LOCAL AND GLOBAL WELL-POSEDNESS FOR THE HIGHER-ORDER NLS & DNLS HIERARCHY EQUATIONS

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Joseph Adams
aus Kingston upon Thames, Großbritannien

Düsseldorf, Februar 2025

aus dem Institut für Mathematik
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Axel Grünrock

2. Prof. Dr. Jürgen Saal

Tag der mündlichen Prüfung: 26. März 2025

Abstract

The subject of this thesis is the well-posedness theory of two hierarchies of higher-order nonlinear dispersive partial differential equations (PDEs).

The first of these hierarchies is the nonlinear Schrödinger (NLS) hierarchy, which is anchored in the classical cubic NLS equation

$$i\partial_t u + \partial_x^2 u \pm |u|^2 u = 0, \qquad u(x,0) = u_0(x).$$

It is well known that, being completely integrable, it possesses an infinite number of conservation laws. We start by deriving a general representation of the structure of the higher-order Hamiltonian PDEs associated with these conservation laws.

Using tools from Fourier analysis, bi- and tri-linear refinements of Strichartz estimates are derived. These can subsequently be used to prove local well-posedness of the Cauchy problems associated with all higher-order equations in the NLS hierarchy in the non-periodic setting. As data spaces we cover the classical L^2-based Sobolev spaces $H^s(\mathbb{R})$, which, despite being a natural choice of data space, turn out not to be well suited for achieving well-posedness close to critical regularity. We therefore also consider alternative classes of data spaces: the Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$, $s \in \mathbb{R}$, $2 \geq r > 1$, and the modulation spaces $M^s_{2,p}(\mathbb{R})$, $s \in \mathbb{R}$, $2 \leq p < \infty$.

Combined with a-priori estimates derived from the complete integrability of the hierarchy equations taken from the literature we are able to extend our L^2-based local solutions to global in time solutions.

Concluding this first part of the thesis we also prove that, within the framework of tools we are using (deriving well-posedness with fixed-point methods), we have obtained optimal results up to the endpoint, the critical regularity, in our respective classes of data spaces. Furthermore we show that fixed-point methods are not applicable in the periodic setting.

In the second part of the thesis we consider a different hierarchy of PDEs based on the infinite number of conservation laws of the derivative nonlinear Schrödinger (dNLS) equation

$$i\partial_t u + \partial_x^2 u - i\partial_x(|u|^2 u) = 0, \qquad u(x,0) = u_0(x),$$

another completely integrable model. We again derive a workable representation of the higher-order PDEs in this dNLS hierarchy, and even go one step further by determining, for a subset of nonlinear terms appearing in the higher-order equations, the coefficients appearing in front of them. This is necessary to establish the effectiveness of a gauge-transformation that will rid the equation of what we call 'bad' cubic terms in the nonlinearity. A similar detailed analysis of the coefficients of a hierarchy of integrable PDEs has not yet been undertaken.

The local well-posedness theory for the dNLS hierarchy equations follows similar arguments as for the NLS hierarchy equations, reusing the smoothing estimates from earlier. However, an additional difficulty arises due to an extra derivative in nonlinear terms of the equations. The smoothing estimates alone are not sufficient to derive well-posedness. We resort to additional smoothing resulting from the structure of the nonlinear terms, expressed through their resonance relation.

For the extension of the local L^2-based well-posedness results to global in time ones, no equivalent to the NLS hierarchy a-priori estimates is available. We therefore have to resort to proving that it is possible to derive sufficient a-priori estimates from the conservation laws associated with the dNLS hierarchy equations.

Finally we again establish optimality of our results within the framework we are using and rule out the possibility of developing a similar theory for the periodic setting (using fixed-point methods).

Contents

1. Introduction

In recent decades the study of Cauchy problems for nonlinear dispersive partial differential equations (PDEs) has become a popular field of research within analysis. Combining methods from harmonic analysis, PDE theory, more recently probability theory and the inverse scattering transform. The goal of this thesis is to contribute to this field, studying the well-posedness theory of two hierarchies of higher-order nonlinear dispersive PDEs.

A dispersive PDE is an evolution equation that is characterised by (its linear part) enabling plane wave solutions of different frequencies to propagate at different speeds. This has the effect, that if one starts with highly concentrated initial data, over time, the dispersion will cause the solution to spread out. Unfortunately this dispersive effect is not enough to have a strong regularising effect on solutions unlike, say, parabolic PDEs, making studying them difficult.

As often dispersive PDEs are motivated by physical applications, it is natural to consider a Hilbert space setting for their initial data. The most common choice of data space is the scale of classical Sobolev spaces $H^s(X)$, $s \in \mathbb{R}$, $X \in \{\mathbb{R}^n, \mathbb{T}^n\}$.[1]

Overall goal, or measure of improvement in result, in the study of well-posedness of dispersive PDEs is to require as little regularity of the initial data as possible, while still deriving well-posedness results. This is motivated by the fact that natural a-priori estimates, which often also have physical interpretation, are usually available at low regularities. Mass (L^2 norm), energy (H^1 norm) or momentum ($H^{\frac{1}{2}}$ norm) conservation laws are typical examples of such. In addition, physically motivated dispersive PDEs are often Hamiltonian equations, where it is natural to consider the equation posed with initial data in its corresponding energy space, as in other PDE disciplines.

Originating in a physical context and having a plethora of conservation laws, makes it unsurprising that such equations are also often invariant with respect to symmetry transformations. Of particular importance for us is invariance under a scaling symmetry. These give further motivation to consider the low-regularity well-posedness question of these equations, because such a scaling symmetry allows one to define a notion of critical regularity. This is the regularity $s \in \mathbb{R}$ at which the data space H^s (more precisely its homogeneous variant) is also invariant under the same scaling symmetry as the equation. The critical regularity is interpreted as the point where one expects to lose control over the lifespan of the solution with respect to the size of the initial data. In turn, this is the point where one expects a dispersive PDE to be ill-posed in any lower regularity space. Regularities above the scaling critical are deemed subcritical and below are said to be supercritical.

Typical examples of dispersive PDEs are the Korteweg-de Vries (KdV) equation, its modified variant (mKdV), higher-dimensional generalisations of these (e.g. Kadomtsev-Petviashvili and Novikov-Veselov equations), the cubic nonlinear Schrödinger (NLS) equation and the derivative nonlinear Schrödinger (dNLS) equation. The latter two of which we will be building on in this thesis.

1.1. **Deriving well-posedness.** Initial studies of the well-posedness question for nonlinear dispersive PDEs, see [10, 11], relied on what has become known as the energy method. For this technique ε-parabolic regularisation is used together with energy estimates that, uniformly in $\varepsilon > 0$, give control over the H^s norm of a solution of the regularised equation. Constructing solutions u_ε for such a regularised equation is simple, since strong smoothing effects derived from presence of the parabolic term are now available. Last step is to send $\varepsilon \to 0$ and proving, using the

[1]We will abbreviate classical Sobolev spaces as H^s, $s \in \mathbb{R}$, for the remainder of this introduction to signify independence of the underlying geometry.

energy estimates, that the solution sequence converges and the limiting solution $u := \lim_{\varepsilon \to 0} u_\varepsilon$ also satisfies the original equation.

An advantage of this technique is that it works independently of the underlying geometry of the equation, quickly deriving local well-posedness in both periodic and non-periodic settings at sufficiently high initial regularity of the data. Also one usually obtains unconditional[2] well-posedness results.

The sufficiently high regularity is dictated though by the Sobolev embedding theorem and the number of derivatives appearing in the nonlinearity of the equation. This means that the lower limit for local well-posedness using the energy method is usually comparatively high (especially for higher dimensional equations). Taking the KdV equation as an example, well-posedness using the energy method can be derived for $s > \frac{3}{2}$, far from its critical regularity at $s = -\frac{3}{2}$. For other equations, like the cubic NLS equation, the gap is smaller, with the lower limit for local well-posedness using the energy method being $s > \frac{1}{2}$ and $s = -\frac{1}{2}$ its critical regularity.

With the energy method providing unsatisfactory results, with regard to the gap between well-posedness and critical regularity or available a-priori estimates, one turns to employing the dispersive effects of the PDE. Such dispersive effects can be formalised through the use of space-time norms as in Strichartz estimates derived with, for example, Fourier analysis or stationary phase methods.

Strichartz estimates, which capture a time-averaged version of dispersion provided by the equation using $L_t^p L_x^q$-norms, can then be used in a fixed-point argument to derive well-posedness for the equation. This works for semi-linear dispersive equations, where the nonlinearity can still be thought of as a perturbation of the linear equation. If the nonlinear effects become too strong though (i.e. the equation is of quasi-linear character) this approximation breaks down.

Such smoothing estimates may be further refined to bi- or in general multilinear estimates, that capture the interaction of multiple solutions. These multilinear refinements of Strichartz estimates allow one to capture smoothing effects that may not be present in a linear evolution, but are present, say, in the interaction of two (frequency separated) solutions of the equation.

In some cases, prominently the KdV and dNLS equations, (multilinear) smoothing effects expressed in mixed L^p-norms alone do not suffice in order to derive well-posedness. This warrants the introduction of Bourgain (also called Fourier-restriction) spaces $X_{s,b}$, $s, b \in \mathbb{R}$, which are also used in the context of a fixed-point argument.

These spaces are defined in a way adapted to the linear part of the equation. Besides use of smoothing effects already known for the equation, they allow one to peek into the structure of a nonlinear term and use additional smoothing by way of the so called resonance relation. Use of the resonance relation allows one to distinguish between, for example, polynomial nonlinear terms of the same order, that only differ in their distribution of complex conjugates. They are also helpful in situations, notably periodic problems, where there are no smoothing effects because of the compactness of the underlying domain. For our investigation of the dNLS hierarchy equations, despite the availability of Strichartz type smoothing effects, we will need to utilise this additional advantage of Bourgain spaces in order to deal with the many derivatives present in their nonlinear terms.

In comparison with the energy method use of fixed-point methods has an additional advantage beyond a lower regularity threshold for the initial data: the estimates proven for use in the application of a fixed-point theorem are also strong

[2]Unconditional well-posedness refers to the solution of an evolution equation being unique in the space of continuous functions with values in the data space, rather than in some continuous subspace of this.

enough to establish smoothness of the flow of the PDE. With the energy method the flow's mere continuity is derived. However unconditional well-posedness now has to be derived with an additional argument and is no longer automatic.

These different levels of regularity of the flow became part of the overall notion of well-posedness that was to be proven for a given dispersive PDE. In turn ill-posedness could be established by analysing the minimal regularity of the initial data required that well-posedness of a PDE necessitates. Ruling out the applicability of fixed-point methods as described above, such ill-posedness results have become commonplace for underlining the optimality of well-posedness results. We also resort to this approach to show that we have achieved best possible well-posedness results in our chosen data spaces in the non-periodic setting and completely rule out the applicability of fixed-point methods in the periodic setting.

Two approaches for ruling out smoothness of the flow are common. Formally differentiating the flow and proving the necessary estimates that would establish its k-times continuous differentiability are, in general, false, was pioneered in [14]. Another is using soliton solutions[3] that are often available for dispersive PDEs. Here one aims to construct a situation where two pieces of initial data only differing in phase, but not in support, have respective solutions that propagate at different speeds so that their supports separate fast enough. Hence initially the difference between solutions is small, but quickly enlarges as the solutions separate. First demonstrated in [60], this can be used to rule out uniform continuity of the flow, a stronger result than C^k ill-posedness. We employ both techniques in this thesis.

1.2. **Alternative function spaces.** Though the classical Sobolev spaces are a natural choice for the initial data of dispersive PDEs (or physically motivated models in general), they are not however always well-suited for achieving well-posedness close to or in the critical space. Often there will be a gap in regularity between the best possible well-posedness result on the H^s scale (indicated by complementary ill-posedness results as described at the end of the preceding subsection) and the Sobolev space of critical regularity. Moreso, for some equations, prominently the Benjamin-Ono equation, no well-posedness result in Sobolev spaces can be achieved at all using fixed-point methods, as the equation is ill-posed in any Sobolev space [66, 73] in the sense that the flow cannot be C^2 or even uniformly continuous.

This necessitates the consideration of alternative scales of function spaces, that at least incorporate the classical Sobolev spaces within them, but also allow for improved well-posedness results. Examples of such spaces that have been successfully employed in the dispersive PDE community are weighted Sobolev spaces, Besov spaces, Fourier-Lebesgue spaces and modulation spaces. Except for the weighted spaces, which are an unfortunate choice, because they are ususally not compatible with off-the-shelf smoothing/Strichartz estimates and are most often employed when using the inverse scattering transform, these examples are (in general) not Hilbert spaces. This is a tradeoff that is made in order to achieve well-posedness closer to a space that is comparable with the Sobolev space of critical regularity of the equation. More precisely, some of these scales of function spaces (or their homogeneous variants) will also be compatible with transformations of scale, which enables one to pinpoint a space of critical regularity for these scales of spaces too.

Another disadvantage of moving to these alternative spaces is that a-priori estimates derived from conservation laws, which are often available for dispersive PDEs, usually do not transfer to these spaces easily. Leading to a much less well understood global well-posedness theory in these alternative spaces.

[3]These are exact solutions of the nonlinear PDE that are highly concentrated in support around a point, that, through the flow of the PDE, is pushed through space.

Pioneering this approach in the context of low-regularity well-posedness theory of dispersive PDEs were [17, 93] where nonlinear Schrödinger equations were considered.

Sometimes though transitioning to these alternative spaces can lead to improved smoothing estimates, as was already observed in [26], where an analogue of Strichartz estimates for what amounts to Fourier-Lebesgue spaces were derived. Only much later were these then actually employed in the well-posedness theory of the NLS equation [93].

Since this phenomenon of a large gap between the optimal well-posedness on the H^s scale and the critical regularity is also prevalent in the NLS and dNLS hierarchies, we will also consider well-posedness in Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$, $s \in \mathbb{R}$, $2 \geq r > 1$, and modulation spaces $M^s_{2,p}(\mathbb{R})$, $s \in \mathbb{R}$, $2 \leq p < \infty$. These have been known to lead to much improved results for closely related equations [18, 19, 38–40, 63, 77, 80, 93].

1.3. **Completely integrable models.** The examples of dispersive PDEs mentioned above are not just of interest because they serve as models in physics, but because they also possess rich algebraic structure. Being a completely integrable model (a term of art that has no exact definition) usually entails the equation having a Lax pair, infinitely many conservation laws and soliton solutions. This is true of all examples mentioned thus far.

It was discovered in the 1960s, while trying to solve the KdV equation using the inverse scattering transform, that the KdV equation possesses an infinite number of conservation laws [1, 25, 70]. KdV being a Hamiltonian equation, it is only natural to also consider the higher-order PDEs resulting from these (higher-order) conserved quantities. This was later also investigated for the cubic nonlinear Schrödinger equation, where a general method of associating a hierarchy of Hamiltonian PDEs to a completely integrable model was developed. These may be interpreted as higher-order corrections to the original equation. It is these associated equations that we mean by higher-order hierarchy equations.

Determined goal of this thesis is to establish local well-posedness, and in certain cases to extend this to global well-posedness, for the higher-order equations in the hierarchies associated with the cubic NLS and dNLS equations in spaces that are as close as possible to what constitutes critical regularity in the non-periodic setting[4]. Since we are able to derive that the higher-order hierarchy equations are ill-posed in classical Sobolev-spaces of regularities far exceeding their critical regularity, we also turn to employ both Fourier-Lebesgue spaces and modulation spaces in order to close this gap. Using these spaces we are able to prove well-posedness of the higher-order hierarchy equations in the entire subcritical range of parameters of these scales of spaces and complement these results with ill-posedness results (both in the sense of C^k and uniform continuity, as discussed above) showing we have achieved best possible bounds on the lowest regularity necessary for local well-posedness.

1.4. **The NLS hierarchy.** Starting point of our investigation in this thesis will be the cubic NLS equation

$$i\partial_t u + \partial_x^2 u \pm |u|^2 u = 0, \qquad u(x,0) = u_0(x),$$

which is a model for various wave propagation phenomena in nonlinear optics and plasmas, see [85] for an introduction to the topic. Its well-posedness theory is already well understood, having been studied over more than the past 30 years.

[4]As it turns out, the higher-order equations are ill-posed in the periodic setting in classical Sobolev-spaces, making an approach with fixed-point methods unfeasible.

We give [37, 76] and the references therein as an introduction to the state of the art results regarding NLS well-posedness.

As implied above, the cubic NLS equation is a completely integrable model. It has ample well-known conservation laws, including the L^2-norm of a solution u,

$$\text{momentum:} \quad \int_{\mathbb{R}} u\partial_x\overline{u}\mathrm{d}x \quad \text{and energy:} \quad \int_{\mathbb{R}} |\partial_x u|^2 \pm \frac{1}{2}|u|^4\mathrm{d}x.$$

The cubic NLS equation is the Hamiltonian PDE associated with its energy conservation law given above. Using the techniques developed in [1, 5, 25, 82] we extend this to an infinite family of conserved quantities, which in turn, by associating a corresponding Hamiltonian PDE leads to what we refer to in the title of this thesis as the NLS hierarchy.

We begin our well-posedness analysis by deriving new multilinear refinements of smoothing estimates for the higher-order hierarchy equations. These, together with well-known estimates from the relevant literature[5], allow us to prove local well-posedness with a smooth flow for all higher-order hierarchy equations. Looking, at first, only towards classical Sobolev spaces, we establish that the jth NLS hierarchy equation (with $j = 1$ corresponding to the cubic NLS equation itself) is locally well-posed in $H^s(\mathbb{R})$ for $s \geq \frac{j-1}{2}$, $j \geq 2$. On this scale of function spaces this result is optimal[6] in the sense that for any lower regularity of initial data the flow cannot be uniformly continuous any more. In order to prove this we construct explicit families of soliton solutions for the higher-order equations.

Using a-priori estimates from the literature, that are derived from the equations' complete integrability, we are able to extend these solutions (for the full range $s \geq \frac{j-1}{2}$) globally in time.

Considering the scaling critical regularity for all hierarchy equations is $s = -\frac{1}{2}$ this is a considerable gap between our well-posedness result and the space of critical regularity. Thus we move on to Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$. The critical space on this alternative scale of function spaces is $\hat{H}^0_1(\mathbb{R})$. Our local well-posedness results in this case covers the entire subcritical range of parameter $1 < r \leq 2$ and $s \geq \frac{j-1}{r'}$, again with a smooth flow. Complementing our result in the same way as above with failure of uniform continuity of the flow (for $1 < r \leq 2$ and $-\frac{1}{r'} < s < \frac{j-1}{r'}$) we show that our result is optimal up to the endpoint.

Finally we also look at the higher-order hierarchy equations posed with initial data in modulation spaces $M^s_{2,p}(\mathbb{R})$. Here there is a less well-defined notion of criticality, as there is no homogenous variant of modulation spaces, but comparisons (via embedding theorems) with Fourier-Lebesgue spaces suggest, that these spaces become critical for $p \to \infty$, s staying fixed.

Our analysis results in local well-posedness with a smooth flow for the higher-order hierarchy equations with data in $M^s_{2,p}(\mathbb{R})$, so long as $s = \frac{j-1}{2}$ and $2 \leq p < \infty$. Again covering what can be considered the entire subcritical range, up to the endpoint. We complement this positive result with a theorem proving failure of uniform continuity of the flow for any lower initial regularity, that is $0 \leq s < \frac{j-1}{2}$ and $2 \leq p \leq \infty$. Again, this uses the previously constructed families of soliton solutions for the higher-order equations.

Additionally, our arguments are general enough for allowing us to establish well-posedness for a great many variants on the actual hierarchy equations. For example,

[5]Here we are referring to Kato smoothing and maximal function estimates.

[6]We prove this ill-posedness result not for the actual jth hierarchy equation but for *an* equation that has the same structure (though possibly different coefficients) as the hierarchy equations. This is harmless with regard to establishing the optimality of our result as we do not take the structure/choice of coefficients into account when proving our well-posedness result. Information on the coefficients of the hierarchy equations is simply not available.

we can allow for much lower initial regularity of the data (in either the Fourier-Lebesgue or modulation space setting) while still retaining the same well-posedness result, if one eliminates cubic nonlinear terms from the equations. We leave the details to later sections.

1.5. **The dNLS hierarchy.** Having dealt with the NLS hierarchy equations, in the second part of this thesis we move on to another completely integrable hierarchy, which is associated with the dNLS equation

$$i\partial_t u + \partial_x^2 u - i\partial_x(|u|^2 u) = 0, \qquad u(x,0) = u_0(x).$$

This equation is a model for the propagation of Alfvén waves in magnetized plasma and of ultra-short pulses in optical fibers. It has (among others) an associated conservation law that can be interpreted as the energy of a solution

$$\int_{\mathbb{R}} |\partial_x u|^2 + \frac{3}{2}\mathrm{Im}(|u|^2 u \partial_x \overline{u}) + \frac{1}{2}|u|^6 \mathrm{d}x.$$

The reader may consult [4, 6, 72, 91] as an introduction to the origin and derivation of the equation.

Between the cubic NLS equation and the dNLS equation there is already a considerable additional difficulty in their respective analysis regarding local well-posedness. Specifically we are referring to the fact that the dNLS equation itself is not amenable to a direct treatment with (even multilinearly refined) smoothing estimates[7]. One first has to use a gauge-transformation[8]

$$\mathcal{G}_\pm : u(x,t) \mapsto v(x,t) \coloneqq \exp\left(\pm i \int_{-\infty}^{x} |u(y,t)|^2 \mathrm{d}y\right) u(x,t)$$

in order to remove a 'bad' cubic nonlinear term from the equation. The dNLS equation in terms of an unknown function u transforms to its gauged version

$$i\partial_t v + \partial_x^2 v + iv^2 \partial_x \overline{v} + \frac{1}{2}|v|^4 v = 0, \qquad v(x,0) = v_0(x),$$

for the unknown function v. Here the previously present $|u|^2\partial_x u$ is now missing. One now instead has to deal with the new quintic term $|v|^4 v$, though this is an innocuous term, because it is of higher order than cubic. In applying the gauge-transformation one also loses some (knowledge of) regularity of the flow for the Cauchy problem of the dNLS equation. As it is only well-known that the gauge-transformation is Lipschitz continuous between relevant function spaces, a well-posedness result for the gauge-transformed equation, even with smooth flow, only translates to well-posedness with Lipschitz continuous flow for the dNLS equation itself.

Specifically, one only knows of the Lipschitz continuity of the gauge-transformation between Sobolev spaces, or more generally Fourier-Lebesgue spaces. We extend the literature by providing the first proof of the Lipschitz continuity of the gauge-transformation between certain modulation spaces. This is a fact that has seen use in well-posedness results in modulation spaces for the dNLS equation in the literature, but so far no proof had been provided.

This downgrade in regularity of the flow is usually only a minor problem since supplementary ill-posedness results concern the failure of uniform continuity of the flow, a strictly weaker property than its Lipschitz continuity.

---

[7]The smoothing estimates between the NLS (hierarchy) equation and the dNLS (hierarchy) equation do not differ, since they share their linear part.

[8]Our well-posedness results all concern the non-periodic case. Therefore we focus on this also when discussing the gauge-transformation. There is however also a more complicated gauge-transformation available for the periodic case, that essentially achieves the same, see [47].

A similar problem arises when developing a local well-posedness theory for the higher-order dNLS hierarchy equations. These also contain (higher-order versions) of such 'bad' cubic terms[9]. Unfortunately it is far from obvious if this/any gauge-transformation is capable of similarly amending the higher-order hierarchy equations in order to be able to develop a satisfactory well-posedness theory.

After again using techniques from the inverse scattering literature, we derive a general form of the hierarchy equations associated with the conservation laws of the dNLS equation. This is in line with our work regarding the NLS hierarchy equations. Though this time we have to go a step further and also deduce specific knowledge of the coefficients appearing in front of 'bad' cubic nonlinear terms in the higher-order dNLS hierarchy equations. Such fine-grained insight into the structure of a hierarchy of completely integrable equations is yet to be found in the literature.

Knowledge of these coefficients then allows us to prove that the classic gauge-transformation, as it does for the dNLS equation, also eliminates all 'bad' cubic terms in the higher-order hierarchy equations, replacing them with innocuous higher-order terms that can be dealt with more easily.

This preparation of the dNLS hierarchy equations takes up a considerable chunk of the second part of this thesis.

Having set the stage we are able to tackle the local well-posedness problem for the gauge-transformed higher-order dNLS hierarchy equations. Though still, mere use of the refined smoothing estimates we derived in the context of the NLS hierarchy equations does not suffice in order to close a fixed-point argument. We resort to utilising the additional smoothing present in the nonlinear equations, derived from the specific structure of the nonlinear terms by way of their resonance relation. These additional arguments beyond what was used for the NLS hierarchy equations then allow us to establish local well-posedness with a Lipschitz continuous flow for all higher-order dNLS hierarchy equations.

On the scale of Sobolev spaces there is a considerable gap again between our local well-posedness result and the scaling critical regularity $s = 0$. Specifically, we establish local well-posedness with initial data in $H^s(\mathbb{R})$ for $s \geq \frac{j}{2}$. A failure of thrice continuous differentiability of the flow for regularities of initial data $s < \frac{j}{2}$ suggests this result to be optimal, if one relies on fixed-point methods.

To close this gap we again move on to Fourier-Lebesgue spaces where we are able to cover the entire subcritical range[10] $s > \frac{1}{2} + \frac{j-1}{r'}$ and $2 \geq r > 1$ with our local well-posedness theory. This is again complemented by a similar ill-posedness result as for the scale of Sobolev spaces at any lower regularity, showing our result is optimal up to the endpoint.

For modulation spaces we are able to establish the local well-posedness with initial data in $M^s_{2,p}(\mathbb{R})$ for $s \geq \frac{j}{2}$ and $2 \leq p < \infty$ with ill-posedness in the sense that the flow can no longer be thrice continuously differentiable at any lower regularity $s < \frac{j}{2}$ and $2 \leq p \leq \infty$.

For the dNLS hierarchy equations no general a-priori estimates at an arbitrary regularity level are available in the literature. That is, at least none that would help us extend our local solutions to global in time solutions. There have been recent efforts [65] where a-priori estimates for the dNLS equation were derived for $0 < s < \frac{1}{2}$. It is reasonable to expect these to also transfer to the higher-order hierarchy equations.

---

[9]The exact definition of 'bad' cubic term is discussed later, suffice to say they are the ones where none of the derivatives appearing in the term fall on $\overline{u}$.

[10]For the dNLS hierarchy equations set in Fourier-Lebesgue spaces the critical space is $\hat{H}_1^{\frac{1}{2}}(\mathbb{R})$. Closest approximation (again via embedding theorems) to a critical space in the scale of modulation spaces is $M^{\frac{j}{2}}_{2,\infty}(\mathbb{R})$.

We turn instead to the L^2-based conservation laws associated with the hierarchy equations. Using the Gagliardo–Nirenberg inequality we construct a-priori estimates at positive integer regularity levels $k \in \mathbb{N}$, enabling us to extend our local solutions in Sobolev spaces $H^k(\mathbb{R})$ to global in time ones.

1.6. **Open questions and recent developments.** To conclude the introduction to this thesis we would like to discuss some recent developments with regard to the well-posedness theory of dispersive hierarchy equations (not just in relation to NLS and dNLS type equations) and also highlight some open questions that we haven't been able to answer entirely within the scope of this thesis.

To begin, we would like to highlight a particular interest of the author: so far only a subset of the coefficients of the higher-order dNLS hierarchy equations have been uncovered. It would certainly be an interesting question to derive the complete coefficient structure of the dNLS hierarchy. A natural extension of this is, if for integrable hierarchies of PDEs there is a general method of deriving the complete coefficient structure.

This would open up the possibility of forming well-posedness results that rely on a more detailed analysis of the interplay between different nonlinear terms of the same order, particularly cubics, which turn out to be the most difficult to deal with. It would be conceivable that, given the right interaction between these terms, one could derive well-posedness results that improve upon the ones presented in this thesis. On the other hand it might allow improved ill-posedness results. We point out that our ill-posedness results regarding the uniform continuity of the flow (for the NLS hierarchy) concern equations that only share the same structure (which nonlinear terms appear in the equation) as the NLS hierarchy equations. This does not weaken our result, since we do not rely on interactions between nonlinear terms to prove well-posedness, but finding families of soliton solutions for the actual hierarchy equations would be an interesting improvement.

Another possibility in extending our results would be to move to a scale of function spaces that combines the advantages of Fourier-Lebesgue spaces and modulation spaces, as was done in [20]. Here the authors improve upon the state of the art well-posedness theory for the complex-valued mKdV equation in the non-periodic case by using a scale of function spaces that interpolates between the two mentioned scales of spaces. Although one has to note that the improvement is on the order of less than an ε in regularity of the initial data.

Moving on to the question of global well-posedness, we have only been able to demonstrate global well-posedness for our hierarchy equations with initial data in L^2-based Sobolev spaces. As was shown in [37] for the dNLS equation, a variant on Bourgain's splitting argument (also known as the Fourier truncation method) can be used to derive global well-posedness for the dNLS equation in Fourier-Lebesgue spaces, when $r < 2$. After initial investigations by the author, this method also seems applicable to the (d)NLS hierarchy equations. Because of the many more nonlinear terms present in the hierarchy though, there is a considerable amount of work to be done in order to derive the necessary estimates for this method to work. Initial calculations suggest that it is reasonable to expect global well-posedness for the fourth-order NLS hierarchy equation in $\hat{H}^1_r(\mathbb{R})$ for $r > \frac{4}{3}$, which would not yet be an improvement over the global result in $H^{\frac{1}{2}}(\mathbb{R})$ if one compares using an embedding theorem.

Generally, interest in the well-posedness of hierarchies of dispersive PDEs is a current topic, with the recent results of [64]. There the authors show well-posedness of the KdV hierarchy equations with a continuous flow in $H^{-1}(\mathbb{R})$ strongly relying on the complete integrability of the KdV equation[11]. A similar program for the (d)NLS hierarchy equations is reasonable, though such results would not immediately be an improvement over the ones presented in this thesis, because of the much lower regularity of the flow (when relying on the complete integrability).

Regarding recent developments in the well-posedness theory for (among others) the fourth-order NLS hierarchy equation in the periodic case we would like to mention [51]. There the author uses a normal form reduction and a cancellation property between nonlinear terms to derive well-posedness in $H^1(\mathbb{T})$. This not only shows that, using alternatives to fixed-point methods, the higher-order hierarchy equations are also amenable to well-posedness results in the periodic case, but also that specific knowledge of the structure of the nonlinear terms can be used to derive such results. This complements points of interest we have raised at the beginning of this section.

Besides the use of a normal form reduction it is also reasonable to expect a refinement of the energy method to be able to derive well-posedness for the higher-order hierarchy equations in the periodic case. This approach was introduced in [66], where Strichartz/smoothing estimates were used in order to lower the regularity requirement for initial data in an application of the energy method. Unfortunately this approach would also suffer from proving only mere continuity of the flow, if a well-posedness result is derived.

[11] In the same paper the authors also derive well-posedness in the periodic case for nth hierarchy equation in $H^{n-2}(\mathbb{T})$.

Part 1. Well-posedness of the NLS hierarchy

The following part of this thesis is an independent paper written by the author. It has since been published in a peer-reviewed journal, see [2]. We reproduce it here as it appears in the published version, with the difference that its bibliography is included in the overall bibliography of this thesis. Its abstract reads:

We prove well-posedness for higher-order equations in the so-called NLS hierarchy (also known as part of the AKNS hierarchy) in almost critical Fourier-Lebesgue spaces and in modulation spaces. We show the jth equation in the hierarchy is locally well-posed for initial data in $\hat{H}^s_r(\mathbb{R})$ for $s \geq \frac{j-1}{r'}$ and $1 < r \leq 2$ and also in $M^s_{2,p}(\mathbb{R})$ for $s = \frac{j-1}{2}$ and $2 \leq p < \infty$. Supplementing our results with corresponding ill-posedness results in Fourier-Lebesgue and modulation spaces shows optimality. Using the conserved quantities derived in [69] we argue that the hierarchy equations are globally well-posed for data in $H^s(\mathbb{R})$ for $s \geq \frac{j-1}{2}$.

Our arguments are based on the Fourier restriction norm method in Bourgain spaces adapted to our data spaces and bi- & trilinear refinements of Strichartz estimates.

2. Introduction

The cubic nonlinear Schrödinger (NLS) equation

$$\begin{cases} i\partial_t u + \partial_x^2 u = \pm 2|u|^2 u \\ u(t=0) = u_0 \end{cases} \tag{2.1}$$

with initial data u_0, has over the past 30 years become one of the canonical objects of study in the well-posedness theory of dispersive PDEs. We direct the interested reader to [12, 16, 88] and the references therein for an overview of developments in its study.

Contemporary research is more and more leaning into the fact that the NLS equation possesses a rich internal structure that may be exploited in order to prove new well-posedness results or a-priori bounds on solutions. We are, of course, referring to the fact that the NLS equation is considered to be a completely integrable system [1, 5, 25, 69, 79] – an exact definition of which though escapes the literature. Usually one considers the fact that there exists an infinite sequence of non-trivial conserved quantities one of the markers of complete integrability. A fact that is also true for the NLS equation. The first few of these conserved quantities are

$$\begin{aligned} &\text{Mass:} && H_0 = \int |u|^2 \,\mathrm{d}x \\ &\text{Momentum:} && H_1 = -i \int u\partial_x \overline{u} \,\mathrm{d}x \\ &\text{Energy:} && H_2 = \int |\partial_x u|^2 \pm |u|^4 \,\mathrm{d}x \end{aligned}$$

More precisely, the NLS equation is a Hamiltonian equation that is induced by its energy H_2. (Induced in what way we will make more precise in Section 3.) This begs the question: do the higher-order conserved quantities $H_3, H_4, \ldots$ also induce any *interesting* dispersive PDE[12]?

