SYMPLECTIC NON-SQUEEZING OF THE KDV FLOW

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

Abstract. We prove two finite dimensional approximation results and a symplectic non-squeezing property for the Korteweg-de Vries (KdV) flow on the circle $\mathbb{T}$. The nonsqueezing result relies on the aforementioned approximations and the finite-dimensional nonsqueezing theorem of Gromov [13]. Unlike the work of Kuksin [21] which initiated the investigation of non-squeezing results for infinite dimensional Hamiltonian systems, the nonsqueezing argument here does not construct a capacity directly. In this way our results are similar to those obtained for the NLS flow by Bourgain [3]. A major difficulty here though is the lack of any sort of smoothing estimate which would allow us to easily approximate the infinite dimensional KdV flow by a finite-dimensional Hamiltonian flow. To resolve this problem we invert the Miura transform and work on the level of the modified KdV (mKdV) equation, for which smoothing estimates can be established.

Contents

J.C. was supported in part by N.S.E.R.C. Grant RGPIN 250233-03 and the Sloan Foundation.
M.K. was supported in part by N.S.F. Grant DMS 9801558.
G.S. was supported in part by N.S.F. Grant DMS 0100345 and by a grant from the Sloan Foundation.
H.T. was supported in part by J.S.P.S. Grant No. 13740087.
T.T. was a Clay Prize Fellow and was supported in part by grants from the Packard Foundation.

1. Introduction

This paper is concerned with the symplectic behavior of the Korteweg-de Vries (KdV) flow

$$u_t + u_{xxx} = 6uu_x; \quad u(0,x) = u_0(x) \tag{1.1}$$

on the circle $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, where $u(t,x)$ is real-valued. In particular we investigate how the flows may (or may not) be accurately approximated by certain finite-dimensional models, and then use such an approximation to conclude a symplectic non-squeezing property. In order to describe the symplectic space involved, and state the result precisely, we need to set notation and recall some previous results describing the well-posedness of the initial value problem (1.1).

On the circle we have the spatial Fourier transform

$$\widehat{u}(k) := \frac{1}{2\pi}\int_0^{2\pi} u(x)\exp(-ikx)\, dx \tag{1.2}$$

for all $k \in \mathbb{Z}$, and the spatial Sobolev spaces

$$\|u\|_{H^s_x} := (2\pi)^{1/2}\|\langle k\rangle^s \widehat{u}\|_{l^2_k}$$

for $s \in \mathbb{R}$, where $\langle k\rangle := (1+|k|^2)^{1/2}$. These are natural spaces for analyzing the KdV flow.

Let P_0 denote the mean operator

$$P_0 u := \frac{1}{2\pi}\int_0^{2\pi} u$$

or equivalently

$$\widehat{P_0 u}(k) = \chi_{k=0}\widehat{u}(k).$$

The KdV flow is mean-preserving, and it will be convenient to work in the case when u has mean zero[1]. Accordingly we define the mean-zero periodic Sobolev spaces H_0^s by

$$H_0^s := \{u \in H_x^s : P_0 u = 0\}$$

endowed with the same norm as H_x^s.

Recent work on the local and global well-posedness theory in H_0^s for (1.1) is basic to our results here. For example, the geometric conclusions from finite-dimensional Hamiltonian dynamics which we ultimately need for our nonsqueezing result can only be applied in the setting of rather rough solutions to the initial value problem (1.1). We now pause to summarize some of the analytical techniques that have been developed for the study of such rough solutions, and the resulting regularity theory (see e.g. [1], [18], [6], and [8], [9]).

1.1. **Summary of local and global well-posedness theory.** If the initial datum u_0 for (1.1) is smooth, then there is a global smooth solution[2] $u(t)$ (see e.g. [26]). We can thus define the non-linear flow map $S_{KdV}(t)$ on $C^\infty(\mathbb{T})$ by $S_{KdV}(t)u_0 := u(t)$. In particular this map is densely defined on every Sobolev space H_0^s.

If $s \geq -1/2$, then the equation (1.1) is globally well-posed in H_0^s. In other words, the flow map $S_{KdV}(t)$ is uniformly continuous (indeed, it is analytic) on H_0^s for times t restricted to a compact interval $[-T, T]$, and for such s we have bounds of the form

$$\sup_{|t| \leq T} \|S_{KdV}(t)u_0\|_{H_0^s} \leq C(s, T, \|u_0\|_{H_0^s}), \tag{1.3}$$

(see [18], [8], [9] (and also Section 9.1 below)). For $s < -1/2$ the flow map $S_{KdV}(t)$ is no longer uniformly continuous [6] (see also [19]) or analytic [4], so from the point of view which requires a uniformly continuous flow in time, the Sobolev space $H_0^{-1/2}$ is the endpoint space for the KdV flow. Coincidentally, this space is also a natural phase space for which KdV becomes a Hamiltonian flow; we will have more to say about this at the end of the introduction. Note however that if one asks only that the flow be continuous in time, then global well-posedness for (1.1) has been established for all $s \geq -1$ in [16] using inverse scattering methods. Combining mapping properties of the Miura Transform and the result in [27], local well-posedness of (1.1) in H_0^s with a (not uniformly) continuous flow map holds for $-5/8 < s < -1/2$.

To obtain many of the local and global well-posedness results mentioned above, one iterates in a certain spacetime Banach space Y^s (defined in (3.1) below; this space is a variant of the $X^{s,b}$ spaces used for instance in [1], [18]), which has the

[1] One can easily pass from the mean zero case to the general mean case by a Galilean transformation $u(t,x) \to u(t, x - P_0(u)t) - P_0(u)$.

[2] This result can also be obtained by inverse scattering methods, since the KdV equation is completely integrable. However, our methods here do not use inverse scattering techniques, although the special algebraic structure of KdV (in particular, the Miura transform [24]) is certainly exploited.

same regularity as H^s in the sense that one has the embedding[3]

$$\|u\|_{L^\infty_t H^s_x} \lesssim \|u\|_{Y^s}.$$

The nonlinearity is then placed in a companion space Z^s (see (3.2) below), which is related to Y^s via an energy estimate of the form

$$\|\eta(t)u\|_{Y^s} \lesssim \|u(t_0)\|_{H^s} + \|u_t + u_{xxx}\|_{Z^s},$$

for any time t_0, and any bump function η supported near t_0. (We will elaborate more upon these spaces and estimates in Section 3). The local well-posedness theory[4] for the KdV equation (1.1) then hinges on the bilinear estimate

$$\|(uv)_x\|_{Z^s} \lesssim \|u\|_{Y^s}\|v\|_{Y^s} \tag{1.4}$$

whenever u, v are mean-zero functions and $s \geq -1/2$ (see [18], [8], [9]).

To pass from local well-posedness to global well-posedness one needs to obtain long-time bounds on the H^s_0 norm. For $-\frac{1}{2} \leq s < 0$, this has been achieved by means of the "I-method", constructing an almost conserved quantity comparable to the H^s norm; see [8], [9], or Section 9.1.

1.2. **Low frequency approximation of KdV.** The KdV flow (1.1) is, formally at least, a Hamiltonian flow on an infinite-dimensional space. In order to rigorously apply results from symplectic geometry, we must approximate this infinite-dimensional flow by a finite-dimensional flow. Furthermore, in order to apply these geometric tools, we need that the finite-dimensional flow is itself Hamiltonian.

We begin with a negative result. Suppose that we wish to study the KdV flow for data u_0 whose Fourier transform is supported on $[-N, N]$ for some large fixed N, and specifically to approximate the KdV flow by a finite-dimensional model. A first guess for such a model might be the flow

$$u_t + u_{xxx} = P_{\leq N}(6uu_x); \quad u(0) = u_0, \tag{1.5}$$

where $P_{\leq N}$ is the Fourier projection to frequencies $\leq N$:

$$\widehat{P_{\leq N}u}(k) = \chi_{|k|\leq N}\widehat{u}(k).$$

Denote the flow map associated to (1.5) by $S_{P_{\leq N}KdV}(t)$. This flow has several advantageous properties; for instance, $S_{P_{\leq N}KdV}(t)$ is a symplectomorphism on the space $P_{\leq N}H_0^{-1/2}$, associated with a natural symplectic structure (see next subsection). Since $P_{\leq N}H_0^{-1/2}$ is a finite dimensional space, it is easy to see (e.g. using L^2 norm conservation and Picard iteration) that this flow $S_{P_{\leq N}KdV}$ is globally smooth and well-defined. In [3], the NLS flow $iu_t + u_{xx} = |u|^2u$ was similarly truncated, and it was shown that the truncated flow was a good approximation to the original (infinite dimensional) flow. Unfortunately, the same result does not apply for KdV:

[3] In this paper we use $A \lesssim B$ to denote an estimate of the form $A \leq CB$, where the implicit constant C may depend on certain parameters such as s which we will specify later in the paper. Similarly, $A \ll B$ denotes $B \geq CA$ for some such universal constant C.

[4] Strictly speaking, in order to handle large initial data one must also generalize this estimate to circles $\mathbb{R}/2\pi\lambda\mathbb{Z}$ of arbitrarily large period, in order to apply rescaling arguments to make the data small again. See [8], [9], or Section 9.1.

Theorem 1.1. *Let $k_0 \in \mathbb{Z}^*$, $T > 0$, $A > 0$. Then for any $N \gg C(A, T, k_0)$ there exists initial data u_0 with $\|u_0\|_{H_0^{-1/2}} \leq A$ and supp $(\widehat{u_0}) \subset \{|k| \leq N\}$ such that*

$$|(\widehat{S_{KdV}(T)}u_0)(k_0) - (\widehat{S_{P_{\leq N}KdV}(T)}u_0)(k_0)| \geq c(T, A, k_0) \tag{1.6}$$

for some $c(T, A, k_0) > 0$.

In other words, $S_{P_{\leq N}KdV}$ does not converge to S_{KdV} even in a weak topology.

We prove this negative result in Section 8. Basically, the problem is that the multiplier $\chi_{[-N,N]}$ corresponding to $P_{\leq N}$ is very rough, and this creates significant deviations between S_{KdV} and $S_{P_{\leq N}KdV}$ near the Fourier modes $k = \pm N$. In cubic equations such as mKdV (see (1.9) below) or the cubic nonlinear Schrödinger equation, these deviations would stay near the high frequencies $\pm N$, but in the quadratic KdV equation these deviations create significant fluctuations near the frequency origin, eventually leading to failure of weak convergence in (1.6).

Of course there are several obvious ways to modify the finite-dimensional flow (1.5) in an attempt to find an effective approximation to the KdV flow for data with Fourier transform supported on $[-N, N]$, but at least a little bit of care is needed when considering these modifications. We let $b(k)$ be the restriction to the integers of a real even bump function adapted to $[-N, N]$ which equals 1 on $[-N/2, N/2]$, and consider the evolution

$$u_t + u_{xxx} = B(6uu_x); \quad u(0) = u_0 \tag{1.7}$$

where

$$\widehat{Bu}(k) = b(k)\widehat{u}(k).$$

Let S_{BKdV} denote the flow map associated to (1.7). Observe that this is a finite-dimensional flow on the space $P_{\leq N}H_0^s$. Unfortunately, S_{BKdV} is not a symplectomorphism, but we will explain in (1.27) below how by conjugating a flow of the form (1.7) with a simple multiplier operator we will arrive at our desired finite dimensional symplectomorphism on $P_{\leq N}H^{-\frac{1}{2}}(\mathbb{T})$ that well-approximates the full KdV flow at low frequencies. This desired symplectomorphism is labelled $S_{KdV}^{(N)}(t)$ in (1.27) below[5], and once the aforementioned approximation properties are established, the nonsqueezing result will follow almost immediately after quoting the finite dimensional nonsqueezing result of Gromov [13].

The first step in the argument is to show we can approximate S_{KdV} by S_{BKdV} in the strong H_0^s topology:

Theorem 1.2. *Fix $s \geq -1/2$, $T > 0$, and $N \gg 1$. Let $u_0 \in H_0^s$ have Fourier transform supported in the range $|k| \leq N$. Then*

$$\sup_{|t| \leq T} \|P_{\leq N^{1/2}}(S_{BKdV}u_0(t) - S_{KdV}(t)u_0)\|_{H_0^s} \leq N^{-\sigma}C(s, T, \|u_0\|_{H_0^s})$$

for some $\sigma = \sigma(s) > 0$.

[5]The equation which defines this flow is given in (7.1) below.

In particular, we can accurately model the KdV evolution for band-limited initial data by a finite-dimensional flow, at least for frequencies $|k| \leq N^{1/2}$.

The well-posedness statement (1.3) gives Theorem 1.2 for all $0 \leq N \leq C(s, T, \|u_0\|_{H_0^s})$, hence our proof needs only to consider $N \geq C(s, T, \|u_0\|_{H_0^s})$. This turns out to be the most interesting case from the point of view of the nonsqueezing applications of this approximation theorem which we take up below.

Theorem 1.2 can be viewed as a statement that one can (smoothly) truncate the KdV evolution at the high frequencies without causing serious disruption to the low frequencies, in spite of the obstruction posed by Theorem 1.1. Our second main result (proven in Section 5) is in a similar vein:

Theorem 1.3. *Fix $s \geq -1/2$, $T > 0, N \geq 1$. Let $u_0, \tilde{u}_0 \in H_0^s$ be such that $P_{\leq 2N} u_0 = P_{\leq 2N} \tilde{u}_0$ (i.e. u_0 and $\tilde{u}_0$ agree at low frequencies). Then we have,*

$$\sup_{|t| \leq T} \|P_{\leq N}(S_{KdV}(t)\tilde{u}_0 - S_{KdV}(t)u_0)\|_{H_0^s} \leq N^{-\sigma} C(s, T, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$$

for some $\sigma = \sigma(s) > 0$.

By the same reasoning made following Theorem 1.2, we may assume in the proof of Theorem 1.3 that $N \geq C(s, T, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$.

The point of Theorem 1.3 is that changes to the initial data at frequencies $\geq 2N$ do not significantly affect the solution at frequencies $\leq N$, as measured in the strong H_0^s topology. This is in stark contrast to the negative result in Theorem 1.1. The point is that there is some delicate cancellative structure in the KdV equation which permits the decoupling of high and low frequencies, and this structure is destroyed by projecting the KdV equation crudely using (1.5).

To prove Theorem 1.2 and Theorem 1.3, we shall need to exploit the subtle cancellation mentioned in the previous paragraph in order to avoid the obstructions arising from Theorem 1.1. We do not know how to do this working directly with the KdV flow. Rather, we are able to prove estimates which explicitly account for this subtle structure in KdV by using the *Miura transform* $u = \mathbf{M}v$, defined by

$$u = \mathbf{M}v := v_x + v^2 - P_0(v^2). \tag{1.8}$$

As discovered in [24], this transform allows us to conjugate the KdV flow to the *modified Korteweg-de Vries* (mKdV) flow

$$v_t + v_{xxx} = F(v); \quad v(x, 0) = v_0(x) \tag{1.9}$$

where the non-linearity $F(v)$ is given by

$$F(v) := 6(v^2 - P_0(v^2))v_x. \tag{1.10}$$

The modified KdV equation has slightly better smoothing properties[6] than the ordinary KdV equation, and in addition the process of inverting the Miura transform adds one degree of regularity (from $H_0^{-1/2}$ to $H_0^{1/2}$). In particular, the types of counterexamples arising in Theorem 1.1 do not appear in the mKdV setting, and

[6]See Section 4, in particular Theorem 4.3.

by proving a slightly more refined trillinear estimate than those found in e.g. [9] (see in particular Theorem 4.3 below) we are able to prove the above two theorems by passing to the mKdV setting using the Miura transform. Of course, in order to close the argument we will need some efficient estimates on the invertibility of the Miura transform; we set up these estimates (which may be of independent interest) in Section 2.

1.3. **Application to symplectic non-squeezing.** We can apply the above approximation results to study the symplectic behavior of KdV in a natural phase space $H_0^{-1/2}(\mathbb{T})$. Before doing so, we recall some context and results from previous works. We are following here especially the exposition from [15, 22].

Definition 1.4. *Consider a pair* $(\mathbb{H}, \omega)$ *where* ω *is a symplectic form*[7] *on the Hilbert space* $\mathbb{H}$. *We say* $(\mathbb{H}, \omega)$ *is the symplectic phase space of a PDE with Hamiltonian* $H[u(t)]$ *if the PDE can be written in the form,*

$$\dot{u}(t) = J\nabla H[u(t)]. \tag{1.11}$$

Here J is an almost complex structure[8] on $\mathbb{H}$, which is compatible with the Hilbert space inner product $\langle \cdot, \cdot \rangle$. That is, for all $u, v \in \mathbb{H}$,

$$\omega(u, v) = \langle Ju, v \rangle. \tag{1.12}$$

The notation ∇ in (1.11) denotes the usual gradient with respect to the Hilbert space inner product,

$$\langle v, \nabla H[u] \rangle \equiv dH[u](v) \tag{1.13}$$

$$\equiv \frac{d}{d\epsilon}\Big|_{\epsilon=0} H[u + \epsilon v]. \tag{1.14}$$

One easily checks that an equivalent way to write the PDE corresponding to the Hamiltonian $H[u(t)]$ in $(\mathbb{H}, \omega)$ is

$$\dot{u}(t) = \nabla_\omega H[u(t)] \tag{1.15}$$

where the symplectic gradient $\nabla_\omega H[u]$ is defined in analogy with (1.13),

$$\omega(v, \nabla_\omega H[u]) = dH[u](v). \tag{1.16}$$

For example, on the Hilbert space $H_0^{-\frac{1}{2}}(\mathbb{T})$, we can define the symplectic form

$$\omega_{-\frac{1}{2}}(u, v) := \int_{\mathbb{T}} u(x) \partial_x^{-1} v(x)\, dx \tag{1.17}$$

where $\partial_x^{-1} : H_0^{-1/2}(\mathbb{T}) \to H_0^{1/2}(\mathbb{T})$ is the inverse to the differential operator ∂_x defined via the Fourier transform by

$$\widehat{\partial_x^{-1} f}(k) := \frac{1}{ik} \widehat{f}(k).$$

---

[7]That is, a nondegenerate, antisymmetric form $\omega : \mathbb{H} \times \mathbb{H} \to \mathbf{C}$. We identify in the usual way $\mathbb{H}$ and it's tangent space $T_x\mathbb{H}$ for each $x \in \mathbb{H}$.

[8]That is, a bounded, anti-selfadjoint operator with $J^2 = -(\text{identity})$.

The KdV flow (1.1) is then *formally* the Hamiltonian equation in $(H_0^{-1/2}(\mathbb{T}), \omega_{-\frac{1}{2}})$ corresponding to the (densely defined) Hamiltonian

$$H[u] := \int_{\mathbb{T}} \frac{1}{2} u_x^2 + u^3 dx. \tag{1.18}$$

Indeed, working formally[9] we have for any $v \in H_0^{-\frac{1}{2}}(\mathbb{T})$,

$$\begin{aligned}
\frac{d}{d\epsilon}\Big|_{\epsilon=0} H[u+\epsilon v] &= \int_{\mathbb{T}} u_x v_x + 3u^2 v dx \\
&= \int_{\mathbb{T}} (-u_{xx} + 3u^2) v dx \\
&= \int_{\mathbb{T}} \partial_x^{-1}(-u_{xxx} + 6uu_x) v dx \\
&= -\int_{\mathbb{T}} (-u_{xxx} + 6uu_x) \partial_x^{-1} v dx \\
&= \omega_{-\frac{1}{2}}(u_{xxx} - 6uu_x, v) \\
&= \omega_{-\frac{1}{2}}(v, -u_{xxx} + 6uu_x).
\end{aligned}$$

Comparing (1.15)-(1.16) with (1.1), we see KdV is indeed the Hamiltonian PDE corresponding to $H[u]$ on the infinite dimensional symplectic space $(H_0^{-\frac{1}{2}}, \omega_{-\frac{1}{2}})$. In particular, the flow maps $S_{KdV}(t)$ are, formally, symplectomorphisms on $H_0^{-1/2}(\mathbb{T})$.

