ is well-defined on Σ, and each Φ_t maps Σ to itself. Note that from part (4) of Theorem 13.1, any two solutions to (1.1) that belong to $\mathcal{BO}^T$ and agree at one time must coincide, thus u will be unique for each fixed $f \in \Sigma$. Now fix a Gibbs measure ν; to prove the invariance of ν, we only need to show that

$$\nu(\Phi_t(E)) \geq \nu(E) \tag{13.19}$$

for each Borel subset E and each $|t| \leq 1$, since the rest can be done by iteration.

Define

$$\Sigma_A = \left\{ f \in \Sigma : \sup_{|t| \leq 2} \|\Phi_t f\|_{Z_1} \leq A \right\}$$

for each A. Then $\Sigma = \bigcup_A \Sigma_A$, so we only need to prove (13.19) assuming $E \subset \Sigma_A$ for some A. By iteration, it suffices to prove (13.19) when $E \subset \{f : \|f\|_{Z_1} \leq A\}$ and $|t| \leq t(A)$. Next, we introduce on the set $\{f : \|f\|_{Z_1} \leq A\}$ the metric

$$d(f, g) = \|\langle n \rangle^{-s^6 + r}(f - g)\|_{l^p},$$

making it a complete separable metric space. By a well-known theorem in measure theory, the restriction of ν to this set is a finite Borel measure on this metric space, and thus is regular (meaning every Borel set can be approximated from the inside by compact sets). Therefore we may further assume E is compact with respect to the metric d. Recall the solution flow $\{\Phi_t^N\}$ for (1.6); for each N we have

$$\nu_N(\{g : \Pi_N g = \Phi_t^N(\Pi_N h),\ h \in E\}) \geq \nu_N(E) \tag{13.20}$$

by the invariance of ν_N° under the flow Φ_t^N. To prove (13.19) it thus suffices to show

$$\limsup_{N \to \infty} \{g : \Pi_N g = \Phi_t^N(\Pi_N h),\ h \in E\} \subset \Phi_t(E), \tag{13.21}$$

since we already know that the total variation of $\nu_N - \nu$ tends to zero.

Now suppose for some $g \in \mathcal{V}$ we have a subsequence $N_k \uparrow \infty$ and $h^{N_k} \in E$ such that $\Pi_{N_k} g = \Phi_t^{N_k}(\Pi_{N_k} h^{N_k})$ for each k. By compactness we may assume $h^{N_k} \to h$ with respect to the metric d for some $h \in E$. Since every function involved here is bounded in Z_1 norm by $O_A(1)$, and we are assuming $|t| \leq t(A)$, we may use Propositions 12.5 and 13.1, as well as the limit

$$\|\langle \partial_x \rangle^{-s^5}(h^{N_k} - h)\|_{Z_1} \lesssim d(h^{N_k}, h) \to 0 \tag{13.22}$$

to conclude that

$$\begin{aligned}\|\langle \partial_x \rangle^{-s^5}(\Phi_t h - \Pi_{N_k} g)\|_{Z_1} \leq \|\langle \partial_x \rangle^{-s^5}(\Phi_t h - \Phi_t^{N_k}\Pi_{N_k} h)\|_{Z_1} \\ + \|\langle \partial_x \rangle^{-s^5}(\Phi_t^{N_k}\Pi_{N_k} h - \Phi_t^{N_k}\Pi_{N_k} h^{N_k})\|_{Z_1} \to 0.\end{aligned}$$

This implies $g = \Phi_t h \in \Phi_t(E)$, so the proof is complete. □

13.4. Modified continuity

In this section we prove Theorem 1.4; note that this modified continuity statement is not needed in the proof of Theorems 1.2 and 1.3.

To prove part (1), noting that $u \in \mathcal{BO}^T$, we find that

$$u^* \in Y_2^T \subset C_t^0([-T, T] \to Z_1)$$

using the notation in Proposition 7.3. Recall from (7.28) that $\Delta_n(t) = \int_0^t \delta_n(t')\, dt'$ and

$$\delta_n(t) = \frac{1}{2}\sum_{k=1}^{n} |w_k|^2 = \frac{1}{2}\sum_{k=1}^{n} |u_k|^2 + R$$

for $n > 0$, where R does not grow with n (this is easy using $w = \mathbb{P}_+(Mu)$ and the assumptions about u). Now, if it were not for the logarithmic factor on the right hand side of (8.16) which these factors satisfy, $\Delta_n(t)$ would be continuous in t *uniformly* in n, and u would be in $C_t^0([-T, T] \to Z_1)$; this shows that we may disregard R and pretend that Δ_n is defined as in (1.7), and this proves part (1).