Yes, in fact the fourth conserved quantity

$$H_3 = i \int \partial_x u \partial_x^2 \overline{u} + 3|u|^2 u \partial_x \overline{u} \,\mathrm{d}\lambda$$

[12] The mass H_0 and momentum H_1 also induce PDE, namely of phase shifts and of translations. Though these are not dispersive and thus are of no interest to us.

induces the also well-known modified Korteweg-de-Vries (mKdV) equation[13]

$$\begin{cases} \partial_t u + \partial_x^3 u = \pm 2\partial_x(|u|^2 u) \\ u(t=0) = u_0 \end{cases} \tag{2.2}$$

see [23, 55, 69] for an overview.

The next higher-order equation is not quite as well known as the NLS and/or mKdV equations, though it has also appeared independently in the literature [28, 29].

To the author's best knowledge there is no complete description of all conserved quantities of NLS available. It is though a simple, but tedious, task to calculate them. See Appendix A, where we list more of the conserved quantities and their associated equations.

We want to mention at this point, that the pattern of even-numbered conserved quantities H_{2k}, $k \in \mathbb{N}$ inducing NLS-like equations and odd-numbered ones H_{2k+1}, $k \in \mathbb{N}$ inducing mKdV-like equations continues [38, 69]. This sequence of NLS-like equations is what is referred to in the title of this paper as the NLS hierarchy[14]. We will give a more precise definition of the NLS hierarchy in Section 3.

Aim of this paper is to deal with questions of low-regularity well-posedness for the NLS hierarchy equations in classical Sobolev spaces $H^s(\mathbb{R})$, Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$ (sometimes written as $\mathcal{F}L^{s,r'}(\mathbb{R})$, where the integrability exponent is conjugated) defined by the norm

$$\|u\|_{\hat{H}^s_r} = \|u\|_{\mathcal{F}L^{s,r'}} = \|\langle\xi\rangle^s \hat{u}\|_{L^{r'}}$$

and modulation spaces $M^s_{2,p}(\mathbb{R})$ defined by the norm

$$\|u\|_{M^s_{2,p}} = \|\|\square_n u\|_{H^s}\|_{\ell^p_n(\mathbb{Z})}$$

with a family of isometric decomposition operators $(\square_n)_{n\in\mathbb{Z}}$. Precise definitions of the function spaces and an overview of associated embeddings are given in Section 2.2.

While we embrace the rich integrability structure of these equations for their derivation and conservation laws, we will not be making use of their integrability to argue our local well-posedness results. This has the advantage that our arguments work for a rather large class of equations, but the disadvantage that we also cannot utilise any special structure that may be present in the NLS hierarchy equations, that could aid the well-posedness.

Moreover our arguments will be based on the contraction mapping principle in versions of Bourgain spaces $X_{s,b}$ adapted to our data spaces, in combination with bi- and trilinear refinements of Strichartz estimates.

Acknowledgements. This work is part of the author's PhD thesis. He would like to greatly thank his advisor, Axel Grünrock, for suggesting this line of problems and his continued and ongoing support.

The author would also like to thank the second anonymous reviewer for his suggested corrections and additions. In particular asking a question about the influence of the distribution of complex conjugates in quintic and higher-order terms in the Fourier-Lebesgue space setting that lead to Remarks 4.5 and 6.4.

---

[13] There is a caveat to this that is discussed in Appendix A. In short, when looking at the complex mKdV equations a slightly different nonlinearity is produced when one follows the construction of the NLS hierarchy in [5], as we do. When looking at the real valued mKdV equation there is no discrepancy.

[14] The NLS hierarchy is a part of what is often called the AKNS hierarchy, after the names of the authors that played a considerable role in developing the inverse scattering transform, see [1].

2.1. **Organisation of the paper.** In the next and final subsection of this introduction we will establish the general notation and function spaces that we will be using throughout the rest of this paper. The acquainted reader may skip immediately to Section 3.

Following that we will define exactly what we mean by NLS hierarchy (and its generalisations) in Section 3.

We give an overview of prior work related to the well-posedness study of hierarchies of PDEs, a statement of our main results and a discussion of these in Section 4.

In Section 5 we collect general smoothing estimates based on the dispersion present in the equations we are dealing with. This includes linear estimates we will be citing from the literature, some new bilinear estimates adapted to the case of higher-order Schrödinger equations, so-called Fefferman-Stein estimates (which generalize Strichartz estimates to the Fourier-Lebesgue spaces we will be using), and trilinear estimates. The new bi- and trilinear refinements as well as the Fefferman-Stein estimates are based on [38].

Then in Section 6 we will follow up with the nonlinear estimates needed to prove Theorems 4.1 and 4.3. First we deal with estimates regarding well-posedness in Fourier-Lebesgue spaces, following up with the same for modulation spaces.

Finally in Section 7 we will deal with the question of ill-posedness. In this section we will see that, on the line, our methods lead to optimal results in the framework that we use. Also we will deal with the fact that a fixed-point theorem based approach cannot work in the same generality on the torus, as it does on the line.

2.2. **Notation and function spaces.** We use the notation $A \lesssim B$ to mean $A \leq CB$ for a constant $C > 0$ independent of A and B, and $A \sim B$ denotes $A \lesssim B$ and $A \gtrsim B$, while $A \ll B$ means $A \leq \varepsilon B$ for a *small* constant $\varepsilon > 0$. For a given real number $a \in \mathbb{R}$ we will denote by $a+$ and $a-$ the numbers $a + \varepsilon$ and $a - \varepsilon$ for an arbitrarily small $\varepsilon > 0$, respectively. The so called Japanese brackets denote the quantity $\langle x \rangle = (1 + x^2)^{\frac{1}{2}}$.

We use the following conventions regarding the Fourier-transform: the Fourier-transform of a function $u : \mathbb{R}_x \times \mathbb{R}_t \to \mathbb{C}$ with respect to the space-variable x is given by

$$\mathcal{F}_x u(\xi, t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u(x, t) e^{ix\xi} \, \mathrm{d}x.$$

The Fourier-transform with respect to the time-variable is defined analogously, though the Fourier-variable corresponding to t shall be called τ. We will also use the notation $\hat{u}$ to denote the Fourier-transform with respect to either one or both of those variables, but it will be clear from context which of those cases we are referring to, specifically from the use of spatial- and time-Fourier variables rather than their physical-space counterparts.

For two functions f and g we use the notation

$$\int_* f(\xi_1) g(\xi_2) \, \mathrm{d}\xi_1 = \int_{\mathbb{R}} f(\xi_1) g(\xi - \xi_1) \, \mathrm{d}\xi_1$$

to represent the integral under the convolution constraint $\xi = \xi_1 + \xi_2$. This generalises naturally to an arbitrary number of functions.

Given $s \in \mathbb{R}$ we define the Bessel potential operator J^s through its Fourier transform $\mathcal{F} J^s u(\xi) = \langle \xi \rangle^s \hat{u}(\xi)$ for a function u, and similarly the Riesz potential operator I^s as $\mathcal{F} I^s u(\xi) = |\xi|^s \hat{u}(\xi)$.

Next we define the frequency projections that we will be utilizing. Given a dyadic number $N \in 2^{\mathbb{N}}$ let P_N denote the Littlewood-Paley projector onto the (spatial) frequencies $\{\xi \in \mathbb{R} \mid |\xi| \sim N\}$. The special case P_1 shall mean the projector onto the (spatial) frequencies $\{\xi \in \mathbb{R} \mid |\xi| \lesssim 1\}$. We direct the reader to [34] for a reference on Littlewood-Paley theory.

For $n \in \mathbb{Z}$, let the uniform (or isometric) frequency decomposition operators $\square_n$ be defined by

$$\widehat{\square_n f}(\xi) = \psi(\xi - n)\hat{f}(\xi),$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is a smooth cut-off function with the properties $\operatorname{supp}\psi \subset [-\frac{1}{4}, \frac{5}{4}]$ and $\psi(\xi) \equiv 1$ on $[0,1]$.

For these operators it is well known that, for any $1 \leq q \leq p \leq \infty$, one has

$$\|P_N f\|_{L^p} \lesssim N^{\frac{1}{q}-\frac{1}{p}} \|f\|_{L^q} \quad \text{and} \quad \|\square_n f\|_{L^p} \lesssim \|f\|_{L^q}.$$

When dealing with estimates of products of frequency localized functions, to simplify notation, we will adhere to the following convention: for $n \in \mathbb{Z}$ or a dyadic number $N \in 2^{\mathbb{N}}$ we write $u_n = \square_n u$ or $u_N = P_N u$ respectively. Complex conjugation has higher precedence than this notation, so that $\overline{u}_n = \overline{(u_{-n})}$. Different indices on different factors will not cause confusion, as we will not mix dyadic and uniform frequency localisation. Also, for ease of presentation, subscripts referring to frequency localisation may suppress other indices of functions, i.e. using $u_\ell \overline{u}_m u_n$ to refer to $(\square_\ell u_1)(\square_m \overline{u}_2)(\square_n u_3)$.

Next let us define the Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R}^n)$ (also referred to as hat-spaces for obvious reasons), for $s \in \mathbb{R}$ and $1 \leq r \leq \infty$, to be the subspace of functions $u \in \mathcal{S}'(\mathbb{R}^n)$ such that the norm

$$\|u\|_{\hat{H}^s_r} = \|\langle\xi\rangle^s \hat{u}\|_{L^{r'}(\mathbb{R}^n)}$$

is finite. In the case $s = 0$ one may resort to the slightly different notation $\hat{H}^0_r = \widehat{L^r}$. And similarly we define the modulation space $M^s_{q,p}(\mathbb{R}^n)$, for $s \in \mathbb{R}$ and $1 \leq p, q \leq \infty$, as the subspace of functions $u \in \mathcal{S}'(\mathbb{R}^n)$ such that the norm

$$\|u\|_{M^s_{q,p}} = \|\langle n\rangle^s \|\square_n u\|_{L^q(\mathbb{R})}\|_{\ell^p_n(\mathbb{Z})}$$

is finite.

Though we will not be exhaustive with the properties that these spaces have, we do want to emphasize an embedding connecting Fourier-Lebesgue and modulation spaces. For $p \geq 2$ one has $M^s_{2,p}(\mathbb{R}) \supset \hat{H}^s_{p'}(\mathbb{R})$. This embedding can be utilised to gain a notion of criticality in modulation spaces (that are otherwise not well-behaved with respect to transformations of scale because of the isometric frequency decomposition). Also we mention, that in the periodic setting these data spaces actually coincide, i.e. $M^s_{2,p}(\mathbb{T}) = \hat{H}^s_{p'}(\mathbb{T})$.

Furthermore we note, that both Fourier-Lebesgue and modulation spaces behave in a natural way with respect to complex interpolation and duality. Let $\theta \in [0,1]$, $s, s_0, s_1 \in \mathbb{R}$ and $1 < r, r_0, r_1, p, p_0, p_1 \leq \infty$. Then for $s = (1-\theta)s_0 + \theta s_1$ one has the following interpolation identities

$$\left[\hat{H}^{s_0}_{r_0}, \hat{H}^{s_1}_{r_1}\right]_{[\theta]} = \hat{H}^s_r \ \text{ for } \frac{1}{r} = \frac{1-\theta}{r_0} + \frac{\theta}{r_1} \text{ as well as}$$
$$\left[M^{s_0}_{2,p_0}, M^{s_1}_{2,p_1}\right]_{[\theta]} = M^s_{2,p} \ \text{ for } \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

as long as $(p_0, p_1) \neq (\infty, \infty)$. Under the additional constraint that $p < \infty$ the following duality relationships

$$\left(\hat{H}^s_r\right)' \cong \hat{H}^{-s}_{r'} \text{ and } \left(M^s_{2,p}\right)' \cong M^{-s}_{2,p'}$$

also hold. We mention [8, 27] as references for embedding, duality and interpolation results regarding modulation spaces.

In order to prove local well-posedness we have a necessity for spaces that are more well-adapted to performing a contraction mapping argument. In [12,13] Bourgain introduced the now almost classical $X_{s,b}$ spaces dependent on a phase function $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $s, b \in \mathbb{R}$, defined by the norm

$$\|u\|_{X_{s,b}} = \|\langle\xi\rangle^s \langle\tau - \varphi(\xi)\rangle^b \hat{u}\|_{L^2_{xt}}.$$

Using these spaces to study the local well-posedness of dispersive PDEs has since become known as the Fourier restriction norm method. It was later refined and built upon in [32,58,59] to arrive at its current use state.

In connection with the $X_{s,b}$ spaces we also define the operator Λ^b through its Fourier transform as $\mathcal{F}\Lambda^b u(\xi,\tau) = \langle\tau - \phi(\xi)\rangle^b \hat{u}(\xi,\tau)$ for a function u. The quantity $\sigma = \tau - \phi(\xi)$ is referred to as the modulation.

In the following, we define $X_{s,b}$ spaces adapted to the Fourier-Lebesgue $\hat{H}^s_r$ and modulation spaces $M^s_{2,p}$ we will be using as data spaces. For papers dealing in the same spaces see e.g. [39,40,77].

For $s, b \in \mathbb{R}$ and $1 \le r \le \infty$, we denote the Bourgain spaces adapted to Fourier-Lebesgue spaces by $\hat{X}^r_{s,b}$. They are defined as the subspace of $\mathcal{S}'(\mathbb{R}^2)$ induced by the norm

$$\|u\|_{\hat{X}^r_{s,b}} = \|\langle\xi\rangle^s \langle\tau - \varphi(\xi)\rangle^b \hat{u}\|_{L^{r'}_{xt}} = \|J^s \Lambda^b u\|_{\widehat{L^r_{xt}}},$$

so that the classical $X_{s,b}$ spaces can be recovered by setting $r = 2$. Note the lack of inverse Fourier transformation. Recall that for $1 \le r \le \infty$ we have the following embedding:

$$\hat{X}^r_{s,b} \hookrightarrow C(\mathbb{R}; \hat{H}^s_r(\mathbb{R})) \quad \text{if } b > \frac{1}{r}.$$

The contraction mapping argument leading to well-posedness will be carried out in their respective time restriction norm spaces

$$\hat{X}^r_{s,b}(\delta) = \left\{ u = \tilde{u}|_{\mathbb{R}\times[-\delta,\delta]} \mid \tilde{u} \in \hat{X}^r_{s,b} \right\}$$

endowed with the norm

$$\|u\|_{\hat{X}^r_{s,b}(\delta)} = \inf \left\{ \|\tilde{u}\|_{\hat{X}^r_{s,b}} \mid \tilde{u} \in \hat{X}^r_{s,b},\, \tilde{u}|_{\mathbb{R}\times[-\delta,\delta]} = u \right\}.$$

Similarly, for $s, b \in \mathbb{R}$ and $1 \le p \le \infty$ we define the Bourgain spaces adapted to modulation spaces $X^p_{s,b}$ (note the missing circumflex compared to the Fourier-Lebesgue based spaces). In this case they are the subspace of $\mathcal{S}'(\mathbb{R}^2)$ induced by the norm

$$\|u\|_{X^p_{s,b}} = \left\| \langle n\rangle^s \|\square_n u\|_{X_{0,b}} \right\|_{\ell^p_n}.$$

Again, $p = 2$ corresponds to the classical case. The embedding giving us the persistence property is paralleled by

$$X^p_{s,b} \hookrightarrow C(\mathbb{R}; M^s_{2,p}(\mathbb{R})) \quad \text{if } b > \frac{1}{2}.$$

for $1 \le p \le \infty$. In the same fashion as for the Fourier-Lebesgue adapted spaces, we have time restriction norm spaces $X^p_{s,b}(\delta)$.

Remark 2.1. *We fix $q = 2$ in the modulation space setting, because of the lack of available good (i.e. time independent) linear estimates in the $q \neq 2$ case, see* [8,63].

Having defined the spaces we will be using it is time to mention some of their properties. Among other things what makes Bourgain spaces useful is the ability to transfer estimates of free solutions in (mixed) L^p spaces or their Fourier-Lebesgue cousins $\widehat{L^r}$ to estimates in $\hat{X}^r_{s,b}$ spaces. This is commonly known as a transfer principle. For a proof in the classical spaces we direct the reader to the self-contained exposition in [35]. The arguments for transferring (multi)linear estimates to the Fourier-Lebesgue variants $\hat{X}^r_{s,b}$ are contained within [36].

Also contained in [36] is a general local well-posedness theorem for $\hat{X}^r_{s,b}$ spaces. A similar result, though for the modulation space variants $X^p_{s,b}$, can be found in [77], though which can easily be derived from the classics [32, 33]. Using these general well-posedness theorems we will establish our well-posedness theorems with mere proofs of necessary multilinear estimates.

As we will also be using complex multilinear interpolation and duality arguments we shall state the relevant properties of our solution, and data spaces. For this let $\theta \in [0,1]$, $s, s_0, s_1, b, b_0, b_1 \in \mathbb{R}$ and $1 < r, r_0, r_1, p, p_0, p_1 \leq \infty$. Then for $s = (1-\theta)s_0 + \theta s_1$ and $b = (1-\theta)b_0 + \theta b_1$ one has the following complex interpolation relations

$$\left[\hat{X}^{r_0}_{s_0,b_0}, \hat{X}^{r_1}_{s_1,b_1}\right]_{[\theta]} = \hat{X}^r_{s,b} \quad \text{when } \frac{1}{r} = \frac{1-\theta}{r_0} + \frac{\theta}{r_1} \text{ and}$$
$$\left[X^{p_0}_{s_0,b_0}, X^{p_1}_{s_1,b_1}\right]_{[\theta]} = X^p_{s,b} \quad \text{when } \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1},$$

at least if $(p_0, p_1) \neq (\infty, \infty)$. Moreover, with respect to the L^2 inner-product, their dual spaces are given by

$$\left(\hat{X}^r_{s,b}\right)' \cong \hat{X}^{r'}_{-s,-b} \quad \text{and} \quad \left(X^p_{s,b}\right)' \cong X^{p'}_{-s,-b}$$

if one imposes the additional constraint $p < \infty$.

Finally we recall some common inequalities that will be useful in piecing together multilinear estimates that we can establish in L^2-based $X_{s,b}$ spaces:

$$\|u_N\|_{X^q_{s,b}} \lesssim N^{\max(0,\frac{1}{q}-\frac{1}{p})} \|u_N\|_{X^p_{s,b}} \quad \text{and} \quad \sum_{N \geq 1} N^{0-} \|u_N\|_{X^p_{s,b}} \lesssim \|u\|_{X^p_{s,b}}. \tag{2.3}$$

3. The NLS hierarchy in detail

In describing what we refer to as the NLS hierarchy we most closely follow [5], where the general structure of nonlinear evolution equations that arise as zero-curvature conditions is described. Though there are many more good references for this topic (see for example [25, 79]), the chosen work [5] concisely contains all the details we need about the NLS hierarchy.

3.1. **From linear scattering to NLS.** We start out in a geometric context, where we have an $N \times N$ matrix of differential one-forms Ω depending on a so-called spectral parameter $\zeta \in \mathbb{C}$. For this matrix one can express a linear scattering problem [5, eq. (1.1)]

$$\mathrm{d}v = \Omega v. \tag{3.1}$$

Associated with this scattering problem is the zero-curvature (or integrability) condition [5, eq. (1.2)]

$$0 = \mathrm{d}\Omega - \Omega \wedge \Omega, \tag{3.2}$$

which for the right choice of Ω will result in the NLS hierarchy equations (and many other classical dispersive PDE).

In particular, as in [5, eq. (1.3)], we will use the Ansatz $\Omega = (\zeta R_0 + P)\,\mathrm{d}x + Q(\zeta)\,\mathrm{d}t$ with

$$R_0 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 0 & q \\ r & 0 \end{pmatrix}.$$

We leave the choice of Q open for now, but will refer back to it at a later point.

After a lengthy calculation, that we will not reproduce for brevities sake, it is established that the zero-curvature condition (3.2) can under our Ansatz be equivalently expressed as [5, eq. (2.3.5)]

$$\frac{\mathrm{d}}{\mathrm{d}t} u = J \frac{\delta}{\delta u} \mathcal{H}, \tag{3.3}$$

where $u = \left(\begin{smallmatrix} r \\ q \end{smallmatrix}\right)$ is a vector of the "potentials"[15], $J = -2\left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$, $\frac{\delta}{\delta r}$ is a functional derivative and $\mathcal{H}$ is *the* Hamiltonian of the system, defined by

$$\mathcal{H} = 2\sum_{n=0}^{\infty} \alpha_n(t) I_{n+1}. \tag{3.4}$$

In this sum the I_{n+1} represent the sequence of conserved quantities of our system, i.e. up to constant factor, what was referred to in the introduction as H_n. With [5, eqns. (3.1.6) and (3.1.7)] we are given explicit expressions for calculating these conserved quantities recursively

$$I_n = \int_{\mathbb{R}} qY_n \,\mathrm{d}x \quad \text{and} \quad Y_{n+1} = \frac{1}{2i}\left[\partial_x Y_n - r\delta_{0,n} + q\sum_{k=1}^{n-1} Y_{n-k}Y_k\right] \tag{3.5}$$

with $Y_0 = 0$.

The $\alpha_n(t)$ are the choice of Q we left open previously. Referring again to [5], the α_n control the weight of each individual flow (induced each by I_{n+1}) in the overall equation (3.3). Thus by choosing the coefficients $\alpha_n(t)$ appropriately we will be able to recover NLS and the other equations that are part of the NLS hierarchy.

It is important to mention that, as we are working under the assumption $r = +\overline{q}$ in the context of the NLS hierarchy, our choice of coefficients $\alpha_n(t)$ are subject to the constraints

$$\alpha_{2n} = -\overline{\alpha_{2n}} \quad \text{and} \quad \alpha_{2n+1} = \overline{\alpha_{2n+1}} \tag{3.6}$$

as layed out in [5, Section 3.2.3].

3.2. **Defining the NLS hierarchy.** Having established the general origin of the NLS hierarchy equations we are now ready to give an exact definition, i.e. fix a choice of $(\alpha_n)_{n\in\mathbb{N}_0}$. From there on we will derive the general structure of the equations in the NLS hierarchy by means of (3.5). This strictly larger class of equations will be very broad in the nonlinearities contained within, but still sufficiently small for us to be able to carry out our further analysis in this generalised context.

Definition. *For $j \in \mathbb{N}$, we define the jth (defocusing) NLS hierarchy equation to be the Hamiltonian equation for the potential $q(x,t)$ in* (3.3)*, where we choose $\alpha_{2j} \equiv -i2^{2j-1}$ and $\alpha_n \equiv 0$ for $n \neq 2j$ in* (3.4)*. We identify occurrences of the potential $r(x,t)$ with the complex conjugate of $q(x,t)$, i.e. $r = +\overline{q}$.*

Remark 3.1. *A few remarks are in order:*

(1) *Note that our choice of α_{2j} aligns with the constraint in* (3.6)*. Since we only have a single non-zero α_n a simple rescaling (and possible time reversion) of the equation would lead to any arbitrary choice of α_{2j} that aligns with* (3.6).
(2) *Since we are only interested in a single component of* (3.3) *we may simplify. The jth NLS hierarchy equation thus reads*

$$q_t = -2^{2j+1}\frac{\delta}{\delta r}\int_{\mathbb{R}} qY_{2j+1}\,\mathrm{d}x \tag{3.7}$$

with Y_{2j+1} defined in (3.5)*, keeping in mind the identification $r = +\overline{q}$.*
(3) *The first NLS hierarchy equation ($j = 1$) corresponds to the classical defocusing cubic NLS equation. In the notation of the previous display it reads*

$$iq_t = -q_{xx} + 2q^2 r = -q_{xx} + 2|q|^2 q.$$

Later we will switch to the more common notation of calling the unknown function u instead of q.

[15]Potentials are what we would usually refer to as the solution of, say, NLS. In the context of NLS we have the additional assumption $r = \pm\overline{q}$. They are referred to as potentials in [5], as they are the objects along which scattering happens in (3.1).

(4) Above we only defined the defocusing NLS hierarchy, corresponding to the $+$-sign in (2.1)*. There is also an equivalent focusing NLS hierarchy (that builds on the focusing cubic NLS, corresponding to the $-$-sign in* (2.1)*). Its equations can be derived in the same way, though with the identification $r = -\overline{q}$. This possibility is also mentioned in* [5, Section 3.2.3].

(5) No complete description of the NLS hierarchy, i.e. the choice of coefficients for the nonlinear terms, is known. A lengthy calculation leads to Appendix 7, where we list the first few conserved quantities and the associated equations.

It would certainly be an interesting problem to derive a general formula describing the jth NLS hierarchy equation in detail.

(6) Instead of a choice of $(\alpha_k)_k$, where only even numbered α_k are non-zero, going the opposite route and having only a single α_k non-zero with k uneven results in the real mKdV hierarchy.

There is a caveat to this, that is also discussed in Appendix A, where the identification $r = \pm\overline{q}$ does not lead to the (de)focusing complex mKdV hierarchy. Using $r = q$ (which is also a compatible choice with the model, see [5, Section 3.2.2]*) one arrives at the real mKdV hierarchy, which was discussed in* [38]*. This fact is also mentioned in* [64, Appendix B].

(7) Contained within this calculus of hierarchies is another well-known one, the KdV hierarchy. Choosing $r = 1$ (which is also a compatible choice in this model, see [5, Section 3.2.1]*) results in its equations. This is also remarked in* [64, Appendix B].

Having defined the NLS hierarchy equations we may now reason about their general structure. We claim the following proposition.

Proposition 3.2. *For $n \in \mathbb{N}$ the terms Y_n have the following properties:*

(1) Y_n is a sum of monomials in q, r and their derivatives.

(2) The polynomial Y_n is homogeneous in the order of monomials, where the order (of a monomial) is defined as the sum of the total number of derivatives and number of factors in the monomial.

(3) In every monomial of Y_n the total number of factors of r and its derivatives is one greater than the total number of factors q and its derivatives.

(4) The coefficients of Y_n are an integer multiple of $(2i)^{-n}$.

(5) In Y_n there is a single monomial with only one factor. It is $(2i)^{-n}\partial_x^{n-1} r$.

Proof. All of the claims in this proposition are trivially true for $Y_1 = \frac{-1}{2i} r$. For all higher-order Y_n they follow inductively using the recursion formula (3.5). □

It is only a small step from the polynomials Y_n to the conserved quantities I_n and their associated, via (3.7), evolution equations. Having derived the properties of Y_n mentioned in Proposition 3.2 we are ready to state the general structure of the NLS hierarchy equations. In doing so we switch back to the more common notation of calling the unknown solution u (instead of q).

Theorem 3.3. *For $j \in \mathbb{N}$, there exist coefficients $c_{k,\alpha} \in \mathbb{Z}$ for every $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$, for $1 \le k \le j$, such that the jth NLS hierarchy equation can be written as*

$$i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = \sum_{k=1}^{j} \sum_{\substack{\alpha \in \mathbb{N}_0^{2k+1} \\ |\alpha| = 2(j-k)}} c_{k,\alpha} \partial_x^{\alpha_1} u \prod_{\ell=1}^{k} \partial_x^{\alpha_{2\ell}} \overline{u} \partial_x^{\alpha_{2\ell+1}} u. \tag{3.8}$$

Proof. Of course, we heavily rely on the structure of Y_{2j+1} established in the preceding proposition.

First we deal with the linear part of equation (3.8): all monomials part of Y_{2j+1} have a coefficient, that is an integer multiple of $(2i)^{-(2j+1)} = -i2^{-2j-1}$. Keeping in mind, that the "leading term" of Y_{2j+1} is $\partial_x^{2j} r$ and reminding the reader of the formula for calculating functional derivatives: For a smooth function $f : \mathbb{C}^{N+1} \to \mathbb{C}$ and a functional

$$F[\phi] = \int_{\mathbb{R}} f(\phi, \partial_x \phi, \partial_x^2 \phi, \dots, \partial_x^N \phi)\,\mathrm{d}x \quad \text{one has} \quad \frac{\delta F}{\delta \phi} = \sum_{k=0}^{N} (-1)^k \partial_x^k \frac{\partial f}{\partial(\partial_x^k \phi)}, \tag{3.9}$$

we may now establish, that the linear part of the equation must read

$$i\partial_t u + (-1)^{j+1} \partial_x^{2j} u = 0$$

and we may ignore the rest of the coefficients of the nonlinearity, as they are only integers.

Using (2) and (3) from Proposition 3.2 and the fact that (3.9) reduces the number of factors $\partial_x^k r$, for a $0 \leq k \leq 2j$, by one, it is clear that the nonlinear terms must have between three and $2j+1$ factors. Of these, now there must be one more factor u (or its derivatives) compared to $\overline{u}$ (or its derivatives), as the functional derivative reduces the number of factors $\overline{u}$ (or its derivatives) by one.

The homogeneity of these nonlinear terms fixes the number of total derivatives to $2(j-k)$, if there are $2k+1$ factors.

Since (3.8) covers all possible nonlinearities that fulfil these restrictions we have established the claim of this theorem. ☐

In our later dealings we will not be relying on any more information about the structure of the NLS hierarchy equations than is given in the previous theorem. Thus it makes sense to give a name to this general class of equations.

Definition. *For $j \in \mathbb{N}$, we call an equation a (higher-order) NLS-like equation, if there exist coefficients $c_{k,\alpha} \in \mathbb{Z}$ for every $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$, for $1 \leq k \leq j$, such that the equation can be written as* (3.8).

Remark 3.4. *In a previous remark we mentioned the possibility of differentiating between the defocusing and focusing NLS hierarchy. Since the difference between the two is solely in the distribution of signs in the nonlinearity, both the defocusing and focusing NLS hierarchy equations are higher-order NLS-like equations, according to the above definition.*

Remark 3.5. *A natural question is whether there are further hierarchies of dispersive PDE arising as zero-curvature conditions* (3.2)*, possibly stemming from a different Ansatz than $\Omega = (\zeta R_0 + P)\,\mathrm{d}x + Q(\zeta)\,\mathrm{d}t$, the one we used.*

Indeed, this question is discussed in a follow-up paper [82] *to* [5]*, where the Ansatz $\Omega = (\zeta^2 R_0 + \zeta P)\,\mathrm{d}x + Q(\zeta)\,\mathrm{d}t$ is used to derive the derivative nonlinear Schrödinger (dNLS) equation*

$$i\partial_t u + \partial_x^2 u = \pm i \partial_x(|u|^2 u).$$

and more generally its associated hierarchy of PDEs.

The dNLS equation itself is an interesting object of study in the field of dispersive PDE, see for example [24, 44, 61, 65] *for some recent results and the references therein. The additional derivative in the nonlinearity, compared to the NLS equation* (2.1)*, introduces considerable difficulty in its analysis.*

A paper dealing with the well-posedness theory of the dNLS hierarchy equations is in preparation by the author.

Remark 3.6. *Having established the structure of NLS-like equations* (3.8)*, we would like to note their associated critical regularity $s_c(j)$. This will guide us as a heuristic on our investigation of the well-posedness theory of said equations.*

In line with the scaling law of NLS, a solution u of an NLS-like equation is invariant under the transformation of scale $u_\lambda(x,t) = \lambda u(\lambda x, \lambda^{2j} t)$, i.e. u_λ is a solution of the same equation, but now with initial data $u_{0,\lambda} = \lambda u_0(\lambda x)$.

This leads to all NLS-like equations being critical in the same space $\dot{H}^{-\frac{1}{2}}$ in the family of L^2-based Sobolev spaces as NLS itself, i.e. $s_c(j) = -\frac{1}{2}$. In fact, this is also true for the mKdV hierarchy [38].

Remark 3.7. *As it will turn out though, no positive well-posedness results will be possible using the contraction mapping principle near the critical regularity in L^2 based Sobolev spaces. All our well-posedness results in the scale of spaces H^s will be at fairly high regularity, supplemented by corresponding ill-posedness results to show optimality.*

Thus we turn to other scales of function spaces, in which we may keep this notion of criticality, though are able to obtain positive well-posedness results for the whole sub-critical range of spaces. In particular, we turn to Fourier-Lebesgue spaces $\hat{H}^s_r$ and modulation spaces $M^s_{2,p}$. Utilising these spaces for initial data has become commonplace for dispersive equations, as they allow to widen the class of functions for which well-posedness may be proven, inching further towards criticality. See [19, 24, 36–40, 63, 77] *for some examples where these spaces were successfully deployed.*

Especially for Fourier-Lebesgue spaces there is a well-defined notion of homogeneous space, in which one may ask the question of critical regularity for our NLS hierarchy equations. Using the equations' invariance, mentioned in Remark 3.6, we establish that all NLS-like equations are critical in the spaces $\hat{H}^s_r$ for $s_c(j,r) = -\frac{1}{r'}$ for $1 \le r \le \infty$.

For modulation spaces though there is a much less clear notion of criticality, as the spaces are not invariant under transformations of scale, due to the isometric decomposition operators $(\square_n)_{n\in\mathbb{Z}}$. Often the embedding $M^s_{2,r'} \supset \hat{H}^s_r$, for $r \le 2$, is used as guidance in the absence of criticality. Even under this notion though, we are unable to establish well-posedness with our techniques in or near the space $M^0_{2,\infty}$ (which corresponds to the critical case), paralleling results already known for the mKdV equation [19, 77].

3.3. Generalising further.

Our later well-posedness arguments sometimes do not even rely on the particular structure of the nonlinearity in (3.8), regarding the complex conjugation of factors. It is only when cubic nonlinear terms are involved, or when we are in Fourier-Lebesgue spaces, that the number of complex conjugated factors in the nonlinearity is of importance for our analysis.

We thus generalise further to an even larger class of equations.