That the KdV flow arises as a Hamiltonian flow from a symplectic structure as described above was discovered by Gardner and Zakharov-Faddeev (see [12, 31]). A second structure was given by Magri [23] using $\int u^2 dx$ as Hamiltonian, but it is not as convenient as the first structure for our strategy to prove nonsqueezing. Roughly speaking, it seems the symplectic form in this second structure could possibly be used to establish a nonsqueezing property - in the $H^{-\frac{3}{2}}$ topology - of a finite dimensional analog of (1.1). However, since the well-posedness theory, and the accompanying estimates, for the full KdV flow do not presently exist at such rough norms, we do not see how we could approximate the full KdV flow in a space as rough as $H^{-\frac{3}{2}}$ with a finite dimensional flow. The first structure described above allows us to adopt this strategy in the space $H_0^{-\frac{1}{2}}$, within which we do have well-posedness. (See below for references for this approach to proving nonsqueezing for PDE. See e.g [25, 11] for more details and history of the various symplectic structures for KdV.)

For any $u_* \in H_0^{-1/2}(\mathbb{T})$, $r > 0$, $k_0 \in \mathbb{Z}^*$, and $z \in \mathbf{C}$, we consider the infinite-dimensional ball

$$\mathbf{B}^\infty(u_*; r) := \{u \in H_0^{-1/2}(\mathbb{T}) : \|u - u_*\|_{H_0^{-1/2}} \leq r\}$$

and the infinite-dimensional cylinder

$$\mathbf{C}_{k_0}^\infty(z; r) := \{u \in H_0^{-1/2}(\mathbb{T}) : |k_0|^{-1/2} |\widehat{u}(k_0) - z| \leq r\}.$$

[9] By the word 'formally', we mean here that no attempt is made to justify various differentiations or integration by parts. Later, when we localize the space $H_0^{-\frac{1}{2}}$ and Hamiltonian in frequency and write down the corresponding equations, the reader can carry out the analogous computation where the justification of the necessary calculus will be evident.

The final result of this paper is the following symplectic non-squeezing theorem,

Theorem 1.5. *Let* $0 < r < R$, $u_* \in H_0^{-1/2}(\mathbb{T})$, $k_0 \in \mathbb{Z}^*$, $z \in \mathbf{C}$, *and* $T > 0$. *Then*

$$S_{KdV}(T)(\mathbf{B}^\infty(u_*; R)) \not\subseteq \mathbf{C}^\infty_{k_0}(z; r).$$

In other words, there exists a global $H_0^{-1/2}(\mathbb{T})$ *solution* u *to* (1.1) *such that*

$$\|u(0) - u_*\|_{H_0^{-1/2}} \leq R$$

and

$$|k_0|^{-1/2} |\widehat{u(T)}(k_0) - z| > r.$$

Note that no smallness conditions are imposed on u_*, R, z, or T.

Roughly speaking, this Theorem asserts that the KdV flow cannot squash a large ball into a thin cylinder. Notice that the balls and cylinders can be arbitrarily far away from the origin, and the time T can also be arbitrary. Note though that this result is interesting even for $u_* = 0, z = 0$ and smooth initial data u_0, as it tells us that the flow cannot at any time uniformly squeeze the ball $\mathbf{B}^\infty(0; R)$ even at a fixed frequency k_0. By Theorem (1.5), the well-posedness theory for KdV reviewed above, and density considerations,we know that for any $T, r < R$, there will be some initial data $u_0 \in \mathbf{B}^\infty(0; R)$ for which[10] $|\widehat{u}(k_0, T)| > |k_0|^{\frac{1}{2}} r$. (See [5], page 96 for the same discussion in the context of a nonlinear Klein-Gordon equation.) A second immediate application of Theorem 1.5 to smooth solutions was highlighted in a different context already in [21], namely that such smooth solutions of (1.1) cannot uniformly approach some asymptotic state: for any neighborhood $\mathbf{B}^\infty(u_0; R)$ of the initial data in $H^{-\frac{1}{2}}(\mathbb{T})$ and for any time t, the diameter of the set $S_{KdV}(t)(\mathbf{B}^\infty(u_0; R))$ cannot be less than R.

The motivation for Theorem 1.5, and an important component of its proof, is the finite-dimensional nonsqueezing theorem of Gromov [13] (see also subsequent extensions in [14], [15]). The extension to the infinite-dimensional setting provided by a nonlinear PDE seems nontrivial. The program was initiated by Kuksin [21], [22] for certain equations where the nonlinear flow is a compact perturbation of the linear flow. That the KdV equation doesn't meet this requirement can be seen by an argument involving simple computations similar to those supporting Theorem 1.1 which are detailed in Section 8 below: fix $\sigma \ll 1$ and for each integer $N \geq 1$ consider initial data,

$$u_{0,N}(x) := \sigma N^{\frac{1}{2}} \cos(Nx).$$

Clearly the set $\{u_{0,N} : N = 1, 2, \ldots\}$ is bounded in $H_0^{-\frac{1}{2}}$. However, when one computes the second iterate[11] $u_N^{[2]}$ one sees that it differs from the linear evolution

[10] We are using here the statement of the Theorem only in the case $u_* = 0, z = 0$. Of course one gets a similar conclusion to the one we draw here, but with different weights and a different initial data set, by simply using the L^2 conservation and time reversability properties of the flow. That is, for any $R > r$, there is data $\tilde{u}_0 \in \{\|f\|_{L^2(\mathbb{T})} \leq R\}$ such that the evolution $\tilde{u}$ of this data satisfies $|\widehat{\tilde{u}}(k_0, T)| > r$.

[11] See in particular equation (8.2) for the notation used here, and if necessary Section 8 for what we hope is a sufficiently detailed discussion to allow the reader to reproduce the elementary computations we quote here.

of $\widehat{u_N^{[0]}}$ at frequency $k = N$ in that,

$$\widehat{u_N^{[2]}}(N,t) - \widehat{u_N^{[0]}}(N,t) \sim N^{\frac{1}{2}}\sigma^3 e^{iN^3 t}. \tag{1.19}$$

By the local well-posedness theory we know, assuming σ is sufficiently small compared to t, that the difference between the second iterate and the actual nonlinear evolution $u_N(t)$ of the data $u_{0,N}$ satisfies,

$$\|u_N(t) - u_N^{[2]}(t)\|_{H_0^{-\frac{1}{2}}(\mathbb{T})} \lesssim \sigma^4. \tag{1.20}$$

Together, (1.19) and (1.20) show that if $\{N_k\}$ is a sequence of integers relatively prime to one-another[12], then

$$\widehat{u_{N_k}}(N_l,t) - \widehat{u_{N_k}^{[0]}}(N_l,t) \sim \delta_{k,l}\cdot\sigma^3\cdot N_k^{\frac{1}{2}} e^{iN_k^3 t}.$$

Hence the set $\{u_{N_k}(t) - u_{N_k}^{[0]}(t)\}$ has no limit point in $H_0^{-\frac{1}{2}}(\mathbb{T})$.

The nonsqueezing results of Kuksin were extended to certain stronger nonlinearities by Bourgain [3, 5] - for instance [3] treats the cubic non-linear Schrödinger flow on $L^2(\mathbb{T})$. In these works, the full solution map is shown to be well-approximated by a finite dimensional flow constructed by cutting the solution off to frequencies $|k| \leq N$ for some large N. The nonsqueezing results in [3, 5] follow then from a direct application of Gromov's finite dimensional nonsqueezing result to this approximate flow.

The argument we follow here for the KdV flow is similar to the work in [3, 5], but seems to require a bit more care. The complication seems to us to be somehow rooted in the counterexample of Theorem 1.1, which clearly exhibits that a sharp cut-off is not appropriate in constructing the approximating flow, but which seems also to be subtly related to the fact that the estimates necessary to approximate the full KdV flow by a more gradually truncated flow are unavailable to us when we work directly with the KdV equation. We have already sketched how we will deal with this difficulty (that is, by passing to the modified KdV equation) in the discussion which followed Theorem 1.3 above.

We now provide some details of the previous paragraph's sketch, in particular we indicate the difficulties that arise when one tries to repeat the argument in [3, 5].

Let $N \geq 1$ be an integer. By simply restricting the form $\omega_{-\frac{1}{2}}$, the space $(P_{\leq N} H_0^{-1/2}(\mathbb{T}), \omega_{-\frac{1}{2}})$ is a $2N$-dimensional real symplectic space and hence by general arguments (see e.g. Proposition 1 in [15]) is symplectomorphic to the standard space $(\mathbb{R}^{2N}, \omega_0)$. We will make explicit use of such an equivalence below: any $u \in P_{\leq N} H_0^{-1/2}(\mathbb{T})$ is determined completely by

$$\begin{aligned}(\mathrm{Re}(\widehat{u}(1)), \ldots, \mathrm{Re}(\widehat{u}(N)), \mathrm{Im}(\widehat{u}(1)), \ldots, \mathrm{Im}(\widehat{u}(N)))\\ \equiv (e_1(u), \ldots, e_n(u), f_1(u), \ldots, f_N(u)) \in \mathbb{R}^{2N}.\end{aligned} \tag{1.21}$$

[12]Note (for example by examining the iterates and using well-posedness) that $\widehat{u}_N(t)$ is supported only at frequencies which are integer multiples of N.

In terms of the coordinates (1.21) the form $\omega_{-\frac{1}{2}}$ defined in (1.17) can be written using the Plancherel theorem as,

$$\begin{aligned}\omega_{-\frac{1}{2}}(u,v) &= \sum_{\substack{k=-N\\k\neq 0}}^{N} \widehat{u}(-k)\frac{1}{ik}\widehat{v}(k)\\ &= \sum_{k=1}^{N}\frac{1}{ik}\left(\widehat{u}(-k)\widehat{v}(k)-\widehat{u}(k)\widehat{v}(-k)\right)\\ &= \sum_{k=1}^{N}\frac{2}{k}(\mathrm{Im}(\widehat{v}(k)\overline{\widehat{u}(k)}))\\ &= \sum_{k=1}^{N}\frac{2}{k}(e_k(u)\cdot f_k(v)-e_k(v)\cdot f_k(u)).\end{aligned}$$

Write Γ for the $N\times N$ matrix $\Gamma \equiv \mathrm{diag}(1,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{3}},\dots\frac{1}{\sqrt{N}})$, $\Lambda \equiv \mathrm{diag}(\Gamma,\Gamma)$, and $u=(\vec{e}(u),\vec{f}(u))\in\mathbb{R}^{2N}$ for the coordinates in $P_{\leq N}H_0^{-1/2}(\mathbb{T})$, we summarize the discussion above by saying,

$$\omega_{-\frac{1}{2}}(u,v)=\omega_0(\Lambda(\vec{e}(u),\vec{f}(u)),\Lambda(\vec{e}(v),\vec{f}(v))), \tag{1.22}$$

where as before we've written ω_0 for the standard symplectic form on $\mathbb{R}^{2N}$. In other words,

$$\Lambda:(P_{\leq N}H_0^{-1/2}(\mathbb{T}),\omega_{-\frac{1}{2}})\to(R^{2N},\omega_0)$$

is a symplectomorphism.

Following [3], our goal is to find a flow which satisfies three conditions: it should be finite dimensional - that is, map $P_{\leq N}H^{-\frac{1}{2}}(\mathbb{T})$ into itself; it should be a symplectic map for each t; and it should well-approximate the full flow $S_{KdV}(t)$ in a sense that we will make rigorous momentarily. For now, we write $S^{(N)}_{\text{Good!}}(t)$ for this yet to be determined flow.

$$\begin{array}{ccc}(P_{\leq N}H_0^{-\frac{1}{2}},\omega_{-\frac{1}{2}}) & \xrightarrow{\Lambda} & (\mathbb{R}^{2N},\omega_0)\\ {\scriptstyle S^{(N)}_{\text{Good!}}(t)}\big\downarrow & & \\ (P_{\leq N}H_0^{-\frac{1}{2}},\omega_{-\frac{1}{2}}) & \xrightarrow[\Lambda]{} & (\mathbb{R}^{2N},\omega_0)\end{array} \tag{1.23}$$

Note then that the map,

$$\Lambda\circ S^{(N)}_{\text{Good!}}(t)\circ\Lambda^{-1}:(\mathbb{R}^{2N},\omega_0)\longrightarrow(\mathbb{R}^{2N},\omega_0) \tag{1.24}$$

is likewise a symplectomorphism to which we can apply the finite dimensional theory of symplectic capacity (see [13], and e.g. [15]). One defines, for any $\vec{x}_*\in\mathbb{R}^{2N}, u_*^{(N)}\in P_{\leq N}H_0^{-1/2}(\mathbb{T})$, $r>0$, $0<|k_0|\leq N$, and $z\in\mathbf{C}$, the finite-dimensional balls in $P_{\leq N}H_0^{-1/2}(\mathbb{T}), R^{2N}$, respectively, by the notation,

$$\mathbf{B}^N(u_*^{(N)};r):=\{u^{(N)}\in P_{\leq N}H_0^{-1/2}(\mathbb{T}):\|u^{(N)}-u_*^{(N)}\|_{H_0^{-1/2}}\leq r\} \tag{1.25}$$

$$B(\vec{x}_*,r):=\{\vec{x}\in\mathbb{R}^{2N}:|\vec{x}-\vec{x}_*|\leq r\}. \tag{1.26}$$

and the finite-dimensional cylinders in the same spaces by,

$$\mathbf{C}^N_{k_0}(z;r) := \{u^{(N)} \in P_{\leq N} H_0^{-1/2}(\mathbb{T}) : |k_0|^{-1/2} |\widehat{u^{(N)}}(k_0) - z| \leq r\}$$
$$\mathbf{C}_{k_0}(z;r) := \{(\vec{e}, \vec{f}) \in \mathbb{R}^{2N} : |(e_{k_0} + \sqrt{-1} f_{k_0}) - z| \leq r)\}.$$

From [13], (see also e.g. Theorem 1, Page 55 in the exposition [15]) we have the finite-dimensional analogue of Theorem 1.5:

Theorem 1.6 ([13])**.** *Assume that for some $R, r \geq 0, z \in \mathbf{C}, 0 \leq k_0 \leq N, \vec{x}_* \in \mathbb{R}^{2N}$ there is a symplectomorphism ϕ defined on $B(\vec{x}_*, R) \subset (\mathbb{R}^{2N}, \omega_0)$ so that*

$$\phi(B(\vec{x}_*, R)) \subset C_{k_0}(z; r).$$

Then necessarily $r \geq R$.

We apply this theorem to the symplectomorphism $\Lambda \circ S^{(N)}_{\text{Good!}} \circ \Lambda^{-1}$ defined in (1.24) above to conclude,

Theorem 1.7. *Let $N \geq 1$, $0 < r < R$, $u^{(N)}_* \in P_{\leq N} H_0^{-1/2}(\mathbb{T})$, $0 < |k_0| \leq N$, $z \in \mathbf{C}$, and $T > 0$. Let $S^{(N)}_{Good!}(T) : P_{\leq N} H_0^{-1/2}(\mathbb{T}) \to P_{\leq N} H_0^{-1/2}(\mathbb{T})$ be any symplectomorphism. Then*

$$S^{(N)}_{Good!}(T)(\mathbf{B}^N(u^{(N)}_*; R)) \not\subseteq \mathbf{C}^N_{k_0}(z; r).$$

To deduce Theorem 1.5 from Theorem 1.7, one would like to let $N \to \infty$ and show that the flow $S^{(N)}_{\text{Good!}}(T)$ converged to $S_{KdV}(T)$ in some weak sense. More precisely, one would need,

Condition 1.8. *Let $k_0 \in \mathbb{Z}^*$, $T > 0$, $A > 0$, $0 < \varepsilon \ll 1$. Then there exists an $N_0 = N_0(k_0, T, \varepsilon, A) > |k_0|$ such that*

$$|k_0|^{-1/2} |\widehat{S_{KdV}(T)} u_0(k_0) - \widehat{S^{(N)}_{Good!}(T)} u_0(k_0)| \ll \varepsilon$$

for all $N \geq N_0$ and all $u_0 \in \mathbf{B}^N(0, A)$.

Once we find a finite dimensional symplectic flow $S^{(N)}_{\text{Good!}}(t)$ for which Condition 1.8 holds, it is an easy matter to conclude Theorem 1.5. Indeed, let r, R, u_*, k_0, z, T be as in that Theorem, and choose $0 < \varepsilon < (R - r)/2$. The ball $\mathbf{B}^\infty(u_*; R)$ is contained in some ball $\mathbf{B}^\infty(0; A)$ centered at the origin. We choose $N \geq N_0(k_0, T, \varepsilon, A)$ so large that $\|u_* - P_{\leq N} u_*\|_{H_0^{-1/2}} \leq \varepsilon$. From Theorem 1.7 we can find initial data $u^{(N)}_0 \in P_{\leq N} H^{-\frac{1}{2}}(\mathbb{T})$ satisfying $\|u^{(N)}_0 - P_{\leq N} u_*\|_{H_0^{-1/2}} \leq R - \varepsilon$, and hence by the triangle inequality,

$$\|u^{(N)}_0 - u_*\|_{H_0^{-1/2}} \leq R,$$

and so that at time T we have,

$$|k_0|^{-1/2} |\widehat{S^{(N)}_{\text{Good!}}(T)} u^{(N)}_0(k_0) - z| > r + \varepsilon.$$

If we then apply Condition 1.8 and the triangle inequality we obtain Theorem 1.5 with $u_0 := u_0^{(N)}$,

$$
\begin{aligned}
&|k_0|^{-1/2}|z - \widehat{S_{KdV}(T)} u_0^{(N)}(k_0)| \geq \\
&|k_0|^{-1/2}\left| |z - \widehat{S^{(N)}_{\mathrm{Good!}}(T)} u_0^{(N)}(k_0)| - |\widehat{S_{KdV}(T)} u_0^{(N)}(k_0) - \widehat{S^{(N)}_{\mathrm{Good!}}(T)} u_0^{(N)}(k_0)| \right| \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad > r + \epsilon - \epsilon \quad = r.
\end{aligned}
$$

It remains to define the flow $S^{(N)}_{\mathrm{Good!}}(t)$. One might first try to follow Bourgain's treatment of several different Hamiltonion PDE, notably the cubic NLS flow on $L^2(\mathbb{T})$ (see [3], [5]). Note that the Hamiltonian $H[u]$ (1.18) is well-defined on $(P_{\leq N} H_0^{-1/2}(\mathbb{T}), \omega_{-\frac{1}{2}})$, and the equation giving the corresponding Hamiltonian flow on this space can be computed as before to be (1.5), which can be viewed either as a PDE or as a system of $2N$ ODE. The maps $S_{P_{\leq N} KdV}(t)$ are therefore symplectomorphisms, but from Theorem 1.1 we know that Condition 1.8 fails.

We proceed instead by using a flow of the form (1.7) as follows: Theorem 1.2 tells us that for any multiplier $\tilde{B}$ of the form described in (1.7), the finite dimensional flow $S_{\tilde{B} KdV}$ provides a good approximation to the low frequency behavior of KdV However, the flows $S_{\tilde{B} KdV}$ are not symplectomorphisms, and hence cannot be candidates for our flow $S^{(N)}_{\mathrm{Good!}}(t)$ in the discussion above. Fortunately, there is a quick cure for this hiccup using the approximation given by Theorem 1.3 as follows: we will define a symplectic, finite dimensional flow $S^{(N)}_{KdV}(t)$ on $P_{\leq N} H_0^{-\frac{1}{2}}$ so that the following diagram commutes.

$$
\begin{array}{ccc}
u_0 \in P_{\leq N} H_0^{-\frac{1}{2}} & \xrightarrow{\;B\;} & B u_0 \\
{\scriptstyle S^{(N)}_{KdV}(t)} \big\downarrow & & \big\downarrow {\scriptstyle S_{B^2 KdV}(t)} \\
S^{(N)}_{KdV}(t) u_0 & \xrightarrow[B]{} & w(t)
\end{array}
\tag{1.27}
$$

We write explicitly the PDE defining this flow in (7.1) below. To show that $S^{(N)}_{KdV}(t)$ well approximates $S_{KdV}(t)$ at frequency k_0, and hence qualifies as our choice of $S^{(N)}_{\mathrm{Good!}}(t)$, we will simply spell out the following: Theorem 1.3 allows us to replace $S_{B^2 KdV}(t)$ on the right side of (1.27) with $S_{KdV}(t)$; and our choice $N \gg |k_0|$ allows us to ignore both the mappings on the top of (1.27) (again, by Theorem 1.3) and the bottom of (1.27) (by the definition of B, this is the identity at frequency k_0). We give the details in section 7 below.