For part (2) we need another probabilistic argument. Recall that $u = \Phi f = (\Phi_t f)_t$ is defined for $f \in \Sigma$ which is equipped with the Gaussian measure ρ. In order to use part (1) we just proved, we will define

$$\widetilde{\delta_n}(t) = \sum_{k=1}^{n}\left(|u_k(t)|^2 - \frac{1}{4\pi k}\right) = \sum_{d \leq \lfloor \log_2 n \rfloor} \widetilde{\delta_{(d)}}(t) + R,$$

where

$$\widetilde{\delta_{(d)}}(t) = \sum_{0<k\sim 2^d}\left(|u_k(t)|^2 - \frac{1}{4\pi k}\right), \tag{13.23}$$

and $\widetilde{\Delta_n}$ similarly, where R is already bounded in n and can be neglected. We only need to prove continuity in any interval $[-T, T]$; for simplicity assume $T = 1$. If we define

$$Y_{(d)} = \int_{-1}^{1} |\widetilde{\delta_{(d)}}(t)|^2\, dt$$

as a random variable on Σ for each d, then part (2) will follow if we can show that

$$\limsup_{d\to\infty} 2^{d/2} Y_{(d)} \le 1 \tag{13.24}$$

for ρ-almost all $f \in \Sigma$. Fix one Gibbs measure ν; we have

$$\mathbb{E}_\nu(2^d Y_{(d)}) \lesssim \int_{-1}^{1} \big[\mathbb{E}_\nu\big(\exp\big(2^{d/2}\widetilde{\delta_{(d)}}(t)\big)\big) + \mathbb{E}_\nu\big(\exp(-2^{d/2}\widetilde{\delta_{(d)}}(t))\big)\big]\,dt. \tag{13.25}$$

Now *using the invariance of* ν, we only need to consider $t = 0$; also we will study only the first term. Since $(\mathbb{E}_\nu H)^2 \lesssim \mathbb{E}_\rho H^2$ by Cauchy–Schwarz, this is bounded by

$$\mathbb{E}_\omega\bigg(\exp\bigg(2^{d/2}\sum_{0<k\sim 2^d}\frac{|g_k(\omega)|^2-1}{2\pi k}\bigg)\bigg),$$

which can be easily computed and is $O(1)$ due to our choice of parameters. Then (13.24) follows by standard measure-theoretic arguments (for any ν, and thus for ρ).

Acknowledgments. The author would like to thank Alexandru Ionescu for his encouragement and constructive suggestions regarding the gauge transform; he would also like to thank Tadahiro Oh for introducing the problem and suggesting the treatment used in [5], and Xuecheng Wang for helpful discussions about the proof of Proposition 3.6.

References

[1] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. Geom. Funct. Anal. **3**, 209–262 (1993) Zbl 0787.35098 MR 1215780

[2] Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. Comm. Math. Phys. **166**, 1–26 (1994) Zbl 0822.35126 MR 1309539

[3] Bourgain, J.: On the Cauchy and invariant measure problem for the periodic Zakharov system. Duke Math. J. **76**, 175–202 (1994) Zbl 0821.35120 MR 1301190

[4] Bourgain, J.: Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. Comm. Math. Phys. **176**, 421–445 (1996) Zbl 0852.35131 MR 1374420

[5] Bourgain, J.: Periodic Korteweg–de Vries equation with measures as initial data. Selecta Math. (N.S.) **3**, 115–159 (1997) Zbl 0891.35138 MR 1466164

[6] Bourgain, J.: Nonlinear Schrödinger equations. In: Hyperbolic Equations and Frequency Interactions (Park City, UT, 1995), IAS/Park City Math. Ser. 5, Amer. Math. Soc., 3–157 (1999) Zbl 0952.35127 MR 1662829

[7] Bourgain, J.: Global Solutions of Nonlinear Schrödinger Equations. Amer. Math. Soc. Colloq. Publ. 46, Amer. Math. Soc. (1999) Zbl 0933.35178 MR 1691575

[8] Burq, N., Planchon, F.: On well-posedness for the Benjamin–Ono equation. Math. Ann. **340**, 497–542 (2008) Zbl 1148.35074 MR 2357995

[9] Burq, N., Thomann, L., Tzvetkov, N.: Long time dynamics for the one dimensional nonlinear Schrödinger equation. Ann. Inst. Fourier (Grenoble) **63**, 2137–2198 (2013) Zbl 06325429 MR 3237443

[10] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. I. Local theory. Invent. Math. **173**, 449–475 (2008) Zbl 1156.35062 MR 2425133