Definition. *We call an equation a generalised (higher-order) NLS-like equation, if for $j \in \mathbb{N}$ there exist coefficients $c_{k,\alpha,b} \in \mathbb{R}$ for every $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$ and $b \in \{+,-\}^{2k+1}$, for $1 \le k \le j$, such that it can be written as*

$$i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = \sum_{k=1}^{j} \sum_{\substack{b\in\{\pm\}^{2k+1} \\ \alpha\in\mathbb{N}_0^{2k+1} \\ |\alpha|=2(j-k)}} c_{k,\alpha,b}\partial_x^{\alpha_1} v_{b_1} \prod_{\ell=1}^{k} \partial_x^{\alpha_{2\ell}} v_{b_2} \partial_x^{\alpha_{2\ell+1}} v_{b_3}, \tag{3.10}$$

where each $v_\pm$ is to be identified with u or $\overline{u}$ respectively.

In short, allowing arbitrary complex conjugation in the nonlinearity of a NLS-like equation leads to the definition of generalised NLS-like equation.

Remark 3.8. *Note that the behaviour of the equations under transformations of scale does not change with this generalisation. Thus we keep the previously established critical regularity* $s_c(j,r) = \frac{1}{r} - 1$ *in the family of Fourier-Lebesgue spaces* $\hat{H}^s_r$ *as laid out in Remarks 3.6 and 3.7.*

4. Statement of results

4.1. **Prior work on higher-order equations.** Before we move on to state our main results, let us review related work. We try to be brief and thus focus on results concerning only higher-order NLS/(m)KdV equations. Giving a complete account of the history of well-posedness theory for the NLS and (m)KdV equations is beyond our scope, though we will mention some important comparative results in the two sections following the current.

Already in [83] global existence of solutions to the jth KdV hierarchy equation was proven, with data in high regularity Sobolev spaces H^k, $k \geq j$, using a-priori estimates provided by the structure of the hierarchy equations, together with parabolic regularisation. Positive results could be achieved in both geometries $\mathbb{R}$ and $\mathbb{T}$, though due to the techniques used, full on well-posedness was not proven, as uniqueness was left unclear.

Later, in [56, 57], well-posedness even for a more general class of higher-order KdV like equations was proven. This was still at a comparatively high level of regularity for the initial data and was achieved using a gauge-transformation combined with linear smoothing estimates. As data spaces the weighted spaces $H^k(\mathbb{R}) \cap H^m(|x|^2\,\mathrm{d}x)$, for $k, m \in \mathbb{N}$ large enough, were used. It was already noted in [56] that one can drop the weight, if only cubic or higher-order terms appear in the nonlinearity.

The weighted spaces (or similar alternative spaces, like Fourier-Lebesgue ones) though turn out to be indispensable in the study of the KdV hierarchy using the contraction mapping principle. This was shown in [81], where it was established that the higher-order equations ($j \geq 2$) of the KdV hierarchy cannot have twice continuously differentiable flow[16]. In the same work it was also proven (using the contraction mapping principle), that higher dispersion KdV-like equations with quadratic nonlinearities are locally well-posed in an intersection of $H^s(\mathbb{R})$, for $s > 2j + \frac{1}{4}$, with a weighted Besov space.

Most closely resembling our results and techniques is [38], where well-posedness for the mKdV hierarchy equations (mentioned above) was derived in Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$, for $s(j,r) = \frac{2j-1}{2r'}$ with $1 < r \leq 2$, inching right up to the critical endpoint space $\widehat{L^1}(\mathbb{R})$. These results were established using a contraction mapping argument in appropriate versions of Bourgain spaces that we use too. A partial transfer of these results to higher-order KdV type equations was possible, and appears natural due to the Miura map.

A positive result, again independent of the underlying geometry, was also proven in [52]. Here the authors established well-posedness for all higher-order ($j \geq 2$) KdV hierarchy equations in H^s for $s > 4j - \frac{9}{2}$. The result relies on a modified energy estimate using lower order correction terms for the energy, thus it isn't susceptible to the barrier when trying to prove well-posedness using the contraction mapping principle mentioned above.

In recent years it has also become more fashionable to utilise the underlying integrability structure of the equations in order to derive a-priori estimates. We mention [69], where a-priori estimates for solutions of both the mKdV and NLS

[16]Technically this consequence for the KdV hierarchy was noted [38], as [81] deals only with quadratic nonlinearities. The non-quadratic nonlinear terms though are well-behaved, so failure of smoothness of the flow carries over to the KdV hierarchy.

equations were derived, building on the earlier works [67, 68] by the same authors. See also [62] for a general approach to conservation laws for integrable PDE.

Most recently published was the seminal work [64], where, relying on the integrability structure of the equations, the well-posedness of the entire KdV hierarchy in the space $H^{-1}(\mathbb{R})$ was proven, as well as in $H^{j-2}(\mathbb{T})$ for the jth equation (with dispersion order $2j+1$) in the hierarchy.

Focusing on just a single equation of the NLS hierarchy (besides NLS itself), there are only few papers dealing with low regularity well-posedness. In [28, 29] the author derives global well-posedness for data in $H^s(\mathbb{R})$ for $s \geq 4$ an integer[17]. More recently in [49] it was proven that the fourth order equation is locally well-posed in $H^s(\mathbb{R})$ for $s \geq \frac{1}{2}$ under a non-resonance condition on the coefficients of the nonlinearity. Managing to improve to local well-posedness in $H^s(\mathbb{R})$ for all $s \geq \frac{1}{2}$ (without a non-resonance condition) for generalised fourth-order NLS-like equations we mention [48, Theorem 1.3]. The same paper also contains some results on fourth-order dNLS-like equations.

There also exists a rich body of literature that deals with equations that are referred to as higher-order Schrödinger equations, but differ fundamentally from what we refer to as NLS-like equations. Usually only the order of dispersion is increased or one generalises to a higher power nonlinearity $|u|^{p-1}u$, $p > 3$, compared to NLS, specifically without increasing the number of derivatives in the nonlinearity. We note the introduction of an ever increasing number of derivatives in the nonlinearity makes the analysis considerably more difficult, compared to merely upping the dispersion[18]; this is what we focus on as our goal is covering (at least) the equations contained within the NLS hierarchy itself.

4.2. **Main results.** As we have now established, dealing with higher-order (or higher dispersion) equations is nothing new. Though what is missing from the literature is a low-regularity well-posedness theory dealing with (generalised) higher-order NLS-like equations (i.e. that mixes higher dispersion with an appropriate number of derivative in the nonlinearity).

We hope to close this gap, at least partially, with the following theorems. For this, consider a general Cauchy problem

$$\begin{cases} i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = F(u), \\ u(t=0) = u_0 \end{cases}. \tag{4.1}$$

Theorem 4.1. *Let $j \geq 2$ and (4.1) be a higher-order NLS-like equation (3.8). Then*

(1) if $1 < r \leq 2$ and $s \geq \frac{j-1}{r'}$, the Cauchy problem (4.1) for $u_0 \in \hat{H}^s_r(\mathbb{R})$ is locally well-posed in the analytic sense,

(2) if $1 < r \leq 2$, $s > -\frac{1}{r'}$ and additionally $c_{1,\alpha} = 0$ for all $\alpha \in \mathbb{N}_0^3$ (i.e. the equation contains no cubic nonlinear terms), the Cauchy problem (4.1) for $u_0 \in \hat{H}^s_r(\mathbb{R})$ is locally well-posed in the analytic sense.

For $j = 1$ this result corresponds to well-posedness of NLS in Fourier-Lebesgue spaces and is already known [37]. The case of periodic initial data was dealt with by different authors in [76].

Remark 4.2. *If we restrict ourselves to the classic Sobolev spaces $H^s(\mathbb{R})$ only, we can generalise further in Theorem 4.1 part (2). Because Proposition 6.2 allows an*

[17]We suspect there to be a typo in the the cited works as the fourth order NLS hierarchy equation given there differs slightly from the ones given by us in Appendix A.

[18]Increasing just the power in the nonlinearity, at constant dispersion and if one remains in the realm of algebraic nonlinearities, also leads to more well-behaved equations. This is mirrored by our Theorem 4.1 part (2).

arbitrary number of factors in the nonlinear terms to be complex conjugates, it is also true that any generalised higher-order NLS-like equation (3.10) *that contains no cubic terms in the nonlinearity is locally well-posed in* $H^s(\mathbb{R})$ *for* $s > -\frac{1}{2}$.

Besides Fourier-Lebesgue spaces we are also able to prove a general well-posedness result for modulation spaces $M^s_{2,p}$. In the following theorem we rely less on the distribution of complex conjugates in the nonlinearity compared with Theorem 4.1. The attentive reader will note, that Theorem 4.3 deals with any generalised higher-order NLS-like equation, so long as the cubic term corresponds to the usual $u\overline{u}u$, ignoring derivatives.

Theorem 4.3. *Let* $j \geq 2$ *and* (4.1) *be a generalised higher-order NLS-like equation* (3.10)*, where* $c_{1,\alpha,b} = 0$ *for all* $b \neq (+,-,+)$ *and* $\alpha \in \mathbb{N}_0^3$*. Then for* $2 \leq p < \infty$ *and* $s = \frac{j-1}{2}$*, the Cauchy problem* (4.1) *for* $u_0 \in M^s_{2,p}(\mathbb{R})$ *is locally well-posed in the analytic sense.*

Again, for $j = 1$ this result is essentially[19] already known [41, 63, 80] and in the periodic case from [76].

Remark 4.4. *For Theorem 4.3 a similar second part as with Theorem 4.1 could be stated, though here seems of much less value. It would be that, if* (4.1) *is a generalised higher-order NLS-like equation, but contains no cubic terms (i.e.* $c_{1,\alpha,b} = 0$ *for all* $b \in \{\pm\}^3$ *and* $\alpha \in \mathbb{N}_0^3$*) and* $s > \frac{1}{4k} - \frac{2k+1}{2k}\frac{1}{p}$*, the Cauchy problem* (4.1) *for* $u_0 \in M^s_{2,p}(\mathbb{R})$ *is locally well-posed in the analytic sense.*

Remark 4.5. *Of note is the differing influence of the distribution of complex conjugates on the well-posedness results we state in the above theorems. To quickly recap: for the cubic terms the canonical* $|u|^2u$ *(ignoring derivatives) is necessary in both Fourier-Lebesgue and modulation space settings. For the higher-order nonlinear terms though the distribution of complex conjugates can be chosen arbitrarily in the modulation space setting, whereas in Fourier-Lebesgue spaces the canonical distribution was stated necessary in Theorem 4.1.*

This is more restrictive than would be necessary considering our proof of Theorem 4.1, in particular Proposition 6.3. Looking into the details, one finds that in fact also in the Fourier-Lebesgue space setting an arbitrary distribution of complex conjugates for the higher-order nonlinear terms okay. More details on the necessary changes to the arguments given in the proof of Proposition 6.3 are given in Remark 6.4.

For proving our well-posedness theorems we rely on multilinear estimates in $\hat{X}^r_{s,b}$ (see Proposition 6.1 and Corollary 6.5) and $X^p_{s,b}$ (see Proposition 6.7 and Corollary 6.9) spaces which combined with the contraction mapping theorem lead to local well-posedness in these spaces. Using such estimates to derive local well-posedness results is a well-known technique, so we omit the specifics. They were pioneered in [12, 13] and we direct the interested reader to [35, 36] for a self-contained review of such techniques in more contemporary notation.

Contrasting the positive results above, we are also able to prove the following ill-posedness results for initial data in Fourier-Lebesgue spaces.

Theorem 4.6. *For* $j \geq 2$*,* $1 < r \leq 2$ *and* $-\frac{1}{r'} < s < \frac{j-1}{r'}$ *there exists a NLS-like equation (i.e. choice of coefficients* $c_{k,\alpha,b} \in \mathbb{R}$*) such that for the Cauchy problem* (4.1) *the flow-map* $S : \hat{H}^s_r(\mathbb{R}) \times (-T, T) \to \hat{H}^s_r(\mathbb{R})$ *cannot be uniformly continuous on bounded subsets.*

[19] Particularly for large $p \geq 3$ the continuity of the solution is an issue. This was pointed out in [80], where at least for $1 < p < 3$ continuity of the solutions was established.

And with initial data in modulation spaces, the situation is similar.

Theorem 4.7. *For $j \geq 2$, $2 \leq p \leq \infty$ and $0 \leq s < \frac{j-1}{2}$ there exists a NLS-like equation (i.e. choice of coefficients $c_{k,\alpha,b} \in \mathbb{R}$) such that for the Cauchy problem* (4.1) *the flow-map $S : M^s_{2,p}(\mathbb{R}) \times (-T,T) \to M^s_{2,p}(\mathbb{R})$ cannot be uniformly continuous on bounded subsets.*

Thus far we have only stated results about the well-posedness theory on the line $\mathbb{R}$. Regarding the torus $\mathbb{T}$, it seems no positive result is possible without additional arguments, like renormalizing the equation or moving to a weaker sense of well-posedness.

Theorem 4.8. *For any $1 \leq r \leq \infty$ and $s \in \mathbb{R}$ the flow-map $S : \hat{H}^s_r(\mathbb{T}) \times (-T,T) \to \hat{H}^s_r(\mathbb{T})$ of the Cauchy problem for the fourth-order NLS hierarchy equation ($j = 2$) cannot be three times continuously differentiable.*

Looking at lower regularities only, we may generalise to large j as well. In this case the flow becomes even more irregular:

Theorem 4.9. *For $j \geq 2$, $1 \leq r \leq \infty$ and $s < j-1$ there exists a NLS-like equation (i.e. choice of coefficients $c_{k,\alpha,b} \in \mathbb{R}$) such that for the Cauchy problem* (4.1) *the flow-map $S : \hat{H}^s_r(\mathbb{T}) \times (-T,T) \to \hat{H}^s_r(\mathbb{T})$ cannot be uniformly continuous on bounded subsets.*

4.3. **Global well-posedness for the NLS hierarchy.**

Theorem 4.10. *The solutions constructed in Theorem 4.1 for initial data $u_0 \in H^s(\mathbb{R})$, for $s \geq \frac{j-1}{2}$, extend globally in time.*

Proof. In order to prove this theorem we rely on the scale of conserved quantities constructed in [69]. Specifically referring to Theorem 1.1 and Corollary 1.2 therein, there exist conserved quantities, for all $s > -\frac{1}{2}$, such that the norm of a solution remains bounded if the norm of the initial data was finite under the flow of NLS and complex mKdV.

We must argue that the same holds for all flows in the NLS hierarchy. Combined with our local result in Theorem 4.1 this will prove the Theorem. Referencing the construction of the conserved quantities in [69, eqns. (2.12) and (2.13)], one notices that these solely rely on the so-called transmission coefficient. This quantity arises in the scattering problem we reference in (3.1), translating between two Jost solutions of $\partial_x v = (\zeta R_0 + P)v$, see [5, eqns. (2.1.6) and (2.1.25)ff.] and [69, eqns. (2.5)ff.].

Key insight is, that the transmission coefficient is always the same, independent of which equation in the hierarchy one is interested in. This is also reflected in the fact that our choice of Q (see the paragraph after (3.5)) does not influence the transmission coefficient. The importance of the transmission coefficient for at least polynomial conservation laws was also already recognized in [5, eqn. (2.1.29)]. □

4.4. **Discussion.**

Before moving on we would like to discuss our positive and negative results laid out in the preceding subsection.

First let us mention that our Theorem 4.6 establishes that, within the realm of the technique we utilise, our well-posedness result in Theorem 4.1 is optimal. In the sense that no direct application of the contraction mapping theorem will lead to well-posedness at a lower initial regularity $s \in \mathbb{R}$ than stated in Theorem 4.1, since this would lead to the flow being analytic.

This of course does not preclude the possibility of different arguments, more heavily relying on the integrability of the hierarchy, similar to [64] for the KdV hierarchy, leading to well-posedness in $H^s(\mathbb{R})$ for some $s < \frac{j-1}{2}$.

We extend previous results regarding the fourth order equation: in [49] it was shown that the fourth order NLS equation is locally well-posed in $H^s(\mathbb{R})$, for $s \geq \frac{1}{2}$, under a non-resonance condition on the coefficients in the nonlinearity. This non-resonance condition could be removed by different authors in [48]. We are also able to remove this condition (using different underlying function spaces) and extend the well-posedness result to all higher-order Schrödinger equations. Also, the global result in [28, 29] we extend all the way down to our local result using the a priori estimates from [69].

Not included in our well-posedness result is the critical space on our scale of function spaces $\widehat{L^1}(\mathbb{R})$. Though this comes at no surprise as this space contains some nasty initial data, including the Dirac delta δ_0. For this, shown in [60, Theorem 1.5], it is known that no suitable notion of solution may be defined in the case of NLS. We mention the ongoing effort of extending well-posedness results (under weakened continuity assumptions on the flow) to ever greater spaces comparable to the critical $H^{-\frac{1}{2}}$ or $\widehat{L^1}$. See [7] for recent developments and an overview.

In connection with Theorem 4.6 we would also like to mention, that our arguments do not establish ill-posedness for the *actual* NLS hierarchy equations. Rather looking at a set of related equations, the first of which we give in (7.4).

Next we argue our interest in the other scale of function spaces that we deal with, modulation spaces. Recall that in [41, 80] it was shown, that NLS is locally well-posed in $M^0_{2,p}(\mathbb{R})$ for $2 \leq p < \infty$. This exhausts the entire subcritical range suggested by the scaling heuristic (where the critical space is $H^{-\frac{1}{2}}$) and the embedding $M^s_{2,p} \supset \hat{H}^s_{p'}$ for $p \geq 2$. In the Fourier-Lebesgue space setting similar results were shown in [37], establishing local well-posed of NLS in $\widehat{L^r}$ for $1 < r < \infty$.

For mKdV local well-posedness was also established in almost critical Fourier-Lebesgue spaces. Specifically in [40] it was shown that mKdV is locally well-posed in $\hat{H}^s_r(\mathbb{R})$ for $s = \frac{1}{2} - \frac{1}{2r}$ and $2 \geq r > 1$. In the modulation space setting though a gap of a quarter[20] derivative between the scaling heuristic and the optimal result appears. To be exact, in [19, 77], it was shown that mKdV is locally well-posed in $M^{\frac{1}{4}}_{2,p}(\mathbb{R})$ for $2 \leq p \leq \infty$ and that this is optimal in the sense that the flow fails to be uniformly continuous for $s < \frac{1}{4}$.

Our Theorem 4.3 parallels this development for the higher-order equations, i.e. for every step to the next equation in the NLS hierarchy another half-derivative regularity of the initial data is necessary for our positive result[21].

Moving on to results for the torus $\mathbb{T}$, Theorem 4.8 establishes that no well-posedness results may be established using the contraction mapping principle directly in the data spaces we use, at least for the next higher-order equation. This is in stark contrast to NLS, where well-posedness in $L^2(\mathbb{T})$ was established in [12].

It is reasonable to believe that the further NLS hierarchy equations are ill-posed in a similar manner and do not allow direct treatment with the contraction mapping principle. Even worse though at low regularities: here Theorem 4.9 establishes a milder form of ill-posedness, but in this case for an NLS-like equation of arbitrarily high (dispersion) order.

[20] It is a quarter of a derivative keeping in mind we accept the embedding $M^0_{2,\infty} \supset \widehat{L^1}$ as our guidance for criticality in the modulation space setting.

[21] Note that the half-derivative increase is for stepping from one NLS hierarchy equation to the next. Looking also at the mKdV hierarchy equations in modulation spaces would be an interesting feat. The author expects that well-posedness would be achieved in modulation spaces differing by a quarter derivative from the corresponding NLS hierarchy results.

In such cases, where the proper model fails to have a well-behaved local theory, it sometimes helps to look at a renormalized/gauge-transformed version of the equation. For example, with NLS below $L^2(\mathbb{T})$, considering the so-called Wick ordered NLS equation

$$i\partial_t u + \partial_x^2 u = \pm\left(|u|^2 - \frac{1}{\pi}\int_{\mathbb{T}} |u|^2 \,\mathrm{d}x\right) u \tag{4.2}$$

has lead to some success. See [74] for a review. Transitioning to a renormalised equation (via a gauge-transformation) is also a common approach with the derivative NLS equation [37, 39, 46, 78, 86].

For our NLS hierarchy equations such a renormalisation might also lead to positive well-posedness results. Though it is not clear if such an approach would yield well-posedness only for the NLS hierarchy equations or for a general class, like in our results on the line.

Another viable path to approaching well-posedness on the torus (but also on the line in $H^s(\mathbb{R})$ for some $s < \frac{j-1}{2}$) is to rely on the integrability of the equation, as was done for the KdV hierarchy in [64]. This has the disadvantage of definitely not working for similar, but non-integrable, variants of higher-order NLS-like equations.

5. Linear and multilinear smoothing estimates

In the following section we will be collecting and proving smoothing estimates for free solutions of equation (4.1), i.e. with $F = 0$. To shorten notation, consider solutions $u(x,t) = e^{(-1)^j t\partial_x^{2j}} u_0(x)$, $v(x,t) = e^{(-1)^j t\partial_x^{2j}} v_0(x)$ and $w(x,t) = e^{(-1)^j t\partial_x^{2j}} w_0(x)$ with initial data u_0, v_0 and w_0 respectively when proving estimates involving free solutions. Likewise u, v and w will refer to functions in appropriate $X_{s,b}$ space variants when talking about estimates in these spaces.

5.1. Linear estimates.

The following linear estimates are essentially known in the literature. Our proof of Proposition 6.2 relies heavily upon them.

Proposition 5.1. *Let $b > \frac{1}{2}$, then the following inequalities hold*

(1) *for $2 \leq q \leq \infty$ and $\sigma > \frac{1}{2} - \frac{2j}{q}$*

$$\|u\|_{L_x^\infty L_t^q} \lesssim \|u\|_{X_{\sigma,b}} \tag{5.1}$$

(2) *for $2 \leq p \leq \infty$ and $\sigma = -\frac{2j-1}{2}(1 - \frac{2}{p})$*

$$\|u\|_{L_x^p L_t^2} \lesssim \|u\|_{X_{\sigma,b}} \tag{5.2}$$

(3) *for $4 \leq p \leq \infty$ and $\sigma > \frac{1}{2} - \frac{1}{p}$*

$$\|u\|_{L_x^p L_t^\infty} \lesssim \|u\|_{X_{\sigma,b}} \tag{5.3}$$

Proof. These linear estimates are interpolated variants of a Kato-type local smoothing estimate (for (5.1) and (5.2)) and a maximal function estimate (for (5.3)).

From [53, Theorem 4.1] we know for large frequencies

$$\|(\mathrm{id} - P_1)u\|_{L_x^\infty L_t^2} \lesssim \|(\mathrm{id} - P_1)u_0\|_{H^\sigma}, \quad \text{for } \sigma = -\frac{2j-1}{2}. \tag{5.4}$$

For small frequencies we may use a Sobolev-embedding in the space variable, where we may ignore the loss of derivatives. So we also know (5.4) without the projector $(\mathrm{id} - P_1)$. Using the transfer principle on this bound and interpolating with the trivial bounds

$$\|u\|_{L_{xt}^\infty} \lesssim \|u\|_{X_{\frac{1}{2}+,b}} \quad \text{and} \quad \|u\|_{L_{xt}^2} \lesssim \|u\|_{X_{0,0}} \quad \text{for } b > \frac{1}{2} \tag{5.5}$$

results in estimates (5.1) and (5.2) above respectively.

For the maximal function estimate we cite [53, Theorem 2.5], where

$$\|(\mathrm{id} - P_1)u\|_{L_x^4 L_t^\infty} \lesssim \|(\mathrm{id} - P_1)u_0\|_{H^{\frac{1}{4}}}$$

is established, again only for high frequencies. The same estimate was also independently found in [84]. Taking care of low frequencies as above and interpolating with the first bound in (5.5) results in estimate (5.3) above. □

5.2. **Bilinear estimates.** Before we can go about proving any bilinear estimates, we will first define the bilinear operators which we will use. We will need two bilinear operators, the estimates for which will also differ if complex conjugation is applied to one of the factors, since our phase function is even. This is in contrast to the mKdV hierarchy in [38], where the phase function is odd, and thus $X_{s,b}$ norms are invariant under complex conjugation.

So for $j \in \mathbb{N}$ and $1 \le p \le \infty$ define the pair of bilinear operators $I^{\pm}_{p,j}$ by their Fourier transform:

$$\mathcal{F}_x I^{\pm}_{p,j}(f,g)(\xi) = c \int_* k_j^{\pm}(\xi_1,\xi_2)^{\frac{1}{p}} \hat{f}(\xi_1)\hat{g}(\xi_2)\,\mathrm{d}\xi_1.$$

Their symbol is given by

$$k_j^{\pm}(\xi_1,\xi_2) = |\xi_1 \pm \xi_2|(|\xi_1|^{2j-2} + |\xi_2|^{2j-2}).$$

Comparing with the linear estimates above, this bilinear operator is a refinement in the sense that we now have access to the symbol of the not-quite-derivative $|\xi_1 \pm \xi_2|$. The following proposition establishes a corresponding estimate:

Proposition 5.2. *Let $j \ge 1$ and $1 \le q \le r_1, r_2 \le p \le \infty$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r_1} + \frac{1}{r_2}$. Then one finds*

$$\|\mathcal{F}_x I^{\pm}_{p,j}(u, v_\pm)(\xi,\cdot)\|_{\widehat{L^p_t}} \lesssim (|\hat{u}_0|^{p'} * |\hat{v}_0|^{p'}(\xi))^{\frac{1}{p'}}$$

and

$$\|I^{\pm}_{p,j}(u,v_\pm)\|_{\widehat{L^q_x}\widehat{L^p_t}} \lesssim \|u_0\|_{\widehat{L^{r_1}_x}} \|v_0\|_{\widehat{L^{r_2}_x}},$$

where $v_+ = \overline{v}$ and $v_- = v$.

Proof. We will only write down the details for the +-case, the proof of the −-case is similar and we omit the details. Let us begin by calculating the Fourier transform only in the space-variable:

$$\mathcal{F}_x I^{+}_{p,j}(u,\overline{v})(\xi,t) = \int_* k_j^{+}(\xi_1,\xi_2)^{\frac{1}{p}} e^{it(\xi_1^{2j} - \xi_2^{2j})} \hat{u}_0(\xi_1)\hat{\overline{v}}_0(\xi_2)\,\mathrm{d}\xi_1.$$

And now for the complete Fourier transform, substituting $x = \xi_1 - \frac{\xi}{2}$

$$\mathcal{F} I^{+}_{p,j}(u,\overline{v})(\xi,t) = \int_* k_j^{+}(\xi_1,\xi_2)^{\frac{1}{p}} \delta(\tau - \xi_1^{2j} + \xi_2^{2j})\hat{u}_0(\xi_1)\hat{\overline{v}}_0(\xi_2)\,\mathrm{d}\xi_1 \tag{5.6}$$

$$= \int_* k_j^{+}\left(\frac{\xi}{2} + x, \frac{\xi}{2} - x\right)^{\frac{1}{p}} \delta(\tau - g(x))\hat{u}_0\left(\frac{\xi}{2} + x\right)\hat{\overline{v}}_0\left(\frac{\xi}{2} - x\right)\mathrm{d}x \tag{5.7}$$

$$= \int_* \left(\sum_n \frac{\delta(x - x_n)}{|g'(x_n)|}\right) k_j^{+}\left(\frac{\xi}{2} + x, \frac{\xi}{2} - x\right)^{\frac{1}{p}} \hat{u}_0\left(\frac{\xi}{2} + x\right)\hat{\overline{v}}_0\left(\frac{\xi}{2} - x\right)\mathrm{d}x, \tag{5.8}$$

where the sum $\sum_n$ is over the simple solutions of the equation $\tau - g(x) = 0$ involving the function

$$g(x) = \left(\frac{\xi}{2}+x\right)^{2j} - \left(\frac{\xi}{2}-x\right)^{2j} = \sum_{k=0}^{2j} \binom{2j}{k} \left(\frac{\xi}{2}\right)^{2j-k} (x^k - (-x)^k)$$
$$= 2\sum_{l=1}^{j} \binom{2j}{2l-1} \left(\frac{\xi}{2}\right)^{2(j-l)+1} x^{2l-1} = \xi x \sum_{l=1}^{j} \binom{2j}{2l-1} \left(\frac{\xi}{2}\right)^{2(j-l)} x^{2(l-1)}.$$

By our choice in the definition of the symbol of our bilinear operator we have the following lower bound on the absolute value of the derivative of $g(x)$

$$|\partial_x g(x)| \sim |\xi| \sum_{l=1}^{j} \binom{2j}{2l-1} \left(\frac{\xi}{2}\right)^{2(j-l)} x^{2(l-1)} \gtrsim k_j^+\left(\frac{\xi}{2}+x, \frac{\xi}{2}-x\right).$$

Now $g(x)$, as a sum of monotone functions, only admits a single (real) solution of $\tau - g(x) = 0$. Calling this solution $y \in \mathbb{R}$ we can bound, except on a ξ set of measure zero

$$(5.8) \lesssim k_j^+\left(\frac{\xi}{2}+y, \frac{\xi}{2}-y\right)^{-\frac{1}{p'}} \hat{u}_0\left(\frac{\xi}{2}+y\right) \hat{\overline{v}}_0\left(\frac{\xi}{2}-y\right).$$

In order to now calculate the $L^{p'}_\tau$-norm of this expression we substitute the measure $\mathrm{d}\tau = |g'(y)|\,\mathrm{d}y$, since we have $\tau = g(y)$, which causes the symbol of the operator to disappear and we arrive at

$$\|\mathcal{F}_x I^+_{p,j}(u,\overline{v})(\xi,\cdot)\|_{\widehat{L^p_t}} = \int_{\mathbb{R}} \left|\hat{u}_0\left(\frac{\xi}{2}+y\right)\hat{v}_0\left(\frac{\xi}{2}-y\right)\right|^{p'} \mathrm{d}y = |\hat{u}_0|^{p'} * |\hat{v}_0|^{p'}(\xi).$$

This proves our first claim. In order to now extend this to an $\widehat{L^q_x}\widehat{L^p_t}$ result we make use of Young's convolution inequality. For this choose $\rho' = \frac{q'}{p'}$, $\rho_k = \frac{r_k'}{p'}$ for $k \in \{1,2\}$, so that $\frac{1}{\rho} = \frac{1}{\rho_1} + \frac{1}{\rho_2}$. Then

$$\|I^+_{p,j}(u,\overline{v})\|_{\widehat{L^q_x}\widehat{L^p_t}} \lesssim \left(\int \left||\hat{u}_0|^{p'} * |\hat{v}_0|^{p'}(\xi)\right|^{\frac{q'}{p'}} \mathrm{d}\xi\right)^{\frac{1}{q'}} = \||\hat{u}_0|^{p'} * |\hat{v}_0|^{p'}\|_{L^{\rho'}_\xi}$$
$$\lesssim \left[\||\hat{u}_0|^{p'}\|_{L^{\rho_1'}_\xi} \||\hat{v}_0|^{p'}\|_{L^{\rho_2'}_\xi}\right]^{\frac{1}{p'}} = \|u_0\|_{\widehat{L^{r_1}_x}} \|v_0\|_{\widehat{L^{r_2}_x}}$$

as claimed and the proof is complete. □

Using the transfer principle for $\hat{X}^r_{s,b}$ spaces mentioned in Section 2.2 we may now conclude:

Corollary 5.3. *Let $1 \le q \le r_1, r_2 \le p < \infty$ and $b_i > \frac{1}{r_i}$. Then we have*

$$\|I^\pm_{p,j}(u,v_\pm)\|_{\widehat{L^q_x}\widehat{L^p_t}} \lesssim \|u\|_{\hat{X}^{r_1}_{0,b_1}} \|v\|_{\hat{X}^{r_2}_{0,b_2}}, \tag{5.9}$$

where $v_+ = \overline{v}$ and $v_- = v$.

Throughout dealing with the cubic terms we will also make heavy use of inequalities that can be interpreted as the duals of those in (5.9). For this view the bilinear operators as maps

$$u \mapsto I^\pm_{p,j}(u, v_\pm), \qquad \hat{X}^{r_1}_{0,b_1} \to \widehat{L^q_x}\widehat{L^p_t}$$

i.e. as a multiplication with $v_\pm$ with operator norm $\lesssim \|v\|_{X^{r_2}_{0,b_2}}$. By duality we also have the continuity, except in the endpoint case, of the map

$$w \mapsto I^{\pm,*}_{p,j}(w, v_\mp), \qquad \widehat{L^{q'}_x}\widehat{L^{p'}_t} \to \hat{X}^{r_1'}_{0,-b_1}$$

with the same upper bound for the operator norm. Note how we now multiply with $v_\mp$ instead of $v_\pm$. A straightforward calculation gives the associated symbols of the operators $I^{\pm,*}_{p,j}$ as

$$k_j^{+,*}(\xi_1,\xi_2) = |\xi_1|(|\xi_1|^{2j-2}+|\xi_2|^{2j-2}), \quad k_j^{-,*}(\xi_1,\xi_2) = |\xi_1+2\xi_2|(|\xi_1|^{2j-2}+|\xi_2|^{2j-2}).$$

We collect the new estimates in the following

Corollary 5.4. *Let $1 < q \leq r_1, r_2 \leq p < \infty$ with $\frac{1}{p}+\frac{1}{q} = \frac{1}{r_1}+\frac{1}{r_2}$ and $b_i > \frac{1}{r_i}$. Then the estimate*

$$\|I^{\pm,*}_{p,j}(u, v_\mp)\|_{\hat{X}^{r_1'}_{0,-b_1}} \lesssim \|u\|_{\widehat{L^{q'}_x}\widehat{L^{p'}_t}} \|v\|_{\hat{X}^{r_2}_{0,b_2}} \tag{5.10}$$

holds. If alternatively $0 \leq \frac{1}{\rho'} \leq \frac{1}{r'}$ and $\beta < -\frac{1}{\rho'}$ we have

$$\|I^{\pm,*}_{\rho',j}(u, v_\mp)\|_{\hat{X}^{r}_{0,\beta}} \lesssim \|u\|_{\widehat{L^{r}_{xt}}} \|v\|_{\hat{X}^{\rho'}_{0,-\beta}}. \tag{5.11}$$

In both cases $v_+ = \overline{v}$ and $v_- = v$.