Acknowledgements. This work was conducted at UCLA. The authors are indebted to Tom Mrowka for his detailed explanation of symplectic non-squeezing.

2. Inverting the Miura transform

As described in the introduction above, our work here on the KdV equation relies on the continuity and invertibility properties of the Miura transform $u = \mathbf{M} v$, where $\mathbf{M}$ is defined by (see [24]),

$$\mathbf{M}v := v_x + v^2 - P_0(v^2).$$

The additional $P_0(v^2)$ term here is necessary to make the mean of $\mathbf{M}v$ vanish. Let $S_{mKdV}(t)$ denote the flow associated to the mKdV equation (1.9). Then we have the intertwining relationship

$$\mathbf{M}S_{mKdV}(t) = S_{KdV}(t)\mathbf{M}. \tag{2.1}$$

To see this, we suppose that v solves the mKdV equation (1.9), and set $u := \mathbf{M}v$. Then one easily checks,

$$\begin{aligned} u_t + u_{xxx} - 6uu_x &= (\partial_x + 2v)v_t + (\partial_x + 2v)v_{xxx} + 6v_x v_{xx} \\ &\quad - 6(v_x + v^2 - P_0(v^2))(v_{xx} + 2vv_x) \\ &= (\partial_x + 2v)(v_t + v_{xxx} - 6v^2 v_x + 6P_0(v^2)v_x) \\ &= 0. \end{aligned}$$

Heuristically, the Miura transform acts like a derivative operator ∂_x, and in particular we expect it to be a locally bilipschitz bijection from H_0^s to H_0^{s-1}. The purpose of this section is to make this heuristic rigorous for the range $s \geq 1/2$. (See also [17], which studies the Miura transform for the larger range $s \geq 0$.)

In what follows we shall make frequent use of the well-known Sobolev multiplication law

$$\|uv\|_{H^s(\mathbb{T})} \lesssim \|u\|_{H^{s_1}(\mathbb{T})}\|v\|_{H^{s_2}(\mathbb{T})} \tag{2.2}$$

whenever $s \leq \min(s_1, s_2)$ and $s \leq s_1 + s_2 - \frac{1}{2}$, with at least one of the two inequalities being strict.

From (2.2) it is clear that $\mathbf{M}$ is a locally Lipschitz[13] map from $H_0^s(\mathbb{T})$ to $H_0^{s-1}(\mathbb{T})$ for $s \geq 1/2$ (in fact $s > 0$ would suffice). The main result of this section is to invert this statement:

Theorem 2.1. *Let $s \geq 1/2$. Then the map $\mathbf{M}$ is a bijection from $H_0^s(\mathbb{T})$ to $H_0^{s-1}(\mathbb{T})$, and the inverse map $\mathbf{M}^{-1}$ is a locally Lipschitz map from $H_0^{s-1}(\mathbb{T})$ to $H_0^s(\mathbb{T})$.*

Proof. We shall focus on the endpoint case $s = 1/2$. We shall see at the end of the proof that the higher regularity cases $s > 1/2$ then follow from the endpoint case and standard elliptic regularity theory. We remark that the arguments here (based on a variational approach) are unrelated to the rest of the paper and can be read independently.

Since the linearization $v \mapsto v_x$ of the Miura transform $\mathbf{M}$ is clearly bilipschitz from $H_0^{1/2}(\mathbb{T})$ to $H_0^{-1/2}(\mathbb{T})$ it is tempting to treat the lower order terms $v^2 - P_0(v^2)$ as perturbations to be iterated away. This works well if v and $\mathbf{M}(v)$ are small, however for large v it appears that iterative techniques alone cannot obtain

[13]By this we mean that $\mathbf{M}$ is Lipschitz on every ball in $H_0^s(\mathbb{T})$, with a Lipschitz constant depending on the ball.

this result[14]. Indeed, we shall need to also rely on variational techniques, and in particular we will use the well-known connection between the Miura transform and the spectral theory of Schrödinger operators. The key identity here is

$$(\frac{d}{dx}+v)(-\frac{d}{dx}+v) = -\frac{d^2}{dx^2}+(v_x+v^2) = -\frac{d^2}{dx^2}+\mathbf{M}(v)+P_0(v^2). \tag{2.3}$$

We shall work entirely with the smooth functions in $H_0^{1/2}(\mathbb{T})$ and $H_0^{-1/2}(\mathbb{T})$, and obtain bilipschitz bounds for $\mathbf{M}$ on these functions; it will then be clear from standard limiting arguments that one has bilipschitz bounds in general.

Let $u \in H_0^{-1/2}(\mathbb{T})$ be smooth. We consider the problem of finding a smooth function $v \in H_0^{1/2}(\mathbb{T})$ with $\mathbf{M}v = u$, showing this v is unique and of estimating v in terms of u. This will be achieved by studying the self-adjoint Schrödinger operator $L = L_u$ defined by

$$L := -\frac{d^2}{dx^2} + u(x)$$

and the associated energy functional $E[\phi] = E_u[\phi]$ defined on $H^1(\mathbb{T})$ by

$$E[\phi] := \langle L\phi, \phi\rangle = \int_{\mathbb{T}} \phi_x^2(x) + u(x)\phi^2(x)\; dx.$$

Since L is a self-adjoint elliptic operator on a compact manifold $\mathbb{T}$, it has a discrete spectrum $\lambda_1 \leq \lambda_2 \leq \ldots$ with $\lambda_n \to +\infty$. In particular we have a lowest eigenvalue $\lambda_1 = \lambda_1(u) \in \mathbb{R}$, and a non-zero (real-valued) eigenfunction ϕ_1 with $L\phi_1 = \lambda_1\phi_1$. *A priori* ϕ_1 is only in $H^1(\mathbb{T})$, but since u is smooth one can use the equation $L\phi_1 = \lambda_1\phi_1$ to deduce that ϕ_1 is also smooth.

Our analysis here shall rely solely on λ_1. It is interesting to note that the work in [4], which is at a similar level of scaling to $H_0^{-1/2}(\mathbb{T})$, uses the entire spectrum λ_n of the operator L.

From construction of $E[\phi]$ we observe that

$$E[\phi] \geq \lambda_1 \int_{\mathbb{T}} \phi^2 \tag{2.4}$$

for all $\phi \in H^1(\mathbb{T})$, with equality attained if and only if ϕ is a λ_1 eigenfunction of L. (As we shall see, λ_1 is an isolated eigenvalue, so equality only occurs when $\phi = c\phi_1$ for some c.) Thus λ_1 can be described in a variational manner.

Since $u \in H_0^{-1/2}(\mathbb{T})$ we see that $E[1] = 0$, thus λ_1 must be non-positive. If $u \not\equiv 0$ then 1 is not an eigenfunction, and so λ_1 becomes strictly negative.

We now claim that ϕ_1 cannot vanish anywhere. If it had a double zero at some point, i.e. $\phi_1(x_0) = \partial_x\phi_1(x_0) = 0$, then from the second-order ODE $L\phi_1 = \lambda_1\phi_1$

[14] However, iterative techniques do allow us to bootstrap low regularity estimates to high regularity estimates, basically because $\mathbf{M}$ is elliptic and v lies above the critical regularity $H^{-1/2}$ for $\mathbf{M}$ (and for mKdV). The strategy of this argument will be to use variational estimates to obtain a preliminary estimate in very rough norms, and use iteration to improve this to estimates in the correct norms $H_0^{1/2}(\mathbb{T})$ and $H_0^{-1/2}(\mathbb{T})$.

and the Picard existence theorem for ODE we see that $\phi_1 \equiv 0$, a contradiction. Now suppose that ϕ_1 had a simple zero at x_0, so in particular ϕ_1 changed sign. Let $\phi_1 = \phi_1^+ + \phi_1^-$ denote the positive and negative components of ϕ_1. An integration by parts shows that

$$E[\phi_1^+] = \int_{\phi_1>0} L\phi_1(x)\phi_1(x) = \lambda_1 \int \phi_1^+(x)^2.$$

This implies that ϕ_1^+ is a λ_1 eigenfunction of L, which contradicts the fact that all such eigenfunctions are smooth[15]. Thus ϕ_1 is nowhere vanishing; without loss of generality we may take ϕ_1 to be positive and L^2-normalized (which uniquely identifies ϕ_1). If we now define v to be the logarithmic derivative of ϕ_1

$$v(x) := \frac{\partial_x \phi_1(x)}{\phi_1(x)}$$

then v is smooth and we have

$$v_x = \frac{\partial_{xx}\phi_1}{\phi_1} - (\frac{\partial_x \phi_1}{\phi_1})^2 = u - \lambda_1 - v^2$$

(since $L\phi_1 = \lambda_1 \phi_1$) and hence

$$u = v_x + v^2 + \lambda_1.$$

Taking means of both sides we see that

$$-\lambda_1 = P_0(v^2) \tag{2.5}$$

and hence $u = \mathbf{M}v$.

This shows existence of v such that $u = \mathbf{M}v$. Observe from (2.3) and an integration by parts that

$$E[\phi] = \int (\phi_x - v\phi)^2 \, dx - P_0(v^2) \int \phi^2; \tag{2.6}$$

from this and (2.5) we immediately see that (2.4) holds (which we already knew), and that equality occurs if and only if $\phi_x = v\phi$, or in other words if ϕ is a constant multiple of $\exp(\partial_x^{-1} v)$. In particular this shows that v is unique, for if we had $\mathbf{M}v = \mathbf{M}\tilde{v}$ then the above argument yields $\exp(\partial_x^{-1} v)$ is a constant multiple of $\exp(\partial_x^{-1}\tilde{v})$, which implies $v = \tilde{v}$ if $v, \tilde{v}$ both lie in $H_0^{1/2}(\mathbb{T})$.

We have now shown that $\mathbf{M}$ is smooth, locally Lipschitz, and bijective on smooth functions with mean zero. To extend this to $H_0^{1/2}(\mathbb{T})$ and $H_0^{-1/2}(\mathbb{T})$ we need some *a priori* estimates on $\mathbf{M}^{-1}$ in these norms.

Let $u \in H_0^{-1/2}(\mathbb{T})$ and $v \in H_0^{1/2}(\mathbb{T})$ be smooth functions such that $u = \mathbf{M}v$. For this discussion we will allow implicit constants to depend on the $H_0^{-1/2}(\mathbb{T})$ norm of u. Write $U := \partial_x^{-1} u$, thus $\|U\|_{H_0^{1/2}(\mathbb{T})} \lesssim 1$. We observe from integration by parts,

[15]Alternatively, one can smooth ϕ_1^+ at the zeroes of ϕ_1 to contradict (2.4).

Hölder, Sobolev, and Gagliardo-Nirenberg that

$$\begin{aligned} E[\phi] &= \int \phi_x^2 + \int u\phi^2 \\ &= \|\phi\|_{\dot H^1}^2 - 2\int U\phi\phi_x \\ &\geq \|\phi\|_{\dot H^1}^2 - C\|U\|_{L^4_x}\|\phi\|_{L^4_x}\|\phi_x\|_{L^2_x} \\ &\geq \|\phi\|_{\dot H^1}^2 - C\|U\|_{H_0^{1/2}(\mathbb{T})}\|\phi\|_{H_x^{1/2}}\|\phi\|_{\dot H^1} \\ &\geq \|\phi\|_{\dot H^1}^2 - C\|\phi\|_{L^2}^{1/2}\|\phi\|_{H^1}^{1/2}\|\phi\|_{\dot H^1}. \end{aligned}$$

In particular we have the coercivity bound

$$E[\phi] + C\|\phi\|_{L^2}^2 \gtrsim \|\phi\|_{H^1}^2$$

for all $\phi \in H^1(\mathbb{T})$. Applying this to $\phi = \phi_1$ in particular and recalling the upper bound on λ_1 we obtain the eigenvalue bound

$$-C \leq \lambda_1 \leq 0 \tag{2.7}$$

and the preliminary eigenfunction bound

$$\|\phi_1\|_{H^1} \lesssim 1.$$

From (2.2) and the $H_0^{-1/2}(\mathbb{T})$ bound on u we thus have

$$\|u\phi_1\|_{H^{-1/2}} \lesssim 1$$

which by the eigenfunction equation $L\phi_1 = \lambda_1\phi_1$ implies the better eigenfunction bound

$$\|\phi_1\|_{H^{3/2}} \lesssim 1. \tag{2.8}$$

Now we estimate v. From (2.5) and (2.7) we have the preliminary bound

$$\|v\|_{L^2} \lesssim 1;$$

since $u = \mathbf{M}v$, we thus have

$$\|v_x - u\|_{L^1} \lesssim 1.$$

Since L^1 and $H_0^{-1/2}(\mathbb{T})$ both embed into $H^{-3/4}$ (for instance) we thus have by Sobolev that

$$\|v\|_{L^4} \lesssim \|v\|_{H_0^{1/4}} \lesssim \|v_x\|_{H^{-3/4}} \lesssim 1.$$

Returning once again to the equation $u = \mathbf{M}v$, we thus have

$$\|v_x - u\|_{L^2} \lesssim 1$$

which then implies

$$\|v\|_{H_0^{1/2}(\mathbb{T})} \lesssim 1. \tag{2.9}$$

In particular we have

$$\|\partial_x^{-1}v\|_{L^\infty} \lesssim \|\partial_x^{-1}v\|_{H_0^{3/2}(\mathbb{T})} \lesssim \|v\|_{H_0^{1/2}(\mathbb{T})} \lesssim 1,$$

and thus $\exp(\partial_x^{-1}v)$ is bounded above and below. Since ϕ_1 is a constant multiple of $\exp(\partial_x^{-1}v)$, we thus see from (2.8) that

$$|\phi_1(x)| \sim 1 \text{ for all } x \in \mathbb{T}. \tag{2.10}$$

We have obtained good bounds for $v = \mathbf{M}^{-1}(u)$ and for the ground state ϕ_1. We now establish that $\mathbf{M}^{-1}$ is Lipschitz for smooth v in a given bounded subset of $H_0^{1/2}$. From the inverse function theorem and the fact (from (2.2)) that $\mathbf{M}$ is a locally uniformly C^2 map from $H_0^{1/2}$ to $H_0^{-1/2}$, it suffices to show that the derivative map $\mathbf{M}'(v) : H_0^{1/2} \to H_0^{-1/2}$ is uniformly invertible for v in this set.

A direct computation shows

$$\mathbf{M}'(v)(w) = (1 - P_0)(\partial_x + 2v)w$$

We shall invert this explicitly.

Lemma 2.2. *We have*

$$\mathbf{M}'(v)^{-1} = A[\exp(-2\partial_x^{-1} v)]\partial_x^{-1} A[\exp(2\partial_x^{-1} v)]$$

where for any positive function $V \in H^{3/2}(\mathbb{T})$, $A[V] : H_0^{\pm 1/2}(\mathbb{T}) \to H_0^{\pm 1/2}(\mathbb{T})$ is the operator

$$A[V](w) := Vw - \frac{V}{P_0(V)} P_0(Vw).$$

We recommend that the reader think of $\mathbf{M}'(v)$ and $\mathbf{M}'(v)^{-1}$ as perturbations of ∂_x and ∂_x^{-1} respectively.

Proof. We have

$$\begin{aligned} \mathbf{M}'(v) &= (1 - P_0) \exp(-2\partial_x^{-1} v)\partial_x \exp(2\partial_x^{-1} v) \\ &= (1 - P_0) \exp(-2\partial_x^{-1} v)\partial_x (1 - P_0) \exp(2\partial_x^{-1} v). \end{aligned}$$

Also, observe that $A[V]$ is the inverse of $(1 - P_0)V^{-1}$ on $H_0^{\pm 1/2}(\mathbb{T})$. The claim follows. □

Since $H^{3/2}$ is a Banach algebra (by (2.2)), we have

$$\| \exp(\pm 2\partial_x^{-1} v)\|_{H^{3/2}} \lesssim \exp(C\|\partial_x^{-1} v\|_{H^{3/2}}) \lesssim \exp(C\|v\|_{H_0^{1/2}(\mathbb{T})}) \lesssim 1. \tag{2.11}$$

Thus from Lemma 2.2 we see that $\mathbf{M}'(v)^{-1}$ is uniformly bounded from $H_0^{-1/2}$ to $H_0^{1/2}$.

We have now proven Theorem 2.1 at the endpoint $s = 1/2$. We now sketch how one can use elliptic regularity theory to bootstrap this to higher regularities $s > 1/2$.

Let us first show the boundedness of $\mathbf{M}^{-1}$ from H_0^{s-1} to H_0^s for smooth functions. In other words, if $u = \mathbf{M}v$ is smooth, we wish to show that

$$\|v\|_{H_0^s} \lesssim C(\|u\|_{H_0^{s-1}}).$$

From the $H^{1/2}$ theory we already know that

$$\|v\|_{H_0^{1/2}} \lesssim C(\|u\|_{H_0^{s-1}}).$$

Suppose for the moment that $1/2 < s < 3/2$. We write

$$\begin{aligned}\|v\|_{H_0^s} &\lesssim \|v_x\|_{H_0^{s-1}}\\ &\lesssim \|\mathbf{M}v\|_{H_0^{s-1}} + \|(1-P_0)v^2\|_{H_0^{s-1}}\\ &\lesssim \|u\|_{H_0^{s-1}} + \|v^2\|_{H^{s-1}}.\end{aligned}$$

If $s < 3/2$, then by (2.2) we see that $\|v^2\|_{H^{s-1}} \lesssim \|v\|^2_{H_0^{1/2}} \lesssim C(\|u\|_{H_0^{s-1}})$, which establishes boundedness. By iterating this type of argument again one can cover the case $3/2 \le s < 5/2$, and so forth until we obtain boundedness for all $s > 1/2$. The local Lipschitz property for $\mathbf{M}^{-1}$ is proven similarly and is left to the reader. □

From the above Theorem, the analyticity of $\mathbf{M}$, and the inverse function theorem we see in fact that $\mathbf{M}^{-1}$ is locally uniformly C^m as a map from $H_0^{s-1}(\mathbb{T})$ to $H_0^s(\mathbb{T})$, for any integer m and any $s \ge 1/2$.

3. The Fourier restriction spaces Y^s and Z^s

In view of the results of the last section, we see that to analyze the KdV flow in the H_0^{s-1} topology it will suffice to analyze the mKdV flow in the H_0^s topology. We now review the basic machinery (from [1], [18], [8], [9]) for doing so.

If $u(x,t)$ is a function on the cylinder $\mathbb{T}\times\mathbb{R}$ with mean zero at every time, and $s, b \in \mathbb{R}$, we define the $X^{s,b} = X^{s,b}(\mathbb{T}\times\mathbb{R})$ norm by

$$\|u\|_{X^{s,b}} := \|\widehat{u}(k,\tau)\langle k\rangle^s \langle \tau - k^3\rangle^b\|_{L^2_{\tau,k}}$$

where $L^2_{\tau,k}$ is with respect to Lebesgue measure $d\tau$ in the τ variable and counting measure in the k variable, $\langle x\rangle^2 \equiv (1+|x|^2)$, and the space-time Fourier transform $\widehat{u}(k,\tau)$ is given for $k \in \mathbb{Z}^*$, $\tau \in \mathbb{R}$ by

$$\widehat{u}(k,\tau) := \int_{\mathbb{T}\times\mathbb{R}} e^{-2\pi i(xk+t\tau)} u(x,t)\, dxdt.$$

We use the same notation here as for the purely spatial Fourier transform (1.2), relying on context to distinguish the two.