[11] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. II. A global existence result. Invent. Math. **173**, 477–496 (2008) Zbl 1187.35233 MR 2425134

[12] Burq, N., Tzvetkov, N.: Probabilistic well-posedness for the cubic wave equation. J. Eur. Math. Soc. **16**, 1–30 (2014) Zbl 1295.35387 MR 3141727

[13] Colliander, J., Oh, T.: Almost sure well-posedness of the periodic cubic nonlinear Schrödinger equation below $L^2(\mathbb{T})$. Duke Math. J. **161**, 367–414 (2012) Zbl 1260.35199 MR 2881226

[14] Deng, Y.: Two dimensional NLS equation with random radial data. Anal. PDE **5**, 913–960 (2012) Zbl 1264.35212 MR 3022846

[15] Deng, Y., Tzvetkov, N., Visciglia, N.: Invariant measures and long time behavior for the Benjamin–Ono equation III. Comm. Math. Phys., submitted

[16] Ionescu, A., Kenig, C.: Global well-posedness of the Benjamin–Ono equation in low-regularity spaces. J. Amer. Math. Soc. **20**, 753–798 (2007) Zbl 1123.35055 MR 2291918

[17] Iorio, J.: On the Cauchy problem for the Benjamin–Ono equation. Comm. Partial Differential Equations **11**, 1031–1081 (1986) Zbl 0608.35030 MR 1104110

[18] Lebowitz, J., Rose, R., Speer, E.: Statistical mechanics of the nonlinear Schrödinger equation. J. Statist. Phys. **50**, 657–687 (1988) Zbl 1084.82506 MR 0939505

[19] Molinet, L.: Global well-posedness in the energy space for the Benjamin–Ono equation on the circle. Math. Ann. **337**, 353–383 (2007) Zbl 1140.35001 MR 2262788

[20] Molinet, L.: Global well-posedness in L^2 for the periodic Benjamin–Ono equation. Amer. J. Math. **130**, 635–683 (2008) Zbl 1157.35001 MR 2418924

[21] Molinet, L., Ribaud, F.: Well-posedness in H^1 for generalized Benjamin–Ono equations on the circle. Discrete Contin. Dynam. Systems **23**, 1295–1311 (2009) Zbl 1165.35302 MR 2461852

[22] Nahmod, A., Oh, T., Rey-Bellet, L., Staffilani, G.: Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS. J. Eur. Math. Soc. **14**, 1275–1330 (2012) Zbl 1251.35151 MR 2928851

[23] Oh, T.: White noise for KdV and mKdV on the circle. In: Harmonic Analysis and Nonlinear Partial Differential Equations, RIMS Kôkyûroku Bessatsu B18, Res. Inst. Math. Sci., Kyoto, 99–124 (2010) Zbl 1211.35243 MR 2762393

[24] Tao, T.: Global well-posedness of the Benjamin–Ono equation in $H^1(\mathbb{R})$. J. Hyperbolic Differential Equations **1**, 27–49 (2004) Zbl 1055.35104 MR 2052470

[25] Tao, T.: Nonlinear Dispersive Equations: Local and Global Analysis. CBMS Reg. Conf. Ser. Math. 106, Amer. Math. Soc., Providence, RI (2006) Zbl 1106.35001 MR 2233925

[26] Tzvetkov, N.: Invariant measures for the nonlinear Schrödinger equation on the disc. Dynam. Partial Differential Equations **3**, 111–160 (2006) Zbl 1142.35090 MR 2227040

[27] Tzvetkov, N.: Invariant measures for the defocusing nonlinear Schrödinger equation. Ann. Inst. Fourier (Grenoble) **58**, 2543–2604 (2008) Zbl 1171.35116 MR 2498359

[28] Tzvetkov, N.: Construction of a Gibbs measure associated to the periodic Benjamin–Ono equation. Probab. Theory Relat. Fields **146**, 481–514 (2010) Zbl 1188.35183 MR 2574736

[29] Tzvetkov, N., Visciglia, N.: Gaussian measures associated to the higher order conservation laws of the Benjamin–Ono equation. Ann. Sci. École Norm. Sup. (4) **46**, 249–299 (2013) Zbl 06216387 MR 3112200

[30] Tzvetkov, N., Visciglia, N.: Invariant measures and long time behaviour for the Benjamin–Ono equation. Int. Math. Res. Notices **2014**, 4679–4714 (2014) Zbl 1301.35141 MR 3257548

[31] Tzvetkov, N., Visciglia, N.: Invariant measures and long time behaviour for the Benjamin–Ono equation II. J. Math. Pures Appl. (9) **103**, 102–141 (2015) Zbl 06374914 MR 3281949