Proof. The first estimate follows from above arguments, for the second inequality we first mention the endpoint of Young's convolution inequality

$$\|uv_\mp\|_{\widehat{L^r_{xt}}} \lesssim \|u\|_{\widehat{L^r_{xt}}} \|v\|_{\widehat{L^\infty_{xt}}}.$$

which we will use in the form

$$\|I^{\pm,*}_{\infty,j}(u, v_\mp)\|_{\hat{X}^{r}_{0,0}} \lesssim \|u\|_{\widehat{L^{r}_{xt}}} \|v\|_{\hat{X}^{\infty}_{0,0}}. \tag{5.12}$$

Setting $q = r_1 = r_2 = p = r'$ in (5.10) results in

$$\|I^{\pm,*}_{r',j}(u, v_\mp)\|_{\hat{X}^{r}_{0,-b}} \lesssim \|u\|_{\widehat{L^{r}_{xt}}} \|v\|_{\hat{X}^{r'}_{0,b}} \quad \text{for } b > \frac{1}{r'}. \tag{5.13}$$

Now applying Stein's interpolation theorem between (5.12) and (5.13) results in the desired bound (5.11). □

5.3. **Fefferman-Stein estimate.** For later interpolation arguments we need a generalization of the Fefferman-Stein [26] inequality for higher-order phase functions.

Proposition 5.5. *Let $4 < q < \infty$ and $\frac{1}{r} = \frac{1}{2} + \frac{1}{q}$. For $\sigma = \frac{j-1}{2}$ one has*

$$\|I^\sigma u\|_{L^4_t L^q_x} \lesssim \|u_0\|_{\widehat{L^r_x}}.$$

Proof. We at first assume, that $\hat{u}_0(\xi) = \chi_{(0,\infty)}(\xi)\hat{u}_0(\xi)$. Furthermore let $v = I^\sigma u$, then

$$\|I^\sigma u\|^4_{L^4_t L^q_x} = \||v|^2\|^2_{L^2_t L^{\frac{q}{2}}_x} \lesssim \|I^\varepsilon |v|^2\|^2_{L^2_{xt}} = \|\mathcal{F} I^\varepsilon |v|^2\|^2_{L^2_{xt}},$$

where we have $\varepsilon = \frac{1}{2} - \frac{2}{q}$. Calculating the Fourier transform and substituting $x = \xi_1 - \frac{\xi}{2}$ we get

$$\mathcal{F}(I^\varepsilon v\overline{v})(\xi,\tau) \sim \int_{\mathbb{R}} |\xi|^\varepsilon \delta(g(x)-\tau) \hat{u}_0\left(\frac{\xi}{2}+x\right) \overline{\hat{u}}_0\left(\frac{\xi}{2}-x\right) \mathrm{d}x. \tag{5.14}$$

In order to rid ourselves of the Dirac delta present in the integral we derive a lower bound on the derivative of its argument:

$$g(x) = \xi_1^{2j} - \xi_2^{2j} = \sum_{l=1}^{j} \binom{2j}{2l-1} \left(\frac{\xi}{2}\right)^{2(j-l)+1} x^{2l-1}$$

$$|g'(y)| \sim |\xi| \sum_{l=1}^{j} (2l-1) \binom{2j}{2l-1} \left(\frac{\xi}{2}\right)^{2(j-l)} x^{2l} \gtrsim |\xi| y^{2(j-1)} \tag{5.15}$$

In (5.15) y refers to the single real solution that $g(x)-\tau=0$ admits, as a sum of monotone functions. With this we can simplify (5.14) to

$$\mathcal{F}(I^\varepsilon v\overline{v})(\xi,\tau)\lesssim|\xi|^{\varepsilon-\frac{1}{2}}\frac{y^{-(j-1)}}{\sqrt{|g'(y)|}}\hat{u}_0\left(\frac{\xi}{2}+y\right)\hat{\overline{u}}_0\left(\frac{\xi}{2}-y\right).$$

Thanks to our assumed condition on the support of u_0 we only have a contribution if $\frac{\xi}{2}+y\geq 0$ and $-\frac{\xi}{2}+y\geq 0$ which allows us to write $2y=(\frac{\xi}{2}+y)+(-\frac{\xi}{2}+y)=|\frac{\xi}{2}+y|+|\frac{\xi}{2}-y|$. Thus we control the arguments of $\hat{u}_0$ and $\hat{\overline{u}}_0$ and with that the derivatives on these terms via y.

$$\lesssim\frac{|\xi|^{\varepsilon-\frac{1}{2}}}{\sqrt{|g'(y)|}}\mathcal{F}_x(I^{-\frac{j-1}{2}}u_0)\left(\frac{\xi}{2}+y\right)(\mathcal{F}_xI^{-\frac{j-1}{2}}\overline{u}_0)\left(\frac{\xi}{2}-y\right)$$

Piecing the $L^2_{\xi\tau}$-norm together and substituting the measure $\mathrm{d}\tau=g(y)\,\mathrm{d}y$ and $z_\pm=y\pm\frac{\xi}{2}$ gives

$$\begin{aligned}\|\mathcal{F}I^\varepsilon|v|^2\|^2_{L^2_{xt}}&\lesssim\int\frac{|\xi|^{2\varepsilon-1}}{|g'(y)|}\left|\mathcal{F}_x(I^{-\frac{j-1}{2}}u_0)\left(\frac{\xi}{2}+y\right)(\mathcal{F}_xI^{-\frac{j-1}{2}}\overline{u}_0)\left(\frac{\xi}{2}-y\right)\right|^2\mathrm{d}\xi\,\mathrm{d}\tau\\&\lesssim\int|z_+-z_-|^{2\varepsilon-1}|\hat{u}_0(z_+)\hat{u}_0(z_-)|^2\,\mathrm{d}z_+\,\mathrm{d}z_-.\end{aligned}$$

An application of the Hardy-Littlewood-Sobolev inequality requires $0<1-2\varepsilon<1$ and $\frac{4}{r'}+1-2\varepsilon=2$, which is equivalent to $4<q<\infty$ and $\frac{1}{r}=\frac{1}{2}+\frac{1}{q}$. So HLS gives us the desired upper bound. The support condition on $\hat{u}_0$ can be lifted by noting that both norms on the left and right hand side of the inequality are invariant with respect to complex conjugation. □

Interpolating the above proposition with the endpoint of the Riemann-Lebesgue lemma $\|u\|_{L^\infty_{xt}}\lesssim\|u\|_{\widehat{L^\infty_{xt}}}$ gives

Corollary 5.6. *Let* $\frac{1}{r}=\frac{2}{p}+\frac{1}{q}$, $0<\frac{1}{q}<\frac{1}{4}$ *and* $0\leq\frac{1}{p}\leq\frac{1}{4}$. *Then one finds that*

$$\|I^{\frac{2(j-1)}{p}}u\|_{L^p_tL^q_x}\lesssim\|u_0\|_{\widehat{L^r_x}}.$$

The diagonal case $p=q=3r$ is of special interest and the only one we will make use of. Using the transfer principle we have the estimate

$$\|I^{\frac{2(j-1)}{3r}}u\|_{L^{3r}_{xt}}\lesssim\|u\|_{\hat{X}^r_{0,b}}\tag{5.16}$$

as long as $b>\frac{1}{r}$ and $0\leq\frac{1}{r}<\frac{3}{4}$.

5.4. **Trilinear estimates.** Particularly in the realm of $r\ll 2$ we rely on a trilinear refinement of a Strichartz type estimate in order to derive our local well-posedness result. Specifically we rely on it in proving the trilinear estimates leading to Theorems 4.1 and 4.3. Though in contrast to the mKdV hierarchy, we may prove our trilinear estimate in a more general setting, not relying on a specific frequency constellation; see [38, Section 3.2]. This parallels the $j=1$ case, see for example [37].

Proposition 5.7. *Let* $1<p_1<p<p_0<\infty$, $p<p_0'$, $\frac{3}{p}=\frac{1}{p_0}+\frac{2}{p_1}$ *and* $\frac{2}{p_1}<1+\frac{1}{p}$. *Then we have the estimate*

$$\|uv\overline{w}\|_{\widehat{L^p_{xt}}}\lesssim\|u_0\|_{\widehat{L^{p_0}_x}}\|I^{-\frac{j-1}{p}}v_0\|_{\widehat{L^{p_1}_x}}\|I^{-\frac{j-1}{p}}w_0\|_{\widehat{L^{p_1}_x}}.\tag{5.17}$$

Proof. We begin by taking the Fourier transform in both space- and time-variable of the product $uv\overline{w}$ and substituting $\xi_{2,3}=\frac{\xi-\xi_1}{2}\pm x$

$$\mathcal{F}(uv\overline{w})(\xi,\tau)\sim\int_*\delta(g(\xi_1;x)-\tau)\hat{u}_0(\xi_1)\hat{v}_0\left(\frac{\xi-\xi_1}{2}+x\right)\hat{\overline{w}}_0\left(\frac{\xi-\xi_1}{2}-x\right)\mathrm{d}\xi_1\,\mathrm{d}\xi_2,$$

where in the argument of the Dirac delta

$$g(\xi_1;x) = \xi_1^{2j} + \xi_2^{2j} - \xi_3^{2j} = \xi_1^{2j} + \sum_{k=0}^{2j} \binom{2j}{k} \left(\frac{\xi-\xi_1}{2}\right)^{2j-k} (x^k - (-x)^k)$$
$$= \xi_1^{2j} + (\xi - \xi_1) \sum_{l=1}^{j} \binom{2j}{2l-1} \left(\frac{\xi-\xi_1}{2}\right)^{2(j-l)} x^{2l-1}.$$

As a sum of monotone functions $g(\xi_1;x)$ only admits a single (real) solution with respect to x of $g(x) - \tau = 0$, which we will call $y \in \mathbb{R}$. We can bound the derivative of g from below at this root by

$$|g'(\xi_1;y)| = |\xi - \xi_1| \sum_{l=1}^{j} (2l-1) \binom{2j}{2l-1} \left(\frac{\xi-\xi_1}{2}\right)^{2(j-l)} y^{2(l-1)}$$
$$\gtrsim |\xi - \xi_1|(|\xi-\xi_1|^{2(j-1)} + y^{2(j-1)}).$$

Having estimated $|g'(\xi_1,y)|$ we may move back to proving our trilinear estimate. An application of Hölder's inequality splits the integral into two parts:

$$\mathcal{F}(uv\overline{w})(\xi,\tau) = \int \frac{\hat{u}_0(\xi_1)\hat{v}_0(\frac{\xi-\xi_1}{2}+y)\hat{\overline{w}}_0(\frac{\xi-\xi_1}{2}-y)}{|g'(\xi_1;y)|}\,\mathrm{d}\xi_1 \tag{5.18}$$

$$\leq \left(\int \frac{|\hat{u}_0(\xi_1)|^p\,\mathrm{d}\xi_1}{|\xi-\xi_1|^{(1-\theta)p}}\right)^{\frac{1}{p}} \left(\int \frac{|\hat{v}_0(\frac{\xi-\xi_1}{2}+y)\hat{\overline{w}}_0(\frac{\xi-\xi_1}{2}-y)|^{p'}|\xi-\xi_1|^{p'}}{|\xi-\xi_1|^{\theta p'}|g'(\xi_1,y)|^{p'}}\,\mathrm{d}\xi_1\right)^{\frac{1}{p'}}. \tag{5.19}$$

To estimate the first factor in (5.19) we use the weak Young inequality to deal with the $L_\xi^{p'}$-norm

$$\||\hat{u}_0|^p * |\cdot|^{(\theta-1)p}\|^{\frac{1}{p}}_{L_\xi^{\frac{p'}{p}}} \lesssim \left(\||\hat{u}_0|^p\|_{L_\xi^{\frac{p_0'}{p}}} \||\cdot|^{(\theta-1)p}\|_{L_\xi^{\frac{1}{(\theta-1)p},\infty}}\right)^{\frac{1}{p}} \lesssim \|u_0\|_{\widehat{L_x^{p_0}}}.$$

Its application calls for

$$0 < (1-\theta)p < 1, \quad 1 < \frac{p_0'}{p} < \frac{1}{1-(1-\theta)p}, \quad \theta = \frac{1}{p_0'}$$

which are all fulfilled thanks to our requirements for the Hölder exponents.

Moving on to the second factor in (5.19), where we rely on our bound on the derivative $|g'(\xi_1;y)| \gtrsim |\xi-\xi_1|(|\xi-\xi_1|^{2(j-1)} + y^{2(j-1)})$, we may estimate

$$\left(\int \frac{|\hat{v}_0(\frac{\xi-\xi_1}{2}+y)\hat{\overline{w}}_0(\frac{\xi-\xi_1}{2}-y)|^{p'}|\xi-\xi_1|^{p'}}{|\xi-\xi_1|^{\theta p'}|g'(\xi_1,y)|^{p'}}\,\mathrm{d}\xi_1\right)^{\frac{1}{p'}} \tag{5.20}$$

$$\lesssim \left(\int \frac{|(\mathcal{F}_x I^{-\frac{j-1}{p}} v_0)(\frac{\xi-\xi_1}{2}+y)(\mathcal{F}_x I^{-\frac{j-1}{p}} w_0)(\frac{\xi-\xi_1}{2}-y)|^{p'}\,\mathrm{d}\xi_1}{|\xi-\xi_1|^{\theta p'-1}|g'(\xi_1,y)|}\right)^{\frac{1}{p'}}. \tag{5.21}$$

Now taking the $L_\tau^{p'}$-norm of the preceding line and then substituting both the measure $\mathrm{d}\tau = g'(\xi_1;y)\,\mathrm{d}y$ and $z_\pm = \frac{\xi-\xi_1}{2} \pm y$ we arrive at

$$\left(\int \frac{|(\mathcal{F}_x I^{-\frac{j-1}{p}} v_0)(z_+)(\mathcal{F}_x I^{-\frac{j-1}{p}} w_0)(z_-)|^{p'}\,\mathrm{d}z_+\,\mathrm{d}z_-}{|z_+ + z_-|^{\theta p'-1}}\right)^{\frac{1}{p'}} \tag{5.22}$$

$$\lesssim \quad \|I^{-\frac{j-1}{p}} v_0\|_{\widehat{L_x^{p_1}}} \|I^{-\frac{j-1}{p}} w_0\|_{\widehat{L_x^{p_1}}}, \tag{5.23}$$

where we used the Hardy-Littlewood-Sobolev inequality, noting that $\theta = \frac{3}{p'} - \frac{2}{p_1'} \in (0,1)$ by our conditions on the Hölder exponents and thus that $\theta p' - 1 \in (0,1)$, $\frac{2}{p_1'} + \theta p' - 1 = 2$ and $p_1' > 1$. This concludes the proof of the trilinear estimate. □

In order for this trilinear estimate to actually be useful (we want the same $\widehat{L^r_x}$-norm on all factors) we must interpolate this estimate with the Fefferman-Stein inequality from the previous subsection.

Corollary 5.8. *Let $1 < r \leq 2$, then there exist $s_0, s_1 \geq 0$ such that $s_0 + 2s_1 = \frac{2(j-1)}{r}$ and*

$$\|uv\overline{w}\|_{\widehat{L^r_{xt}}} \lesssim \|I^{-s_0}u_0\|_{\widehat{L^r_x}} \|I^{-s_1}v_0\|_{\widehat{L^r_x}} \|I^{-s_1}w_0\|_{\widehat{L^r_x}}.$$

In addition, if $b > \frac{1}{r}$ then

$$\|uv\overline{w}\|_{\widehat{L^r_{xt}}} \lesssim \|I^{-s_0}u\|_{\hat{X}^r_{0,b}} \|I^{-s_1}v\|_{\hat{X}^r_{0,b}} \|I^{-s_1}w\|_{\hat{X}^r_{0,b}}. \tag{5.24}$$

Proof. Using Hölder's inequality we derive

$$\begin{aligned} \|uvw\|_{L^2_{xt}} &\lesssim \|u\|_{L^{3q_0}_{xt}} \|v\|_{L^{3q_1}_{xt}} \|w\|_{L^{3q_1}_{xt}} & (5.25)\\ &\lesssim \|I^{-\frac{2(j-1)}{3q_0}}u_0\|_{\widehat{L^{q_0}_x}} \|I^{-\frac{2(j-1)}{3q_1}}v_0\|_{\widehat{L^{q_1}_x}} \|I^{-\frac{2(j-1)}{3q_1}}w_0\|_{\widehat{L^{q_1}_x}}, & (5.26) \end{aligned}$$

where $q_0, q_1 > \frac{4}{3}$ are chosen such that $\frac{1}{2} = \frac{1}{3q_0} + \frac{2}{3q_1}$. Furthermore interpolating with the trilinear estimate (5.17) leads to the additional constraints $\frac{1}{r} = \frac{1-\theta}{p} + \frac{\theta}{2} = \frac{1-\theta}{p_0} + \frac{\theta}{q_0} = \frac{1-\theta}{p_1} + \frac{\theta}{q_1}$. The derivative gain on the factors is thus $s_0 = \frac{2(j-1\theta)}{3q_0}$ on the first and $s_1 = 2(j-1)(\frac{1-\theta}{2p} + \frac{\theta}{3q_1})$ on the other two, for a grand total of $s_0 + 2s_1 = \frac{2(j-1)}{r}$ as claimed. □

Remark 5.9. *It is at this point we would like to discuss the applicability of our estimates, particularly Corollary 5.8, to other problems only tangentially related to NLS-like equations. We refer to the recently published work* [15], *in which the cubic fractional Schrödinger equation (fNLS)*

$$i\partial_t u = I^\alpha u + |u|^2 u \tag{5.27}$$

was studied on both the real line and the torus[22]. *There, the local well-posedness in $H^s(\mathbb{R})$ for $\frac{2-\alpha}{4} \leq s < 0$ with $\alpha > 2$ and in $H^s(\mathbb{T})$ for the same range of regularities was established. The local solutions could be extended globally in time for the range $\frac{2-\alpha}{4} \leq s < 0$ on the line and for $\frac{2-\alpha}{6} \leq s < 0$ on the circle.*

In [15, Remark 1.12] *the question of well-posedness of* (5.27) *in Fourier-Lebesgue spaces was posed. Assuming, as is usual, the resonant interaction high × high × high → low is the culprit, our trilinear estimate from Corollary 5.8 suggests that* (5.27) *is well-posed in $\hat{H}^s_r(\mathbb{R})$ for $\frac{2-\alpha}{3r} \leq s$, $1 < r \leq 2$ and $\frac{\alpha}{2} \in \mathbb{N}_{\geq 2}$. This would already cover a big chunk of the subcritical regime up to $s_c(r) = \frac{2-\alpha r}{2r}$, where $r \to 1$.*

[22] On the torus the equation stated above (5.27) is in fact not well-behaved at negative Sobolev regularities $s < 0$. In order to achieve positive results on the circle the equation has to be renormalised to

$$i\partial_t u = I^\alpha u + \left(|u|^2 - \frac{1}{\pi}\int_{\mathbb{T}} |u|^2 \,\mathrm{d}x\right) u$$

using a gauge-transformation to eliminate a certain set of resonant interactions.

6. Well-posedness results

Now we have all our smoothing estimates together we can deal with the necessary multilinear estimates that lead to Theorems 4.1 and 4.3. We separate out the cases dealing with Fourier-Lebesgue and modulation spaces.

For both families of spaces the cubic nonlinear terms are strictly less well behaved, so dealing with them requires separate analysis. In contrast the quintic and higher-order terms are more tame and we are thus able to prove a general multilinear estimate for these.

The latter estimates, specifically Corollaries 6.5 and 6.9, we establish by multilinear interpolation between an $X_{s,b}$ (corresponding to the case $r = 2$ or equivalently $p = 2$) and an (almost) endpoint estimate in the respective class of spaces.

6.1. Multilinear estimates in $\hat{X}^r_{s,b}$ spaces.

6.1.1. *Estimates for cubic nonlinearities.*

Proposition 6.1. *Let $1 < r \le 2$, $s = \frac{j-1}{r'}$, $\alpha \in \mathbb{N}_0^3$ with $|\alpha| = 2(j-1)$ then there exist $b' > -\frac{1}{r'}$, $b > \frac{1}{r}$ and one has*

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}. \tag{6.1}$$

Proof. We divide the proof into different cases, depending on the size of the interacting frequencies.

(1) **Low frequency case** $|\xi_{max}| \le 1$: Here, using the trivial estimate suffices, since $s \ge 0$:

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \|u_1 \overline{u_2} u_3\|_{\widehat{L^r}_{xt}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\widehat{L^{3r}}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}.$$

(2) **Non-resonant interaction** $|\xi_{max}| \gg |\xi_{min}|$: If there is at least one *small* frequency then without loss of generality we may assume that $|\xi_1 + \xi_2| \gtrsim |\xi_1|$ (otherwise swap the factors u_1 and u_3). This in turn allows us to estimate $k_j^+(\xi_1, \xi_2) \gtrsim |\xi_1|^{2j-1}$ and $k_j^{+,*}(\xi_1 + \xi_2, \xi_3) \gtrsim |\xi_1|^{2j-1}$. Applied to the quantity to be estimated this gives

$$\begin{aligned}
\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} &\lesssim \|(J^{s+2(j-1)} u_1) \overline{u_2} u_3\|_{\hat{X}^r_{0,b'}} \\
&\lesssim \|I^+_{r,j}(J^{s+\frac{2j-1}{r'}-1} u_1, \overline{u_2}) u_3\|_{\hat{X}^r_{0,b'}} \\
&\lesssim \|I^{+,*}_{\rho',j}(I^+_{r,j}(J^{s+(2j-1)(\frac{1}{r'}-\frac{1}{\rho'})-1} u_1, \overline{u_2}), u_3)\|_{\hat{X}^r_{0,b'}},
\end{aligned}$$

where ρ' is to be chosen later, according to the constraints set forth in the following. First, we want to assume $(2j-1)(\frac{1}{r'} - \frac{1}{\rho'}) - 1 \le 0$, which allows us to reshuffle the derivatives and apply estimate (5.11):

$$\begin{aligned}
&\lesssim \|I^{+,*}_{\rho',j}(I^+_{r,j}(J^s u_1, \overline{u_2}), J^{(2j-1)(\frac{1}{r'}-\frac{1}{\rho'})-1} u_3)\|_{\hat{X}^r_{0,b'}} \\
&\lesssim \|I^+_{r,j}(J^s u_1, \overline{u_2})\|_{\widehat{L^r_{xt}}} \|J^{(2j-1)(\frac{1}{r'}-\frac{1}{\rho'})-1} u_3\|_{\hat{X}^{\rho'}_{0,-b'}}
\end{aligned}$$

For this to hold we must have $1 < r < \infty$, $\infty \ge \rho' \ge r'$ and $b' < -\frac{1}{\rho'}$. Now for the first factor we may apply estimate (5.9) on the condition that $b > \frac{1}{r}$ and for the second factor we use a Sobolev-embedding style estimate assuming that $b' + b > -\frac{1}{\rho'}$ and $\frac{2(j-1)}{r'} - \frac{2(j-1)}{\rho'} < s$. This is also the point where our argument breaks down for the classic cubic NLS, with $s = 0$. After choosing ρ' appropriately the proof for this case is complete.

(3) **Resonant interaction** $|\xi_1| \sim |\xi_2| \sim |\xi_3| \gtrsim 1$: Now we may utilise our trilinear smoothing estimate. As is mentioned above we do not rely on a specific frequency constellation (their signs, see [38, Section 3.2]) for its application, so choosing $s_0, s_1 \geq 0$ so that (5.24) is applicable we may directly estimate

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \|(J^{s+s_0} u_1)(J^{s+s_1} \overline{u_2})(J^{s+s_1} u_3)\|_{\widehat{L^r_{xt}}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}},$$

which concludes the proof. □

6.1.2. *Estimates for quintic and higher-order nonlinearities.* The following proposition is the $X_{s,b}$ estimate we will later interpolate with, as mentioned in the beginning of this section. Because its proof does not rely on the specific number of factors that are complex conjugates it is responsible for the remark following Theorem 4.1.

Proposition 6.2. *Let $2 \leq k \leq j$, $s > -\frac{1}{2}$, $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$. Then there exists a $b' > -\frac{1}{2}$ such that for all $b > \frac{1}{2}$ with $b' + 1 > b$ one has*

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i\|_{X_{s,b'}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{X_{s,b}}. \tag{6.2}$$

Additionally for an arbitrary subset of the factors on the left hand side these may be replaced with their complex conjugates.

Proof. Without loss of generality assume that the frequencies are sorted in descending order of magnitude i.e., $|\xi_1| \geq |\xi_2| \geq \ldots \geq |\xi_{2k+1}|$. We distinguish two cases for the magnitude of the resulting frequency $|\xi|$.

(1) $|\xi| \sim |\xi_1|$. Here we can make proper use of the $-\frac{1}{2}+$ derivatives that lie on the product. First we apply the dual form of Kato's smoothing estimate (5.2) and redistribute derivatives, introducing $\delta > 0$, in order to at a later point use the maximal function estimate (5.3). After using Hölder's inequality, we make use of (5.2) again (this time literally). Finally we apply the maximal function estimate, where the magnitude of δ ensures we had previously gained enough derivatives:

$$\begin{aligned}
\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i\|_{X_{s,b'}} &\lesssim \|J^{\frac{2j-1}{2}-}(J^{2(j-k)+s-\frac{2j-1}{2}+\delta+} u_1) \prod_{i=2}^{2k+1} J^{-\frac{\delta}{2k}} u_i\|_{X_{0,b'}} \\
&\lesssim \|(J^{2(j-k)+s-\frac{2j-1}{2}+\delta+} u_1) \prod_{i=2}^{2k+1} J^{-\frac{\delta}{2k}} u_i\|_{L_x^{1+} L_t^2} \\
&\lesssim \|J^{\frac{2j+1}{2}-2k+s+\delta+} u_1\|_{L_x^\infty L_t^2} \prod_{i=2}^{2k+1} \|J^{-\frac{\delta}{2k}} u_i\|_{L_x^{2k(1+\varepsilon)} L_t^\infty} \\
&\lesssim \|u_1\|_{X_{s,b}} \prod_{i=2}^{2k+1} \|J^{\frac{1}{2}-\frac{1}{2k(1+\varepsilon)}-\frac{\delta}{2k}} u_i\|_{X_{0,b}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{X_{s,b}}
\end{aligned}$$

This holds as long as $\delta + 1 < 2k$ and $\frac{1}{2} - \frac{1}{2k(1+\varepsilon)} - \frac{\delta}{2k} < s = -\frac{1}{2}+$, which can be achieved by choosing $\varepsilon > 0$ sufficiently small.

(2) $|\xi| \ll |\xi_1|$. In this case we argue there must be at least one factor that also has large frequency magnitude compared to ξ_1, since $|\xi|$ is small. Thus we know $|\xi_1| \sim |\xi_2|$. Though there must also be another factor with comparatively small frequency magnitude, because if all frequencies had comparable magnitude the resulting frequency ξ must also be large since we have an uneven number of

factors. Hence also $|\xi_1| \gg |\xi_{2k+1}|$. We now argue

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i\|_{X_{s,b'}} \lesssim \|(J^{j-1}u_1)(J^{-\frac{1}{2}}u_{2k+1})(J^{j-1}u_2)\prod_{i=3}^{2k} J^{-1+\frac{1}{4(k-1)}}u_i\|_{X_{s,b'}}$$

$$\lesssim \|I^{\pm}_{2,j}(J^{-\frac{1}{2}}u_1, J^{-\frac{1}{2}}u_{2k+1})(J^{j-1}u_2)\prod_{i=3}^{2k} J^{-1+\frac{1}{4(k-1)}}u_i\|_{L^{1+}_{xt}},$$

where we used the Sobolev embedding theorem and may freely make use of the bilinear operator $I^{+}_{2,j}$ since $|\xi_1 \pm \xi_{2k+1}| \sim |\xi_1|$. Next, setting $r = 2(k-1)(2+\varepsilon)$, we use Hölder's inequality

$$\lesssim \|I^{+}_{2,j}(J^{-\frac{1}{2}}u_1, J^{-\frac{1}{2}}u_{2k+1})\|_{L^2_{xt}}\|J^{j-1}u_2\|_{L^\infty_x L^{2+}_t}\prod_{i=3}^{2k}\|J^{-1+\frac{1}{4(k-1)}}u_i\|_{L^r_x L^\infty_t}$$

For the first factor we used the bilinear estimate (5.9), for the second the interpolated Kato's smoothing (5.2) and for the rest the maximal function estimate (5.3), in order to arrive at our desired bound.

For the latter estimate to lead us into the correct $X_{s,b}$-space we need

$$-1+\frac{1}{4(k-1)}+\frac{1}{2}-\frac{1}{4} = -\frac{1}{2}+\frac{1}{4(k-1)}-\frac{1}{2(k-1)(2+\varepsilon)} < s = -\frac{1}{2}+$$

which can be achieved by choosing $\varepsilon > 0$ small enough.

In both cases every factor passes through a norm that is invariant under complex conjugation, or we have the freedom to use $I^{-}_{2,j}$ over $I^{+}_{2,j}$, so fulfilling the additional claim that an arbitrary number of the factors can be complex conjugated is also dealt with. □

Unfortunately, when transitioning to Fourier-Lebesgue spaces, one loses the freedom to choose arbitrarily the number of factors in the nonlinearity that may be complex conjugates of the solution u.

Proposition 6.3. *Let $2 \le k \le j$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$. Then there exists an $r_0 > 1$ such that for all $1 < r < r_0$ and $s > \frac{j-k}{kr'}$ there exists a $b\prime > -\frac{1}{r'}$ such that for all $b > \frac{1}{r}$ with $b' + 1 > b$ one has*

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{2k+1}\|u_i\|_{\hat{X}^r_{s,b}}, \tag{6.3}$$

where exactly k of $v_1, v_2, \dots, v_{2k+1}$ are equal to the complex conjugate of u_i and otherwise just equal to u_i.

Proof. We assume, without loss of generality, that the magnitudes of the frequencies are sorted i.e., $|\xi_1| \ge |\xi_2| \ge \dots |\xi_{2k+1}|$. Distinguish cases based on the number of high-frequency factors that are present in the product:

(1) $|\xi_4| \gtrsim |\xi_1|$. So we have at least four high-frequency factors which is enough for us to make use of the Fefferman-Stein estimate (5.16). We start by choosing $r_0 > 1$ such that $s < \frac{1}{r}$. Next fix $s_1 > \frac{1}{4}(2(j-k)+s+(2k-3)(\frac{1}{r}-s))$ and $s_2 < s - \frac{1}{r} < 0$ fulfilling $4s_1 + (2k-3)s_2 = 2(j-k)+s$. Then we can estimate using the Hausdorff-Young inequality

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i\|_{\hat{X}^r_{s,b'}} \lesssim \|\prod_{i=1}^{4} J^{s_1}v_i \prod_{i=5}^{2k+1} J^{s_2}v_i\|_{\widehat{L^r_{xt}}} \lesssim \|\prod_{i=1}^{4} J^{s_1}v_i \prod_{i=5}^{2k+1} J^{s_2}v_i\|_{L^r_{xt}}$$

$$\lesssim \prod_{i=1}^{4}\|J^{s_1}u_i\|_{L^{4r}_{xt}}\prod_{i=5}^{2k+1}\|J^{s_2}u_i\|_{L^\infty_{xt}}$$

For every factor in the second product we can now use $\|f\|_{L^\infty_{xt}} \lesssim \|f\|_{\widehat{L^\infty_{xt}}}$ followed by a Sobolev style embedding, where we end up with $s_2 + \frac{1}{r} - \frac{1}{\infty}+$ space- and $\frac{1}{r}+$ time-derivatives. The first four factors can be dealt with by the diagonal case of the Fefferman-Stein inequality (5.16). So that we end up in the correct $\hat{X}^r_{s,b}$-norm we need $s > s_1 + \frac{1-2(j-1)}{4r}$, which we can achieve for every $s > \frac{j-k}{kr'}$ (by choosing s_1 near enough $\frac{1}{4}(2(j-k)+s+(2k-3)(\frac{1}{r}-s)))$ as claimed.