We also need the spaces

$$\|u\|_{Y^s} := \|u\|_{X^{s,1/2}} + \|\langle k\rangle^s \widehat{u}\|_{L^2_k L^1_\tau} \tag{3.1}$$

and

$$\|u\|_{Z^s} := \|u\|_{X^{s,-1/2}} + \|\frac{\langle k\rangle^s \widehat{u}}{\langle \tau - k^3\rangle}\|_{L^2_k L^1_\tau}. \tag{3.2}$$

Observe that we have the crude estimate

$$\|u\|_{Z^s} \lesssim \|u\|_{X^{s,0}} = \|u\|_{L^2_t H^s_x} \tag{3.3}$$

which will be useful for controlling quartic or higher order error terms; often we will be localized in time and just estimate $L^2_t H^s$ by $L^\infty_t H^s$. Here and in the sequel, we always allow implicit constants to depend on the exponent s.

We can restrict the space Y^s to a time interval $I \subseteq \mathbb{R}$ in the usual manner as

$$\|u\|_{Y^s_I} := \inf\{\|v\|_{Y^s} : v|_{\mathbb{T}\times I} = u\}.$$

Similarly we can restrict the Z^s norm. In practice we shall work in a fixed time interval (usually $[-T, T]$) and implicitly restrict all of our norms to this interval.

Now we give some embeddings for the Y^s and Z^s spaces. Since the Fourier transform of an L^1 function is continuous and bounded, we have from (3.1) that

$$Y^s \subseteq C_t H^s_x \subseteq L^\infty_t H^s_x. \tag{3.4}$$

We have the "energy estimate",

$$\|\eta(t)v\|_{Y^s} \lesssim \|v(t_0)\|_{H^s_0} + \|v_t + v_{xxx}\|_{Z^s} \tag{3.5}$$

for any $t_0 \in \mathbb{R}$ and any bump function η supported on $[t_0 - C, t_0 + C]$. ([1], see also Lemma 3.1, [9]; see Lemmas 3.1 - 3.3 in [20] for analogous estimates in the nonperiodic context.)

Recall too the main estimate from [9] (see Proposition 1 in that paper), namely,

$$\|(1-P_0)\left(((1-P_0)\prod_{i=1}^k u_i)w_x\right)\|_{Z^s} \lesssim (\prod_{i=1}^k \|u\|_{Y^s})\|w\|_{Y^s} \tag{3.6}$$

for any $s \geq 1/2$ and any integer $k \geq 2$, where the implicit constant depends on k. (We shall only use (3.6) with $k = 2, 3, 4$). This particular estimate is crucial (especially at the endpoint $s = 1/2$) in order to prove the local (and global) well-posedness of the modified KdV equation (1.9) in $H^s_0(\mathbb{T})$ for $s \geq 1/2$.

It would be very convenient if the Z^s on the left-hand side of (3.6) could be replaced by $Z^{s+\sigma}$ for some $\sigma > 0$; this *extra smoothing* estimate would make it easy to ignore the high-frequency components of the evolution and concentrate on the low frequency evolution. Unfortunately it is easy to see (by modifying the examples in [18]) that such estimates fail, especially at $s = 1/2$. Fortunately, as we will see in the next section there are some other ways to improve the trilinear version of (3.6) which will be useful for our approximation results.

4. An improved trilinear estimate

The estimate (3.6) with $k = 2$ allows us to estimate the cubic nonlinearity $F(v)$ defined in (1.10). However for our analysis we shall need a refined version of this estimate.

The first step is to decompose F into "resonant" and "non-resonant" components. In the following analysis we shall always assume that v has mean zero.

We start with the Fourier inversion formula

$$v(x) = \sum_{k\in\mathbb{Z}^*} \widehat{v}(k)\exp(ikx)$$

for $v \in H_0^s$, where $\mathbb{Z}^* := \mathbb{Z}\backslash\{0\}$ is the set of the non-zero integers. A direct computation gives that the Fourier transform of $F(v)$ is

$$\widehat{F(v)}(k) = 6 \sum_{k_1,k_2,k_3 \in \mathbb{Z}^*: k_1+k_2+k_3=k; k_1+k_2 \neq 0} \widehat{v}(k_1)\widehat{v}(k_2) i k_3 \widehat{v}(k_3) \tag{4.1}$$

for all $k \in \mathbb{Z}^*$. The constraint $k_1 + k_2 \neq 0$ arises since we have subtracted the mean $P_0(v^2)$ from v^2 in the definition of $F(v)$. Observe that $F(v)$ is a perfect derivative and so has mean zero and thus no Fourier component at 0.

Lemma 4.1. *We have*

$$F(v) = F_0(v, v, v) + F_{\neq 0}(v, v, v)$$

where the "resonant" trilinear operator F_0 is given by

$$F_0\widehat{(u, v, w)}(k) := -6ik\widehat{u}(k)\widehat{v}(k)\widehat{w}(-k) \tag{4.2}$$

for $k \in \mathbb{Z}^*$, and the "non-resonant" trilinear operator $F_{\neq 0}$ is defined by

$$F_{\neq 0}\widehat{(u, v, w)}(k) := -\sum_{\substack{k_1,k_2,k_3 \in \mathbb{Z}^*: \\ k_1+k_2+k_3=k; \\ (k_1+k_2)(k_1+k_3)(k_2+k_3) \neq 0}} 2i(k_1 + k_2 + k_3)\widehat{u}(k_1)\widehat{v}(k_2)\widehat{w}(k_3) \tag{4.3}$$

for $k \in \mathbb{Z}^*$.

Proof. Consider the right-hand side of (4.1), and break the sum into pieces according to how many of the quantities $k_1 + k_3, k_2 + k_3$ are zero. There is a single term in the sum for which $k_2 + k_3 = k_1 + k_3 = 0$, and the summation in this case is $F_0(v, v, v)$. If just $k_2 + k_3$ is zero, then the total contribution of this case vanishes since the summand in this case is antisymmetric with respect to swapping k_2 and k_3. Similarly if just $k_1 + k_3$ is zero. The remaining portion of the summation can be seen to be $F_{\neq 0}(v, v, v)$ by a symmetrization in k_1, k_2, k_3. □

If $k = k_1 + k_2 + k_3$, then we have the fundamental *resonance identity*

$$k^3 - (k_1^3 + k_2^3 + k_3^3) = 3(k_1 + k_2)(k_1 + k_3)(k_2 + k_3) \tag{4.4}$$

(see e.g. [1]). This justifies the terminology that F_0 is "resonant" but $F_{\neq 0}$ is "non-resonant".

We remark that, if u, v, w are real, then $F_0(u, v, w)$ and $F_{\neq 0}(u, v, w)$ are also real, despite the presence of the imaginary i in the definitions of these quantities. This follows from identities such as $\widehat{u}(-k) = \overline{\widehat{u}(k)}$. We leave the details to the reader. We also remark that eventually these two functions will be estimated in absolute value, so the constants which appear (e.g. the minus signs out front) will play no role.

4.1. **The F_0 (resonant) estimate.** We now give an estimate for F_0. Morally at least, the bound we give follows from the trilinear version of (3.6), but we present an independent proof here for the sake of completeness.

Lemma 4.2. *For any $s \geq 1/2$, and any $u, v, w \in Y^s$ with mean zero, we have*

$$\|F_0(u, v, w)\|_{Z^s} \lesssim \|u\|_{Y^s} \|v\|_{Y^s} \|w\|_{Y^s}. \tag{4.5}$$

Proof. We shall just prove the endpoint case $s = 1/2$, as the general case easily follows (e.g. by using the identity $\partial_x^{s-1/2} F_0(u, v, w) = F_0(\partial_x^{s-1/2} u, v, w)$).

Split $u = \sum_{k \in \mathbb{Z}^*} u_k$, where u_k is a complex-valued function whose spatial Fourier transform is supported on a single frequency k. Observe that

$$F_0(u, v, w) = \sum_k F_0(u_k, v_{-k}, w_k).$$

Thus if we show that

$$\|F_0(u_k, v_{-k}, w_k)\|_{Z^{1/2}} \lesssim \|u_k\|_{1/2,1/2} \|v_{-k}\|_{1/2,1/2} \|w_k\|_{1/2,1/2},$$

then the claim (4.5) follows by summing in k and using Cauchy-Schwartz in u and v (just estimating the w_k term crudely by w).

Fix k, and define the function $G_{u_k}(t)$ by $u_k(x, t) = e^{ikx} e^{ik^3 t} G_{u_k}(t)$, so that

$$\|u_k\|_{X^{s,\delta}} =< k >^s \| < \tau >^\delta \widehat{G}_{u_k}(\tau)\|_{L^2_\tau(\mathbb{R})},$$

similarly for $G_{v_{-k}}$ and G_{w_k}. The claim then collapses (after some translation in frequency space) to the one-dimensional temporal estimate

$$\|G_{u_k} G_{v_{-k}} G_{w_k}\|_{H_t^{-1/2}} \lesssim \|G_{u_k}\|_{H_t^{1/2}} \|G_{v_{-k}}\|_{H_t^{1/2}} \|G_{w_k}\|_{H_t^{1/2}}$$

and

$$\|G_{u_k} \widehat{G_{v_{-k}}} G_{w_k} \langle \tau \rangle^{-1}\|_{L^1_\tau} \lesssim \|G_{u_k}\|_{H_t^{1/2}} \|G_{v_{-k}}\|_{H_t^{1/2}} \|G_{w_k}\|_{H_t^{1/2}}.$$

But both left-hand sides can be estimated by $\|G_{u_k} G_{v_{-k}} G_{w_k}\|_{L^2_t}$, and the claim follows easily from Hölder and Sobolev. □

4.2. **The $F_{\neq 0}$ (nonresonant) estimate.** We now turn to the non-resonant portion $F_{\neq 0}$ of the non-linearity. In analogy with (3.6), (4.5) we have the estimate

$$\|F_{\neq 0}(u, v, w)\|_{Z^s} \lesssim \|u\|_{Y^s} \|v\|_{Y^s} \|w\|_{Y^s} \tag{4.6}$$

for all $s \geq 1/2$ and $u, v, w \in Y^s$ with mean zero. This estimate can be proven by the techniques used to prove (3.6) in [9], but we shall obtain it as a consequence of a slightly stronger version, which we now state.

We first need some Littlewood-Paley notation. If N is an integer power of two, we let P_N denote the dyadic projection operator

$$\widehat{P_N u}(k) = \chi_{N \leq |k| < 2N} \widehat{u}(k).$$

If N_0, N_1, N_2, N_3 are four integer powers of two, we let *soprano*, *alto*, *tenor*, *baritone* be a permutation of the indices $0, 1, 2, 3$ such that

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

Theorem 4.3. *Let N_0, N_1, N_2, N_3 be integer powers of two. Then*

$$\|P_{N_0} F_{\neq 0}(P_{N_1} u, P_{N_2} v, P_{N_3} w)\|_{Z^{1/2}} \lesssim (\frac{N_0}{N_{soprano}})^\sigma N_{tenor}^{-\sigma} \|u\|_{Y^{1/2}} \|v\|_{Y^{1/2}} \|w\|_{Y^{1/2}} \tag{4.7}$$

for some absolute constant[16] $\sigma > 0$.

This means that (4.7) is only sharp when the output frequency N_0 is essentially the highest frequencies, and the two lowest frequencies N_{tenor} and $N_{baritone}$ are $O(1)$. This means that very low Fourier modes can influence high modes, but medium and high modes do not. In addition, the high modes do not have much influence on the low modes[17]. From (4.7) one can easily obtain (4.6) by summing[18] in the N_i.

The estimates (4.5) and (4.7) give some intuition for why it's possible to find a finite dimensional approximation to the mKdV flow - and hence, using the Miura transform, for the KdV flow as well: the only nonlinear interactions for which we now have no sharpened estimates are the resonant interactions coming from F_0 (which doesn't mix frequencies) and the high-low-low interactions in $F_{\neq 0}$. Heuristically, then, we might start believing that if we truncate high frequencies, the evolution will not see much of a difference at low frequencies. In fact, it is possible to use these estimates to prove low frequency approximation theorems for mKdV analogous to Theorems 1.2, 1.3, but we do not write out these results explicitly in this work.

The rest of this section is devoted to the proof of Theorem 4.3. We remark that the computations in this section are not needed elsewhere in the paper, and the reader may wish to take (4.7) for granted on the first pass and move to the next section.

Proof. We begin by reviewing some (non-trivial) estimates from [9].

The proof of (4.7) relies mainly on the trilinear estimate

$$\|u_1 u_2 u_3\|_{L^2_{x,t}} \lesssim \|u_1\|_{X^{0,\frac{1}{2}-\frac{1}{100}}} \|u_2\|_{X^{0,\frac{1}{2}-\frac{1}{100}}} \|u_3\|_{X^{\frac{1}{2}-\frac{1}{100},\frac{1}{2}-\frac{1}{100}}} \tag{4.8}$$

proven in Section 7 of [9]. This estimate can be viewed as a trilinear variant of the $L^6_{x,t}$ Strichartz estimate in [1], and its proof requires a small amount of elementary number theory.

---

[16] The quantity σ shall vary from line to line.

[17] That is, when the soprano and alto dyadic factors were high frequencies and N_0 were low, we have a small first factor on the right side of (4.7).

[18] More precisely, one first observes that the left-hand side of (4.7) vanishes unless $N_{soprano} \sim N_{alto}$. Then one decomposes u, v, w into dyadic pieces and exploits orthogonality of the projections P_N in the Y_s and Z_s spaces. We omit the details.

We will also use the following estimate, which follows relatively quickly from some bounds found in [9],

$$\| \frac{\langle k \rangle^s \widehat{F_{\neq 0}(u,v,w)}(k)}{\langle \tau - k^3 \rangle^{1-\delta}} \|_{L^2_k L^1_\tau} \lesssim \|u\|_{Y^s} \|v\|_{Y^s} \|w\|_{Y^s} \tag{4.9}$$

for all $s \geq 1/2$ and some $\delta > 0$. To establish (4.9), recall Theorem 3 from [9]

$$\| \prod_{i=1}^k u_i \|_{X^{s-1,\frac{1}{2}}} \lesssim \prod_{i=1}^k \|u_i\|_{Y^s}, \tag{4.10}$$

for $s \geq \frac{1}{2}$. We need also equation (9.2) in [9], which also holds when $s \geq \frac{1}{2}$,

$$\| \frac{\langle k \rangle^s \chi_{k \neq 0} \widehat{((1-P_0)u_1 \cdot (1-P_0)u_2)}(k,\tau)}{\langle \tau - k^3 \rangle^{1-\delta}} \|_{L^2_k L^1_\tau} \lesssim \|u_1\|_{X^{s-1,\frac{1}{2}}} \|u_2\|_{X^{s-1,\frac{1}{2}}}. \tag{4.11}$$

Combining these two and writing for the moment

$$\begin{aligned} W(k,\tau) &\equiv \chi_{k \neq 0}(k) \widehat{((1-P_0)u_x \cdot (1-P_0)(vw))}(k,\tau) \\ &= \chi_{k \neq 0}(k) \sum_{\substack{k_1+k_2+k_3=k \\ k_1, k_2+k_3 \neq 0}} ik_1 \widehat{u}(k_1)\widehat{v}(k_2)\widehat{w}(k_3), \end{aligned} \tag{4.12}$$

we conclude by (4.10)

$$\begin{aligned} \| \frac{\langle k \rangle^s W(k,\tau)}{\langle \tau - k^3 \rangle^{1-\delta}} \|_{L^2_k L^1_\tau} &\lesssim \|u_x\|_{X^{s-1,\frac{1}{2}}} \cdot \|vw\|_{X^{s-1,\frac{1}{2}}} \\ &\lesssim \|u\|_{Y^s} \|v\|_{Y^s} \|w\|_{Y^s}. \end{aligned} \tag{4.13}$$

We quickly conclude (4.9) from (4.13): looking at the definition of the norms involved, one sees that without loss of generality we may assume $\widehat{u}(k), \widehat{v}(k), \widehat{w}(k) \geq 0$. Next, by replacing the factor $\widehat{u}(k)$ appearing in (4.13), (4.12) with $\chi_{k_1 \geq 0}(k_1)\widehat{u}(k_1), \chi_{k_1 \leq 0}\widehat{u}(k_1)$, one concludes (4.13) with the function W now replaced by,

$$W_{II}(k,\tau) \equiv \chi_{k \neq 0}(k) \sum_{\substack{k_1+k_2+k_3=k \\ k_1, k_2+k_3 \neq 0}} |k_1| \widehat{u}(k_1)\widehat{v}(k_2)\widehat{w}(k_3).$$

Repeating this argument while interchanging the roles of k_1, k_2, and then k_1, k_3 and summing gives (4.13) with W replaced with,

$$W_{III}(k,\tau) \equiv \chi_{k \neq 0}(k) \sum_{\substack{k_1 \cdot k_2 \cdot k_3 \neq 0 \\ (k_1+k_2)(k_2+k_3)(k_1+k_3) \neq 0}} (|k_1|+|k_2|+|k_3|) \widehat{u}(k_1)\widehat{v}(k_2)\widehat{w}(k_3).$$

By the definition of $F_{\neq 0}$ (4.3) this yields (4.9).

We now begin the proof of (4.7). It will suffice to prove the estimate

$$\|P_{N_0} F_{\neq 0}(P_{N_1}u, P_{N_2}v, P_{N_3}w)\|_{X^{1/2,-1/2}} \lesssim \left(\frac{N_0}{N_{soprano}}\right)^{\sigma} N_{tenor}^{-\sigma} \|u\|_{X^{1/2,1/2}} \|v\|_{X^{1/2,1/2}} \|w\|_{X^{1/2,1/2}}. \tag{4.14}$$

Indeed, this estimate already controls the $X^{1/2,-1/2}$ portion of the $Z^{\frac{1}{2}}$ norm. To control the $L^2_k L^1_\tau$ portion, we observe from Hölder that the left-hand side of (4.14) controls

$$\| \frac{\langle k\rangle^{\frac{1}{2}} (P_{N_0} F_{\neq 0}(\widehat{P_{N_1} u, P}_{N_2} v, P_{N_3} w)(k)}{\langle \tau - k^3\rangle^{1+\delta}} \|_{L^2_k L^1_\tau},$$

and the claim follows by a suitable interpolation with (4.9) (decreasing σ if necessary).

It remains to prove (4.14). By duality this is equivalent to

$$|\int\int u_0 \partial_x^{-1} F_{\neq 0}(u_1, u_2, u_3)\, dxdt| \lesssim (\frac{N_0}{N_{soprano}})^\sigma N_{tenor}^{-\sigma} \|u_0\|_{X^{-3/2,1/2}} \|u_1\|_{X^{1/2,1/2}} \|u_2\|_{X^{1/2,1/2}} \|u_3\|_{X^{1/2,1/2}}$$

where u_i has Fourier support on the region $|k_i| \sim N_i$. We have inserted the ∂_x^{-1} multiplier to cancel the $(k_1 + k_2 + k_3)$ factor in (4.3).

The right-hand side is comparable to

$$(\frac{N_0}{N_{soprano}})^\sigma N_{tenor}^{-\sigma} \frac{(N_0 N_1 N_2 N_3)^{1/2}}{N_0^2} \prod_{j=0}^3 \|u_j\|_{X^{0,1/2}}. \tag{4.15}$$

Note that we may assume $N_{soprano} \sim N_{alto}$ since the left-hand side of (4.14) vanishes otherwise. Hence the right side of (4.15) is bounded below (throwing away the factor $\left(\frac{N_0}{N_{soprano}}\right)^\sigma$) by

$$N_{tenor}^{1/2-\sigma} N_{baritone}^{1/2} N_{soprano}^{-1} \prod_{j=0}^3 \|u_j\|_{X^{0,1/2}}$$

Taking spacetime Fourier transforms and taking advantage of the frequency localization, we thus reduce to showing

(4.16)

$$|\sum_{\substack{k_0, k_1, k_2, k_3 \in \mathbb{Z}^* \\ k_0 + k_1 + k_2 + k_3 = 0; \\ (k_1+k_2)(k_2+k_3)(k_2+k_3) \neq 0}} \int \prod_{j=0}^3 \widehat{u_j}(k_j, \tau_j)\; d\tau| \lesssim N_{tenor}^{1/2-\sigma} N_{baritone}^{1/2} N_{soprano}^{-1} \prod_{j=0}^3 \|u_j\|_{X^{0,1/2}}$$

where $d\tau$ is integration over the three-dimensional space $\{(\tau_0, \tau_1, \tau_2, \tau_3) \in \mathbb{R}^4 : \tau_0 + \tau_1 + \tau_2 + \tau_3 = 0\}$ with measure $d\tau := \delta(\tau_0 + \tau_1 + \tau_2 + \tau_3) \prod_{j=0}^3 d\tau_j$. We remark that the above estimate is now symmetric with respect to permutations of k_0, k_1, k_2, k_3.

Without loss of generality we may assume that the $\widehat{u}_j$ are all non-negative. The next step is to exploit the implicit $\langle \tau_j - k_j^3\rangle^{1/2}$ denominators. From the fundamental identity (4.4),

$$\sum_{j=0,1,2,3} \tau_j - k_j^3 = -\sum_{j=0}^{3} k_j^3 = 3(k_1+k_2)(k_2+k_3)(k_1+k_3) \tag{4.17}$$

we see that

$$\begin{aligned}\sup_{j=0,1,2,3} \langle \tau_j - k_j^3 \rangle &\gtrsim |k_1+k_2||k_2+k_3||k_1+k_3| \\ &= |k_{soprano}+k_{baritone}||k_{alto}+k_{baritone}||k_{tenor}+k_{baritone}|.\end{aligned}$$

By symmetry we may assume that the supremum on the left-hand side is attained when $j=0$.

Lemma 4.4. *We have*

$$|k_{soprano}+k_{baritone}||k_{alto}+k_{baritone}||k_{tenor}+k_{baritone}|N_{baritone} \gtrsim N_{soprano}^2 \tag{4.18}$$

Proof. We have four cases:

Case 1: $N_{baritone} \ll N_{tenor} \ll N_{alto}$. Then the left-hand side of (4.18) is comparable to $N_{soprano}^2 N_{tenor} N_{baritone}$.

Case 2: $N_{baritone} \sim N_{tenor} \ll N_{alto}$. Then the left-hand side of (4.18) is at least $\gtrsim N_{soprano}^2 N_{tenor}$.

Case 3: $N_{baritone} \ll N_{tenor} \sim N_{alto}$. Then the left-hand side of (4.18) is comparable with $N_{soprano}^3 N_{baritone}$.

Case 4: $N_{baritone} \sim N_{tenor} \sim N_{alto}$. Then at least one of k_1+k_2, k_2+k_3, k_1+k_3 must have magnitude $\sim N_{soprano}$ (since they sum to $-2k_0$). Since the other two factors have magnitude at least 1, the left-hand side of (4.18) is $\gtrsim N_{soprano}^2$. □

From this lemma, we have

$$\langle \tau_0 - k_0^3 \rangle \gtrsim N_{soprano}^2 N_{baritone}^{-1}.$$

Thus to prove (4.16) it will suffice to show that

$$\Big| \sum_{\substack{k_0,k_1,k_2,k_3 \in \mathbb{Z}^* : \\ k_0+k_1+k_2+k_3=0; \\ (k_1+k_2)(k_2+k_3)(k_2+k_3) \neq 0}} \int N_{tenor}^{-1/2+\delta} \langle \tau_0 - k_0^3 \rangle^{1/2} \prod_{j=0}^{3} \widetilde{u_j}(k_j,\tau_j)\, d\tau \Big| \lesssim \prod_{j=0}^{3} \|u_j\|_{X^{0,1/2}}.$$

At least one of k_1,k_2,k_3 is $O(N_{tenor})$; by symmetry let's suppose it's k_3. Then we can bound $N_{tenor}^{1/2-\delta}$ by $k_3^{1/2-\delta}$, and then by undoing the Fourier transform and doing some substitutions the estimate becomes

$$\Big| \int\int v_0 v_1 v_2 v_3 \, dxdt \Big| \lesssim \|v_0\|_{X^{0,0}} \|v_1\|_{X^{0,1/2}} \|v_2\|_{X^{0,1/2}} \|v_3\|_{X^{1/2-\sigma,1/2}}.$$

But this follows directly from (4.8) if σ is small enough. This proves (4.7). □

5. Proof of Theorem 1.3: KdV low frequencies are stable under high frequency perturbations of data.

We now prove Theorem 1.3. Fix s, T, u_0, $\tilde{u}_0$.

We have no upper bound on the time T, and so in particular we cannot hope to control the flow $S_{KdV}(t)$ on the entire interval $[-T, T]$ by a single application of the local well-posedness theory. On the other hand, because of the uniform bounds (1.3) we see that we can divide $[-T, T]$ into a bounded number $C(s, T, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$ of time intervals such that the local well-posedness theory can be used on each interval. It will thus suffice to prove a local-in-time version of Theorem 1.3; more precisely, it will suffice to show

Proposition 5.1. *Fix $s \geq -1/2$, and $N' \geq 1$. Let $u_0, \tilde{u}_0 \in H_0^s$ be such that $P_{\leq N'} u_0 = P_{\leq N'} \tilde{u}_0$. Then, if T' is sufficiently small depending on s, $\|u_0\|_{H_0^s}$, and $\|\tilde{u}_0\|_{H_0^s}$, we have*

$$\sup_{|t| \leq T'} \|P_{\leq N' - (N')^{1/2}} (S_{KdV}(t)\tilde{u}_0 - S_{KdV}(t)u_0)\|_{H_0^s} \leq (N')^{-\sigma} C(s, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$$

for some $\sigma = \sigma(s) > 0$.

The exponent $1/2$ in $(N')^{1/2}$ is not particularly important here; any exponent between 0 and 1 would suffice.

To see how this proposition implies the theorem, first recall that we may assume that N is large, $N \geq C(s, T, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$, since the claim in Theorem 1.3 trivially follows from (1.3) otherwise. (This same remark also applies of course in Proposition 5.1, allowing us to assume $N' \geq C(s, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$ there too.) From (1.3) we may divide $[-T, T]$ into $C(s, T, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s})$ time intervals, such that on each interval (a time-translated version of) Proposition 5.1 holds. Consider for example the first such time interval $[0, T']$ on the positive real axis. We start with $N' := 2N$ and apply Proposition 5.1, to get,

$$\sup_{t \in [0, T']} \|P_{\leq N' - (N')^{1/2}} (S_{KdV}(t)\tilde{u}_0 - S_{KdV}(t)u_0)\|_{H_0^s} \leq (N')^{-\sigma} C(s, \|u_0\|_{H_0^s}, \|\tilde{u}_0\|_{H_0^s}).$$

Before moving on to the next subinterval, modify $S_{KdV}(T')\tilde{u}_0$ on frequencies $|k| \leq N' - (N')^{\frac{1}{2}}$ to agree with $S_{KdV}(T')u_0$. By the local-well posedness theory and the triangle inequality, we can proceed as on the first subinterval, decrementing N' by $(N')^{1/2}$ each time we apply Proposition 5.1, to obtain Theorem 1.3 if N (and hence N') is sufficiently large.

It remains to prove Proposition 5.1. Henceforth we allow our implicit constants to depend on s, $\|u_0\|_{H_0^s}$, and $\|\tilde{u}_0\|_{H_0^s}$.

Define,

$$v_0 := \mathbf{M}^{-1} u_0; \quad v(t) := S_{mKdV}(t) v_0;$$

$$\tilde{v}_0 := \mathbf{M}^{-1} \tilde{u}_0; \quad \tilde{v}(t) := S_{mKdV}(t) \tilde{v}_0;$$

from Theorem 2.1 we thus have

$$\|v_0\|_{H_0^{s+1}}, \|\tilde{v}_0\|_{H_0^{s+1}} \leq C$$

while from (2.1) we have

$$S_{KdV}(t)u_0 = \mathbf{M}v(t); \quad S_{KdV}(t)\tilde{u}_0 = \mathbf{M}\tilde{v}(t).$$

Our task is thus to show that

$$\sup_{|t|\leq T'} \|P_{\leq N'-(N')^{1/2}}(\mathbf{M}\tilde{v}(t) - \mathbf{M}v(t))\|_{H_0^s} \leq C(N')^{-\sigma}. \tag{5.1}$$

Henceforth we allow the quantity $\sigma > 0$ to vary from line to line.

We first investigate the discrepancy between $\tilde{v}$ and v at time 0.

Lemma 5.2. *With $v_0, \tilde{v}_0$ defined as above, we have,*

$$\|P_{\leq N'}(\tilde{v}_0 - v_0)\|_{H_0^{s+1}} \leq C(N')^{-\sigma}.$$

Proof. From the definitions and our assumptions on $u_0, \tilde{u}_0$ we have

$$P_{\leq N'}(\mathbf{M}\tilde{v}_0 - \mathbf{M}v_0) = 0.$$

On the other hand, from Theorem 2.1 we have

$$\|P_{\leq N'}(\tilde{v}_0 - v_0)\|_{H_0^{s+1}} \leq C\|\mathbf{M}P_{\leq N'}\tilde{v}_0 - \mathbf{M}P_{\leq N'}v_0\|_{H_0^s}.$$

Thus by the triangle inequality, it will suffice to show the commutator estimate,

$$\|\mathbf{M}P_{\leq N'}v_0 - P_{\leq N'}\mathbf{M}v_0\|_{H_0^s} \leq C(N')^{-\sigma} \tag{5.2}$$

and similarly for $\tilde{v}_0$.

Clearly it will suffice just to consider v_0. From the definition (1.8) of the tranform $\mathbf{M}$ and the fact that P_0, $P_{\leq N'}$ and ∂_x all commute, we have

$$\begin{aligned}\mathbf{M}P_{\leq N'}v_0 - P_{\leq N'}\mathbf{M}v_0 =&(1 - P_0)[(P_{\leq N'}v_0)^2 - P_{\leq N'}(v_0)^2] \\ =&(1 - P_{\leq N'})[(P_{\leq N'}v_0)^2] \\ &- (P_{\leq N'} - P_0)[((1 - P_{\leq N'})v_0)((1 + P_{\leq N'})v_0)].\end{aligned}$$

But the last two terms have an H_0^s norm of $O((N')^{-\sigma})$ for some $\sigma > 0$; this can be seen by the Sobolev multiplication law (2.2), the H^{s+1} bound on v_0, and the estimate

$$\|(1 - P_{\leq N'})v\|_{H^s} \lesssim N^{-\sigma}\|v\|_{H^{s+\sigma}}$$

to extract the $(N')^{-\sigma}$ decay from the high frequency projection $1 - P_{\leq N'}$. The claim follows. □

We still have to prove (5.1). It will suffice to show that

$$\sup_{|t|\leq T'} \|P_{\leq N'-(N')^{1/2}}(\tilde{v}(t) - v(t))\|_{H_0^{s+1}} \leq C(N')^{-\sigma}. \tag{5.3}$$

This is basically because the commutator of $\mathbf{M}$ with $P_{\leq N'-(N')^{1/2}}$ is small thanks to the argument in the proof of Lemma 5.2. We omit the details as they are very similar to those in Lemma 5.2.

From Lemma 5.2 we see that $\tilde{v}_0$ and v_0 are almost identical at low frequencies $|k| \leq N'$. In fact, because the solution map $S_{mKdV}(t)$ is locally Lipschitz[19] in H_0^{s+1}, we may assume that

$$P_{\leq N'}(\tilde{v}_0 - v_0) = 0, \tag{5.4}$$

since the general case then follows by modifying $\tilde{v}_0$ (or v_0) by a small amount in H_0^{s+1} and using the Lipschitz property.

Henceforth we assume (5.4), so that the low frequency ($|k| \leq N'$) portions of $\tilde{v}(t)$ and $v(t)$ are identical at time 0. Our task is to prove (5.3), which asserts that the slightly lower frequency ($|k| \leq N' - (N')^{1/2}$) portions of $\tilde{v}(t)$ and $v(t)$ are still very close together at later times. This will be achieved primarily through the improved trilinear estimate (4.7).

In what follows we assume all our spacetime norms are restricted to the time interval $[-T', T']$.

From the local well-posedness theory of mKdV (See[20] (3.5), (4.5), (4.6), or [1], [20], [9]) we have the local estimates

$$\|v\|_{Y^{s+1}} + \|\tilde{v}\|_{Y^{s+1}} \leq C \tag{5.5}$$

if the time T' is chosen sufficiently small depending on the H_0^{s+1} norms of v_0, $\tilde{v}_0$.

The frequency interval $[N' - (N')^{1/2}, N']$ contains $O((N')^{1/4})$ intervals of the form $[M, M + (N')^{1/4}]$. By orthogonality and the pigeonhole principle, we see that there must exist one of these intervals $[M, M + (N')^{1/4}]$ such that

$$\|(P_{\leq M+(N')^{1/4}} - P_{\leq M})v\|_{Y^{s+1}} + \|(P_{\leq M+(N')^{1/4}} - P_{\leq M})\tilde{v}\|_{Y^{s+1}} \leq C(N')^{-\sigma}. \tag{5.6}$$

Fix this M. We split

$$v = v_{lo} + v_{med} + v_{hi}$$

where

$$v_{lo} := P_{\leq M}v; \quad v_{med} = (P_{\leq M+(N')^{1/4}} - P_{\leq M})v; \quad v_{hi} := (1 - P_{\leq M+(N')^{1/4}})v.$$

Thus from (5.5), (5.6) we have

$$\|v_{lo}\|_{Y^{s+1}}, \|v_{hi}\|_{Y^{s+1}} \leq C; \quad \|v_{med}\|_{Y^{s+1}} \leq C(N')^{-\sigma}. \tag{5.7}$$

Applying $P_{\leq M}$ to (1.9) and using Lemma 4.1, we see that v_{lo} obeys the equation

$$(\partial_t + \partial_{xxx})v_{lo} = P_{\leq M}F_0(v, v, v) + P_{\leq M}F_{\neq 0}(v, v, v).$$

From the definition (4.2) of the resonant operator F_0, we see that

$$P_{\leq M}F_0(v, v, v) = F_0(v_{lo}, v_{lo}, v_{lo}).$$

The situation for $F_{\neq 0}$ is more complicated as this nonlinearity will mix v_{lo}, v_{med}, v_{hi} together. Define an *error term* to be any quantity with a Z^{s+1} norm of $O((N')^{-\sigma})$.

[19] Since we are assuming T' to be small this follows directly from the local well-posedness theory.

[20] Strictly speaking, when the data v_0, $\tilde{v}_0$ has large H_0^{s+1} norm one has to first rescale the torus by a suitable scaling parameter λ in order to close the iteration, but this has no significant effect on our argument. The details are carried out in [9], [8].

From (5.7) and (4.6) we see that any term in $F_{\neq 0}(v,v,v)$ involving v_{med} is an error term.

Now let us consider the terms which involve v_{hi}. A typical term is

$$P_{\leq M} F_{\neq 0}(v_{lo}, v_{lo}, v_{hi}).$$

We can dyadically decompose this as

$$\sum_{N_0, N_1, N_2, N_3} P_{N_0} P_{\leq M} F_{\neq 0}(P_{N_1} v_{lo}, P_{N_2} v_{lo}, P_{N_3} v_{hi}).$$

Such a term can be estimated using the frequency separation between v_{lo} and v_{hi}: for the summand to be nonzero, we need $N_1, N_2 \leq M$, and $N_3 \geq M + (N')^{\frac{1}{2}}$. Using the notation in the definition of $F_{\neq 0}$ (4.3), we also need $|k_1 + k_2 + k_3| \sim N_0 \leq M$, hence we must clearly also have $N_{tenor} \gtrsim (N')^{1/4}$. From our non-resonant estimate (4.7), the bounds (5.7) above, and a summation of the dyadic indices N_j (conceding some powers of $\log N'$ if necessary) we thus see that this term is an error term. A similar argument shows that any other term involving v_{hi} will also be an error term. Thus we see that v_{lo} obeys the equation

$$(\partial_t + \partial_{xxx}) v_{lo} = F_0(v_{lo}, v_{lo}, v_{lo}) + P_{\leq M} F_{\neq 0}(v_{lo}, v_{lo}, v_{lo}) + \text{error terms.} \tag{5.8}$$

By similar reasoning, the function $\tilde{v}_{lo} := P_{\leq M} \tilde{v}$ also obeys the same equation (but with slightly different error terms, of course). Since $\tilde{v}_{lo}(0) = v_{lo}(0)$, we thus see from the standard local well-posedness theory[21] that

$$\|\tilde{v}_{lo} - v_{lo}\|_{Y^{s+1}} \leq C(N')^{-\sigma} \tag{5.9}$$

which by (3.4) implies (5.3) as desired. This proves Theorem 1.3.

6. Proof of Theorem 1.2: *BKdV* approximates *KdV* at low frequencies.

We now prove the more difficult of our KdV approximation theorems, namely Theorem 1.2. The proof here is definitely in the same spirit as that of Theorem 1.3, in that we show two flows remain close by showing that their mKdV analogues remain close. However, the proof will be more complicated since one of the flows being studied is S_{BKdV} (see (1.7)), and the standard Miura transform $\mathbf{M}$ defined by (1.8) seems an inappropriate tool with which to pull the S_{BKdV} flow back to an *mKdV*-type evolution, as it introduces a v_x^2 type nonlinearity on the right side of (1.9) which is too rough for us to estimate. Instead, we introduce a *modified*

[21] A rough sketch of what we have in mind here is: write G for that portion of the nonlinearity on the right side of (5.8) not involving the error terms, and note

$$\tilde{v}_{lo} - v_{lo} = \int_0^t e^{i(t-\tau)\xi^3} \left(G(\tilde{v}_{lo}) - G(v_{lo}) + \text{error terms}\right) d\tau.$$

Writing $G(\tilde{v}_{lo}) - G(v_{lo}) = \int_0^1 DG(\theta \tilde{v}_{lo} + (1-\theta) v_{lo})(\tilde{v}_{lo} - v_{lo})) d\theta$, we use (3.5), (4.5), (4.6), and the fact that by scaling, we may assume that the data for $v_{lo}, \tilde{v}_{lo}$ are small in Y^{s+1} to conclude (5.9).

Miura transform $\mathbf{M}_B$. This strategy is illustrated in (6.1), where we have written S_{BmKdV} for the flow which intertwines $\mathbf{M}_B$ and $BKdV$ in the sense that

$$\mathbf{M}_B \circ S_{BmKdV}(t) \circ \mathbf{M}_B^{-1} \equiv S_{BKdV}(t).$$

$$\begin{array}{ccc} v_0 & \xrightarrow{S_{mKdV}(t)} & v(t) \\ \mathbf{M}\big\downarrow & & \big\downarrow\mathbf{M} \\ u_0 & \xrightarrow{S_{KdV}(t)} & u(t) \\ u_0 & \xrightarrow[S_{BKdV}(t)]{} & \tilde{u}(t) \\ \mathbf{M}_B\big\uparrow & & \big\uparrow\mathbf{M}_B \\ \tilde{v}_0 & \xrightarrow{S_{BmKdV}(t)} & \tilde{v}(t) \end{array} \tag{6.1}$$

We can summarize the proof of Theorem 1.2 (using the same notation as in (6.1), which will be defined momentarily!) by saying that $u(t), \tilde{u}(t)$ are shown to be close at low frequencies by showing that $\tilde{v}(t), v(t)$ are likewise close.

We now turn to the details. Fix $s \geq -1/2$, $T > 0$, $N \gg 1$, B, and $u_0 \in H_0^s$; our implicit constants may depend on s, T, and $\|u_0\|_{H_0^s}$. We work exclusively in the time interval $[-T, T]$.

Let $\tilde{u}(t) := S_{BKdV}(t)u_0$ denote the evolution of the flow (1.7). Our task is to show that

$$\sup_{|t|\leq T} \|P_{\leq N^{1/2}}(S_{KdV}(t)u_0 - \tilde{u}(t))\|_{H_0^s} \lesssim N^{-\sigma}. \tag{6.2}$$

We first claim (in analogy with (1.3)) the bound

$$\sup_{|t|\leq T} \|\tilde{u}(t)\|_{H_0^s} \lesssim 1, \tag{6.3}$$

if N is large enough. This bound is achieved by a repetition of the arguments in [8]. As it is somewhat technical and uses techniques different from those elsewhere in this paper (notably the "I-method"), we defer the proof of (6.3) to an Appendix.

We may assume from (6.3) and the local well-posedness theory[22] that u_0, and hence $\tilde{u}$, is smooth.