(2) $|\xi| \sim |\xi_1| \gg |\xi_2|$. With only a single high-frequency factor v_i we must distinguish if it is a complex conjugate or not. Without loss of generality we assume $v_1 = \overline{u_1}$ and that (since we know exactly k of the factor are complex conjugates) we are dealing with a product of the form $\overline{u_1}(\prod_{i=2}^{2k-3} v_i)u_{2k-2}u_{2k-1}u_{2k}\overline{u_{2k+1}}$ (omitting the derivatives). The arguments for the alternate cases is similar, we omit the details. Having only $|\xi_1|$ large gives us control over the symbols of $I^{-,*}_{\rho',j}$ and $I^+_{r,j}$ when applied as in

$$
\begin{aligned}
&\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i\|_{\hat{X}^r_{0,b'}}\\
&\lesssim \|I^+_{r,j}(J^{2(j-k)-\frac{2j-1}{r}}\overline{u_1}, u_{2k})\overline{u_{2k+1}} \prod_{i=2}^{2k-1} v_i\|_{\hat{X}^r_{0,b'}}\\
&\lesssim \|I^{-,*}_{\rho',j}(I^+_{r,j}(J^{2(j-k)-\frac{2j-1}{r}-\frac{2j-1}{\rho'}}\overline{u_1}, u_{2k}) \prod_{i=2}^{2k-1} v_i, \overline{u_{2k+1}})\|_{\hat{X}^r_{0,b'}}\\
&\lesssim \|I^{-,*}_{\rho',j}(I^+_{r,j}(J^{\frac{2(j-k)}{r'}-\frac{2(j-1)}{\rho'}}\overline{u_1}, u_{2k}) \prod_{i=2}^{2k-1} J^{-\frac{1}{r}-}v_i, J^{-\frac{1}{r}+\frac{1}{\rho'}-}\overline{u_{2k+1}})\|_{\hat{X}^r_{0,b'}}
\end{aligned}
$$

Now choosing $\rho \sim r$ such that $\frac{2(j-k)}{r'} - \frac{2(j-1)}{\rho'} \le 0$ we get for a $b' < -\frac{1}{\rho'}$

$$
\begin{aligned}
&\lesssim \|I^+_{r,j}(\overline{u_1}, u_{2k}) \prod_{i=2}^{2k-1} J^{-\frac{1}{r}-}v_i\|_{\widehat{L^r_{xt}}} \|J^{-\frac{1}{r}+\frac{1}{\rho'}-}u_{2k+1}\|_{\hat{X}^{\rho'}_{0,-b'}}\\
&\lesssim \|I^+_{r,j}(\overline{u_1}, u_{2k})\|_{\widehat{L^r_{xt}}} \prod_{i=2}^{2k-1} \|J^{-\frac{1}{r}-}v_i\|_{\widehat{L^\infty_{xt}}} \|J^{-\frac{1}{r}+\frac{1}{\rho'}-}u_{2k+1}\|_{\hat{X}^{\rho'}_{0,-b'}}
\end{aligned}
$$

Using a the bilinear estimate (5.9), a Sobolev style embedding and Young's inequality we arrive at the desired upper bound, at least in the case $s = 0$.

(3) $|\xi_1| \sim |\xi_2| \gg |\xi_3|$ or $|\xi_1| \sim |\xi_3| \gg |\xi_4|$.
subcase: $v_1 = u_1$ and $v_2 = u_2$. If there are two or three high-frequency factors we proceed similarly as to the case where there is only a single one, though parenthesizing differently with the bilinear operators. Here further cases can be made depending on if the high-frequency factors are complex conjugates or not, though these are remedied by using $I^-_{r,j}$ rather than $I^+_{r,j}$ and vice versa (dito for the dual operators). The arguments are very similar to the preceding cases, so we omit the details.

We proved the inequality for $s = 0$ in the latter two cases, thus it also holds for every $s \ge 0$.

□

Remark 6.4. *Let us discuss what influence the distribution of complex conjugates has on the estimate proven in Proposition 6.3. In the first case, where we have 'enough', that is four or more, high-frequency factors, whether the terms in the nonlinearity are complex conjugates or not is irrelevant. Inspecting the proof for the subsequent cases, where there are three or fewer high-frequency factors, we point* *out that $2k-2$ of the factors pass through a $\widehat{L^\infty_{xt}}$ norm and thus, if these are complex conjugates or not is irrelevant.*

Also in these cases, since u_1 is a high-frequency factor and u_{2k} has low frequency, which of the symbols of either bilinear operators $I^{\pm}_{r,j}$ we gain does not matter. Hence we are not restricted in the sense that the 'partner' of u_1 in the application of $I^{\pm}_{r,j}$ has a complex conjugate or not. (This is also independent of whether u_1 is a complex conjugate, because $I^{\pm}_{r,j}$ passed through a $\widehat{L^r_{xt}}$ norm.)

What would remain to argue is why one also has free choice to apply either of the dual bilinear operators $I^{\pm,*}_{\rho',j}$ and hence again, that if the 'partner' (u_{2k+1} in the argument given in the proof) is a complex conjugate or not, is irrelevant. This is slightly more delicate and one must vary the 'partner' in application of the dual bilinear operator between u_{2k+1} and one of the other high-frequency factors, if the total frequency of $I^{+}_{r,j}(v_1, v_{2k})\prod_{i=2}^{2k-1} v_i$ (ignoring derivatives) is small. (This product having small frequency can only happen in case there are multiple (but fewer than four) high-frequency factors.) In such a case the symbol of, say, $I^{+,*}_{\rho',j}$ is small and one can thus not fully exploit the gain in derivative this operator would offer. To remedy this one can swap out u_{2k+1} with one of the high-frequency factors besides u_1 to ensure the symbol of both bilinear operators is large again.

We deem adding such a case by case analysis to the proof of Proposition 6.3 would distract from the overall argument, so we leave working out further details to the reader.

Finally we may use multilinear interpolation to interpolate between the estimates in Propositions 6.2 and 6.3 in order to establish the corollary from which Theorem 4.1 follows.

Corollary 6.5. *Let $2 \leq k \leq j$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$. Then for $1 < r \leq 2$ and $s > -\frac{1}{r'}$ there exists a $b' > -\frac{1}{r'}$ such that for all $b > \frac{1}{r}$ we have*

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{\hat{X}^r_{s,b}}, \tag{6.4}$$

where exactly k of $v_1, v_2, \dots, v_{2k+1}$ are equal to the complex conjugate of u_i and otherwise just equal to u_i.

6.2. **Multilinear estimates in $X^p_{s,b}$ spaces.** Before we dive into the proofs that will lead to Proposition 6.7 and Corollary 6.9, which in turn imply Theorem 4.3, we would like to give the reader a run down of extra conventions we will be using when dealing with estimates of frequency localised functions. As in the previous section, we will be proving our estimates separately for different frequency constellations on a case by case basis.

Let us first mention that, even though we are in modulation spaces, we will not need the added control the associated uniform frequency localisation may give us. In particular we will only rely on this additional control in the resonant case for the cubic nonlinear term. For all other cases a more common dyadic frequency decomposition will suffice, which we may sum to arrive in the correct modulation space using, for example, (2.3).

Furthermore, in order to save vertical space and give a more compact presentation of our estimates, we will play loose with the description of the set over which we will be summing in some cases. Implicitly it is understood that we are always summing over all dyadic frequencies $N, N_1, N_2, \dots$ or integer frequencies $n, n_1, n_2, \dots$ that appear in the expression we want to estimate, subject to the restrictions implied by the case we are currently estimating. An example of the suppression of information in a sum, would be the following two sums being equivalent

$$\sum_{N_1 \gtrsim N_3} \int_{\mathbb{R}^2} u_{N_1} \overline{u_{N_2}} u_{N_3} \overline{v_N} \, \mathrm{d}x \, \mathrm{d}t = \sum_{\substack{N, N_1, N_2, N_3 \geq 1 \\ N_1 \gtrsim N_3}} \int_{\mathbb{R}^2} P_{N_1} u_1 P_{N_2} \overline{u_2} P_{N_3} u_3 P_N \overline{v} \, \mathrm{d}x \, \mathrm{d}t,$$

where additional we have made clear the convention mentioned in Section 2.2 that indices denoting frequency decomposition may suppress other indices.

We also introduce the notation ξ_{max}, ξ_{min} and N_{max}, N_{min} referring to the largest and smallest element of the sets of all frequencies $\{|\xi_i| \mid 1 \leq i \leq 2k+1\}$ and of all dyadic frequencies $\{N_i \mid 1 \leq i \leq 2k+1\}$, where $2k+1$ is the total number of factors in a nonlinear term.

One last ingredient: the following lemma will help us piece together uniform-frequency localized functions. It had previously appeared in [77, eq. (2.7)], without proof, but we include its proof here for the reader's convenience.

Lemma 6.6. *Let $(a_m)_{m\in\mathbb{Z}}$ and $(b_n)_{n\in\mathbb{Z}}$ be two sequences. Then for $1 \leq p \leq \infty$ and every $\varepsilon > 0$ one has*

$$\sum_{\substack{m,n \in \mathbb{Z} \\ m \neq n}} \frac{a_m b_n}{|m-n| \langle n \rangle^{\varepsilon}} \lesssim_{\varepsilon} \|a_m\|_{\ell^p_m(\mathbb{Z})} \|b_n\|_{\ell^{p'}_n(\mathbb{Z})}.$$

Proof. We apply Hölder's inequality and Young's convolution inequality

$$\begin{aligned}
\sum_{\substack{m,n \in \mathbb{Z} \\ m \neq n}} \frac{a_m b_n}{|m-n| \langle n \rangle^{\varepsilon}} &= \sum_{m \in \mathbb{Z}} a_m \sum_{n \in \mathbb{Z}} \frac{b_n}{\langle n \rangle^{\varepsilon}} \cdot \frac{\chi_{m \neq n}}{|m-n|} \\
\lesssim \|a_m\|_{\ell^p_m} \left\| \frac{b_\cdot}{\langle \cdot \rangle^{\varepsilon}} * \frac{\chi_{\cdot \neq 0}}{|\cdot|} \right\|_{\ell^{p'}} &\lesssim \|a_m\|_{\ell^p_m} \|b_n \langle n \rangle^{-\varepsilon}\|_{\ell^q_n} \|\chi_{n \neq 0} |n|^{-1}\|_{\ell^r_n} \\
\lesssim \|a_m\|_{\ell^p_m} \|b_n\|_{\ell^{p'}_n} \|\langle n \rangle^{-\varepsilon}\|_{\ell^{\tilde{q}}_n} &\|\chi_{n \neq 0} |n|^{-1}\|_{\ell^r_n} \lesssim_{\varepsilon} \|a_m\|_{\ell^p_m} \|b_n\|_{\ell^{p'}_n},
\end{aligned}$$

where $1 + \frac{1}{p'} = \frac{1}{q} + \frac{1}{r}$ and $\frac{1}{q} = \frac{1}{p'} + \frac{1}{\tilde{q}}$. The last inequality becomes true, if we choose $\tilde{\varepsilon} > 0$ small enough and then set $\frac{1}{r} = \frac{1}{1+\tilde{\varepsilon}}$, as well as $\frac{1}{\tilde{q}} = \frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}$. ☐

6.2.1. *Estimates for cubic nonlinearities.* In the proof of the following Proposition 6.7 we assume $s = \frac{j-1}{2}$, though because of the inequality $\langle \xi \rangle \lesssim \langle \xi_1 \rangle \langle \xi_2 \rangle \langle \xi_3 \rangle$ for $\xi = \xi_1 + \xi_2 + \xi_3$ the derived estimate also holds true for $s > \frac{j-1}{2}$.

Proposition 6.7. *Let $j \geq 2$, $2 \leq p < \infty$, $s = \frac{j-1}{2}$, $\alpha \in \mathbb{N}_0^3$ with $|\alpha| = 2(j-1)$. Then there exist $b' < 0$ and $b' + 1 > b > \frac{1}{2}$ such that one has*

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{X^p_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{X^p_{s,b}}.$$

Proof. Again, the proof is a case by case analysis of different frequency interactions. We prove the estimate in each case by duality:

(1) **Low frequency case** $|N_{max}| \lesssim 1$: In this case we may deduce that the frequency of the product N is also small. So we use Hölder's inequality, Sobolev

embeddings and (2.3) for the sum

$$\sum_{N_{max},N\lesssim 1}\int_{\mathbb{R}^2}\partial_x^{\alpha_1}u_{N_1}\partial_x^{\alpha_2}\overline{u_{N_2}}\partial_x^{\alpha_3}u_{N_3}N^s\overline{v_N}\,\mathrm{d}x\,\mathrm{d}t$$
$$\lesssim \sum_{N_{max},N\lesssim 1}N_{max}^{s+2(j-1)+1+}\|u_{N_1}\|_{L^2_{xt}}\|u_{N_2}\|_{L^\infty_t L^2_x}\|u_{N_3}\|_{L^\infty_t L^2_x}\|v_N\|_{L^2_{xt}}$$
$$\lesssim \sum_{N_{max},N\lesssim 1}N_{max}^{s+2(j-1)+1+\frac{3}{2}-\frac{3}{p}+}\|v_N\|_{X^{p'}_{0,-b'}}\prod_{i=1}^{3}\|u_{N_i}\|_{X^p_{s,b}}$$

This is a finite sum (remember, our dyadic frequencies are $N_i\in 2^{\mathbb{N}}$), so we may bound the final expression by our desired $\|v\|_{X^{p'}_{0,-b'}}\prod_{i=1}^3\|u_i\|_{X^p_{s,b}}$.

(2) **Non-/Semi-resonant interaction** $N_{max}\gg N_{min}$: Here there are two subcases to be dealt with, depending on which frequencies are of similar magnitude to N_{max}, but with opposite sign, if any. The arguments in both cases are the same (just with the roles of some of the factors interchanged), so we will only present one of the cases.

Say we have $|\xi_{max}|=|\xi_1|\gg|\xi_3|=|\xi_{min}|$. Then either $|\xi_1+\xi_2|\sim|\xi_1|$ or $|\xi_1+\xi|\sim|\xi_1|$. In the former case, both $|\xi_1+\xi_2|$ and $|\xi_3+\xi|$ are comparable to $|\xi_{max}|$ and in the latter it is both $|\xi_1+\xi|$ and $|\xi_2+\xi_3|$ that are comparable. For other choices of ξ_{max} and ξ_{min} one may argue similarly.

Observe the argument for the case with $|\xi_{max}|=|\xi_1|\gg|\xi_3|=|\xi_{min}|$ and $|\xi_1+\xi_2|\sim|\xi_1|$: first we use Hölder's inequality

$$\sum_{N_1\gg N_3}\int_{\mathbb{R}^2}\partial_x^{\alpha_1}u_{N_1}\partial_x^{\alpha_2}\overline{u_{N_2}}\partial_x^{\alpha_3}u_{N_3}N^s\overline{v_N}\,\mathrm{d}x\,\mathrm{d}t \tag{6.5}$$
$$\lesssim \sum_{N_1\gg N_3}N_{max}^{s+2(j-1)}\|u_{N_1}\overline{u_{N_2}}\|_{L^2_{xt}}\|u_{N_3}\overline{v_N}\|_{L^2_{xt}} \tag{6.6}$$

Next we would like to apply our bilinear estimate (5.9) with $q=p=2$ to both terms in the L^2 norm. Though because we are estimating by duality simply using (5.9) as-is would leave us with v_N in the wrong space $X_{0,b}$ for $b>\frac{1}{2}$. To remedy this we interpolate (5.9) with the much simpler bound

$$\|I^+_{2,j}(u,\overline{v})\|_{L^2_{xt}}\lesssim\|(J^\sigma u)(J^\sigma v)\|_{L^2_{xt}} \tag{6.7}$$
$$\lesssim\|J^\sigma u\|_{L^\infty_t L^2_x}\|J^\sigma v\|_{L^2_t L^\infty_x}\lesssim\|u\|_{X_{\sigma,\frac{1}{2}+}}\|v\|_{X_{\sigma+\frac{1}{2}+,0}}, \tag{6.8}$$

where $\sigma=2j-1$ and we used Hölder's inequality and Sobolev embeddings. Using our interpolated bound we may proceed with estimating (6.6):

$$\lesssim\sum_{N_1\gg N_3}N_{max}^{s-1}\|I^+_{2,j}(u_{N_1},\overline{u_{N_2}})\|_{L^2_{xt}}\|I^+_{2,j}(u_{N_3}\overline{v_N})\|_{L^2_{xt}}$$
$$\lesssim\sum_{N_1\gg N_3}N_{max}^{s-1+}(N_1N_2N_3)^{-s+}\|v_N\|_{X^{p'}_{0,-b'}}\prod_{i=1}^3 N_i^{0-}\|u_{N_i}\|_{X_{s,b}}$$
$$\lesssim\sum_{N_1\gg N_3}N_1^{-\frac{1}{2}-\frac{1}{p}+}(N_2N_3)^{-s+\frac{1}{2}-\frac{1}{p}+}N^{0-}\|v_N\|_{X^{p'}_{0,-b'}}\prod_{i=1}^3 N_i^{0-}\|u_{N_i}\|_{X^p_{s,b}}.$$

At this point it becomes important, that $j\neq 1$, because otherwise $s=0$ and we wouldn't be able to sum up. For $j\geq 2$ though, one has $s\geq\frac{1}{2}$ so that $N_1^{-\frac{1}{2}-\frac{1}{p}+}(N_2N_3)^{-s+\frac{1}{2}-\frac{1}{p}+}\lesssim 1$ and we can close our argument with a final application of (2.3).

(3) **Resonant interaction** $N_{max} \sim N_{min}$: Here we will have to utilize the added control modulation spaces give us with the unit cube decomposition. We distinguish between the following subcases:

1. $\forall (i,j) : |\xi_i + \xi_j| \gtrsim |\xi_i - \xi_j|$: This means that all frequencies have the same sign. Since we have separate control over the symbols $|\xi_i + \xi_j|$ and $|\xi_i - \xi_j|$ we may argue simpler than in [77]. The estimate in this subcase may be proven analogously to the non-/semi-resonant case.

2. $|\xi_1 - \xi_2| \geq |\xi_1 + \xi_2|$:

2.1. $|\xi_1 + \xi_2| \lesssim 1$ and $\min(|\xi_2 + \xi_3|, |\xi_2 - \xi_3|) \lesssim 1$: Without loss of generality we will assume $|\xi_2 - \xi_3| \lesssim 1$, the other case may be argued analogously. So here we have the following frequencies for the individual factors and their product

$$n_1 = -\ell + \mathcal{O}(1), \quad n_2 = \ell, \quad n_3 = \ell + \mathcal{O}(1), \quad n = \ell + \mathcal{O}(1)$$

for a fixed $\ell \in \mathbb{Z}$. We may restrict ourselves to proving the diagonal case, where $-n_1 = n_2 = n_3 = n = \ell$ hold exactly. This is because after having established the inequality for the diagonal case, the general case may be proven by switching to a different family of isometric decomposition operators $(\tilde{\Box}_n)_{n\in\mathbb{Z}}$ and using the inequality for the diagonal case. We omit the details.

After using Hölder's inequality we use our trilinear estimate (5.24) to bound the contribution in this case:

$$\sum_{\ell\in\mathbb{Z}} \langle \ell \rangle^{s+2(j-1)} \int_{\mathbb{R}^2} u_{-\ell}\overline{u}_\ell u_\ell \overline{v}_\ell \,\mathrm{d}x\,\mathrm{d}t \lesssim \sum_{\ell\in\mathbb{Z}} \langle \ell \rangle^{s+2(j-1)} \|u_{-\ell}\overline{u}_\ell u_\ell\|_{L^2_{xt}} \|v_{-\ell}\|_{L^2_{xt}}$$
$$\lesssim \sum_{\ell\in\mathbb{Z}} \langle \ell \rangle^{s+2(j-1)} \|u_{-\ell}\overline{u}_\ell u_\ell\|_{L^2_{xt}} \|v_{-\ell}\|_{L^2_{xt}} \lesssim \sum_{\ell\in\mathbb{Z}} \|u_{-\ell}\|^3_{X_{s,b}} \|v_{-\ell}\|_{X_{0,-b'}}$$

Using the trivial embeddings $\ell^2 \supset \ell^{p'}$ and $\ell^{3p} \supset \ell^p$ we arrive at our desired bound:

$$\lesssim \|v_{-\ell}\|_{X^{p'}_{0,-b'}} \prod_{i=1}^{3} \|u_i\|_{X^{3p}_{s,b}} \lesssim \|v\|_{X^{p'}_{0,-b'}} \prod_{i=1}^{3} \|u_i\|_{X^{p}_{s,b}}$$

2.2. $|\xi_1 + \xi_2| \lesssim 1$ and $|\xi_2 \pm \xi_3| \gg 1$: In this case we have the following frequencies:

$$n_1 = -\ell + \mathcal{O}(1), \quad n_2 = \ell, \quad n_3 = m + \mathcal{O}(1), \quad n = m + \mathcal{O}(1).$$

for fixed $\ell, m \in \mathbb{Z}$. Also we may note, that $|\ell \pm m| \gtrsim 1$, as well as $|m| \sim |\ell|$ because we are in a resonant case. By symmetry we may additionally assume $|m + \ell| \geq |m - \ell|$. Again it suffices to deal with the diagonal case, where $-n_1 = n_2 = \ell$ and $n_3 = n = m$ exactly. As usual we begin with an application of Hölder's inequality:

$$\sum_{\ell,m\in\mathbb{Z}} \langle m \rangle^{s+2(j-1)} \int_{\mathbb{R}^2} u_{-\ell}\overline{u}_\ell u_m \overline{v}_m \,\mathrm{d}x\,\mathrm{d}t \lesssim \sum_{\ell,m\in\mathbb{Z}} \langle m \rangle^{s+2(j-1)} \|u_m\overline{u}_\ell\|_{L^2_{xt}} \|u_{-\ell}\overline{v}_m\|_{L^2_{xt}}$$

Being left in a similar situation to (6.6), we argue with the same interpolated inequality (between (5.9) and (6.8)) to arrive at

$$\lesssim \sum_{\ell,m\in\mathbb{Z}} \frac{\langle m \rangle^{s+2(j-1)} \|u_m\|_{X_{0,b}} \|u_{-\ell}\|_{X_{0,b}} \|u_{-\ell}\|_{X_{0,b}} \|v_{-m}\|_{X_{0,-b'}}}{\langle m \rangle^{2j-2-} \sqrt{|m-\ell| \cdot |\ell + m|}}.$$

Here we may use $|m+\ell| \geq |m-\ell|$ and then apply Lemma 6.6, which is again reliant on the fact $s > 0$:

$$\lesssim \sum_{\ell,m\in\mathbb{Z}} \frac{1}{|m-\ell|\langle \ell \rangle^{2s-}} \|u_m\|_{X_{s,b}} \|u_{-\ell}\|_{X_{s,b}} \|u_{-\ell}\|_{X_{s,b}} \|v_{-m}\|_{X_{0,-b'}}$$
$$\lesssim \Big\| \|u_{-\ell}\|_{X_{s,b}} \|u_{-\ell}\|_{X_{s,b}} \Big\|_{\ell_\ell^{\frac{p}{2}}} \cdot \Big\| \|u_m\|_{X_{s,b}} \|v_{-m}\|_{X_{0,-b'}} \Big\|_{\ell_m^{\frac{p}{p-2}}}$$

Finally for the first factor we utilise Hölder's inequality, for the second we send $\|u_m\|_{X_{s,b}}$ to $X^\infty_{s,b}$ and then use the embeddings $\ell^\infty \supset \ell^p$ and $\ell^{\frac{p}{p-2}} \supset \ell^{p'}$ to arrive at our desired bound for this case.

2.3. $\forall i \neq j : |\xi_i \pm \xi_j| \gg 1$: This subcase starts similarly to the preceding one, where we first apply Hölder's inequality and then our interpolated bilinear estimate (between (5.9) and (6.8)) in order to place v_n in the correct space $X_{0,-b'}$.

$$\begin{aligned}
&\sum_{n_1+n_2+n_3=n} |n|^{s+2(j-1)} \int_{\mathbb{R}^2} u_{n_1}\overline{u}_{n_2}u_{n_3}\overline{v}_n \,\mathrm{d}x\,\mathrm{d}t \\
\lesssim &\sum_{n_1+n_2+n_3=n} |n|^{s+2(j-1)} \|u_{n_1}\overline{u}_{n_2}\|_{L^2_{xt}} \|u_{n_3}\overline{v}_n\|_{L^2_{xt}} \\
\lesssim &\sum_{n_1+n_2+n_3=n} \frac{|n|^{s+}\|u_{n_1}\|_{X_{0,b}}\|u_{-n_2}\|_{X_{0,b}}\|u_{n_3}\|_{X_{0,b}}\|v_{-n}\|_{X_{0,-b'}}}{\sqrt{|n_1+n_2|\cdot|n_3+n|}}
\end{aligned}$$

Now at least one of $|n_1+n_2|$ or $|n_3+n|$ is comparable to $|n|$, so assuming without loss, that $|n_3+n| \sim |n|$ we may split the factor $|n|^{s+}$ and apply Hölder's inequality:

$$\begin{aligned}
\lesssim &\sum_{n_1+n_2+n_3=n} \frac{\|u_{n_1}\|_{X_{s,b}}\|u_{-n_2}\|_{X_{s,b}}}{\sqrt{|n_1+n_2|}|n|^{\frac{1}{2}-}} \frac{\|u_{n_3}\|_{X_{s,b}}\|v_{-n}\|_{X_{0,-b'}}}{|n|^{j-1}} \\
\lesssim &\sup_{n,n_3}\left(\sum_{n_2}\langle n_2\rangle^{-1+}\|u_{n_1}\|_{X_{s,b}}\|u_{-n_2}\|_{X_{s,b}}\right)\cdot\sum_{n,n_3}\frac{\|u_{n_3}\|_{X_{s,b}}\|v_{-n}\|_{X_{0,-b'}}}{|n_3\pm n|\langle n\rangle^{0+}} \\
\lesssim &\|v\|_{X^{p'}_{0,-b'}}\prod_{i=1}^{3}\|u_i\|_{X^p_{s,b}},
\end{aligned}$$

where in the final step we used Lemma 6.6 again.

3. $|\xi_2 - \xi_3| \geq |\xi_2 + \xi_3|$: One can deal with this case in the same way as the previous with the roles of ξ_1 and ξ_3 swapped. □

6.2.2. *Estimates for quintic and higher-order nonlinearities.*

Proposition 6.8. *Let $2 \leq k \leq j$, $s > \frac{1}{4k}$, $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$. Then there exist $b' < 0$ and $b' + 1 > b > \frac{1}{2}$ such that one has*

$$\|\prod_{i=1}^{2k+1}\partial_x^{\alpha_i}u_i\|_{X^\infty_{s,b'}} \lesssim \prod_{i=1}^{2k+1}\|u_i\|_{X^\infty_{s,b}}. \tag{6.9}$$

Additionally for an arbitrary subset of the factors on the left hand side these may be replaced with their complex conjugates.

Proof. In the proof of this proposition we again assume that the frequencies of the factors in the nonlinearity are ordered in decending order $|\xi_1| \geq |\xi_2| \geq \cdots \geq |\xi_{2k+1}|$. There are essentially two cases to be dealt with, depending on if ξ_1 is cancelled out by ξ_2 or not. We estimate both cases by duality:

(1) $|\xi_1| \sim |\xi|$: Here ξ_1 is not cancelled by ξ_2, but the factor corresponding to the product v_N must thus have high frequency. The contribution from this case may be bounded by first using Hölder's inequality

$$\begin{aligned}
&\sum_{N\sim N_1}\int_{\mathbb{R}^2} N^s\overline{v}_N N_1^{2(j-k)}\prod_{i=1}^{2k+1}u_{N_i}\,\mathrm{d}x\,\mathrm{d}t \\
\lesssim &\sum_{N\sim N_1} N_1^{s+2(j-k)}\|v_N\|_{L_x^{\infty-}L_t^2}\|u_{N_1}\|_{L_x^\infty L_t^2}\prod_{i=2}^{2k+1}\|u_{N_i}\|_{L_x^{2k+}L_t^\infty}
\end{aligned}$$

Now we use Kato's inequality (5.2) for both v_N and u_{N_1} and the maximal function estimate (5.3) $2k$ times for the remaining u_{N_i}.

$$\lesssim \sum N_1^{s+1-2k+} \|v_N\|_{X_{0,-b'}} \|u_{N_1}\|_{X_{0,b}} \prod_{i=2}^{2k+1} N_i^{\frac{1}{2}-\frac{1}{2k+}+} \|u_{N_i}\|_{X_{0,b}}$$
$$\lesssim \left(\prod_{i=1}^{2k+1} \|u\|_{X^\infty_{s,b}} \right) \|v\|_{X^1_{0,-b'}} \sum N_1^{\frac{3}{2}-2k+} \prod_{i=2}^{2k+1} N_i^{1-\frac{1}{2k}-s+}.$$

Finally we make use of the embedding $\ell^2 \supset \ell^1$ and (2.3), where we lose half a derivative using the endpoint estimate. The last term is summable, since we may distribute the $2k-\frac{3}{2}-$ derivatives gain from the first factor and $1-\frac{1}{2k}-(1-\frac{3}{4k}-)-s+<0$ can be achieved for $s>\frac{1}{4k}$.

(2) $|\xi_1| \gg |\xi|$: In this case we must have $|\xi_1| \sim |\xi_2|$. To bound this case's contribution we use a Sobolev-embedding for the factor v_N and Kato's inequality (5.2) for the two high frequency factors u_{N_1} and v_N after an application of Hölder's inequality.

$$\sum_{N \ll N_1} \int_{\mathbb{R}^2} N^s \overline{v}_N N_1^{2(j-k)} \prod_{i=1}^{2k+1} u_{N_i} \,\mathrm{d}x\,\mathrm{d}t$$
$$\lesssim \sum N_1^{s+2(j-k)} \|v_N\|_{L^2_x L^{\infty-}_t} \|u_{N_1}\|_{L^\infty_x L^{2+}_t} \|u_{N_2}\|_{L^\infty_x L^{2+}_t} \prod_{i=3}^{2k+1} \|u_{N_i}\|_{L^{2(2k-1)}_x L^\infty_t}$$
$$\lesssim \sum N_1^{s+1-2k+} \|v_N\|_{X_{0,-b'}} \|u_{N_1}\|_{X_{0,b}} \|u_{N_2}\|_{X_{0,b}} \prod_{i=3}^{2k+1} N_i^{\frac{1}{2}-\frac{1}{2(2k-1)}+} \|u_{N_i}\|_{X_{0,b}}$$

For all other factors we applied the maximal function estimate (5.3). We close this case by (2.3) for the u_{N_i} and using the embedding $\ell^2 \supset \ell^1$ for the factor v_N.

$$\lesssim \left(\prod_{i=1}^{2k+1} \|u\|_{X^\infty_{s,b}} \right) \|v\|_{X^1_{0,-b'}} \sum N_1^{2-2k-s+} \prod_{i=3}^{2k+1} N_i^{1-\frac{1}{2(2k-1)}-s+}$$

The final sums converge, because for every $i=3,4,\dots,2k+1$ we have an additional gain of $\frac{2(k-1)+s}{2k-1}-$ derivatives and one can easily check that

$$1-\frac{1}{2(2k-1)}-s-\frac{2(k-1)+s}{2k-1}+<0 \iff s>\frac{1}{4k}.$$

□

Again, as in Corollary 6.5, the following corollary is derived from a multilinear interpolation between Proposition 6.2 and the endpoint estimate in Proposition 6.8 we just proved.

Corollary 6.9. *Let $2 \le k \le j$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2(j-k)$. Then for $2 \le p \le \infty$, $s > \frac{1}{4k} - \frac{2k+1}{2kp}$, and $b'+1 > b > \frac{1}{2}$ we have*

$$\| \prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i \|_{X^p_{s,b'}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{X^p_{s,b}}. \tag{6.10}$$

Additionally for an arbitrary subset of the factors on the left hand side these may be replaced with their complex conjugates.

7. Ill-posedness results on $\mathbb{R}$ and $\mathbb{T}$

After now dealing with the positive results regarding the NLS hierarchy in this paper, let us now move focus to negative results. First we will establish Theorems 4.6 and 4.7, that shows our Theorems 4.1 and 4.3 to be optimal in the framework we are using. To do so we first exhibit a family of solutions to equations of type (3.7).

Lemma 7.1. *For $j \geq 2$ let us choose*

$$\delta_0 = \sum_{n=0}^{j} (-1)^{n+1} N^{2(j-n)} \binom{2j}{2n} \quad \textit{and} \quad c_0 = \sum_{n=0}^{j-1} (-1)^n N^{2(j-n)-1} \binom{2j}{2n+1}$$

and set $u_N(x,t) = \exp(i(Nx + \delta_0 t)) \operatorname{sech}(x - c_0 t)$. Then for every $N > 0$ the function u_N is a solution of a higher-order NLS-like equation (3.10).

Before we prove this Lemma, let us note that the one-parameter family u_N of solutions will not suffice for our ill-posedness argument. Luckily, due to the scaling invariances of the equations we are looking at, we can extend this family:

Corollary 7.2. *The family of solutions in Lemma 7.1 can be extended to a two-parameter family $v_{N,\omega}$ of solutions by setting $v_{N,\omega}(x,t) = \omega u_{\frac{N}{\omega}}(\omega x, \omega^{2j} t)$.*

Proof of Lemma 7.1. To simplify notation in the forthcoming proof we will use $f = \operatorname{sech}$. Similarly to the argument in [38] we begin with calculating the time derivatives of our supposed solution:

$$i\partial_t u_N(x,t) = \exp(i(Nx + \delta_0 t))(-\delta_0 f - ic_0 f')$$

Turning to the space derivatives, a slightly more lengthy calculation yields

$$\partial_x^{2j} u_N(x,t) = (-1)^j \exp(i(Nx + \delta_0 t)) \sum_{m=0}^{j} f^{2m} \sum_{n=m}^{j} (-1)^n c_{n,m} N^{2(j-n)} \dots$$
$$\dots \left[\binom{2j}{2n} f - \frac{i}{N} \binom{2j}{2n+1} (2m+1) f' \right],$$

where we have omitted the arguments to f (which are always equal to $x - c_0 t$) and the coefficients $c_{n,m}$ are taken from the identities

$$f^{(2n)}(x) = \sum_{m=0}^{n} c_{n,m} f^{2m+1}(x) \text{ and } f^{2n+1}(x) = \sum_{m=0}^{n} c_{n,m}(2m+1) f^{2m} f'. \tag{7.1}$$

Of these coefficients we will only need to know the exact value $c_{n,0} = 1$. One may easily derive these identities from the well-known fact $f'^2 = f^2 - f^4$ and $f'' = f - 2f^3$.