The Miura transform (1.8) intertwines the KdV flow with the (renormalized) mKdV flow (1.9), (1.10). We seek a similar transform to intertwine the KdV-like flow S_{BKdV} with an mKdV-like flow. It turns out that the correct transform to use is given by

$$\mathbf{M}_B\tilde{v} := \tilde{v}_x + B(1 - P_0)(\tilde{v}^2) = \tilde{v}_x + B(\tilde{v}^2) - P_0(\tilde{v}^2), \tag{6.4}$$

where of course the multiplier B here is that which appears in the flow (1.7) above.

---

[22]The well-posedness theory for KdV from [18] can be applied without substantial change to the BKdV equation (1.7). The presence of the multiplier B on the right hand side presents no difficulty.

As with $\mathbf{M}$, the operator $\mathbf{M}_B$ is a locally Lipschitz map from H_0^{s+1} to H_0^s. We now address the question of invertibility of $\mathbf{M}_B$.

Let $\tilde{v}$ be a function bounded in H_0^{s+1}. We first look at the derivative operator $\mathbf{M}'_B(\tilde{v})$, defined by

$$\mathbf{M}'_B(\tilde{v})f := f_x + 2B(1-P_0)(\tilde{v}f).$$

Lemma 6.1. *Fix $\tilde{v} \in H_0^{s+1}, s \geq -\frac{1}{2}$, and allow the implicit constants in this Lemma to depend on $\|\tilde{v}\|_{H_0^{s+1}}$. If N is sufficiently large, then $\mathbf{M}'_B(\tilde{v})$ is invertible from H_0^s to H_0^{s+1}, in the sense that*

$$\|\mathbf{M}'_B(\tilde{v})^{-1}f\|_{H_0^{s+1}} \lesssim \|f\|_{H_0^s}$$

for all (smooth) f.

Proof. Recall from the proof of Theorem 2.1 that we have the bound

$$\|\mathbf{M}'(\tilde{v})^{-1}f\|_{H_0^{s+1}} \lesssim \|f\|_{H_0^s}. \tag{6.5}$$

We proved this for $s = -1/2$ but it is easy to see the same argument works for $s > -1/2$. From the resolvent identity

$$O^{-1} = A^{-1}(1-(A-O)A^{-1})^{-1}$$

it thus suffices to show that the operator

$$(\mathbf{M}'_B(\tilde{v}) - \mathbf{M}'(\tilde{v}))\mathbf{M}'(\tilde{v})^{-1}$$

is a contraction on H_0^s. Applying (6.5) again, it thus suffices to show the bound

$$\|\mathbf{M}'_B(\tilde{v})f - \mathbf{M}'(\tilde{v})f\|_{H_0^s} \ll \|f\|_{H_0^{s+1}}.$$

But the left-hand side is just

$$\|2(1-B)(\tilde{v}f)\|_{H_0^s} \lesssim N^{-\sigma}\|\tilde{v}f\|_{H^{s+\sigma}} \lesssim N^{-\sigma}\|\tilde{v}\|_{H_0^{s+1}}\|f\|_{H_0^{s+1}} \lesssim N^{-\sigma}\|f\|_{H_0^{s+1}}$$

by (2.2) for some $\sigma > 0$, and the claim follows if N is sufficiently large. □

Corollary 6.2. *Let $R > 0, s \geq -\frac{1}{2}$. If N is large enough depending on R, then there is a map $\mathbf{M}_B^{-1}$ defined on the ball $\mathbf{B}^\infty(0;R) := \{\tilde{u} \in H_0^s : \|\tilde{u}\|_{H_0^s} \leq R\}$ which inverts $\mathbf{M}_B$ and is a Lipschitz map from $\mathbf{B}^\infty(0;R)$ to H_0^{s+1}.*

Remark: Recall that M_B depends on N through the definition of B (see (1.7)).

Proof. Fix R; implicit constants are allowed to depend on R.

Let $\tilde{u} \in \mathbf{B}^\infty(0;R)$. To define $\mathbf{M}_B^{-1}$ at $\tilde{u}$ we of course have to solve the equation

$$\mathbf{M}_B\tilde{v} = \tilde{u}.$$

From Theorem 2.1 we can find a $\tilde{v}_{\text{appr}}$, bounded in H_0^{s+1}, such that

$$\mathbf{M}\tilde{v}_{\text{appr}} = \tilde{u}.$$

We now apply the ansatz $\tilde{v} = \tilde{v}_{\text{appr}} + \tilde{w}$. One easily checks, using (6.4), that $\tilde{w}$ verifies the difference equation

$$\tilde{w}_x + B(1-P_0)(2\tilde{v}_{\text{appr}}\tilde{w} + \tilde{w}^2) = (1-B)(\tilde{v}_{\text{appr}}^2)$$

or equivalently

$$\tilde{w} = \mathbf{M}'_B(\tilde{v}_{\mathrm{appr}})^{-1}(1-B)(\tilde{v}^2_{\mathrm{appr}}) - \mathbf{M}'_B(\tilde{v}_{\mathrm{appr}})^{-1}B(1-P_0)(\tilde{w}^2).$$

Since $\tilde{v}_{\mathrm{appr}}$ is bounded in H^{s+1} we see from Lemma 6.1 and (2.2) that

$$\|\mathbf{M}'_B(\tilde{v}_{\mathrm{appr}})^{-1}(1-B)(\tilde{v}^2_{\mathrm{appr}})\|_{H^{s+1}} \lesssim N^{-\sigma}.$$

A contraction mapping argument again using Lemma 6.1 and (2.2) thus shows that a solution $\tilde{w}$ to the above difference equation exists and obeys the bound

$$\|\tilde{w}\|_{H^{s+1}} \lesssim N^{-\sigma}$$

if N is sufficiently large. In particular we see that $\mathbf{M}_B^{-1}$ exists at $\tilde{u}$ and that $\mathbf{M}_B^{-1}$ is bounded on H^s_0.

The Lipschitz bound now follows from Lemma 6.1 and the inverse function theorem, since $\mathbf{M}_B$ is a smooth map from H^{s+1}_0 to H^s_0. (Equivalently, one can use contraction mapping arguments similar to the one above to show that $\mathbf{M}_B^{-1}$ is uniformly Lipschitz on very small neighbourhoods of $\tilde{u}$, and hence on the whole ball $\mathbf{B}^\infty(0;R)$). □

Thus if N is large enough, the above corollary and (6.3) let us write

$$\tilde{v}(t) \equiv \mathbf{M}_B^{-1}\tilde{u}(t) \tag{6.6}$$

and conclude also that,

$$\sup_{|t|\le T} \|\tilde{v}(t)\|_{H^{s+1}_0} \lesssim 1. \tag{6.7}$$

From the Leibnitz rule we see that

$$\begin{aligned}
\tilde{u}_t &= \mathbf{M}'_B(\tilde{v})\tilde{v}_t\\
\tilde{u}_x &= \mathbf{M}'_B(\tilde{v})\tilde{v}_x = \tilde{v}_{xx} + 2B(\tilde{v}\tilde{v}_x)\\
\tilde{u}_{xxx} &= \mathbf{M}'_B(\tilde{v})\tilde{v}_{xxx} + 6B(\tilde{v}_x\tilde{v}_{xx})\\
\tilde{u}u_x &= (\tilde{v}_x + B(\tilde{v}^2) - P_0(\tilde{v}^2))\mathbf{M}'_B(\tilde{v})\tilde{v}_x\\
&= \tilde{v}_x\tilde{v}_{xx} + 2\tilde{v}_xB(\tilde{v}\tilde{v}_x) + B(\tilde{v}^2)\tilde{v}_{xx} + 2B(\tilde{v}^2)B(\tilde{v}\tilde{v}_x) - \mathbf{M}'_B(\tilde{v})(P_0(\tilde{v}^2)\tilde{v}_x)
\end{aligned}$$

where we have used the fact that $P_0(ff_x) = 0$ for any f. Expanding out (1.7) and canceling the two terms of $6B(\tilde{v}_x\tilde{v}_{xx})$ which appear, we obtain

$$\mathbf{M}'_B(\tilde{v})(\tilde{v}_t+\tilde{v}_{xxx}) = 6B(2\tilde{v}_xB(\tilde{v}\tilde{v}_x)+B(\tilde{v}^2)\tilde{v}_{xx}+2B(\tilde{v}^2)B(\tilde{v}\tilde{v}_x))-B\mathbf{M}'_B(\tilde{v})(6P_0(\tilde{v}^2)\tilde{v}_x).$$

The first term of the right-hand side is roughly $\mathbf{M}'_B(\tilde{v})(6B(B(\tilde{v}^2)\tilde{v}_x))$. Indeed, a computation shows

$$\mathbf{M}'_B(\tilde{v})(6B(B(\tilde{v}^2)\tilde{v}_x)) = 6B(2\tilde{v}_xB(\tilde{v}\tilde{v}_x) + B(\tilde{v}^2)\tilde{v}_{xx}) + 12B(1-P_0)(\tilde{v}B(B(\tilde{v}^2)\tilde{v}_x)).$$

Thus we have

$$\mathbf{M}'_B(\tilde{v})(\tilde{v}_t + \tilde{v}_{xxx} - 6B(B(\tilde{v}^2)\tilde{v}_x) + 6B(P_0(\tilde{v}^2)\tilde{v}_x)) = 12E_1 + 6E_2$$

where the error terms E_1, E_2 are the "commutator expressions"

$$\begin{aligned}
E_1 &:= B(B(\tilde{v}^2)B(\tilde{v}\tilde{v}_x) - (1-P_0)(\tilde{v}B(B(\tilde{v}^2)\tilde{v}_x)))\\
E_2 &:= P_0(\tilde{v}^2)[\mathbf{M}'_B(\tilde{v}), B]\tilde{v}_x
\end{aligned}$$

Thus $\tilde{v}$ obeys the equation

$$\tilde{v}_t + \tilde{v}_{xxx} = 6B((B - P_0)(\tilde{v}^2)\tilde{v}_x) + \mathbf{M}'_B(\tilde{v})^{-1}(12E_1 + 6E_2); \quad \tilde{v}(0) = \tilde{v}_0. \tag{6.8}$$

We have written $S_{BmKdV}(t)$ in figure (6.1) to represent this flow. Since $\tilde{v}$ is smooth, it is *a priori* in the space Y^{s+1} when restricted to the interval $[-T, T]$. We now seek to control the non-linear terms in (6.8).

If it were not for the error terms E_1, E_2, one could obtain bounds of the form

$$\|\tilde{v}\|_{Y^{s+1}} \lesssim 1 \tag{6.9}$$

from (6.7) and the local well-posedness theory for mKdV in [9] (which can easily handle the presence of the order 0 operator B). To deal with the E_1, E_2 terms we use the following estimate.

Lemma 6.3. *We have*

$$\|\mathbf{M}'_B(\tilde{v})(t)^{-1}E_j\|_{Z^{s+1}} \leq CN^{-\sigma}. \tag{6.10}$$

for $j = 1, 2$*, and* $t \in [-T, T]$.

Proof. By (3.3) and Lemma 6.1 (using (6.7), of course) it suffices to show that

$$\|E_j\|_{L_t^\infty H_x^s} \lesssim N^{-\sigma}. \tag{6.11}$$

We first prove this for E_1. Observe that $B(\tilde{v}^2)B(\tilde{v}\tilde{v}_x) = \partial_x \frac{1}{4}(B(\tilde{v}^2))^2$ has mean zero, and so we can factor out a $(1 - P_0)$, and reduce to showing that

$$\|\tilde{w}B(\tilde{v}\tilde{v}_x) - \tilde{v}B(\tilde{w}\tilde{v}_x)\|_{H_x^s} \lesssim N^{-\sigma}$$

where we have used the shorthand $\tilde{w} := B(\tilde{v}^2)$.

By (2.2) we see that $\tilde{w}$ is bounded in $H_x^{s+\sigma}$ for some $\sigma > 0$. From the identity

$$\tilde{w}B(\tilde{v}\tilde{v}_x) - \tilde{v}B(\tilde{w}\tilde{v}_x) = \tilde{w}[B, \tilde{v}]\tilde{v}_x - \tilde{v}[B, \tilde{w}]\tilde{v}_x$$

and another application of (2.2), we see that it suffices to show the commutator estimate

$$\|[B, f]g\|_{H_x^s} \lesssim N^{-\sigma/2}\|f\|_{H_x^{s+\sigma}}\|g\|_{H_x^s}. \tag{6.12}$$

Without loss of generality we may assume that f and g have non-negative Fourier transform. Observe that

$$\widehat{[B, f]g}(k) = \sum_{k_1+k_2=k} (b(k) - b(k_2))\widehat{f}(k_1)\widehat{g}(k_2).$$

The quantity $b(k) - b(k_2)$ is clearly $O(1)$. If $|k_1| \ll N$ then one also obtains a bound of $O(|k_1|/N)$ by the mean-value theorem. Thus we have a universal bound of

$$|b(k) - b(k_2)| \lesssim |k_1|^{\sigma/2}N^{-\sigma/2}.$$

The commutator estimate then reduces to

$$\|(|\partial_x|^{\sigma/2}f)g\|_{H_x^s} \lesssim \|f\|_{H_x^{s+\sigma}}\|g\|_{H_x^s},$$

but this follows from (2.2).

Now we prove (6.11) for E_2. From (6.7) we see that $P_0(\tilde{v}^2)$ is bounded in time, so it suffices to show that

$$\|[\mathbf{M}'_B(\tilde{v}), B]\tilde{v}_x\|_{L_t^\infty H_x^s} \lesssim N^{-\sigma}.$$

Since $[\partial_x, B] = 0$, we have

$$[\mathbf{M}'_B(\tilde{v}), B]\tilde{v}_x = B(1-P_0)(\tilde{v}B\tilde{v}_x) - B^2(1-P_0)(\tilde{v}\tilde{v}_x) = B(1-P_0)[\tilde{v}, B]\tilde{v}_x,$$

and the claim follows from (6.12). □

From this lemma and perturbation theory in the Y^{s+1} spaces (using the local well-posedness theory in [9]) we thus obtain (6.9).

We now repeat the argument from Section 5. Recall the notation from figure (6.1) that $v(t) \equiv S_{mKdV}(t)v_0$. From (1.3), (2.1) and Theorem 2.1 we see that $v(t)$ is uniformly bounded in H^{s+1}. From the local well-posedness theory for mKdV we thus have

$$\|v\|_{Y^{s+1}} \lesssim 1.$$

From this and (6.9), we may find an interval $[M, M+N^{1/4}] \subseteq [N^{1/2}, 2N^{1/2}]$ such that

$$\|(P_{\leq M+N^{1/4}} - P_{\leq M})\tilde{v}\|_{Y^{s+1}} + \|(P_{\leq M+N^{1/4}} - P_{\leq M})v\|_{Y^{s+1}} \lesssim N^{-\sigma}.$$

Fix this M. Set

$$\tilde{v}_{lo} := P_{\leq M}\tilde{v} \quad \text{and} \quad v_{lo}(t) := P_{\leq M}v.$$

By arguing as in the previous section we see that v_{lo} obeys the equation

$$(\partial_t + \partial_{xxx})v_{lo} = F_0(v_{lo}, v_{lo}, v_{lo}) + F_{\neq 0}(v_{lo}, v_{lo}, v_{lo}) + \text{ error terms} \tag{6.13}$$

where the error terms have a Z^{s+1} norm of $O(N^{-\sigma})$. We now claim that $\tilde{v}_{lo}$ obeys the same equation (but with a different set of error terms, of course). Assuming this claim for the moment, note that v_{lo} and $\tilde{v}_{lo}$ have the same initial data, so we obtain,

$$\sup_{|t|\leq T} \|v_{lo}(t) - \tilde{v}_{lo}(t)\|_{H_0^{s+1}} \lesssim N^{-\sigma} \tag{6.14}$$

by perturbation theory. The bound (6.14) implies our goal (6.2) relatively quickly: apply the Miura transform $\mathbf{M}$ (see (1.8)) to the difference on the left side of (6.14), and use the commutator bound (5.2), the fact that $P_{\leq M}\mathbf{M} \equiv P_{\leq M}\mathbf{M}_B$, and $M \geq N^{\frac{1}{2}}$ to conclude that,

$$\begin{aligned} N^{-\sigma} &\gtrsim \|P_{\leq M}\mathbf{M}_B\tilde{v}(t) - P_{\leq M}\mathbf{M}v(t)\|_{H_0^s} \\ &\gtrsim \|P_{\leq N^{\frac{1}{2}}}\tilde{u}(t) - P_{\leq N^{\frac{1}{2}}}u\|_{H_0^s}, \end{aligned} \tag{6.15}$$

as desired (see (6.2)).

It remains to show that $\tilde{v}_{lo}$ verifies (6.13). Applying $P_{\leq M}$ to (6.8) and using Lemma 6.3 we have

$$(\partial_t + \partial_{xxx})\tilde{v}_{lo} = 6P_{\leq M}((B-P_0)(\tilde{v}^2)\tilde{v}_x) + error.$$

By repeating the argument in Section 5 we have

$$6P_{\leq M}((1-P_0)(\tilde{v}^2)\tilde{v}_x) = P_{\leq M}(F_0(\tilde{v},\tilde{v},\tilde{v}) + F_{\neq 0}(\tilde{v},\tilde{v},\tilde{v})) = F_0(\tilde{v}_{lo},\tilde{v}_{lo},\tilde{v}_{lo}) + F_{\neq 0}(\tilde{v}_{lo},\tilde{v}_{lo},\tilde{v}_{lo}) + \text{error terms}.$$

Thus it will suffice to show that

$$P_{\leq M}((1-B)(\tilde{v}^2)\tilde{v}_x) = \text{error terms.} \tag{6.16}$$

For a fixed time t, the spatial Fourier coefficient of the left-hand side at (k,t) is

$$\sum_{k=k_1+k_2+k_3} \chi_{[-M,M]}(k)(1-b(k_1+k_2))\widehat{\tilde{v}}(k_1,t)\widehat{\tilde{v}}(k_2,t)ik_3\widehat{\tilde{v}}(k_3,t).$$

The summand vanishes unless $|k| \leq M \lesssim N^{1/2}$ and $|k_1+k_2| \gtrsim N$, which forces $|k_3| \gtrsim N$.

First consider the contributions of the case when $(k_1+k_2)(k_2+k_3)(k_1+k_3) \neq 0$. We now apply (4.7). By our previous discussion we have $N_0 \lesssim N^{1/2}$ and $N_{soprano} \gtrsim N$, hence we see from (6.9) (writing things in terms of spacetime Fourier transforms instead of spatial Fourier transforms, taking absolute values and discarding the $1-b(k_1+k_2)$ factor) that this contribution is $error$.

It remains to consider the case when $(k_1+k_2)(k_2+k_3)(k_1+k_3) = 0$. By the previous discussion k_1+k_2 cannot be zero, while $|k_3|$ is much larger than $|k|$. Thus the only two cases are when (k_1,k_2,k_3) is equal to $(k,-k_3,k_3)$ or $(-k_3,k,k_3)$, so by symmetry the total contribution to the Fourier coefficient is

$$2\chi_{[-M,M]}(k)\sum_{|k_3|\gtrsim N} ik_3(1-b(k_3-k))\widehat{\tilde{v}}(k,t)\widehat{\tilde{v}}(-k_3,t)\widehat{\tilde{v}}(k_3,t).$$

Combining the k_3 term with the $-k_3$ term, this becomes

$$2\chi_{[-M,M]}(k)\sum_{k_3\gtrsim N} ik_3(b(-k_3-k)-b(k_3-k))\widehat{\tilde{v}}(k,t)\widehat{\tilde{v}}(-k_3,t)\widehat{\tilde{v}}(k_3,t).$$

By the mean-value theorem and the fact that b is even, we have

$$(b(-k_3-k)-b(k_3-k)) = O(|k|/N) = O(N^{-\sigma}).$$

Meanwhile, we have

$$\sum_{k_3\gtrsim N} |k_3||\widehat{\tilde{v}}(-k_3,t)||\widehat{\tilde{v}}(k_3,t)| \lesssim \|\tilde{v}\|^2_{H_0^{s+1}} \lesssim 1.$$

Thus the above Fourier coefficient is $O(N^{-\sigma}|\widehat{\tilde{v}}(k,t)|)$. By (6.7) we thus see that this contribution to (6.16) has an $L_t^\infty H_0^{s+1}$ norm of $O(N^{-\sigma})$. By (3.3) we thus see that this contribution is $error$ as desired. This completes the proof of (6.13) and hence (6.2). This concludes the proof of Theorem 1.2.

7. Proof of Theorem 1.5: Symplectic nonsqueezing of KdV

Let $N \gg 1$, and let b be a symbol adapted to $[-N,N]$ which equals one on $[-N/2,N/2]$, and let B be the associated Fourier multiplier. We begin by considering the modified Hamiltonian H_N on $P_{\leq N}H_0^{-1/2}(\mathbb{T})$, defined by

$$H_N(u) := \int_{\mathbb{T}} -\frac{1}{2}u_x^2 - (Bu)^3\, dx.$$

We compute the Hamiltonian flow on $P_{\leq N} H_0^{-1/2}$ corresponding to H_N. Fix $u, v \in H_0^{-1/2}$. We see that

$$\begin{aligned}\frac{d}{d\varepsilon} H_N(u+\varepsilon v)|_{\varepsilon=0} = &\int_{\mathbb{T}} -u_x v_x - 3(Bu)^2 Bv \, dx \\ =&\{-u_{xxx} + 6B((Bu)(Bu_x)), v\}.\end{aligned}$$

Since $-u_{xxx} + 6B((Bu)(Bu_x))$ is in $P_{\leq N} H_0^{-1/2}$, we conclude as in (1.15), (1.16) that the Hamiltonian flow of H_N on $P_{\leq N} H_0^{-1/2}$ is given by

$$u_t + u_{xxx} = 6B((Bu)(Bu_x)); \quad u(0) = u_0 \in P_{\leq N} H_0^{1/2}(\mathbb{T}). \tag{7.1}$$

Let $S_{KdV}^{(N)}(t)$ denote the flow map associated to this equation; for each t, we observe that $S_{KdV}^{(N)}(t)$ is thus a symplectomorphism on the finite-dimensional symplectic vector space $P_{\leq N} H_0^{-1/2}$. In particular, it obeys Theorem 1.7 (that is, we pick $S_{\text{Good!}}^{(N)} \equiv S_{KdV}^{(N)}$)). To conclude the proof of Theorem 1.5 it thus suffices to show that the flow $S_{KdV}^{(N)}(t)$ obeys the weak approximation property in Condition 1.8:

Proposition 7.1. *Let $k_0 \in \mathbb{Z}^*$, $T > 0$, $A > 0$, $0 < \varepsilon \ll 1$. Then there exists a frequency $N_0 = N_0(k_0, T, \varepsilon, A) \gg |k_0|$ such that*

$$|k_0|^{-1/2} |\widehat{S_{KdV}(T)u_0}(k_0) - \widehat{S_{KdV}^{(N)}(T)u_0}(k_0)| \leq \varepsilon$$

for all $N \geq N_0$ and all $u_0 \in \mathbf{B}^N(0, A)$ (see (1.25) for the definition of this ball).

Proof. We make the transformation $w := Bu$, where u solves (7.1). Applying B to (7.1) we obtain

$$w_t + w_{xxx} = 6B^2(ww_x); \quad w(0) = Bu_0$$

which is (1.7) with B replaced by B^2. Thus we have the intertwining relationship described by (1.27) in the introduction to this paper,

$$BS_{KdV}^{(N)}(t)u_0 = S_{B^2 KdV}(t)Bu_0.$$

In particular, if $N_0 \gg |k_0|$, then $b(k_0) = 1$, so we have

$$\widehat{S_{KdV}^{(N)}(T)u_0}(k_0) = (\widehat{S_{B^2 KdV}(T)Bu_0})(k_0). \tag{7.2}$$

From Theorem 1.3 we have

$$|k_0|^{-1/2} |\widehat{S_{KdV}(T)u_0}(k_0) - \widehat{S_{KdV}(T)Bu_0}(k_0)| \lesssim N^{-\sigma}. \tag{7.3}$$

From Theorem 1.2 we have (if N_0 is large enough, $N_0 \gg k_0$)

$$|k_0|^{-1/2} |\widehat{S_{KdV}(T)Bu_0}(k_0) - \widehat{S_{B^2 KdV}(T)Bu_0}(k_0)| \lesssim N^{-\sigma}, \tag{7.4}$$

where the implicit constants are allowed to depend on T and A. By (7.2), the second term on the left of (7.4) is the same as $\widehat{S_{KdV}^{(N)}(T)u_0}(k_0)$. Combining this observation with (7.3), (7.4), and the triangle inequality, we obtain the desired claim, if N_0 is sufficiently large depending on k_0, T, ε, A. ☐

The proof of Theorem 1.5 is now complete.

8. Proof of Theorem 1.1: $P_{\leq N}KdV$ does not approximate KdV

Informally, the point of this section is that there is absolutely no slack in the bilinear estimate (1.4) at regularity $s = -1/2$ no matter what the frequencies of the various functions are; see the examples in [18]. But to convert the examples for the bilinear estimate to quantitative estimates of the KdV and truncated KdV flow - in particular, to establish that the two flows differ as claimed in Theorem 1.1 - we must do some tedious computation of iterates, which we detail below.

Fix k_0,A, T, for instance $T, A \sim 1$; our implicit constants in this section will be allowed to depend on these parameters. Without loss of generality we may assume that $k_0 > 0$. We let $0 < \sigma \ll 1$ be a small parameter depending on k_0, A, T to be chosen later.

Let $N \gg \sigma^{-100}$ be a large integer. We consider the initial data

$$u_0(x) := \sigma^3 \cos(k_0 x) + \sigma N^{1/2} \cos(Nx).$$

Note that u_0 lies in $P_{\leq N} H_0^{-1/2}(\mathbb{T})$ with norm $O(\sigma)$, and in particular we have $u_0 \in \mathbf{B}^N(0; A)$ if $\sigma \ll 1$ is sufficiently small.

Let u and $u^{(N)}$ be the solutions to the KdV flow (1.1) and truncated KdV flow (1.5) respectively, with initial data $u(0) = u^{(N)}(0) = u_0$. We shall show that, if σ is sufficiently small,

$$|\widehat{u(T)}(k_0) - \widehat{u^{(N)}(T)}(k_0)| \sim \sigma^5, \tag{8.1}$$

which gives (1.6).

To prove (8.1) we need good approximations of u and $u^{(N)}$. To approximate u, we look at the iterates $u^{[j]}$ for $j = 0, 1, 2, \ldots$ defined inductively by $u^{[-1]}(t,x) \equiv 0$ and

$$(\partial_t + \partial_{xxx})u^{[j]} = \partial_x(3(u^{[j-1]})^2); \quad u^{[j]}(0) = u_0. \tag{8.2}$$

From the contraction mapping arguments in [18] (see also [9]) we know that the $u^{[j]}$ converge to u in the $Y^{-1/2}_{[0,T]}$ norm; indeed each iterate is closer to u by a factor of at least $O(\sigma)$ compared to the previous one[23]. A routine calculation yields

$$u^{[0]}(t,x) = \sigma^3 \cos(k_0 x + k_0^3 t) + \sigma N^{1/2} \cos(Nx + N^3 t),$$

and thus

$$\partial_x(3(u^{[0]})^2) = -\frac{3}{2}\sigma^4 N^{3/2} \sin((N+k_0)x + (N^3+k_0^3)t) - \frac{3}{2}\sigma^4 N^{3/2} \sin((N-k_0)x + (N^3-k_0^3)t) + O_Z(\sigma^6)$$

[23]Strictly speaking, this contraction mapping property was only proven for T sufficiently small, but by subdividing $[0, T]$ into a finite number of small intervals one can obtain the same contraction mapping for arbitrary T if σ is sufficiently small depending on T. This naive argument requires $\sigma \ll e^{-CT}$ for some C; the more sophisticated scaling argument in [9] can improve this to $\sigma \ll T^{-1/3-}$, but we will not need this quantitative improvement for our arguments here.

where $O_Z(K)$ denotes a quantity with a $Z^{-1/2}_{[0,T]}$ norm of $O(K)$ (note that we have used the hypothesis $N \gg \sigma^{-100}$ to absorb several terms into this $O_Z(\sigma^6)$ error[24]).

Observe that

$$(\partial_t+\partial_{xxx})\left(-\frac{1}{2}\sigma^4 N^{-1/2}[\cos((N+k_0)x+(N^3+k_0^3)t)-\cos((N+k_0)x+(N+k_0)^3t)]\right)$$
$$= -\frac{3}{2}\sigma^4 N^{3/2}k_0 \sin((N+k_0)x+(N^3+k_0^3)t)+O_Z(\sigma^6)$$

and

$$(\partial_t+\partial_{xxx})\left(\frac{1}{2}\sigma^4 N^{-1/2}(\cos((N-k_0)x+(N^3-k_0^3)t)-\cos((N-k_0)x+(N-k_0)^3t))\right)$$
$$= -\frac{3}{2}\sigma^4 N^{3/2}k_0 \sin((N-k_0)x+(N^3-k_0^3)t)+O_Z(\sigma^6).$$

Combining this with the calculation of $\partial_x(3u^{[0]})^2$ above and using (3.5) we obtain

$$\begin{aligned} u^{[1]}(t,x) =& u^{[0]}(t,x) \\ &+ \left(-\frac{1}{2}\sigma^4 N^{-1/2}k_0^{-1}(\cos((N+k_0)x+(N^3+k_0^3)t)-\cos((N+k_0)x+(N+k_0)^3t))\right) \\ &+ \left(\frac{1}{2}\sigma^4 N^{-1/2}k_0^{-1}(\cos((N-k_0)x+(N^3-k_0^3)t)-\cos((N-k_0)x+(N-k_0)^3t))\right) \\ &+ O_Y(\sigma^6), \end{aligned}$$

where $O_Y(\sigma^6)$ denotes a quantity with a $Y^{-1/2}_{[0,T]}$ norm of $O(\sigma^6)$. In fact, since the $\cos((N\pm k_0)x+(N\pm k_0)^3t)$ terms are already $O_Y(\sigma^6)$ we have

$$u^{[1]}(t,x) = u^{[0]}(t,x)+\frac{1}{2}\sigma^4 N^{-1/2}k_0^{-1}\left(\cos((N-k_0)x+(N^3-k_0^3)t)-\cos((N+k_0)x+(N^3+k_0^3)t)\right)+O_Y(\sigma^6)$$

Using (1.4) to handle any interaction with a factor of σ^6 or better, we obtain

$$\partial_x(3(u^{[1]})^2) = \partial_x(3(u^{[0]})^2)+O_Z(\sigma^6). \tag{8.3}$$

Note that there are two additional, potentially disruptive terms of the form $\pm\frac{3}{2}\sigma^5\sin(k_0x+k_0^3t)$ which appear in the expansion of $\partial_x(3(u^{[1]})^2)$, but they have opposite signs and so cancel[25] each other. From (8.3) and (3.5) we have

$$u^{[2]} = u^{[1]}+O_Y(\sigma^6).$$

From the contraction mapping property of the iteration map we thus have

$$u = u^{[1]}+O_Y(\sigma^6).$$

[24]For example, the term $\sigma^4 N^{\frac{1}{2}}k_0\sin((N+k_0)x+(N^3+k_0^3)t)$ which appears when one calculates $\partial_x(3(u^{[0]})^2)$ is $O_Z(\sigma^6)$, as the space-time Fourier transform of this term is supported a distance approximately N^2 from the cubic $\tau=\xi^3$. Hence when computing the $Z^{-\frac{1}{2}}$ norm of this term, we get a factor of $N^{-1}\ll\sigma^{100}$ from the denominator in the definition of this norm.

[25]This special cancellation seems to be what distinguishes the KdV flow (1.1) from superficially similar flows such as (1.5), and is crucial to obtaining our high-frequency and low-frequency approximation results for this flow. It is instructive to see this cancellation via the renormalized mKdV flow (1.9) by computing iterates for mKdV and then applying the Miura transform to those iterates.

In particular we see that

$$\widehat{u(T)}(k_0) = \widehat{u^{[1]}(T)}(k_0) + O(\sigma^6) = \widehat{u^{[0]}(T)}(k_0) + O(\sigma^6). \tag{8.4}$$

Now we approximate $u^{(N)}$. To do this we construct iterates $\tilde{u}^{[j]}$, $j = 0, 1, 2, \ldots$ for the truncated equation by setting $\tilde{u}^{[0]} := u^{[0]}$ and

$$(\partial_t + \partial_{xxx})\tilde{u}^{[j]} = P_{\leq N}\partial_x(3(\tilde{u}^{[j-1]})^2); \quad \tilde{u}^{[j]}(0) = u_0.$$

By a variant of the local well-posedness theory from [18] (and [9]) we know that $\tilde{u}^{[j]}$ will converge to $u^{(N)}$ in the Y norm. By reviewing the computation of $u^{[1]}(t,x)$, but now bearing in mind the presence of the projection P_N, we obtain for the first iterate,

$$\begin{aligned}\tilde{u}^{[1]}(t,x) &= u^{[0]}(t,x) + \frac{1}{2}\sigma^4 N^{-\frac{1}{2}}\cos((N-k_0)x + (N^3 - k_0^3)t)) + O_Y(\sigma^6)\\ &= u^{[1]}(x,t) + O_Y(\sigma^6).\end{aligned}$$

Comparing this with the formula for $u^{[1]}$ above, we note that the Fourier modes at $\pm(N + k_0)$ are not present here. As a consequence, the analog of (8.3) reads,

$$\partial_x(3(\tilde{u}^{[1]})^2) = \partial_x(3(u^{[0]})^2) + \frac{3}{2}\sigma^5 \sin(k_0 x + k_0^3 t) + O_Z(\sigma^6).$$

Since $(\partial_t + \partial_x^3)(t\sin(k_0x + k_0^3t)) = \sin(k_0x + k_0^3t)$, we can write,

$$\tilde{u}^{[2]} = \tilde{u}^{[1]} + \frac{3}{2}\sigma^5 t\sin(k_0x + k_0^3t) + O_Y(\sigma^6).$$

We can easily check then that,

$$\partial_x(3(\tilde{u}^{[2]})^2) = \partial_x(3(u^{[1]})^2) + O_Z(\sigma^6),$$

hence $\tilde{u}^{[3]} = \tilde{u}^{[2]} + O_Y(\sigma^6)$, which by the contraction mapping property implies that

$$u^{(N)} = \tilde{u}^{[2]} + O_Y(\sigma^6).$$

In particular we see that

$$\widehat{u^{(N)}(T)}(k_0) = \widehat{\tilde{u}^{[2]}(T)}(k_0) + O(\sigma^6) = \widehat{u^{[1]}(T)}(k_0) - \frac{3}{2}iT\sigma^5 e^{ik_0^3T} + O(\sigma^6).$$

Comparing this with (8.4) we obtain (8.1) as desired. This proves Theorem 1.1.

9. Appendix. Proof of (6.3): H^s bound for the BKdV flow

We now prove the bound (6.3) for H_0^s solutions to the KdV-like equation

$$u_t + u_{xxx} = 6B(uu_x); \quad u(0) = u_0$$

with $\|u_0\|_{H_0^s} \lesssim 1$; this bound is needed to complete the proof of Theorem 1.2 and hence Theorem 1.5.

If $s \geq 0$ then this bound follows from L^2 conservation and standard persistence of regularity theory (see e.g. [1]), so we shall assume that $-1/2 \leq s < 0$.

To do so, let us first review (from [8]) how the corresponding bound (1.3) was proven for the KdV flow

$$u_t + u_{xxx} = 6uu_x; \quad u(0) = u_0.$$

9.1. **Review of proof of H^s bound for KdV** (1.3)**.** The idea is to modify the conserved L^2 norm $\int u^2$ to something resembling the H^s norm and which is still approximately conserved. To do this it is convenient to introduce some notation for multilinear forms.

If $n \geq 2$ is an integer, we define a *(spatial) n-multiplier* to be any function $M_n(k_1, \ldots, k_n)$ on the (discrete) hyperplane

$$\Gamma_n := \{(k_1, \ldots, k_n) \in \mathbb{Z}_*^n : k_1 + \ldots + k_n = 0\}.$$

If M_n is a n-multiplier and $u_1, \ldots, u_n$ are functions on $\mathbb{R}/2\pi\mathbb{Z}$, we define the *n-linear functional* $\Lambda_n(M_n; u_1, \ldots, u_n)$ by

$$\Lambda_n(M_n; f_1, \ldots, f_n) := \sum_{(k_1,\ldots,k_n)\in\Gamma_n} M_n(k_1, \ldots, k_n) \prod_{j=1}^n \widehat{f_j}(k_j).$$

We adopt the notation

$$\Lambda_n(M_n; u) := \Lambda_n(M_n; u, \ldots, u).$$

Observe that $\Lambda_n(M_n; f)$ is invariant under permutations of the k_j indices. In particular we have

$$\Lambda_n(M_n; u) = \Lambda_n([M_n]_{sym}; u)$$

where

$$[M_n]_{sym}(k) := \frac{1}{n!} \sum_{\sigma \in S_n} M_n(\sigma(k)) \tag{9.1}$$

is the symmetrization of M_n.

Thus, for instance, we have $\int u^2 = 2\pi\Lambda_2(1; u)$, and more generally $\|u\|_{H_0^s}^2 = 2\pi\Lambda_2(|k_1|^s|k_2|^s; u) = 2\pi\Lambda_2(|k_1|^{2s}; u)$ for $u \in H_0^s$.

Now suppose that u obeys the KdV evolution (1.1), and M_n is a symmetric multiplier. Then we have the differentiation law

$$\frac{d}{dt}\Lambda_n(M_n; u(t)) = \Lambda_n(M_n\alpha_n; u(t)) - 3in\Lambda_{n+1}(M_n(k_1, \ldots, k_{n-1}, k_n+k_{n+1})(k_n+k_{n+1}); u(t)) \tag{9.2}$$

where

$$\alpha_n := k_1^3 + \ldots + k_n^3.$$

(see [8]). Thus for instance we have

$$\begin{aligned}\frac{d}{dt}\Lambda_2(1; u(t)) &= \Lambda_2(\alpha_2; u(t)) - 6i\Lambda_3(k_2 + k_3; u(t)) \\ &= \Lambda_2(k_1^3 + k_2^3; u(t)) - 4i\Lambda_3(k_1 + k_2 + k_3; u(t)) \\ &= 0 - 0,\end{aligned}$$

demonstrating the conservation of the L^2 norm.

Henceforth we shall omit the $u(t)$ from the Λ_n notation for brevity. We also adopt the convenient notation that $k_{ij} := k_i + k_j$, etc., thus for instance $k_{145} = k_1 + k_4 + k_5$. Also we write $m_i := m(k_i)$, $m_{ij} := m(k_{ij})$, etc, and N_i for $|k_i|$, N_{ij} for $|k_{ij}|$, etc.