Now the parameters δ_0 and c_0 were chosen specifically such that the linear part of the equation (3.10) would vanish, so

$$(i\partial_t + (-1)^{j+1} \partial_x^{2j}) u_N(x,t) = \exp(i(Nx + \delta_0 t)) \left(-\sum_{m=1}^{j} f^{2m} \Sigma_m \right), \tag{7.2}$$

where we set

$$\Sigma_m = \sum_{n=m}^{j} (-1)^n c_{n,m} N^{2(j-n)} \left[\binom{2j}{2n} f - \frac{i}{N} \binom{2j}{2n+1} (2m+1) f' \right]$$

for readability.

What is left to argue now, is that the right-hand side of (7.2) can in fact be expressed by inserting our supposed solution u_N into a nonlinear term, that is part of the family described by (3.10).

Though this can be achieved by the same argument that is used at the end of the proof of [38, Lemma 8]. We merely give the two tables of (nonlinear) terms appearing in the double sum (7.2). The rest of the details are left to the reader.

In (7.2) one may notice "that the last term is missing", i. e. there are only $2(j-m)+1$ terms per line, for a total of j^2 in the whole table (as opposed to $(j+1)^2$ terms in [38]):

$$\begin{array}{lcccccc}
 & n=1 & n=2 & n=3 & n=4 & \cdots & n=j \\
m=1 & N^{2(j-1)}f^3 & N^{2(j-1)-1}f^2f' & N^{2(j-2)}f^3 & N^{2(j-2)-1}f^2f' & \cdots & f^3 \\
m=2 & & & N^{2(j-2)}f^5 & N^{2(j-2)-1}f^4f' & \cdots & f^5 \\
\vdots & & & & & \ddots & \vdots \\
m=j & & & & & & f^{2m+1}
\end{array}$$

Finally the nonlinear terms of the resulting equation that u_N will solve is given:

$$\begin{array}{cccccc}
|u|^2\partial_x^{2(j-1)}u & (\partial_x|u|^2)\partial_x^{2(j-1)-1}u & (\partial_x^2|u|^2)\partial_x^{2(j-2)}u & (\partial_x^3|u|^2)\partial_x^{2(j-2)-1}u & \cdots & (\partial_x^{2(j-1)}|u|^2)u \\
 & & |u|^4\partial_x^{2(j-2)}u & (\partial_x|u|^4)\partial_x^{2(j-2)-1}u & \cdots & (\partial_x^{2(j-2)}|u|^4)u \\
 & & & & \ddots & \vdots \\
 & & & & & |u|^{2j}u
\end{array}$$

Note that these align with the expectation of the equation u_N solves belonging to the family described in (3.10). □

Now with knowledge of our family of solutions from Corollary 7.2 we may reuse an argument given in [38, Proposition 1], based upon [60], in order to prove Theorem 4.6.

Proof of Theorem 4.6. The same argument as given in [38, Proposition 1] works here, just that one has to modify the choices made at the start of the proof. We choose $N_1, N_2 \sim N$ but fulfilling $|N_1 - N_2| = \frac{C}{T}N^{sr'-2(j-1)}$ for a constant $C > 0$. (We keep $N \to \infty$ and $\omega = N^{-sr'}$.)

When checking the details the astute reader should note, that we have the bound $-\frac{1}{r'} < s < \frac{j-1}{r'}$ on the regularity of the data and the propagation speed of a solution is of the order of N_k^{2j-1} (instead of N_k^{2j}), for $k = 1, 2$. □

Though our family of solutions is not just useful for proving ill-posedness in Fourier-Lebesgue spaces. We may reuse it again for the proof of Theorem 4.7. We adapt an argument from [77, Lemma 4.1], which is also based on [60], to our situation.

Proof of Theorem 4.7. The proof of this theorem is similar in spirit to that of Theorem 4.6, only that one has to be more careful in estimating the difference of solutions at a time $T > 0$. This is due to the fact, that the argument relies on the separation of (essential) support of two solutions in physical space, but this "conflicts" with the isometric decomposition used in the definition of modulation spaces.

Let us begin by stating some parameter choices that we will use down the line. Since $s < \frac{j-1}{2}$ we can fix a $\theta > 0$ such that $4s - 2(j-1) + 2\theta < 0$. Let $N \gg 1$ and $N_1, N_2 \sim N$ but fulfilling the separation condition $|N_1 - N_2| = \frac{C}{T}N^{2s-2(j-1)+2\theta}$ for a positive time $T > 0$ and constant $C > 0$. Finally let $\omega = N^{-2s}$. Later we will look at the limiting behaviour $N \to \infty$.

The next step is establishing bounds on our family of solutions in modulation spaces. We reuse the same arguments as in [77, eqns. (4.7) through (4.10)] establishing $\|v_{N_k,\omega}(\cdot,t)\|_{M^s_{2,p}} \sim 1$ uniformly in $t \in \mathbb{R}$ and $N, N_1, N_2 \geq 1$.

For the bound on the difference of solutions at time $t = 0$, we may use the embedding $M^s_{s,p} \supset H^s$ and [60, eqn. (3.5)] to estimate

$$\|v_{N_1,\omega}(\cdot,0) - v_{N_2,\omega}(\cdot,0)\|_{M^s_{2,p}} \lesssim N^{2s}|N_1 - N_2| \sim T^{-1}N^{4s-2(j-1)+2\theta},$$

which converges to zero, for $N \to \infty$.

Next up is bounding the difference of solutions at a positive time $T > 0$. This is the point where an extra argument is necessary in the modulation space setting. One resorts to looking at frequency contributions to the norm in the vicinity of N; in $|\xi - N| \ll N^\theta$ to be precise.

Noting our increased propagation speed of the solutions, we may argue analogously to [77, eqn. (4.12)] and establish

$$|\langle \Box_n v_{N_1,\omega}(\cdot,T), \Box_n v_{N_2,\omega}(\cdot,T)\rangle| \lesssim \frac{1}{N^{2(j-1)}|N_1 - N_2|T} \lesssim T^{-1}N^{-2s-2\theta}, \tag{7.3}$$

which we now utilise in said bound on the difference of solutions at $T > 0$. Following along the lines of [77, eqn. (4.14)], but using our new bound (7.3), we may establish

$$\|v_{N_1,\omega}(\cdot,T) - v_{N_2,\omega}(\cdot,T)\|_{M^s_{2,p}} \gtrsim 1 - T^{-1}N^{\frac{2}{p}\theta+2s}N^{-2\theta-2s} = 1 - T^{-1}N^{-2\theta\frac{1}{p'}}.$$

Letting $N \to \infty$ we have thus established the theorem. □

As mentioned above in the discussion of results in the introduction, the equations leading to ill-posedness on $\mathbb{R}$ are not in general the NLS hierarchy equations. This is of course reflected in the statement of Theorem 4.6.

For the interested reader though we give the family of fourth-order equations ($j = 2$) for which a solution was constructed in Lemma 7.1. Let $\lambda \in \mathbb{R}$, then the solution for $j = 2$ that was constructed in Lemma 7.1 solves the equations

$$\begin{aligned} i\partial_t u - \partial_x^4 u =& \lambda|u|^2\partial_x^2 u + (44 - 3\lambda)u^2\partial_x^2\overline{u} + (6\lambda - 80)|\partial_x^2 u|u \\ &+ (56 - 4\lambda)(\partial_x u)^2\overline{u} + (40 - 2\lambda)|u|^4 u. \end{aligned} \tag{7.4}$$

Next we may deal with the Propositions leading to forms of ill-posedness on the torus $\mathbb{T}$, i.e. Theorems 4.8 and 4.9.

Proposition 7.3. *The flow $S : \hat{H}^s_r(\mathbb{T}) \times (-T,T) \to \hat{H}^s_r(\mathbb{T})$ of the fourth-order equation ($j = 2$) in the NLS hierarchy*

$$iu_t - \partial_x^4 u = -2u^2\partial_x^2\overline{u} - 8|u|^2\partial_x^2 u - 4|\partial_x u|^2 u - 6(\partial_x u)^2\overline{u} + 6|u|^4 u$$

cannot be C^3 for any $1 \le r \le \infty$ and $s \in \mathbb{R}$.

Proof. Following an argument by Bourgain [14], assume the flow is indeed thrice continuously differentiable. For a datum $u_0(x) = \delta\phi(x)$, where $\delta > 0$ and $\phi \in H^s(\mathbb{T})$ for any $s \in \mathbb{R}$ are to be chosen later, we will evaluate the third derivative of the flow at the origin. So let u denote the corresponding solution to u_0, then

$$\left.\frac{\partial^3 u}{\partial\delta^3}\right|_{\delta=0} \sim \int_0^t U(t-t')N_3(U(t')u_0)dt',$$

where we have used the notation $N_3(u)$ to denote solely the cubic terms of the nonlinearity and $U(t)$ the linear propagator of the equation.

We may now write the integrand as its Fourier series to arrive at

$$= \int_0^t \sum_{\substack{k\in\mathbb{Z} \\ k_1+k_2+k_3=k}} e^{ikx}e^{i(t-t')k^4}e^{it'(k_1^4-k_2^4+k_3^4)}\hat{\phi}(k_1)\overline{\hat{\phi}(-k_2)}\hat{\phi}(k_3)n_3(k_1,k_2,k_3)dt'$$

$$= \sum_{\substack{k\in\mathbb{Z} \\ k_1+k_2+k_3=k}} e^{ikx+itk^4}\hat{\phi}(k_1)\overline{\hat{\phi}(-k_2)}\hat{\phi}(k_3)n_3(k_1,k_2,k_3)\int_0^t e^{-it'(k^4-k_1^4+k_2^4-k_3^4)}dt'.$$

Here $n_3(k_1,k_2,k_3) = (k_1+k_2)^2 + \frac{3}{2}(k_1+k_3)^2$ is the symbol corresponding to the terms in N_3. We may now choose $\hat{\phi}(k) = k^{-s}(\delta_{k,N} + \delta_{k,N_0})$, where $N_0 \ll N$. The choice of the N_0 parameter is not important as long as it is, say, fixed. For all further calculations the reader may assume $N_0 = 1$. We then observe $\|\phi\|_{\hat{H}^s_r} \sim 1$ independent of the two parameters.

Inserting this into the above expression we note that it suffices to look at the terms that produce a resulting frequency of $k = N$. There are three such choices for the tuple (k_1,k_2,k_3), namely $(N,-N,N)$, $(N,-N_0,N_0)$ and $(N_0,-N_0,N)$. Note that for each of these three choices the resonance relation $k^4 - k_1^4 + k_2^4 - k_3^4$ cancels and the integral in the formula above is equal to t and the symbol of our nonlinearity has size on the order of N^2.

These frequency choices thus produce Fourier coefficients (at frequency N) on the order of tN^{2-3s} (for the first one) and tN^{2-s} (for the second and third). The remaining five frequency constellations cannot cancel these contributions as they are of lower order in N.

This leaves us with the following lower bound for the Sobolev norm of the operator that is the derivative of the flow:

$$\left\| \frac{\partial^3 u}{\partial \delta^3}\Big|_{\delta=0} \right\|_{\hat{H}^s_r}^{r'} \gtrsim N^{sr'} \cdot t^{r'} N^{(2-s)r'}(1 + N^{-2sr'}) \geq t^{r'} N^{2r'}$$

for $1 < r \leq \infty$. If $r = 1$ we still have a lower bound of tN^2 though with a simpler argument. Letting $N \to \infty$ we can now see, that the flow cannot be C^3 for any $s \in \mathbb{R}$. □

The previous proposition shows that an approach with (just) a fixed-point theorem to prove well-posedness must fail at any regularity in $\hat{H}^s_r(\mathbb{T})$. As is stated in Theorem 4.9 the situation is much more dire at lower regularities. Its proof lies in the following proposition.

Proposition 7.4. *Let $j \in \mathbb{N}$, $1 \leq r \leq \infty$ and $s < j-1$. The flow $S : \hat{H}^s_r(\mathbb{T}) \times (-T,T) \to \hat{H}^s_r(\mathbb{T})$ of the Cauchy problem*

$$i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = |u|^2 \partial_x^{2j-2} u \quad \text{with} \quad u(t=0) = u_0 \in \hat{H}^s_r(\mathbb{T}), \tag{7.5}$$

cannot be uniformly continuous on bounded sets.

We want to point out, that equation (7.5) is in fact a higher-order NLS-like equation according to (3.10). More so it even fits the structure of an NLS hierarchy equation (3.8), though it is unlikely to be one because of its simple nonlinearity.

Proof of Proposition 7.4. We follow a similar argument to the one used in, for example, [75, Appendix A.2].

The reader may verify that our equation (7.5) has the two-parameter family of solutions

$$u_{N,a}(x,t) = N^{-s} a \exp(i(Nx - N^{2j}t + N^{2j-2-2s}|a|^2 t)).$$

We fix $a \in \mathbb{R}$ at two different values and will only deal with the two solution families $u_n(x,t) = u_{N_n,1}(x,t)$ and $\tilde{u}_n(x,t) = u_{N_n,1+\frac{1}{n}}(x,t)$ depending on $n \in \mathbb{N}$. N_n will be chosen later. We find that

$$\|u_n(\cdot,0)\|_{\hat{H}^s_r}, \|\tilde{u}_n(\cdot,0)\|_{\hat{H}^s_r} \lesssim 1 \qquad \text{and} \qquad \|u_n(\cdot,0) - \tilde{u}_n(\cdot,0)\|_{\hat{H}^s_r} \sim \frac{1}{n},$$

where the implicit constant is independent of $n \in \mathbb{N}$. Now choosing

$$t_n = \frac{\pi N_n^{2s+2-2j}}{(1+\frac{1}{n})^2 - 1}$$

and N_n large enough, such that $t_n \leq \frac{1}{n}$, we may then observe that

$$\|u_n(\cdot, t_n) - \tilde{u}_n(\cdot, t_n)\|_{\hat{H}^s_r} = \left|\exp(iN_n^{2j-2-2s}(1 - (1 + \frac{1}{n})^2)t_n) - (1 + \frac{1}{n})\right| = 2 + \frac{1}{n}.$$

Letting $n \to \infty$ this shows that the flow is not uniformly continuous. Such a choice is possible, if $2s + 2 - 2j < 0$ or equivalently $s < j - 1$ as stated. □

Appendix A. The first few NLS hierarchy equations

For the reader's convenience and future reference we will list the first few conserved quantities I_k derived from (3.5) and their associated nonlinear evolution equations (3.7) in terms of the potentials q and r. In this form both the focusing and defocusing variants of the (NLS) hierarchy can be derived by the identifications $r = +\overline{q}$ or $r = -\overline{q}$ respectively.

Though we will not just give the even numbered equations, corresponding to the NLS hierarchy, but also those corresponding to the mKdV hierarchy. Using the identification $r = q$ one arrives at the real mKdV hierarchy discussed in [38]. Deriving *a* complex mKdV hierarchy (of which again there is a defocusing and focusing variant) is also possible (again using the identifications $r = \pm\overline{q}$). But there are two problems:

(1) Identifying $r = \pm\overline{q}$ for the equation induced by I_4, see (7.6), does not lead to the well known form of the complex mKdV equation given in (2.2). Rather the nonlinearity is replaced by $\pm 6|u|^2\partial_x u$, up to a choice of α_3. For our local well-posedness theory this does not make a difference, as we are able to estimate both nonlinearities equally well. Though for a treatment relying more on the structure of the equation (e.g. for cancellation properties) this may be a relevant difference.

    When looking at the real mKdV hierarchy, i.e. using $r = q$, this problem does not present itself.

(2) If one wishes to use the identification $r = -\overline{q}$ the compatibility condition for the coefficients α_{2j+1} reads $\alpha_{2j+1} = -\overline{(\alpha_{2j+1})}$, as in (3.6). Meaning α_{2j+1} is imaginary[23] and thus introducing a factor i that is usually not present in complex mKdV-like equations.

    Again, looking at the real mKdV hierarchy this is a non-issue, see also [5, Section 3.2.2].

Not choosing an identification $r = \pm\overline{q}$ or $r = q$ also has the advantage, that we may derive the equations in the KdV hierarchy by setting $r = -1$, see [5, Section 3.2.1]

Finally we note that our conserved quantities may differ from those given elsewhere in the literature, as these are only determined up to (repeated) partial integration and simplification. The equations though only differ up to a choice of α_k.

A similar listing is given in [69, Appendix C] and [64, Appendix C].

(1) $n = 1, 2$. Phase shifts & Group of translations

$$I_1 = -\frac{1}{2i}\int qr\,\mathrm{d}x \qquad \text{and} \qquad I_2 = -\left(\frac{1}{2i}\right)^2\int qr_x\,\mathrm{d}x$$

$$q_t = 2\alpha_0 q \qquad \text{and} \qquad q_t = i\alpha_1 q_x$$

[23] It is non-zero, as otherwise this would lead to a trivial equation.

(2) $n = 3$. cubic nonlinear Schrödinger equation

$$I_3 = \left(\frac{1}{2i}\right)^3 \int q_x r_x + q^2 r^2 \,\mathrm{d}x$$
$$q_t = \frac{\alpha_2}{2}(-q_{xx} + 2q^2 r)$$

(3) $n = 4$. modified Korteweg-de-Vries equation

$$I_4 = \left(\frac{1}{2i}\right)^4 \int q_x r_{xx} + q q_x r^2 + 4q^2 r r_x \,\mathrm{d}x$$
$$q_t = \frac{-\alpha_3}{4}(q_{xxx} - 6qq_x r) \tag{7.6}$$

(4) $n = 5$. fourth order NLS hierarchy equation

$$I_5 = \left(\frac{1}{2i}\right)^5 \int -q_{xx} r_{xx} + q_{xx} r^2 + 6qq_x r r_x + 5q^2 r_x^2 + 6q^2 r r_{xx} - 2q^3 r^3 \,\mathrm{d}x$$
$$q_t = \frac{-\alpha_4}{8}(-q_{xxxx} + 8qq_{xx} r + 2q^2 r_{xx} + 4qq_x r_x + 6q_x^2 r - 6q^3 r^2)$$

(5) $n = 6$. fifth order mKdV hierarchy equation

$$\begin{aligned} I_6 = \left(\frac{1}{2i}\right)^6 \int & - q r_{xxxxxx} + q q_{xxx} r^2 + 8qq_{xx} r r_x + 11 qq_x r_x^2 + 12 qq_x r r_{xx} \\ & + 18 q^2 r_x r_{xx} + 8q^2 r r_{xxx} - 6q^2 q_x r^3 - 16 q^3 r^2 r_x \,\mathrm{d}x \end{aligned}$$

$$\begin{aligned} q_t = \frac{i\alpha_5}{2^4}(& q_{xxxxx} - 10qq_{xxx} r - 10qq_{xx} r_x - 10qq_x r_{xx} - 20q_x q_{xx} r - 10q_x^2 r_x \\ & + 30q^2 q_x r^2) \end{aligned}$$

(6) $n = 7$. sixth order NLS hierarchy equation

$$\begin{aligned} I_7 = \left(\frac{1}{2i}\right)^7 \int & - q r_{xxxxxxx} + q q_{xxxxx} r^2 + 10 qq_{xxx} r r_x + 19 qq_{xx} r_x^2 \\ & + 52 qq_x r_x r_{xx} + 20 qq_{xx} r r_{xx} + 20 qq_x r r_{xxx} + 19 q^2 r_{xx}^2 \\ & + 28 q^2 r_x r_{xxx} + 10 q^2 r r_{xxxx} + 5q^4 r^4 \\ & - 6qq_x^2 r^3 - 8q^2 q_{xx} r^3 - 64 q^2 q_x r^2 r_x - 50 q^3 r r_x^2 - 30 q^3 r^2 r_{xx} \,\mathrm{d}x \end{aligned}$$

$$\begin{aligned} q_t = \frac{\alpha_6}{2^5}(& -q_{xxxxxx} + 12 qq_{xxxx} r + 2q^2 r_{xxxx} + 18 qq_{xxx} r_x + 22 qq_{xx} r_{xx} + 8qq_x r_{xxx} \\ & + 30 q_x q_{xxx} r + 20 q_x^2 r_{xx} + 20 q_{xx}^2 r + 50 q_x q_{xx} r_x + 20 q^4 r^3 \\ & - 20 q^3 r r_{xx} - 50 q^2 q_{xx} r^2 - 10 q^3 r_x^2 - 60 q^2 q_x r r_x - 70 qq_x^2 r^2) \end{aligned}$$

Data availability statement. No data was used for the research described in this article.

Conflict of interest statement. The author declares there to be no conflict of interest associated with this article.

Part 2. Well-posedness of the dNLS hierarchy

The following part of this thesis is an independent paper written by the author, that has been submitted for publication to the Journal of Fourier Analysis and Applications. We reproduce it here as it appears on a pre-print server, see [3], with the difference that its bibliography is included in the overall bibliography of this thesis. Its abstract reads:

We prove well-posedness for higher-order equations in the so-called dNLS hierarchy (also known as part of the Kaup-Newell hierarchy) in almost critical Fourier-Lebesgue and in modulation spaces. Leaning in on estimates proven by the author in a previous instalment [2], where a similar well-posedness theory was developed for the equations of the NLS hierarchy, we show the jth equation in the dNLS hierarchy is locally well-posed for initial data in $\hat{H}^s_r(\mathbb{R})$ for $s \geq \frac{1}{2} + \frac{j-1}{r'}$ and $1 < r \leq 2$ and also in $M^s_{2,p}(\mathbb{R})$ for $s \geq \frac{j}{2}$ and $2 \leq p < \infty$. Supplementing our results with corresponding ill-posedness results in Fourier-Lebesgue and modulation spaces shows optimality.

Our arguments are based on the Fourier restriction norm method in Bourgain spaces adapted to our data spaces and the gauge-transformation commonly associated with the dNLS equation. For the latter we establish bi-Lipschitz continuity between appropriate modulation spaces and that even for higher-order equations 'bad' cubic nonlinear terms are lifted from the equation.

8. Introduction

The derivative nonlinear Schrödinger (dNLS) equation

$$\begin{cases} i\partial_t u + \partial_x^2 u = i\partial_x(|u|^2 u) \\ u(t=0) = u_0 \end{cases} \tag{8.1}$$

with initial data u_0, is a canonical object of study in the field of well-posedness theory for dispersive PDE. It arises as a model in various branches of physics, ranging from the propagation of circularly polarized Alfvén waves in magnetized plasma to the propagation of ultra-short pulses in optical fibers. We direct the interested reader to [4, 6, 72, 91] for an overview of its origins.

Its analysis, in the sense of low-regularity well-posedness, compared with its closely related cousin, the (de)focusing cubic nonlinear Schrödinger (NLS) equation

$$i\partial_t u + \partial_x^2 u = \pm 2|u|^2 u, \tag{8.2}$$

is considered to be strictly more difficult, because of the additional derivative in the nonlinearity. In particular, one of the nonlinear terms $|u|^2\partial_x u$ in (8.1) is much less well behaved than the remaining term $u^2\partial_x\overline{u}$.

One way to absolve the equation of this 'issue' and still be able to achieve well-posedness within the framework of the Fourier restriction norm method, or more generally by fixed-point arguments, is by utilising the gauge-transformation

$$u(x,t) \mapsto v(x,t) \coloneqq \exp\left(-i\int_{-\infty}^{x} |u(y,t)|^2 \,\mathrm{d}y\right) u(x,t) \tag{8.3}$$

which removes the ill-behaved $|u|^2\partial_x u$ by translating (8.1) to the equation

$$i\partial_t v + \partial_x^2 v = -iv^2\partial_x\overline{v} - \frac{1}{2}|v|^4 v \tag{8.4}$$

for an unknown function v. (The initial value is also adapted in an appropriate fashion.) The continuity properties of the gauge-transformation then ensure essentially[24] the equivalence of Cauchy problems associated with both (8.1) and (8.4). See [37, 42, 45, 86] and the references therein, where this approach has successfully been applied in a variety of function spaces.

Though even after transformation, solely using energy or smoothing estimates does not suffice to prove (near optimal) local well-posedness results. As was layed out in [37], for certain frequency constellations one is forced to exploit the resonance relation to eke out a fraction (in the L^2-based setting) of a derivative in order to close a contraction argument. So there is certainly some added complexity when dealing with the dNLS equation in comparison to the NLS equation.

Furthermore, the dNLS equation is a completely integrable system, which entails but is not limited to possessing an infinite hierarchy of conserved quantities and being induced by (one of) the first of these quantities. Subsequent equations may be induced in a similar fashion to produce what we refer to in the title of this paper as *the dNLS hierarchy*. As the NLS equation is also completely integrable, one can analogously look at an NLS hierarchy. (How these conserved quantities are derived for dNLS and what is meant by 'induce' will be made more precise in Section 9.)

Grounded in the recently published paper [2] by the author, in which the well-posedness theory of the NLS hierarchy is studied, the natural question arises what a similar theory would look like for the dNLS hierarchy, keeping in mind its added complexities?

Goal of the present paper is to (at least partially) answer this question. More precisely we will be proving low-regularity well-posedness results for a general class of PDE, encompassing all equations in the dNLS hierarchy, in classical Sobolev spaces $H^s(\mathbb{R})$, Fourier-Lebesgue spaces $\hat{H}^s_r(\mathbb{R})$ (sometimes written as $\mathcal{F}L^{s,r'}(\mathbb{R})$ in the literature) and modulation spaces $M^s_{2,p}(\mathbb{R})$ defined by the norms

$$\|u\|_{\hat{H}^s_r} = \|u\|_{\mathcal{F}L^{s,r'}} = \|\langle \xi \rangle^s \hat{u}\|_{L^{r'}} \quad \text{and} \quad \|u\|_{M^s_{2,p}} = \|\langle n \rangle^s \|\square_n u\|_{L^2}\|_{\ell^p_n(\mathbb{Z})} \tag{8.5}$$

respectively, with a family of isometric decomposition operators $(\square_n)_{n\in\mathbb{Z}}$. We refer to the author's previous work [2, Section 1.2] for precise definitions and an overview of properties, i.e. embeddings, interpolation and duality theory of these function spaces.

While of course we embrace the integrability structure of the dNLS hierarchy equations for their derivation, we will not be making use of it for proving our well-posedness results. Rather we welcome the fact that our techniques enable us to prove well-posedness for a much larger class of PDE (that nevertheless includes the dNLS hierarchy equations), due to their robustness towards changes in the PDE that lead to them no longer being completely integrable.

The techniques we will be using to argue well-posedness are the Fourier restriction norm method in appropriate Bourgain spaces $X_{s,b}$ adapted to our data spaces, together with bilinear refinements of Strichartz estimates. We will also be heavily leaning in on the estimates proven for the NLS hierarchy equations in [2] by the author and general smoothing estimates of Kato type. As a convenience we recall all necessary estimates in Section 11.

8.1. **Notation and function space properties.** As the present paper may be viewed as a continuation or extension of the author's previous work on the NLS hierarchy, we will refrain from (re)defining our notational conventions and instead refer the reader to [2, Section 1.2] for reference on such matters.

[24]Using the gauge-transformation muddies the uniqueness properties of the solution. See Remark 10.4 where this issue is further discussed.

In addition, we will be using some estimates for modulation spaces not yet given in [2] so we will use this opportunity to cite these from the literature. Of particular use will be a Sobolev-type embedding adapted to modulation spaces, a proof of which may be found in [18, Prop. 2.31]: Let $s_1, s_2 \in \mathbb{R}$ and $1 \le p, q_1, q_2 \le \infty$ then

$$\|f\|_{M^{s_1}_{p,q_1}(\mathbb{R}^n)} \lesssim \|f\|_{M^{s_2}_{p,q_2}(\mathbb{R}^n)} \quad \text{if and only if} \quad s_1 - s_2 > \frac{n}{q_2} - \frac{n}{q_1} > 0. \tag{8.6}$$

The other estimates we will be needing are all with regard to multiplication of modulation space functions. We start by mentioning the well known fact that $M_{\infty,1}$ is a Banach-Algebra, see [18, Prop. 4.2]. In fact, as is also mentioned after that Proposition, since $M^s_{p,q}(\mathbb{R}^n)$ continuously embeds into $M_{\infty,1}(\mathbb{R}^n)$, if $q = 1$ and $s \ge 0$, or if $q > 1$ and $s > \frac{n}{q'}$, we know $M^s_{p,q}(\mathbb{R}^n)$ also to be an algebra in those cases.

More generally we have a form of generalised Leibniz rule for modulation spaces: Let $s \ge 0$ and $1 \le p, p_1, p_2, \tilde{p}_q, \tilde{p}_2, q, q_1, q_2, \tilde{q}_1, \tilde{q}_2 \le \infty$, such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{\tilde{p}_1} + \frac{1}{\tilde{p}_2}$ and $\frac{1}{q'} = \frac{1}{q_1'} + \frac{1}{q_2'} = \frac{1}{\tilde{q}_1'} + \frac{1}{\tilde{q}_2'}$, then

$$\|fg\|_{M^s_{p,q}(\mathbb{R}^n)} \lesssim \|f\|_{M^s_{p_1,q_1}(\mathbb{R}^n)} \|g\|_{M_{p_2,q_2}(\mathbb{R}^n)} + \|f\|_{M_{\tilde{p}_1,\tilde{q}_1}(\mathbb{R}^n)} \|g\|_{M^s_{\tilde{p}_2,\tilde{q}_2}(\mathbb{R}^n)}. \tag{8.7}$$

Taking the uniform-decomposition definition of modulation spaces as known, as simple proof is as follows: We rewrite $\Box_m(fg)$ as $\sum_{k+\ell=m} (\Box_k f)(\Box_\ell g)$ using knowledge of the support of convolutions.

$$\begin{aligned}
\|fg\|_{M^s_{p,q}} &= \|\|\langle m\rangle^s \Box_m(fg)\|_{L^p}\|_{\ell^q_m(\mathbb{Z})} \lesssim \|\|\langle m\rangle^s \sum_{k+\ell=m} (\Box_k f)(\Box_\ell g)\|_{L^p}\|_{\ell^q_m(\mathbb{Z})} \\
&\lesssim \| \sum_{k+\ell=m} (\langle k\rangle^s + \langle \ell\rangle^s) \|(\Box_k f)(\Box_\ell g)\|_{L^p}\|_{\ell^q_m(\mathbb{Z})}
\end{aligned}$$

After applying the triangle inequality $\langle m\rangle^s \lesssim \langle k\rangle^s + \langle \ell\rangle^s$ we use Hölder's inequality depending on which weight is present. Finishing the proof with applications of the triangle and Young's inequality we arrive at the desired upper bound.

$$\begin{aligned}
&\lesssim \| \sum_{k+\ell=m} \langle k\rangle^s \|\Box_k f\|_{L^{p_1}} \|\Box_\ell g\|_{L^p_2}\|_{\ell^q_m(\mathbb{Z})} + \|\langle \ell\rangle^s \|\Box_k f\|_{L^{\tilde{p}_1}} \|\Box_\ell g\|_{L^{\tilde{p}_2}}\|_{\ell^q_m(\mathbb{Z})} \\
&\lesssim \|f\|_{M^s_{p_1,q_1}} \|g\|_{M^0_{p_2,q_2}} + \|f\|_{M^0_{\tilde{p}_1,\tilde{q}_1}} \|g\|_{M^s_{\tilde{p}_2,\tilde{q}_2}}.
\end{aligned}$$

For further properties of modulation spaces we recommend consulting [8, 18].

In addition we will be using the classic Gagliardo-Nirenberg inequality in deriving a-priori bounds for the dNLS hierarchy equations. We take advantage of the phrasing from [30]: Let $1 \le r, p, q \le \infty$, $\ell \in \mathbb{N}_0$, $k \in \mathbb{N}$ and $\frac{\ell}{k} \le \theta \le 1$ such that

$$\frac{1}{r} - \frac{\ell}{n} = \theta \left(\frac{1}{p} - \frac{k}{n} \right) + (1-\theta) \frac{1}{q} \tag{8.8}$$

holds. Then one has the inequality

$$\|\nabla^\ell f\|_{L^r(\mathbb{R}^n)} \lesssim \|\nabla^k f\|^\theta_{L^p(\mathbb{R}^n)} \|f\|^{1-\theta}_{L^q(\mathbb{R}^n)} \tag{8.9}$$

under the additional constraints that $\theta < 1$ if $r = \infty$ and $1 < p < \infty$; or f is vanishing at infinity if $q = \infty$, $k < \frac{n}{p}$ and $\ell = 0$.

8.2. **Organisation of the paper.** In Section 9 we will be deriving and defining what is referred to in the title of this paper as the dNLS hierarchy. We will also review what is known about the gauge-transformation associated with the dNLS equation. In addition we will prove its continuity as a map between appropriate modulation spaces and argue that applied to the higher-order dNLS hierarchy equations it also leads to more well-behaved models. We will be referring to these more well-behaved models as gauged dNLS equations and make reference to them in our well-posedness theorems.

Then in Section 10 we quickly review prior work associated with (higher-order) dNLS equations before stating our main results, followed by a discussion of the latter.

Moving towards proofs of the theorems, in Section 11, we give an overview of the linear and multilinear estimates from [2] that we will be using to argue well-posedness for higher-order dNLS hierarchy/gauged dNLS equations, for the reader's convenience. In addition we will be making use of an estimate for the resonance relation which we take from the literature.

The proofs for Theorems 10.5 and 10.6 are contained in Section 12, where first we deal with estimates regarding well-posedness in Fourier-Lebesgue spaces, followed by the same for modulation spaces. The Theorems 10.1 and 10.3 regarding well-posedness of the dNLS hierarchy equations themselves follow from the former and use of the gauge-transformation.

In Section 13 we give proofs of our ill-posedness results associated with higher-order dNLS equations. These show that our well-posedness results are optimal (up to the endpoint) and that within the framework of techniques we are using, no lower threshold of initial regularity of the data is possible, while still achieving local well-posedness results.

To wrap up, in Appendix A we list the first few equations of the dNLS hierarchy together with their gauge-transformed variants where appropriate. This shall serve as a point of reference and give the interested reader an overview of what typical nonlinearities in the hierarchy look like.

Acknowledgements. This work is part of the author's PhD thesis. He would like to greatly thank his advisor, Axel Grünrock, for suggesting this line of problems and his continued and ongoing support.

9. Description of the dNLS hierarchy

Keeping in line with the literature we referenced in [2] describing the derivation of the NLS hierarchy equations, we stick to [5, 82] for the dNLS hierarchy equations. For literature dealing more generally with completely integrable systems we recommend the reader consult [25, 79] and references therein.

In the forthcoming subsections we describe how dNLS and associated higher-order equations arise as a compatibility condition for a linear scattering problem and how these equations are amenable to being recast in a more well-behaved class using the gauge-transformation (8.3). We will also touch on why this transformation leaves the well-posedness question (mostly) intact, specifically we are referring to the regularity of the gauge-transformation itself.

9.1. **Deriving dNLS hierarchy equations.**

The general setting we start out in is a linear scattering problem [5, eq. (1.1)] of the form

$$\mathrm{d}v = \Omega v \tag{9.1}$$

involving an $N \times N$ matrix of differential one-forms Ω depending on a spectral parameter $\zeta \in \mathbb{C}$. Its zero-curvature (also called integrability) condition [5, eq. (1.2)][82, eq. (2.3)] reads

$$0 = \mathrm{d}\Omega - \Omega \wedge \Omega \tag{9.2}$$

and, for appropriate choice of Ω, leads to various well-known nonlinear evolution equations. Choosing the right Ansatz for Ω decides which particular set of equations one manages to derive. In [2] and [5] the Ansatz $\Omega = (\zeta R_0 + P)\,\mathrm{d}x + Q(\zeta)\,\mathrm{d}t$, where the $\mathrm{d}x$ part of Ω depends only linearly on the spectral parameter $\zeta \in \mathbb{C}$, was chosen.

One picks the involved matrices as

$$R_0 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 0 & q \\ r & 0 \end{pmatrix}, \tag{9.3}$$

where we leave Q open for the time being. The entries q and r (which are functions depending on x and t) are referred to as potentials along which the scattering in (9.1) happens.

This Ansatz leads to (for example) the NLS and (m)KdV hierarchies of equations[25], depending again on the particular choice of relation between the two potentials q and r and matrix Q. In order to derive the dNLS hierarchy equations we follow [82, eq. (2.4)] and now instead choose $\Omega = (\zeta^2 R_0 + \zeta P)\,\mathrm{d}x + Q(\zeta)\,\mathrm{d}t$ with the same matrices R_0 and P as previously, again leaving Q unspecified for now.

A prolonged calculation that we will not reproduce for brevity's sake then shows that the compatibility condition (9.2) has an equivalent formulation as a Hamiltonian equation for our two potentials q and r

$$\frac{\mathrm{d}}{\mathrm{d}t} u = J \frac{\delta}{\delta u} \mathcal{H}, \tag{9.4}$$

see [82, eq. (4.11)]. In this equation $u = \left(\begin{smallmatrix} r \\ q \end{smallmatrix}\right)$ is a vector containing our potentials and $J = -2i \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \partial_x$ is an operator (different from the one involved in the derivation of the NLS hierarchy, cf. [2, eq. (2.3)]). What is left is to define the Hamiltonian $\mathcal{H}$ that is namesake to (9.4).

The Hamiltonian $\mathcal{H}$ has a strikingly similar form as for the NLS hierarchy equations

$$\mathcal{H} = \sum_{n=0}^{\infty} \alpha_n(t) I_n, \tag{9.5}$$

see [2, eq. (2.4)] for comparison. The $\alpha_n(t)$ are derived from the choice of Q we left open previously, and the I_n are conserved quantities of the equations in the dNLS hierarchy, in particular dNLS itself. Appropriate choices of the $\alpha_n(t)$ will thus yield the dNLS hierarchy equations, for which (individually) the I_n are the Hamiltonians.

Last thing is to state the individual Hamiltonian I_n: In [82, eqns. (3.3) and (3.4)] we are given explicit expressions for deriving these conserved quantities/Hamiltonians recursively

$$I_n = \int_{\mathbb{R}} q Y_n \,\mathrm{d}x \quad \text{and} \quad Y_{n+1} = \frac{1}{2i}\left[\partial_x Y_n + q \sum_{k=0}^{n} Y_{n-k} Y_k\right] \quad \text{with} \quad Y_0 = -\frac{r}{2i}. \tag{9.6}$$

The resemblance between (9.6) and [2, eq. (2.5)] is undeniable, though the discerning reader will note that the initial condition for this recursion is different, as well as the sum going up to $k = n$ (rather than $k = n - 1$).

For later reference we would like to give a lemma describing elementary properties of the Y_n all of which may be verified by a simple inductive argument, so we omit the proof.

Lemma 9.1. *For $n \in \mathbb{N}$ the terms Y_n have the following properties:*

(1) Y_n is a sum of monomials in q, r and their derivatives.
(2) Y_n as a polynomial is of homogeneous order, where we define the order of a monomial to be sum of twice the total number of derivatives and the number of factors in it. The order of any monomial in Y_n is $2n + 1$.

[25] The astute reader will note, that both the NLS and mKdV equations are embedded within the Ablowitz-Kaup-Newell-Segur (AKNS) hierarchy, a name more commonly used in the inverse scattering community literature, see for example [1, 25].

(3) *Every monomial in Y_n has a total number of factors r, or its derivatives, one greater than the total number of factors q, or its derivatives.*
(4) *The coefficients of the monomials in Y_n are a positive integer multiples of $(-1)^k(2i)^{k-2n-1}$, where k is the total number of derivatives in a given monomial.*
(5) *Y_n has a single term that consists of just one factor, it is $-(2i)^{-n}\partial_x^n r$.*

We are now ready to give the definition, i.e. fix a choice of coefficients α_n in (9.5), of what is referred to in the title of this paper as the dNLS hierarchy.

Definition. *For $j \in \mathbb{N}$ we define the jth dNLS hierarchy equation to be the Hamiltonian equation for the potential $q(x,t)$ in (9.4), where we choose $\alpha_{2j-1} = 2^{2j-1}$ and $\alpha_n = 0$ for $n \neq 2j-1$ in (9.5). We identify occurrences of the potential $r(x,t)$ with the complex conjugate of $q(x,t)$, i.e. $r = +\overline{q}$.*

Having defined what we deem to be the dNLS hierarchy equations we may quickly establish an equivalent theorem to [2, Theorem 2.3] that describes the general form of such an equation. We leave its proof to the reader as it differs only in details from the one in [2].

Note that this is also the point where we switch back to the more common notation of calling the unknown function u (instead of q or r). This is not to be confused with the vector of potentials $u = \binom{r}{q}$ used in (9.4).

Theorem 9.2. *For $j \in \mathbb{N}$ there exist coefficients $c_{k,\alpha} \in \mathbb{Z} + i\mathbb{Z}$ for every $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2j - k - 1$, for $1 \leq k \leq 2j-1$, such that the jth dNLS hierarchy equation may be written as*

$$i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = \sum_{k=1}^{2j-1} \sum_{\substack{\alpha \in \mathbb{N}_0^{2k+1} \\ |\alpha| = 2j-k-1}} c_{k,\alpha}\partial_x \left(\partial_x^{\alpha_1} u \prod_{\ell=1}^{k} \partial_x^{\alpha_{2\ell}}\overline{u}\partial_x^{\alpha_{2\ell+1}} u \right). \tag{9.7}$$

Remark 9.3. *We give some points of interest and remarks:*

(1) Breaking the definitions down in order to better uncover the structure of the dNLS hierarchy equations, we note that for $n = 2j-1$ the jth dNLS equation is given by

$$i\partial_t u = 2\alpha_n \partial_x \frac{\delta}{\delta \overline{u}} \int_{\mathbb{R}} u Y_n \, \mathrm{d}x. \tag{9.8}$$

(2) The main difference between the equations of the NLS and dNLS hierarchies is that the latter has an additional derivative on each nonlinear term. This is what makes its analysis more difficult, as the nonlinear term $|u|^2\partial_x u$ and its higher-order variants (where none of the derivatives fall on the complex conjugated factor $\overline{u}$) are quite ill-behaved. This is the reason we will be using the gauge-transformation, on which we will give more details in the next subsection.

(3) The first dNLS hierarchy equation ($j = 1$) corresponds to the classical dNLS equation (8.1). The higher-order equations, beyond the dNLS and fourth-order ($j = 2$) equation, do not, to the author's best knowledge, appear in the literature. We list the first few equations of the hierarchy in Appendix A. A further (interleaving) sequence of higher-order PDEs (with odd order of dispersion) can be defined and corresponds to non-zero choices of α_n, for $n \neq 2j-1, j \in \mathbb{N}$. We list these in the same appendix.

(4) Choosing the opposing sign convention $r = -\overline{q}$ also leads to a hierarchy of dNLS-like equations. As, in contrast to NLS, there is no meaningful difference between a focusing or defocusing case depending on the sign in front of the nonlinearity, our sign choice is of no significant importance. We fix it merely to have a designated convention for the name and choose to stay in line with the dNLS equation already present in the literature.

(5) Figuring out a non-recursive description of the coefficients involved in the dNLS hierarchy (or even determining, beyond (9.7)*, which nonlinear terms appear at all) is, to the author's best knowledge and in general, an unsolved problem. Such further insight into the nonlinearities may in the future aid phrasing well-posedness results dependent on a non-resonance condition (only fulfilled by the actual hierarchy equations).*

In the following subsection, where we explore the action of the gauge-transformation on the dNLS hierarchy equations, we will at least be able to obtain the coefficients of 'bad' cubic nonlinear terms, where no derivatives fall on the complex conjugated factor $\overline{u}$. These 'bad' cubic terms are the higher-order generalisations of $|u|^2\partial_x u$ from the nonlinearity of dNLS.

(6) Choosing non-zero values for the even numbered coefficients α_{2j} (and zero for all others) leads to a set of equations that have the same linear parts as the equations in the mKdV hierarchy (see Appendix A). It seems these do not appear independently in the literature, but would surely also make for an interesting object of study. Though we do not pursue this in this work.

Remark 9.4. *Now is the right place to establish the critical regularity $s_c(j,r)$ of the dNLS hierarchy equations in Sobolev and more generally Fourier-Lebesgue spaces*[26].

In a similar fashion to dNLS itself, the higher-order dNLS hierarchy equations are also invariant under the transformation of scale $u_\lambda(x,t) = \lambda^{\frac{1}{2}} u(\lambda x, \lambda^{2j} t)$, meaning if u is a solution of a dNLS-like equation with initial data u_0, then so is u_λ with initial data $u_{0,\lambda}(x) = \lambda^{\frac{1}{2}} u_0(\lambda x)$.

This leads to the critical regularity being $s_c(j,r) = \frac{1}{r} - \frac{1}{2}$, i.e. the L^2-norm stays invariant under this transformation on the scale of Sobolev spaces and $\hat{H}_1^{\frac{1}{2}}$ on the scale of Fourier-Lebesgue spaces for $r \to 1$. Our determined goal is to establish well-posedness of the dNLS hierarchy equations in spaces that are very close to these critical spaces.

9.2. **The gauge-transformation.** As is mentioned above there are certain nonlinear terms that appear in the dNLS hierarchy equations that are gravely less well-behaved than their fellows. These are terms like $|u|^2\partial_x u$ from (8.1), where all derivatives that lie on a cubic nonlinear term fall onto one of the factors that is not the complex conjugate of the unknown solution u. As the reader may verify in Appendix A these types of nonlinear terms do in fact crop up in the higher-order equations too.

Before we move on to proving well-posedness results for the dNLS hierarchy equations we must first absolve ourselves of these ill-behaved nonlinear terms. To do this we will be making use of the gauge-transformation that is already a well-known tool in the context of the dNLS equation itself:

$$\mathcal{G}_\pm : u(x,t) \mapsto v(x,t) := \exp\left(\pm i \int_{-\infty}^{x} |u(y,t)|^2 \,\mathrm{d}y\right) u(x,t). \tag{9.9}$$

See [37, 42, 45, 86], for example.

For the dNLS equation the gauge-transformation (9.9) is useful in the following sense: given a function u, it solves the dNLS equation (8.1) if and only if $v(x,t) := \mathcal{G}_-(u)(x,t)$ solves the gauge-transformed dNLS equation (8.4). Vice versa when you apply the gauge-transformation's inverse $\mathcal{G}_+$.

We want to explore how the gauge-transformation can help us in a similar way in order to simplify, or even enable, the well-posedness analysis of higher-order dNLS hierarchy equations. For this we must first find the right notion of 'simpler'

[26] Modulation spaces are not well-behaved under transformations of scale, due to the uniform frequency decomposition involved, thus there is no proper notion of criticality.

equation, which is specific enough in order for us to be able to achieve well-posedness results for and also general enough so that it is a superset of the image of the dNLS hierarchy equations under the gauge-transformation. We find the following definition appropriate.

Definition. *For $j \in \mathbb{N}$ we call a PDE a (jth order) gauged dNLS equation, if there exist coefficients $c_{k,\alpha} \in \mathbb{C}$, for $1 \le k \le 2j$, and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2j-k$, such that $c_{1,\alpha} = 0$ if $\alpha_2 = 0$ and the PDE may be written as*

$$i\partial_t v + (-1)^{j+1}\partial_x^{2j} v = \sum_{k=1}^{2j} \sum_{\substack{\alpha \in \mathbb{N}_0^{2k+1} \\ |\alpha|=2j-k}} c_{k,\alpha} \partial_x^{\alpha_1} u \prod_{i=1}^{k} \partial_x^{\alpha_{2i}} \overline{u} \partial_x^{\alpha_{2i+1}} u. \tag{9.10}$$

The difference between dNLS hierarchy equations and gauged dNLS equations, in their general form, is evidently rather small. The linear parts of the equations coincide for one. Regarding the cubic nonlinear terms, the gauged dNLS equations cannot contain so called ‘bad’ cubic terms that have none of their derivatives fall on the factor $\overline{u}$ in the cubic. This is exactly the advantage the gauge-transformation delivers. With regard to the higher-order nonlinear terms, the small price we have to pay for the elimination of the ‘bad’ cubic terms is that we incur an additional term of the form $|u|^{2j}u$, without any derivatives lying on it.

Remark 9.5. *We point out that transitioning from dNLS hierarchy equations to gauged dNLS equations does not change the notion of criticality, that was investigated in Remark 9.4. This is because we are, at most, leaving a cubic nonlinear term away and are gaining a term of the form $|u|^{2j}u$, that is invariant with respect to the same transformation of scale.*

Our goal for the rest of this subsection will be to establish, that the gauge-transformation does indeed translate between the dNLS hierarchy equations and what we are now referring to as gauged dNLS equations. This will then later allow us, conditioned on the continuity of the gauge-transformation, to prove well-posedness solely for gauged dNLS equations and pull-back these results to the actual equations of interest: the dNLS hierarchy equations. In this spirit we will be proving the following proposition.

Proposition 9.6. *Let $j \ge 2$, $u(x,t)$ be a function and $v(x,t) := \mathcal{G}_-(u)(x,t)$ its gauge-transform. Then u solves the jth order dNLS hierarchy equation if and only if v solves a (corresponding) gauged dNLS equation. And vice versa for the inverse transformation $\mathcal{G}_+$.*

Even though this proposition does not exactly specify *which* gauged dNLS equation v would solve, this proposition is sufficient for our purposes, since our well-posedness theorems are so general as to cover the whole class of gauged dNLS equations.

Relating to proof strategy, we will be investigating the coefficients of the ‘bad’ cubic nonlinear terms in the dNLS hierarchy equations and show that these coincide with those coefficients of ‘bad’ cubic terms that are lifted when one uses the gauge-transformation. We point out that this makes the dNLS hierarchy equations natural, beyond being derived from a completely integrable system, in the sense that their coefficients for ‘bad’ cubic terms are the unique[27] set that are amenable to use of the gauge-transformation.

As was also the case for the NLS hierarchy equations in [2], there is no specific understanding of the coefficients or finer structure of nonlinearities for the

[27]This uniqueness is only up to scaling of the coefficients. What is actually unique is the relationship (quotient) between the coefficients.

higher-order dNLS hierarchy equations present in the literature, to the author's best knowledge. So the following proposition, where the coefficients of at least the 'bad' cubic terms are uncovered, is a first.

Proposition 9.7. *Let $n \geq 1$. For $0 \leq k \leq n$ the coefficient of the cubic nonlinear term $(\partial_x^{n-k}u)\overline{u}(\partial_x^k u)$ is equal to*

$$\frac{4(-1)^{n+1}\alpha_n}{(2i)^{n+2}}\left(\binom{n+2}{k+1} - \delta_{0,k} - \delta_{n,k}\right), \tag{9.11}$$

where $\delta_{a,b}$ is the Kronecker delta.

For $n = 1$ is an easy and well-known result: the coefficient of $|u|^2 u_x$ in the dNLS equation is $2i$. We note that there is some level of redundancy in the statement as the terms $(\partial_x^{n-k}u)\overline{u}(\partial_x^k u)$ and $(\partial_x^k u)\overline{u}(\partial_x^{n-k}u)$ are the same by commutativity. This representation also still contains a choice of coefficients α_n. For the dNLS hierarchy we have made this choice, which seems canonical in relation to the coefficients appearing in the gauge-transformation, see Lemma 9.9.

Remark 9.8. *Figuring out the coefficients of the cubic nonlinear terms in general or of any of the higher-order terms also seems an interesting problem. Though the author finds that more delicate methods must be required in order to uncover these, as there is less of an obvious pattern compared with the 'bad' cubics.*

Proof of Proposition 9.7. We will prove the claim for $n \geq 2$ only, to eliminate some edge-cases. Referring to (9.8), which we now understand for general $n \in \mathbb{N}$, we must ask ourselves: where do the 'bad' cubic terms come from?[28]

Working our way backwards, such 'bad' terms, say $(\partial_x^{n-k}u)\overline{u}(\partial_x^k u)$, for $0 \leq k \leq n$, originate (before applying the derivative ∂_x present in (9.8)) from cubic terms in $\frac{\delta}{\delta \overline{u}}\int_{\mathbb{R}} uY_n \,\mathrm{d}x$ that also have no derivatives lying on $\overline{u}$ and a single derivative fewer in total, for example $(\partial_x^{n-1-k}u)\overline{u}(\partial_x^k u)$, for $0 \leq k \leq n-1$.

Recursing again, past the functional derivative, such cubic terms with $n-1$ total derivatives, but none on $\overline{u}$, can only originate from quartic terms in the integrand of the Hamiltonian, where at least one of the two $\overline{u}$ factors has no derivatives lying upon it. In turn, since we are multiplying with u in the integral, these come from cubic terms in Y_n where at least one of the two factors $\overline{u}$ has no derivatives lying upon it. General form of these terms is then $(\partial_x^{n-1-k}u)\overline{u}(\partial_x^k \overline{u})$, for $0 \leq k \leq n-1$.

To ease notation let $K_n(k)$ refer to the coefficient of $(\partial_x^{n-1-k}u)\overline{u}(\partial_x^k \overline{u})$ in Y_n, for $0 \leq k \leq n-1$. From here on out we will also use the convention $c = \frac{1}{2i}$, as this factor will appear often.

Our initial task is now to determine $K_n(k)$, for $n \geq 1$ and $0 \leq k \leq n-1$. Looking at the recursive definition of Y_{n+1} in (9.6)

$$Y_{n+1} = c\left[\partial_x Y_n + u\sum_{k=0}^{n} Y_{n-k}Y_k\right] \quad \text{with} \quad Y_0 = -c\overline{u}, \tag{9.12}$$

we can determine that cubic terms with coefficients $K_{n+1}(k)$ appear in Y_{n+1} in two ways:

(1) from the first summand in the brackets, if a term in Y_n that also has a factor $\overline{u}$ with no derivatives gets differentiated, by Leibniz' rule,
(2) in the sum, since the whole sum is multiplied with u, if for either $k = 0$ or $k = n$ a factor Y_0 is involved. This is since this is the only Y_n that contains a singular factor $\overline{u}$ and we would like the result to be cubic. We can be

[28]Even though we haven't formally defined what 'bad' cubic terms for mKdV-like equations with an extra derivative are (so where n is even), we will deal with them to be analogues of those for the dNLS hierarchy equations. That is where none of the derivatives in a cubic term fall on $\overline{u}$.

more specific even: a term we are looking for only appears by the product of $\overline{u}$ from Y_0 and a term $\partial_x^{n-1}\overline{u}$ from Y_n, resulting in $u\overline{u}\partial_x^{n-1}\overline{u}$ for both $k = 0$ and $k = n$ in the sum.

Accounting for the coefficients present and any edge-cases, we thus find that our coefficient function $K_n(k)$ fulfils the following recursion relation

$$K_{n+1}(k) = c\begin{cases} K_n(0) & \text{if } k = 0, \\ 2K_n(0) + K_n(1) & \text{if } k = 1, \\ K_n(k-1) + K_n(k) & \text{if } 1 < k < n, \\ 2c^{n+1} + K_n(n-1) & \text{if } k = n, \end{cases}$$

for $n > 1$ and $0 \leq k \leq n-1$. One may easily verify, with the initial condition $K_1(0) = c^3$ being evident, this recursion relation is solved by

$$K_n(k) = c^{n+2}\left(2\binom{n}{k} - \delta_{0,k}\right), \tag{9.13}$$

at least for $n > 1$. Note that the lack of symmetry here is no coincidence, as the terms whose coefficients are described by $K_n(k)$ shuffle derivatives between u and $\overline{u}$ rather than two identical factors u.

Next we must investigate how the functional derivative $\frac{\delta}{\delta\overline{u}}\int_{\mathbb{R}} uY_n \,\mathrm{d}x$ transforms these coefficients of terms in Y_n. For the readers convenience we recall the action of the functional derivative. If

$$F[\phi] = \int_{\mathbb{R}} f(\phi, \partial_x\phi, \partial_x^2\phi, \ldots, \partial_x^N\phi)\,\mathrm{d}x \quad \text{one has} \quad \frac{\delta F}{\delta\phi} = \sum_{k=0}^{N}(-1)^k\partial_x^k\frac{\partial f}{\partial(\partial_x^k\phi)}.$$

So we must take care to account for the fact that every 'bad' quartic term in the Hamiltonian $\int_{\mathbb{R}} uY_n\,\mathrm{d}x$ is counted twice: once for the factor $\overline{u}$ without any derivatives lying upon it and possibly another time if the remaining $\overline{u}$ factor (that may carry derivatives). We will use the symbol $\mathcal{R}$ to account for terms that are not 'bad' cubics and thus are not of importance for our analysis; it may differ from line to line. For $n > 1$ we figure

$$\frac{\delta}{\delta\overline{u}}\int_{\mathbb{R}} uY_n\,\mathrm{d}x = \sum_{k=0}^{n-1}(-1)^k\partial_x^k\frac{\partial(uY_n)}{\partial(\partial_x^k\overline{u})} \tag{9.14}$$

$$= \sum_{k=0}^{n-1}(-1)^k\partial_x^k(K_n(k) + \delta_{0,k})|u|^2(\partial_x^{n-1-k}u) + \mathcal{R} \tag{9.15}$$

Here we must be careful to account for the extra 1 (which we do by introducing $\delta_{0,k}$), which appears when differentiating the term $u(\partial_x^{n-1}u)\overline{u}^2$ in the functional derivative. This nicely cancels with the Kronecker delta in the coefficient function $K_n(k)$. Next we use the classical Leibniz rule and interchange the order of summation:

$$= \sum_{k=0}^{n-1}\sum_{\ell=0}^{k}(-1)^k 2c^{n+2}\binom{n}{k}\binom{k}{\ell}\overline{u}(\partial_x^{n-1-k+k-\ell}u)(\partial_x^\ell u) + \mathcal{R} \tag{9.16}$$

$$= 2c^{n+2}\sum_{\ell=0}^{n-1}\left(\sum_{k=\ell}^{n-1}(-1)^k\binom{n}{k}\binom{k}{\ell}\right)\overline{u}(\partial_x^{n-1-\ell}u)(\partial_x^\ell u) + \mathcal{R} \tag{9.17}$$

$$= 2c^{n+2}(-1)^{n+1}\sum_{\ell=0}^{n-1}\binom{n}{\ell}\overline{u}(\partial_x^{n-1-\ell}u)(\partial_x^\ell u) + \mathcal{R}, \tag{9.18}$$

where in the final step we used a well-known summation identity for binomial coefficients.

This representation of $\frac{\delta}{\delta \overline{u}} \int_{\mathbb{R}} uY_n \,\mathrm{d}x$ we may now use as the right-hand side in the definition of our evolution equations (9.8) in order to determine the coefficients we are interested in. Again we denote terms that are not of interest to us by use of the symbol $\mathcal{R}$, which may change from line to line:

$$\begin{aligned} i\partial_t u &= 2\alpha_n \partial_x \frac{\delta}{\delta \overline{u}} \int_{\mathbb{R}} uY_n \,\mathrm{d}x = \frac{4(-1)^{n+1}\alpha_n}{(2i)^{n+2}} \partial_x \sum_{\ell=0}^{n-1} \binom{n}{\ell} \overline{u} (\partial_x^{n-1-\ell} u)(\partial_x^\ell u) + \mathcal{R} \\ &= \frac{4(-1)^{n+1}\alpha_n}{(2i)^{n+2}} \sum_{\ell=0}^{n-1} \binom{n}{\ell} \overline{u} \left((\partial_x^{n-\ell} u)(\partial_x^\ell u) + (\partial_x^{n-(\ell+1)u})(\partial_x^{\ell+1} u) \right) + \mathcal{R} \\ &= \frac{4(-1)^{n+1}\alpha_n}{(2i)^{n+2}} \left(\sum_{\ell=0}^{n-1} \binom{n}{\ell} \overline{u} (\partial_x^{n-\ell} u)(\partial_x^\ell u) + \sum_{\ell=1}^{n} \binom{n}{\ell-1} \overline{u} (\partial_x^{n-\ell} u)(\partial_x^\ell u) \right) + \mathcal{R}. \end{aligned}$$

The first and last terms of these sums respectively are both of the form $|u|^2 \partial_x^n u$ so we may combine them. All other 'bad' cubics appear in the sums twice by symmetry so we 'fold over' the sum in order to see the actual coefficient. We omit the leading factor for space reasons.

$$\begin{aligned} &(n+1)|u|^2(\partial_x^n u) + \sum_{\ell=1}^{n-1} \binom{n+1}{\ell} \overline{u} (\partial_x^{n-\ell} u)(\partial_x^\ell u) + \mathcal{R} \\ &= (n+1)|u|^2(\partial_x^n u) + \sum_{\ell=1}^{\lfloor \frac{n-1}{2} \rfloor} \left(\binom{n+1}{\ell} + \binom{n+1}{n-\ell} \right) \overline{u} (\partial_x^{n-\ell} u)(\partial_x^\ell u) + \mathcal{R} \end{aligned}$$

Using the identity $\binom{n+1}{\ell} + \binom{n+1}{n-\ell} = \binom{n+2}{\ell+1}$ we have now been able to completely determine the coefficients of the 'bad' cubic terms appearing in the hierarchy equations. Noting that $\binom{n+2}{0+1} = \binom{n+2}{n+1} = n+2 = (n+1) - 1$ one may verify that the representation given in (9.11) is correct. □

Our next step in preparation of the proof of Proposition 9.6 is figuring out which cubic nonlinear terms can be lifted by the gauge-transformation (and which coefficients lead to total cancellation of these terms). For this we will prove the following lemma in which it is established which 'bad' cubic terms are generated by inserting a gauge-transformed function into the linear part of a dNLS hierarchy equation.

Lemma 9.9. *Let $j \in \mathbb{N}$ and u be a solution of the jth dNLS hierarchy equation. We set $v := \mathcal{G}_-(u)$ to be its gauge-transform. The coefficient of the 'bad' cubic term $(\partial_x^{2j-1-\ell} u)\overline{u}(\partial_x^\ell u)$, in terms of u, appearing in $i\partial_t v + (-1)^{j+1}\partial_x^{2j} v$ is*

$$i(-1)^{j+1} \left(\binom{2j+1}{\ell+1} - \delta_{0,\ell} - \delta_{2j-1,\ell} \right). \tag{9.19}$$

Proof. We begin this proof by simple insertion of v into the proposed linear part of a dNLS hierarchy equation and elementary calculation:

$$\begin{aligned} i\partial_t v + (-1)^{j+1}\partial_x^{2j} v = G_u (&i\partial_t u + (-1)^{j+1}\partial_x^{2j} u \\ &+ u \int_{-\infty}^x u_t \overline{u} + u \overline{u_t} \,\mathrm{d}\lambda + i(-1)^j \sum_{k=0}^{2j-1} \partial_x^{2j-1-k} \left(|u|^2 \partial_x^k u \right) + \mathcal{R}). \end{aligned} \tag{9.20}$$

Here we have re-used the symbol $\mathcal{R}$ to denote higher-order terms and non-'bad' cubics and introduced the notation $G_u = \exp\left(-i \int_{-\infty}^x |u(y)|^2 \,\mathrm{d}y \right)$ to simplify matters. Further cubic nonlinear terms may be produced by the integral, but only if the integrand is quadratic in u. Inserting the dNLS hierarchy equation that is solved by u for the terms u_t and $\overline{u_t}$ we see that the integrand is only quadratic for the linear dispersion term in the equation:

$$u\int_{-\infty}^{x} u_t\overline{u} + u\overline{u_t}\,\mathrm{d}\lambda = u\int_{-\infty}^{x} (i(-1)^{j+1}\partial_x^{2j}u)\overline{u} - u(i(-1)^{j+1}\partial_x^{2j}\overline{u})\,\mathrm{d}\lambda + \mathcal{R} \tag{9.21}$$

$$= i(-1)^{j+1}u\int_{-\infty}^{x} (\partial_x^{2j}u)\overline{u} - u(\partial_x^{2j}\overline{u})\,\mathrm{d}\lambda + \mathcal{R}. \tag{9.22}$$

The reader may now inductively verify the fact that this integral can be rewritten as

$$i(-1)^{j+1}u\int_{-\infty}^{x} (\partial_x^{2j}u)\overline{u} - u(\partial_x^{2j}\overline{u})\,\mathrm{d}\lambda = i(-1)^{j+1}u\sum_{k=0}^{2j-1}(-1)^k(\partial_x^{2j-1-k}u)(\partial_x^k\overline{u}).$$

Of the terms in this sum only the first one is a 'bad' cubic term, so when we now return to (9.20) the other terms of the sum may be absorbed into $\mathcal{R}$ and we are left with

$$\begin{aligned}(9.20) = G_u(&i\partial_t u + (-1)^{j+1}\partial_x^{2j}u + i(-1)^{j+1}|u|^2(\partial_x^{2j-1}u) \\ &+ i(-1)^j\sum_{k=0}^{2j-1}\partial_x^{2j-1-k}\left(|u|^2\partial_x^k u\right) + \mathcal{R})\end{aligned} \tag{9.23}$$

$$\begin{aligned}= G_u(&i\partial_t u + (-1)^{j+1}\partial_x^{2j}u + i(-1)^{j+1}|u|^2(\partial_x^{2j-1}u) \\ &+ i(-1)^j\sum_{k=0}^{2j-1}\sum_{\ell=0}^{2j-1-k}\binom{2j-1-k}{\ell}\overline{u}(\partial_x^{2j-1-k-\ell}u)(\partial_x^{k+\ell}u) + \mathcal{R}).\end{aligned} \tag{9.24}$$

Now interchanging the sums and using a well-known identity $\sum_{\ell=k}^{2j-1}\binom{2j-1-k}{\ell-k} = \binom{2j}{\ell}$ for binomial coefficients we go on to write

$$\begin{aligned}= G_u(&i\partial_t u + (-1)^{j+1}\partial_x^{2j}u - i(-1)^j|u|^2(\partial_x^{2j-1}u) \\ &+ i(-1)^j\sum_{\ell=0}^{2j-1}\frac{2j}{2j-\ell}\binom{2j-1}{\ell}\overline{u}(\partial_x^{2j-1-\ell}u)(\partial_x^\ell u) + \mathcal{R}).\end{aligned} \tag{9.25}$$

To account for the symmetry of the cubic terms with derivatives on the factors u we again 'fold-over' this sum so that we may read off the coefficients more comfortably:

$$\begin{aligned}&= G_u(i\partial_t u + (-1)^{j+1}\partial_x^{2j}u - i(-1)^j|u|^2(\partial_x^{2j-1}u) \\ &+ i(-1)^j\sum_{\ell=0}^{\lfloor\frac{2j-1}{2}\rfloor}\left(\frac{2j}{2j-\ell}\binom{2j-1}{\ell} + \frac{2j}{\ell+1}\binom{2j-1}{2j-1-\ell}\right)\overline{u}(\partial_x^{2j-1-\ell}u)(\partial_x^\ell u) + \mathcal{R})\end{aligned} \tag{9.26}$$

$$\begin{aligned}&= G_u(i\partial_t u + (-1)^{j+1}\partial_x^{2j}u \\ &+ i(-1)^j\sum_{\ell=0}^{\lfloor\frac{2j-1}{2}\rfloor}\left(\binom{2j+1}{\ell+1} - \delta_{0,\ell}\right)\overline{u}(\partial_x^{2j-1-\ell}u)(\partial_x^\ell u) + \mathcal{R})\end{aligned} \tag{9.27}$$

These coefficients coincide with the statement of this lemma so the proof is complete. □

Now all ingredients we need for the proof of Proposition 9.6 are set in place.

Proof of Proposition 9.6. There isn't much left to argue: When applying the gauge-transformation $v = \mathcal{G}_-(u)$ and inserting v into the linear part of a dNLS hierarchy equation, the way one recovers which equation v solves is by using that u solves a dNLS hierarchy equation and then rewriting all nonlinear terms in v instead of u by supplementing factors with the exponential function involved in the gauge-transformation and/or adding correctional higher-order terms.