Let $A \gg 1$ be a large number to be chosen later[26], and let $m(k)$ be a multiplier which equals 1 on $[-A, A]$, equals $(|k|/A)^s$ for $|k| \geq 2A$, and is real, even, and smooth in between. We denote the corresponding Fourier multiplier by I:

$$\widehat{Iu}(k) := m(k)\widehat{u}(k),$$

thus I acts like the identity on frequencies $\leq A$ and is smoothing on frequencies $\gtrsim A$. We define the modified energy $E_2(t)$ by

$$E_2(t) := \Lambda_2(m_1 m_2),$$

then one can verify that

$$\|u(t)\|^2_{H^s_0} \lesssim E_2(t) \lesssim A^{-2s}\|u(t)\|^2_{H^s_0}.$$

From (9.2), (9.1) and the fact that $\alpha_2 = 0$ we have

$$\begin{aligned}\frac{d}{dt}E_2(t) &= -6i\Lambda_3(m_1 m_{23} k_{23})\\ &= 6i\Lambda_3(m_1^2 k_1)\\ &= \Lambda_3(M_3)\end{aligned}$$

where M_3 is the 3-multiplier

$$M_3 := 2i(m_1^2 k_1 + m_2^2 k_2 + m_3^2 k_3).$$

Now define the modified energy $E_3(t)$ by

$$E_3(t) := E_2(t) + \Lambda_3(\sigma_3)$$

where $\sigma_3(k_1, k_2, k_3)$ is the 3-multiplier

$$\sigma_3 := -M_3/\alpha_3.$$

This multiplier may appear to be singular at first glance, but we observe that

$$\alpha_3 = k_1^3 + k_2^3 + k_3^3 = 3k_1 k_2 k_3 \tag{9.3}$$

and that M_3 vanishes whenever $k_1 k_2 k_3 = 0$. Then by (9.2), (9.1) we have

$$\begin{aligned}\frac{d}{dt}E_3(t) &= \Lambda_3(M_3) + \Lambda_3(\sigma_3\alpha_3) - -9i\Lambda_4(\sigma_3(k_1, k_2, k_{34})k_{34})\\ &= \Lambda_4(M_4)\end{aligned}$$

where M_4 is the 4-multiplier

$$M_4 := -9i[\sigma_3(k_1, k_2, k_{34})k_{34}]_{sym}.$$

Now define the modified energy $E_4(t)$ by

$$E_4(t) := E_3(t) + \Lambda_4(\sigma_4)$$

[26]Note that the quantity A here represents what was called N in [8], a notational change necessary since in the present paper N represents something else.

where $\sigma_4 k_1, k_2, k_3, k_4)$ is the 4-multiplier

$$\sigma_4 := -M_4/\alpha_4$$

This multiplier may appear to be singular at first glance, but we observe that

$$\alpha_4 = k_1^3 + k_2^3 + k_3^3 + k_4^3 = 3k_{12}k_{13}k_{14} \tag{9.4}$$

(cf. (4.17)), and one can check that M_4 vanishes when $k_{12}k_{13}k_{14} = 0$. Then as before we have that

$$\frac{d}{dt}E_4(t) = \Lambda_5(M_5) \tag{9.5}$$

where

$$M_5 := -12i[\sigma_4(k_1, k_2, k_3, k_{45})k_{45}]_{sym}.$$

We could continue this procedure indefinitely, but E_4 will turn out to be a suitable almost conserved quantity for our purposes. In [8] it was shown (by Gagliardo-Nirenberg type arguments) that E_4 is bounded if and only if $\|u\|_{H_0^s}$ is bounded, so to obtain (1.3) it suffices to control $E_4(t)$. In light of (9.5) it will suffice to control M_5. The key lemma here was the following:

Lemma 9.1. [8] *Let k_1, k_2, k_3, k_4, k_5 be real numbers (not necessarily integer) such that $k_{12345} = 0$. Then $M_5(k_1, \dots, k_5)$ vanishes when $N_1, \dots, N_5 \ll A$. In all other cases we have the bound*

$$|M_5(k_1, \dots, k_5)| \lesssim [\frac{m^2(N_{*45})N_{45}}{(A+N_1)(A+N_2)(A+N_3)(A+N_{45})}]_{sym}$$

where

$$N_{*45} = \min(N_1, N_2, N_3, N_{45}, N_{12}, N_{13}, N_{14}).$$

With this bound and some multilinear Y^s estimates[27], bounds on the growth of $E_4(t)$ was obtained. In particular, if $E_4(T)$ was small for some time T, it was possible to obtain the bound $E_4(T+\delta) = E_4(T) + O(A^{-\frac{5}{2}-})$ for some small time $\delta \sim 1$. Iterating this and using a rescaling argument one could obtain (1.3) for all $s \geq -1/2$ (after choosing A appropriately depending on $\|u_0\|_{H_0^s}$ and T). See [8] for details.

9.2. **Adapting the argument to the BKdV flow.** We now adapt the above argument to the flow (1.7). The main difference will be the appearance of various quantities of the form $b(k_i)$, $b(k_{ij})$, etc., however these factors will play essentially no role in the argument. Accordingly, we write b_i for $b(k_i)$, etc. We shall assume that the frequency parameter N corresponding to b is much larger than the frequency parameter A corresponding to m.

[27]Strictly speaking, in order to handle large data, these estimates had to take place in the large period setting $\mathbb{R}/2\pi\lambda\mathbb{Z}$, as one would need to rescale large data to be small. This causes some unpleasant technical complications in the arguments, and in particular this is why the k_j in the above lemma need to be real (or lie in $\mathbb{Z}_*/\lambda$) rather than integer. See [8], [9] for more details. In this paper we will ignore the large period issue as it does not cause any essential change to the argument.

Suppose $\tilde{u}$ solves (1.7). Then (9.2) now becomes

(9.6)
$$\frac{d}{dt}\Lambda_n(M_n;\tilde{u}(t)) = \Lambda_n(M_n\alpha_n;\tilde{u}(t)) - 3in\Lambda_{n+1}(M_n(k_1,\ldots,k_{n-1},k_n+k_{n+1})b(k_n+k_{n+1})(k_n+k_{n+1});\tilde{u}(t)).$$

Again we define
$$E_2(t) := \Lambda_2(m_1 m_2),$$
then one can verify that
$$\frac{d}{dt}E_2(t) = \Lambda_3(M_3)$$
where M_3 is the 3-multiplier
$$M_3 := 2i(f_1 + f_2 + f_3)$$
and $f(k) := m^2(k)b(k)k$. Observe that f is an odd function with $f'(k) = O(m(k))$ and $f''(k) = O(m(k)/(A + |k|))$ for all k.

We observe the following bounds on M_3.

Lemma 9.2. *If $N_1, N_2, N_3 \ll A$, then $M_3 = 0$. Otherwise, we have*
$$|M_3| \lesssim \max(m_1^2, m_2^2, m_3^2)\min(N_1, N_2, N_3).$$

Proof. [8] When $N_1, N_2, N_3 \ll A$ then $f_i = k_i$ for $i = 1, 2, 3$, and the claim is clear. Otherwise, we use symmetry to assume that $N_1 \sim N_2 \gtrsim N_3$. But then the mean-value theorem and the above bounds on f gives
$$f_2 = -f_{13} = -f_1 + O(m_1^2 N_3),$$
and the claim easily follows. □

Now define the modified energy $E_3(t)$ by
$$E_3(t) := E_2(t) + \Lambda_3(\sigma_3)$$
where $\sigma_3(k_1, k_2, k_3)$ is the 3-multiplier
$$\sigma_3 := -M_3/\alpha_3.$$

From Lemma 9.2, (9.3) we see that σ_3 vanishes when $\max(N_1, N_2, N_3) \ll N$, and we have the bounds

$$|\sigma_3| \lesssim \frac{\max(m_1^2, m_2^2, m_3^2)}{(N + \max(N_1, N_2, N_3))^2} \tag{9.7}$$

otherwise (note that the two largest values of N_j have to be comparable).

By (9.2), (9.1) we have
$$\frac{d}{dt}E_3(t) = \Lambda_4(M_4)$$
where M_4 is the 4-multiplier
$$M_4 := -9i[\sigma_3(k_1, k_2, k_{34})b_{34}k_{34}]_{sym}.$$
Now define the modified energy $E_4(t)$ by
$$E_4(t) := E_3(t) + \Lambda_4(\sigma_4)$$

where $\sigma_4(k_1, k_2, k_3, k_4)$ is the 4-multiplier

$$\sigma_4 := -M_4/\alpha_4.$$

Then as before we have that

$$\frac{d}{dt}E_4(t) = \Lambda_5(M_5)$$

where

$$M_5 := -12i[\sigma_4(k_1, k_2, k_3, k_{45})b_{45}k_{45}]_{sym}.$$

Our aim is to show that this new M_5 still verifies the bounds in Lemma 9.1; the rest of the arguments in [8] will then give the desired bound (6.3) (the presence of the B multiplier having no impact on the local well-posedness theory).

From the definition of σ_4 and M_5, it will suffice to prove the following M_4 bound.

Lemma 9.3. *If* $\max(N_1, N_2, N_3, N_4) \ll A$ *then* M_4 *vanishes. Otherwise, we have*

$$|M_4| \lesssim \frac{|\alpha_4| m^2(N_*)}{(A + N_1)(A + N_2)(A + N_3)(A + N_4)},$$

where $N_* := \min(N_1, N_2, N_3, N_4, N_{12}, N_{13}, N_{14})$.

Proof. When $\max(N_1, N_2, N_3, N_4) \ll A$ then $\sigma_3(k_1, k_2, k_{34})$ and all of its symmetrizations vanish, hence M_4 vanishes. Now we assume that $\max(N_1, N_2, N_3, N_4) \gtrsim A$. By symmetry we may assume that $N_1 \gtrsim N_2 \gtrsim N_3 \gtrsim N_4$, thus $N_1 \sim N_2 \gtrsim A$. From (9.4) we have $|\alpha_4| \sim N_{13}N_{14}N_{34}$.

We divide into several cases depending on the relative sizes of N_2, N_3, N_4.

Case 1: $N_2 \gg N_3 \gg N_4$**.** In this case $|\alpha_4| \sim N_1^2 N_3$, thus we reduce to showing

$$|M_4| \lesssim \frac{m(N_*)^2}{A + N_4}.$$

But from Lemma 9.2 we have

$$|\sigma_3(k_a, k_b, k_{cd})b_{cd}k_{cd}| \lesssim \frac{\min(m_a, m_b, m_{cd})^2}{A + \max(N_a, N_b, N_{cd})} \lesssim m(N_*)^2/(A + N_4)$$

as desired.

Case 2: $N_2 \sim N_3 \gg N_4$**.** In this case $|\alpha_4| \sim N_1^3$, thus we reduce to showing

$$|M_4| \lesssim \frac{m(N_*)^2}{A + N_4}.$$

One then proceeds as in Case 1.

Case 3: $N_2 \gg N_3 \sim N_4$**.** In this case $|\alpha_4| \sim N_1^2 N_{34}$, thus we reduce to showing

$$|M_4| \lesssim \frac{m(N_*)^2 N_{34}}{(A + N_3)^2}.$$

From Lemma 9.2 we have

$$|\sigma_3(k_1, k_2, k_{34})b_{34}k_{34}| \lesssim m(N_*)^2 N_{34}/(A + \max(N_1, N_2, N_{34}))^2$$

which is acceptable. Similarly

$$|\sigma_3(k_3, k_4, k_{12})b_{12}k_{12}| \lesssim m(N_*)^2 N_{12}/(A + \max(N_3, N_4, N_{12}))^2$$

is acceptable since $N_{12} = N_{34}$. It thus suffices to show that

$$|\sigma_3(k_1, k_3, k_{24})b_{24}k_{24}+\sigma_3(k_1, k_4, k_{23})b_{23}k_{23}+\sigma_3(k_2, k_3, k_{14})b_{14}k_{14}+\sigma_3(k_2, k_4, k_{13})b_{13}k_{13}| \lesssim \frac{m(N_*)^2 N_{34}}{(A + N_3)^2}.$$

We expand out σ_3 using (9.3), and replace k_1 by $-k_{234}$ throughout, and reduce to showing

$$|-\frac{b_{24}(f_3 + f_{24} - f_{234})}{k_{234}k_3} - \frac{b_{23}(f_4 + f_{23} - f_{234})}{k_{234}k_4} + \frac{b_{23}(f_2 + f_3 - f_{23})}{k_2k_3} + \frac{b_{24}(f_2 + f_4 - f_{24})}{k_2k_4}| \lesssim \frac{m(N_*)^2 N_{34}}{(A + N_3)^2}.$$

From the mean-value theorem we have $b_{23} = b_2 + O(N_3/N_2) = b_2 + O(N_3/(A + N_3))$. Similarly $b_{24} = b_2 + O(N_3/(A + N_3))$. Let us consider the contribution of the $O(N_3/(A + N_1))$ errors. It will suffice to show that

$$-\frac{f_3 + f_{24} - f_{234}}{k_{234}k_3} + \frac{f_2 + f_4 - f_{24}}{k_2k_4}$$

and

$$-\frac{f_4 + f_{23} - f_{234}}{k_{234}k_4} + \frac{f_2 + f_3 - f_{23}}{k_2k_3}$$

are both $O(\frac{m(N_*)^2 N_{34}}{N_3(A+N_3)})$. By the $k_3 \leftrightarrow k_4$ symmetry it suffices to estimate the former expression. From the mean-value theorem we have

$$\frac{1}{k_{234}k_3} = \frac{1}{(k_2 + k_{34})(-k_4 + k_{34})} = -\frac{1}{k_2k_4} + O(\frac{N_{34}}{N_2N_4^2}).$$

By Lemma 9.2, the contribution of the error term $O(\frac{N_{34}}{N_2N_4^2})$ is bounded by

$$m(N_*)^2 N_3 O(\frac{N_{34}}{N_2N_4^2}),$$

which is acceptable. Thus it suffices to show that

$$\frac{f_3 + f_{24} - f_{234}}{k_2k_4} + \frac{f_2 + f_4 - f_{24}}{k_2k_4} = O(\frac{m(N_*)^2 N_{34}}{N_3(A + N_3)}).$$

But from the mean-value theorem we have

$$f(k_2) - f(k_{234}) + f(k_3) - f(k_3 - k_{34}) = O(m(N_*)^2 N_{34}),$$

and the claim follows by dividing by k_2k_4.

Case 4: $N_2 \sim N_3 \sim N_4$. Observe this case is essentially symmetric in the indices $1, 2, 3, 4$. By definition of M_4, σ_3, α_3 we have

$$\begin{aligned}|M_4| &\sim |[\frac{(f_1 + f_2 + f_{34})b_{34}}{k_1k_2}]_{sym}| \\ &\sim N_1^{-4}|[(f_1 + f_2 + f_{34})b_{34}k_3k_4]_{sym}|.\end{aligned}$$

Our task is thus to show that

$$[(f_1 + f_2 + f_{34})b_{34}k_3k_4]_{sym} = O(m(N_*)^2 N_{12}N_{23}N_{13}).$$

Since $b_{34} = b_{12}$, it will suffice by symmetry to show that

$$(f_1 + f_2 + f_{34})k_3k_4 + (f_3 + f_4 + f_{12})k_1k_2 = O(m(N_*)^2 N_{12}N_{23}N_{13}).$$

Observe the identity

$$k_3k_4 - k_1k_2 = k_3k_4 + k_{234}k_2 = k_{23}k_{24}$$

hence we can write the left-hand side as

$$(f_1 + f_2 + f_3 + f_4)k_1k_2 + (f_1 + f_2 + f_{34})k_{23}k_{24}$$

(since $f_{34} = -f_{12}$). By Lemma 9.2, the second term is $O(m(N_*)^2 N_{34}N_{23}N_{24})$ which is acceptable. Thus it will suffice to show that

$$f_1 + f_2 + f_3 + f_4 = O(m(N_*)^2 N_{12}N_{23}N_{13}/N_1^2).$$

Since $k_{12}+k_{13}+k_{23} = -2k_4$, we see that at least one of N_{12}, N_{13}, N_{23} is comparable to N_1. Without loss of generality we may take $N_{23} \sim N_1$. We now write the left-hand side as

$$f(k_1) - f(k_1 - k_{12}) - f(k_1 - k_{13}) + f(k_1 - k_{12} - k_{13})$$

and using the double mean-value theorem[28] (since $f'' = O(N_1^{-1})$ here), to conclude the argument. □

References

[1] J. Bourgain, *Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, Part II*, Geometric and Funct. Anal. **3** (1993), 209-262.

[2] J. Bourgain, *Aspects of longtime behaviour of solutions of nonlinear Hamiltonian evolution equations*, GAFA **5** (1995), 105–140.

[3] J. Bourgain, *Approximation of solutions of the cubic nonlinear Schrödinger equations by finite-dimensional equations and nonsqueezing properties*, Int. Math. Res. Notices, **1994**, no. 2, (1994), 79–90.

[4] J. Bourgain, *Periodic Korteweg de Vries equation with measures as initial data*, Selecta Math. (N.S.) **3** (1997), 115-159.

[5] J. Bourgain, *Global solutions of nonlinear Schrödinger equations*, AMS Publications, 1999.

[6] M. Christ, J. Colliander, T. Tao, *Illposedness for canonical defocussing equations below the endpoint regularity*, to appear.

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

[8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Sharp global well-posedness for periodic and non-periodic KdV and mKdV on $\mathbb{R}$ and $\mathbb{T}$*, J. Amer. Math. Soc. **16** (2003), 705–749.

[9] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Multilinear estimates for periodic KdV equations, and applications*, Journ. Funct. Analy. **211** (2004), no. 1, 173–218.

[10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *A refined global well-posedness result for Schrödinger equations with derivatives* Siam J. Math. Anal. **34** (2002), 64-86.

[11] L. Dickey, *Soliton equations and Hamiltonian systems*, World Scientific, 1991.

[12] C.S. Gardner, *Korteweg-de Vries equation and generalizations IV*, J. Math. Phys. **12** (1971), no. 8, 1548–1551.

[13] M. Gromov, *Pseudo-holomorphic curves in symplectic manifolds*, Invent. math., **82** (1985), 307–347.

[28] See e.g. Lemma 4.2 and the preceding definition in [8], or Lemma 2.3 in [10]. One could object that f'' is much larger than N_1^{-1} near the origin. However, since we are only evaluating f at points in the annulus $\{k : |k| \sim N_1\}$, we can smooth out f inside this annulus so that $f'' = O(N_1^{-1})$ throughout the interval $\{k : |k| \lesssim N_1\}$ without affecting the left-hand side.

[14] H. Hofer, E. Zehnder, *A new capacity for symplectic manifolds.*, in Analysis et cetera, Academic Press (1990), 405–428. Edited by P. Rabinowitz and E. Zehnder.

[15] H. Hofer, E. Zehnder, *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser Verlag, 1994.

[16] T. Kappeler, P. Topalov, *Global well-posedness of KdV in* $H^{-1}(\mathbb{T}, \mathbb{R})$, preprint 2003.

[17] T. Kappeler, P. Topalov, *Global fold structure of the Miura map on* L^2, IMRN 2004:39 (2004), 2039–2068.

[18] C. Kenig, G. Ponce, L. Vega, *A bilinear estimate with applications to the KdV equation*, J. Amer. Math. Soc. **9** (1996), 573–603.

[19] C. Kenig, G. Ponce, L. Vega, *On the ill-posedness of some canonical dispersive equations*, Duke Math. J. **106** (2001), no 3, 617–633..

[20] C. Kenig, G. Ponce, L.Vega, *The Cauchy problem for the Korteweg-de Vries equation in Sobolev spaces of negative indices*, Duke Math. J., **71** (1993), 1–21.

[21] S. Kuksin, *Infinite-dimensional symplectic capacities and a squeezing theorem for Hamiltonian PDE's*, CMP **167** (1995), 521–552.

[22] S. Kuksin, *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and Its Applications, **19**, Oxford Univ. Press, 2000.

[23] F. Magri, *A simple model of the integrable Hamiltonian equation*, J. Math. Phys. **19** (1978), no. 5, 1156–1162.

[24] R. Miura, *Korteweg-de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation*, J. Mathematical. Phys. **9** (1968), 1202–1204.

[25] P. Olver, *Applications of Lie groups to differential equations*, Springer, 1997.

[26] A. Sjöberg, *On the Korteweg-de Vries equation: existence and uniqueness*, J. Math. Anal. Appl. **29** (1970), 569–579.

[27] H. Takaoka, Y. Tsutsumi, *Well-posedness of the Cauchy Problem for the modified KdV equation with periodic boundary condition*, Internat. Math. Res. Notices **2004**, no. 56, 3009–3040.

[28] T. Tao, *Multilinear weighted convolution of* L^2 *functions, and applications to non-linear dispersive equations*, Amer. J. Math. **123** (2001), 839–908.

[29] T. Tao, *Global regularity of wave maps I. Small critical Sobolev norm in high dimension*, IMRN **7** (2001), 299–328.

[30] T. Tao, *Global regularity of wave maps II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), 443–544.

[31] V.E. Zakharov, L.D. Faddeev, *The Korteweg-de Vries equation is a completely integrable Hamiltonian System*, Funkz. Anal. Priloz. **5** (1971), no. 4, 18–27.

University of Toronto

University of Minnesota

M.I.T.

Kobe University and the University of Chicago

University of California, Los Angeles