Since we have now found, between Proposition 9.7 and Lemma 9.9, that the coefficients of the dNLS hierarchy equations and the gauge-transformation coincide (we remind the reader that for a dNLS hierarchy equation we set $n = 2j-1$ and our choice of $\alpha_{2j-1} = 2^{2j-1}$), we can be sure of the fact that at least before rewriting the nonlinear terms in terms of v, all 'bad' cubic terms are cancelled by the gauge-transformation. In supplementing cubic terms with the exponential function form the gauge-transformation we do not suddenly turn them 'bad' and higher-order terms that may need to be added (in order to account for cases where the derivative in a gauge-transformed nonlinear terms falls onto the exponential function) are of no concern to us. □

9.3. **Continuity of the gauge-transformation.** After having established that the use of the gauge-transformation absolves us of the most ill-behaved terms in dNLS hierarchy equations, we must also argue that it is compatible with our goal of well-posedness. More precisely we must exhibit its continuity, so that the gauge-transformation may be used to pull-back well-posedness results for gauged dNLS equations to well-posedness for dNLS hierarchy equations that we are actually interested in.

For well-posedness in Fourier-Lebesgue spaces continuity of the gauge-transformation had previously been established in the literature.

Lemma 9.10 ([37, Lemma 3.3 and Remark 3.4])**.** *Let $s \geq \frac{1}{2}$ and $1 < r \leq 2$. Then the gauge-transformation $\mathcal{G}_\pm : \hat{H}^s_r(\mathbb{R}) \to \hat{H}^s_r(\mathbb{R})$ is Lipschitz continuous on bounded sets. The same holds true if the gauge-transformation is viewed as a map $\mathcal{G}_\pm : C(I, \hat{H}^s_r) \to C(I, \hat{H}^s_r)$ for an arbitrary interval $I \subset \mathbb{R}$.*

Though even with well-posedness results for dNLS in modulation spaces already appearing in the literature, see [42], where the gauge-transformation aided in simplifying the equation, the issue of its continuity does not seem to have been tackled. Thus we prove the following Lemma.

Lemma 9.11. *Let $2 \leq p < \infty$ and $s > \frac{1}{2} - \frac{1}{p}$. Then the gauge-transformation $\mathcal{G}_\pm : M^s_{2,p}(\mathbb{R}) \to M^s_{2,p}(\mathbb{R})$ is Lipschitz continuous on bounded sets. Moreover it is also continuous interpreted as a map $\mathcal{G}_\pm : C(I, M^s_{2,p}) \to C(I, M^s_{2,p})$ for an arbitrary interval $I \subset \mathbb{R}$.*

Remark 9.12. *The regularity restriction $s > \frac{1}{2} - \frac{1}{p}$ is only natural since this is necessary for the embedding $M^s_{2,p} \subset L^2$ to hold, which in turn is necessary for the gauge-transformation to be well-defined.*

Proof of Lemma 9.11. In order to simplify notation we will only make the argument for $\mathcal{G}_+$, the minus-case works the same, and we also introduce the notation

$$G_u(x) = \exp\left(i \int_{-\infty}^x |u(y)|^2 \,\mathrm{d}y\right) \quad \text{and} \quad \mathcal{I}(u)(x) = \int_{-\infty}^x |u(y)| \,\mathrm{d}y \tag{9.28}$$

notwithstanding possible t dependence of u, so the gauge-transformation may be written as $\mathcal{G}_+(u)(x) = G_u u(x)$.

We will be following an argument given in [47, Appendix A], thus we will establish an estimate

$$\|(G_v - G_w)u\|_{M^s_{2,p}} \lesssim e^{c\|v\|^2_{M^s_{2,p}} + c\|w\|^2_{M^s_{2,p}}} \|v+w\|_{M^s_{2,p}} \|v-w\|_{M^s_{2,p}} \|u\|_{M^s_{2,p}}. \tag{9.29}$$

With (9.29) we may argue the Lipschitz continuity of $\mathcal{G}_+$ for functions $u, w \in B_r(0) \subset M^s_{2,p}$ as follows

$$\begin{aligned}\|\mathcal{G}_+(u) - \mathcal{G}_+(v)\|_{M^s_{2,p}} &\lesssim \|(G_u - G_v)u\|_{M^s_{2,p}} + \|(G_v - 1)(u-v)\|_{M^s_{2,p}} + \|u-v\|_{M^s_{2,p}} \\ &\lesssim (re^{2cr^2} + re^{cr^2} + 1)\|u-v\|_{M^s_{2,p}} \lesssim_r \|u-v\|_{M^s_{2,p}}.\end{aligned}$$

We are left to argue (9.29). First we use the generalised Leibniz rule for modulation spaces (8.7) which results in

$$\|(G_v - G_w)u\|_{M^s_{2,p}} \lesssim \|G_v - G_w\|_{M^s_{\infty,\tilde{p}}}\|u\|_{M_{2,2}} + \|G_v - G_w\|_{M_{\infty,1}}\|u\|_{M^s_{2,p}}, \quad (9.30)$$

where $\frac{1}{p'} = \frac{1}{2} + \frac{1}{\tilde{p}'}$. Looking at the second term in the sum we must estimate $G_v - G_w$ in the $M_{\infty,1}$ norm. We use the algebra property of this space and the power series expansion of the exponential function to arrive at

$$\begin{aligned}
&\|G_v - G_w\|_{M_{\infty,1}} \\
&\lesssim \|\mathcal{I}(|v|^2 - |w^2|)\|_{M_{\infty,1}} \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{j=0}^{k-1} (c\|\mathcal{I}(|v|^2)\|_{M_{\infty,1}})^j (c\|\mathcal{I}(|w|^2)\|_{M_{\infty,1}})^{k-j-1} \\
&\lesssim \|\mathcal{I}(|v+w||v-w|)\|_{M_{\infty,1}} \exp(c\|\mathcal{I}(|v|^2)\|_{M_{\infty,1}} + c\|\mathcal{I}(|w|^2)\|_{M_{\infty,1}}).
\end{aligned}$$

From here, if we are now able to argue the bilinear estimate

$$\|\mathcal{I}(fg)\|_{M_{\infty,1}} \lesssim \|f\|_{M^s_{2,p}}\|g\|_{M^s_{2,p}}, \quad (9.31)$$

we arrive at our desired (9.29). We look at two cases depending on the magnitude of the frequency of $\mathcal{I}(fg)$ because of the singularity introduced by $\mathcal{I}$ at low frequencies.

(1) **low frequencies**: Since here we only have finitely many terms in the outer ℓ^1 norm we may estimate by L^∞, use Hölder's inequality and a Sobolev-type embedding for modulation spaces (8.6) to arrive at

$$\|P_1\mathcal{I}(fg)\|_{M_{\infty,1}} \lesssim \|\mathcal{I}(fg)\|_{L^\infty} \lesssim \|f\|_{L^2}\|g\|_{L^2} \lesssim \|f\|_{M^s_{2,p}}\|g\|_{M^s_{2,p}}, \quad (9.32)$$

since $s > \frac{1}{2} - \frac{1}{p}$.

(2) **high frequencies**: In this situation we may replace $\mathcal{I}(fg)$ with a Bessel potential operator

$$\begin{aligned}
\|P_{>1}\mathcal{I}(fg)\|_{M_{\infty,1}} &\lesssim \|J^{-1}(fg)\|_{M_{\infty,1}} \lesssim \|fg\|_{M_{\infty,r}} \\
&\lesssim \|f\|_{M_{\infty,\rho}}\|g\|_{M_{\infty,\rho}} \lesssim \|f\|_{M^s_{2,p}}\|g\|_{M^s_{2,p}}
\end{aligned}$$

where we then use Hölder's inequality with $r = \infty-$ in the outer ℓ^1 norm and then Hölder's inequality again in the outer norm, with $\frac{1}{r'} = \frac{2}{\rho'} \Leftrightarrow \frac{1}{\rho} = \frac{1}{2}+$. Finally we use a Sobolev-type embedding for modulation spaces (8.6) which requires $s > \frac{1}{\rho} - \frac{1}{p} = \frac{1}{2} - \frac{1}{p}+$.

Now we turn to the first term in the sum in (9.30). The $M_{2,2} = L^2$ norm of u may again be estimated by $\|u\|_{M^s_{2,p}}$ due to the Sobolev-type embedding (8.6). For the other factor we argue similarly to the above, noting that $M^s_{\infty,\tilde{p}}$ is also an algebra since $s > \frac{1}{2} - \frac{1}{q} = \frac{1}{\tilde{p}'}$, though this time we require a bilinear estimate of the form

$$\|\mathcal{I}(fg)\|_{M^s_{\infty,\tilde{p}}} \lesssim \|f\|_{M^s_{2,p}}\|g\|_{M^s_{2,p}}. \quad (9.33)$$

For low frequencies we may reuse our argument from above, since in that case $\|P_1\mathcal{I}(fg)\|_{M^s_{\infty,\tilde{p}}} \lesssim \|\mathcal{I}(fg)\|_{M_{\infty,1}}$, whereas for high frequencies we argue

$$\|P_{>1}\mathcal{I}(fg)\|_{M^s_{\infty,\tilde{p}}} \lesssim \|fg\|_{M^{s-1}_{\infty,\tilde{p}}} \lesssim \|fg\|_{H^{s-1+s'}} \quad (9.34)$$

where $s' > \frac{1}{\tilde{p}} - \frac{1}{2} = \frac{1}{p}$. Then $s - 1 + s' = -\frac{1}{2}+$ and we may use a Sobolev embedding and Hölder's inequality

$$\lesssim \|fg\|_{L^{1+}} \lesssim \|f\|_{L^{2+}}\|g\|_{L^{2+}} \lesssim \|f\|_{M^s_{2,p}}\|g\|_{M^s_{2,p}}, \quad (9.35)$$

where in the final inequality we used a Sobolev-type embedding for modulation spaces (8.6) again.

The claim of continuity of $\mathcal{G}_\pm$ on $C(I, M^s_{2,p})$ follows by replacing the $M^s_{2,p}$ norms by $L^\infty_t M^s_{2,p}$ norms. We omit the details. □

10. Statement of results

10.1. **Prior work.** Before we state our main results let us quickly review the literature regarding low regularity well-posedness results for the dNLS equation itself as well as the fourth order dNLS hierarchy equation ($j = 2$). To the author's knowledge the other, higher-order, equations part of the dNLS hierarchy do not yet appear in the literature. Giving a complete account of the well-posedness theory (especially concerning results of the inverse scattering community) though is beyond our scope, so we will focus mostly on comparable results to our own.

As is unsurprising the dNLS equation (and variants of it) were first tackled using the energy method, see [89, 90], achieving local well-posedness for initial data in H^s (independent of the underlying geometry) for $s > \frac{3}{2}$.

On the line these results were later improved in [45] to cover both local and global well-posedness (thanks to energy conservation) in $H^1(\mathbb{R})$, under the restriction that the mass of the initial data be smaller than 2π. Already here the gauge-transformation was used in order to make the equation approachable using dispersive PDE techniques.

In parallel it was begun to utilise the dispersive character of the equation[29] in [54], where a variant of Kato smoothing together with a maximal function estimate was used in order to establish small data local well-posedness in $H^{\frac{7}{2}}(\mathbb{R})$.

Using multilinear refinements of smoothing estimates for the Schrödinger propagator together with $X_{s,b}$ spaces the local well-posedness result could be pushed down to $H^{\frac{1}{2}}(\mathbb{R})$. See [86]. In a subsequent paper [87], for $s > \frac{32}{33}$, these newly constructed local solutions were extended globally using the splitting-argument, which was initially developed by Bourgain. It was also recorded that, since the flow fails to be thrice continuously differentiable for $s < \frac{1}{2}$, there was no hope in further improving the local result on the line using the contraction mapping theorem alone.

More dire still, after in [9] it had been established using exact soliton solutions to the dNLS equation, that the flow of the dNLS equation cannot be uniformly continuous for $s < \frac{1}{2}$.

On the front of improvements to global well-posedness, after a refined version of the splitting-argument, today usually referred to as the I-method, had been developed, the global result could be pushed down to almost match the (now known to be optimal, using fixed-point methods) local result. That is, in [21, 22] it was proven that the dNLS equation is globally well-posed in $H^s(\mathbb{R})$ for $s > \frac{1}{2}$, conditioned on a mass below 2π.

Global well-posedness in the endpoint $s = \frac{1}{2}$, under the same mass restriction as previously, was later shown by different authors [71], again using the I-method, but additionally using a resonant decomposition technique to better control a singularity arising from resonant interactions.

Trying to push the local result further towards the scaling critical space, Fourier-Lebesgue spaces were employed in [37], where then local well-posedness was achieved in $\hat{H}^s_r(\mathbb{R})$ for $s \geq \frac{1}{2}$ and $2 \geq r > 1$. This covers the entire scaling sub-critical configuration of parameters.

As modulation spaces moved into focus of the dispersive PDE community these spaces were also employed in order to move well-posedness results closer to the scaling critical space. In [42] local well-posedness for initial data in $M^{\frac{1}{2}}_{2,q}$ for $4 \leq q < \infty$ was proven. Here $M^{\frac{1}{2}}_{2,\infty}$ is understood to be the critical space, even though modulation spaces are not well-behaved under transformations of scale. It is of note, that

[29]Not just the dNLS equation was considered here, but a rather large class with arbitrary polynomial nonlinearity involving derivatives.

in the previously cited work the continuity of the gauge-transform in appropriate modulation spaces was not discussed. We resolve this issue with Lemma 9.11.

The mass restriction though, that had so far been part of all global results, turned out to be a mere technically arising restriction. This, over the course of [94, 95], could be lifted from 2π to 4π using the sharp version of the Gagliardo-Nirenberg inequality for global solutions in $H^1(\mathbb{R})$. This result was later then extended to also cover the full range of possible local results, i.e. in [43] it was shown that, under the lighter mass restriction of 4π solutions with initial data in $H^{\frac{1}{2}}(\mathbb{R})$ extended globally.

The most recent and extensive results concerning the low-regularity well-posedness theory of the dNLS equation were achieved with methods stemming from the equation's complete integrability. Using those techniques it was possible to prove global well-posedness held in the scaling critical space $L^2(\mathbb{R})$ with no restriction on the mass of the initial data [44, 61]. Moreover, those two papers and references therein give a nice, general overview of recent well-posedness results for the dNLS equation achieved with inverse scattering/complete integrability.

Since we are less concerned with results pertaining to periodic initial data we will stick to headlines only. It was only with [47] that a version of the gauge-transformation was found, such that the dNLS equation could be attacked using fixed-point methods on the torus. Here the optimal local well-posedness result could immediately be paralleled (despite the lack of strong smoothing effects of the Schrödinger group), i.e. well-posedness for initial data in $H^{\frac{1}{2}}(\mathbb{T})$ was achieved. The argument used the L^4 Strichartz estimate extensively. Ill-posedness, in the sense of failure of thrice differentiability of the flow below $s = \frac{1}{2}$ is contained in the same work.

Here again, Fourier-Lebesgue spaces (that in the periodic setting coincide with modulation spaces) could be used in order to push the local well-posedness result nearer the scaling critical space. Over the course of [37, 39] well-posedness could be extended to initial data in $\hat{H}^s_r(\mathbb{T})$ for $s = \frac{1}{2}$ and $2 \geq r > \frac{4}{3}$. Covering local well-posedness in the remainder of the subcritical range, that is $\hat{H}^{\frac{1}{2}}_r(\mathbb{T})$ for $r > 1$, was then achieved in [24].

Much fewer works have so far dealt with any higher-order dNLS hierarchy equations. We mention [50], where a well-posedness results covering the fourth-order dNLS equation is proven. Specifically small data local well-posedness for data in $H^s(\mathbb{R})$, $s > 4$, is established.

This was later improved in [48] to small data well-posedness for data in $H^1(\mathbb{R})$. The dNLS hierarchy equation is also explicitly mentioned in this work. Further low-regularity well-posedness results covering the higher-order dNLS hierarchy equations are not present in the literature, to the author's best knowledge.

10.2. **Main results.** First we consider a general Cauchy problem for an evolution equation of the form

$$\begin{cases} i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = F(u) \\ u(t=0) = u_0 \end{cases}, \tag{10.1}$$

where we are able to derive the following well-posedness theorems for data in Fourier-Lebesgue and modulation spaces regarding the dNLS hierarchy.

Theorem 10.1. *Let $j \geq 2$ and (10.1) be the jth dNLS hierarchy equation. If $1 < r \leq 2$ and $s \geq \frac{1}{2} + \frac{j-1}{r'}$, the Cauchy problem (10.1) with initial data $u_0 \in \hat{H}^s_r(\mathbb{R})$ is locally well-posed, with the solution map being Lipschitz continuous on bounded sets.*

For $j = 1$ this theorem corresponds to the well-posedness of the dNLS equation in Fourier-Lebesgue spaces on the line and is already known in the literature [37].

Remark 10.2. *The condition $r \leq 2$ appears naturally in this context, because of the use of the gauge-transformation, the well-definedness of which requires $L^2 \supset \hat{H}^s_r$.*

Theorem 10.3. *Let $j \geq 2$ and (10.1) be the jth dNLS hierarchy equation. Then for $2 \leq p < \infty$ and $s \geq \frac{j}{2}$, the Cauchy problem (10.1) with initial data $u_0 \in M^s_{2,p}(\mathbb{R})$ is locally well-posed with a solution map that is Lipschitz continuous on bounded subsets.*

Again, for $j = 1$ (and $4 \leq p$) the well-posedness of the dNLS equation in modulation spaces on the line can already be found in the literature, see [42], though there the continuity of the gauge-transformation is not discussed.

Remark 10.4. *It was already noted immediately after stating Theorem 3 in [37], that the uniqueness statement in the preceding well-posedness theorems (and in [37]) was to be carefully interpreted. Due to the gauge-transformation, uniqueness of a solution u only holds with respect to other solutions v that fulfil the artificial seeming condition that $\mathcal{G}_- v$ must also solve the associated gauge-transformed equation (corresponding to a dNLS hierarchy equation).*

Most noticeable about these theorems, in comparison with their analogues for the NLS hierarchy equations [2, Theorems 3.1 and 3.2], is the lower regularity of the solution map: being merely Lipschitz continuous rather than analytic. This is due to the fact that Theorems 10.1 and 10.3 are derived from the following well-posedness theorems concerned with gauged dNLS equations. We remind the reader that a gauged dNLS equation contains no 'bad' cubic nonlinear terms, i.e. where no derivatives fall on the complex conjugated term.

Since we only know the gauge-transformation to be a bi-Lipschitz continuous map on bounded sets, see Lemmas 9.10 and 9.11. Hence the pull-back of the solution map is not analytic but merely Lipschitz continuous.

Theorem 10.5. *Let $j \geq 2$ and (10.1) be a gauged dNLS equation. Then*

(1) if $1 < r \leq 2$ and $s \geq \frac{1}{2} + \frac{j-1}{r'}$, the Cauchy problem (10.1) with initial data $u_0 \in \hat{H}^s_r(\mathbb{R})$ is locally well-posed with an analytic solution map,

(2) if additionally (10.1) contains no cubic nonlinear terms, $1 < r \leq 2$ and $s > \frac{1}{r} - \frac{1}{2}$, the Cauchy problem with initial data $u_0 \in \hat{H}^s_r(\mathbb{R})$ is locally well-posed with an analytic solution map.

Theorem 10.6. *Let $j \geq 2$ and (10.1) be a gauged dNLS equation. Then*

(1) if $2 \leq p < \infty$ and $s \geq \frac{j}{2}$, the Cauchy problem (10.1) with initial data $u_0 \in M^s_{2,p}(\mathbb{R})$ is locally well-posed with an analytic solution map,

(2) if additionally (10.1) contains no cubic nonlinear terms and $2 \leq p \leq \infty$, let $k \geq 2$ be the smallest index for which $c_{k,\alpha} \neq 0$ (in (9.10)) for a choice of $\alpha \in \mathbb{N}_0^{2k+1}$. Then for $s > \frac{1}{2} + \frac{1}{4k} - \frac{2k+1}{2kp}$, the Cauchy problem with initial data $u_0 \in M^s_{2,p}(\mathbb{R})$ is locally well-posed with an analytic solution map.

Remark 10.7. *Theorem 10.6 has further extensions: besides the (also called) gauge-invariant (with respect to multiplication with a constant phase-factor $u \mapsto e^{i\theta} u$) distribution of complex conjugates in the nonlinear terms others are possible. Only for the cubic term $|u|^2 u$ is the necessary distribution of complex conjugates with our arguments, ignoring derivatives. For the higher-order terms an arbitrary distribution of complex conjugates is possible.*

A similar, if weaker, statement regarding the arbitrariness of distribution of complex conjugates in the nonlinear terms is true of Theorem 10.5. For example, the *proof of Proposition 12.2 shows that the statement of the theorem still holds true, if only as few as two factors are complex conjugates in a quintic or higher-order nonlinear term.*

Though we do not pursue an in-depth showcasing of which distributions of complex conjugates are covered by our arguments, as the gauge-invariant (see above) nonlinearities are most canonical.

We derive these theorems by means of proving multilinear estimates in $\hat{X}^r_{s,b}$ and $X^p_{x,b}$ spaces for the nonlinear terms in the equations. Definitions of their respective norms are given by $\|f\|_{\hat{X}^r_{s,b}} = \|\langle\tau - \xi^{2j}\rangle^b\langle\xi\rangle^s\mathcal{F}_{x,t}f\|_{L^{r'}_{xt}}$ and $\|f\|_{X^p_{s,b}} = \|\langle n\rangle^s\|\Box_n f\|_{\hat{X}^2_{0,b}}\|_{\ell^p_n(\mathbb{Z})}$, where $(\Box_n)_{n\in\mathbb{Z}}$ is a fixed choice of uniform frequency decomposition operators. Properties of these function spaces were covered in [2, Section 1.2]. Combined with the contraction mapping principle such estimates lead to local well-posedness in Fourier-Lebesgue and modulation spaces respectively. Using such estimates to obtain local well-posedness results is a well-known technique initially investigated in [12, 13]. We omit specific details of the connection between non- or multilinear estimates and well-posedness, but direct the uninitiated reader to [35, 36] for an overview and necessary adaptations in order to deal with Fourier-Lebesgue and modulation spaces (rather than just Sobolev spaces).

In contrast with our well-posedness results given in the preceding theorems, we are also able to derive a number of ill-posedness results regarding the hierarchy equations in conjunction with the techniques that we are utilising. In particular the following two theorems show that no direct application of the contraction mapping theorem can lead to well-posedness for non-periodic initial data below the regularities at which we establish local well-posedness, i.e. our well-posedness results are optimal in this sense.

Theorem 10.8. *For any $j \geq 2$, $1 \leq r \leq \infty$ and $s < \frac{1}{2} + \frac{j-1}{r'}$ the flow map $S : \hat{H}^s_r(\mathbb{R}) \times (-T,T) \to \hat{H}^s_r(\mathbb{R})$ of the Cauchy problem for the jth dNLS hierarchy equation cannot be thrice continuously differentiable.*

Theorem 10.9. *For any $j \geq 2$, $1 \leq p, q \leq \infty$ and $s < \frac{j}{2}$ the flow map $S : M^s_{p,q}(\mathbb{R}) \times (-T,T) \to M^s_{p,q}(\mathbb{R})$ of the Cauchy problem for the jth dNLS hierarchy equation cannot be thrice continuously differentiable.*

Remark 10.10. *The preceding two theorems are phrased for the dNLS hierarchy equations themselves. This turns out to be an unnecessary restriction though. As the proofs will show we are only concerned with cubic nonlinear terms and with that we may also ignore the distribution of derivatives within them. The latter stems from the fact that the ill-posedness result is derived from a high-high-high interaction between the three factors, so that derivatives may be shifted arbitrarily between factors anyway.*

Thus there is still a lot of leeway in phrasing the ill-posedness theorems for more general classes of equations. Since we have not defined a name for this explicit class we refrain from complicating the theorem by trying to be as general as possible in its phrasing. Suffice it to say that our ill-posedness theorems still hold, so long as a cubic nonlinear term (in an equation paralleling (10.1)*) with $2j - 1$ derivatives placed upon it is present in the equation. In particular this also includes the class of gauged dNLS equations.*

Moving from the realm of non-periodic initial data to the periodic problem, we can be sure that no (direct) application of the contraction mapping theorem will lead to any positive results concerning the fourth-order hierarchy equation. Of course, this suggests that a similar result also holds for all higher-order equations. This would mean that merely the dNLS equation itself can be attacked using fixed-point techniques with periodic initial data.

Theorem 10.11. *For any $1 \leq r \leq \infty$ and $s \in \mathbb{R}$ the flow map $S : \hat{H}^s_r(\mathbb{T}) \times (-T, T) \to \hat{H}^s_r(\mathbb{T})$ of the Cauchy problem for the fourth-order (i.e. $j = 2$) dNLS hierarchy equation cannot be thrice continuously differentiable.*

Weakening the regularity requirements for the initial data we are able to showcase that the situation regarding the regularity of the flow map is even worse. This then also generalises to an arbitrary higher-order hierarchy equation, strengthening our conviction that low-regularity well-posedness on the torus is out of reach for any of the hierarchy equations, except for dNLS itself.

Theorem 10.12. *For any $j \geq 2$, $1 \leq r \leq \infty$ and $s < j - 1$ there exists a gauged dNLS equation (i.e. choice of coefficients $c_{k,\alpha}$) such that for the Cauchy problem* (10.1) *the flow map $S : \hat{H}^s_r(\mathbb{T}) \times (-T, T) \to \hat{H}^s_r(\mathbb{T})$ cannot be uniformly continuous on bounded subsets.*

The corresponding proof for our ill-posedness theorems on the torus consist of the Propositions 13.1 and 13.2 respectively.

We point out that these ill-posedness results, seemingly only regarding Cauchy problems in Fourier-Lebesgue spaces, suffice to also rule out well-posedness in modulation spaces on the torus. As in this periodic geometry the two families of function spaces coincide.

10.3. **Global well-posedness for the dNLS hierarchy.** Unfortunately, in contrast with the situation for the NLS hierarchy equations, we do not have a family of conservation laws equivalent to H^s norms for every $s > -\frac{1}{2}$, as were constructed in [69], at our disposal. Hence, for the moment, we are only able to upgrade our local solution to dNLS hierarchy equations to global ones at integer regularity levels. This leads to a discrepancy of at most half a derivative (exactly for the odd-indexed dNLS hierarchy equations) between our local result and the corresponding global continuation result of the solution.

Theorem 10.13. *Let $j \geq 2$. If the initial data $u_0 \in H^{\lceil \frac{j}{2} \rceil}(\mathbb{R})$ has sufficiently small L^2 norm, the solution of the jth dNLS hierarchy equation, constructed in Theorem 10.1, extends globally in time. In other words, one has small mass global well-posedness of the jth hierarchy equation with initial data in $H^{\lceil \frac{j}{2} \rceil}(\mathbb{R})$. The bound on the mass depends on j, but not on the size of the $H^{\lceil \frac{j}{2} \rceil}$-norm of the initial data.*

Proof. We extend the previously constructed local solutions classically by utilising a-priori estimates that we derive from Hamiltonians I_n as in (9.6). We remind the reader, that since we are dealing with a completely integrable hierarchy, the Hamiltonians of the hierarchy equations pairwise (Poisson) commute and are thus conserved along the flow of each other. The statement of this theorem holds true if we manage to derive an a-priori estimate on the H^k norm of a solution (of an arbitrary dNLS hierarchy equation), for $k \in \mathbb{N}$.

Guided by Lemma 9.1, in order to derive an a-priori estimate on the level of H^k we take a closer look at $I_{2k} = \int_{\mathbb{R}} u Y_{2k} \, \mathrm{d}x$, where the 'leading term' (up to constants) is given by $u \partial_x^{2k} \overline{u}$. By partial integration this term becomes equivalent to the homogeneous $\dot{H}^k$ norm. That the L^2 norm is conserved along the dNLS hierarchy equations' flows is well known. So what remains, until we may assert our desired a-priori bound on the H^k norm, is to argue that the other terms in the Hamiltonian I_{2k} cannot interfere with/cancel the leading term $|\partial_x^k u|^2$. That is, so far we have argued

$$|I_{2k}| \gtrsim \|u\|^2_{\dot{H}^k} - |\text{higher order terms}|$$

and still need to ensure that the higher order terms can be controlled by a fraction (less than 1) of $\|u\|_{\dot{H}^k}^2$.

Take such a higher-order term of the Hamiltonian I_{2k}, which in general will be of the form $\prod_{i=0}^{m}(\partial_x^{\alpha_{2i}} u)(\partial_x^{\alpha_{2i+1}}\overline{u})$, for $\alpha \in \mathbb{N}_0^{2m+2}$ with $|\alpha| = 2k - m$ and $1 \leq m \leq 2k$. (In fact, from Lemma 9.1, we know more about the structure: one of the factors u will always be without a derivative placed upon it. But we ignore this additional bit of information at this point.) Since there are strictly less than $2k$ total derivatives, there will be at most a single factor that has more than k derivatives placed upon it. Again, with partial integration, we may adjust such terms of the Hamiltonian so that every term has at most k derivatives lying upon it, with at most a single one with exactly k derivatives.

Now we may apply Hölder's inequality to such higher-order terms in the Hamiltonian (with at most k derivatives on any term) ensuring that, if there exists a factor u with k derivatives placed upon it, we put it in L^2. For the remaining factors with strictly fewer than k derivatives it doesn't matter which L^p they land in, so long as $p \geq 2$ (which is always possible, since we have $2m + 2 \geq 4$ factors).

We are now prepared to apply a special case of the Gagliardo-Nirenberg inequality (8.9). In particular we will be choosing $p = q = 2$, ℓ corresponds to the order of derivatives α_i placed on our factors and in our situation $n = 1$ holds. This leads us to deriving $\theta = \frac{1}{2k} + \frac{\alpha_i}{k} - \frac{1}{rk}$. The inequality then reads

$$\|\partial_x^{\alpha_i} u\|_{L^r} \lesssim \|u\|_{\dot{H}^k}^{\theta} \|u\|_{L^2}^{1-\theta}.$$

Applying this inequality to every factor in a higher-order term $\prod_{i=0}^{2m+1} \|\partial_x^{\alpha_i} u\|_{L^{p_i}}$ we are interested in the resulting exponent for $\|u\|_{\dot{H}^k}$. We may calculate this as follows

$$\sum_{i=0}^{2m+1} \frac{1}{2k} + \frac{\alpha_i}{k} - \frac{1}{p_i k} = \frac{2m+2}{2k} + \frac{|\alpha|}{k} - \frac{1}{k} = \frac{m}{k} + \frac{2k-m}{k} = 2,$$

where we have used the fact $\sum_{i=0}^{2m+1} \frac{1}{p_i} = 1$ and $|\alpha| = 2k - m$. For reasons of homogeneity we know the exponent of $\|u\|_{L^2}$ must thus be $2m$.

Hence at this point we have argued for an a-priori estimate of the form

$$I_{2k} \gtrsim \|u\|_{H^k}^2 (1 - c\|u\|_{L^2}^{2m}) = \|u\|_{H^k}^2 (1 - c\|u_0\|_{L^2}^{2m}) \gtrsim \|u\|_{H^k}^2, \tag{10.2}$$

where c is a fixed constant, depending on the coefficients in the dNLS hierarchy equation (corresponding to the choice of $j \in \mathbb{N}$). The final inequality holds for a sufficiently small bound on the L^2 norm of the initial data. We have thus successfully argued for an a-priori estimate on the H^k norm of solutions of dNLS hierarchy equations, conditioned on a sufficiently small initial mass. □

10.4. **Discussion.** Before moving on to proving our well- and ill-posedness results given in the previous subsection, we would like to discuss their merits and how they fit into the existing literature.

Let us begin by mentioning that our results show, that we have achieved optimal local well-posedness within the realm of our techniques, excluding the respective scaling critical Fourier-Lebesgue and modulation spaces. Specifically Theorems 10.8 and 10.9 rule out the possibility of using fixed-point methods to improve upon the well-posedness theory of the dNLS hierarchy equations beyond what we have achieved. This does not preclude the possibility of using, say, the complete integrability of those equations to lower the regularity threshold on initial data while still achieving local well-posedness. As was already implemented for the dNLS equation itself [44, 61] and recently the KdV hierarchy equations [64]. Though this approach comes with the usual caveat that the flow will be rather irregular, i.e. merely continuous, rather than Lipschitz as in our Theorems 10.1 and 10.3.

On the front of global well-posedness we were able to exploit the Hamiltonians that are conserved along the flow of dNLS hierarchy equations in order to extend our local solutions globally, at least for initial data in Sobolev spaces at integer regularities. This leaves a gap of at most half a derivative between our local and global results. It seems likely that with an application of the (first-generation) I-method it would be possible to close this gap. More generally, extending solutions globally off the scale of Sobolev spaces (i.e. Fourier-Lebesgue or modulation spaces) presents an interesting problem for further research.

We mention at this point that our local theory extends the previous best result concerning the fourth-order dNLS hierarchy equation from [48], lifting the necessity of small data. With Theorem 10.13 we extend these local solutions globally. One point of interest is, that the authors of [48] manage to achieve their result without the use of any gauge-transformation. This seems to stem from their ability to exploit the special position of one of the derivatives in the nonlinearity ∂_x being in front of every product term. See also (9.7) where we have also mentioned this fact. Further research into this possible exploitation may lead to subsequent further improvement of the regularity of the flow (of dNLS hierarchy equations), if one can do without the gauge-transformation.

The worsening of the lower bound for well-posedness by half a derivative in Sobolev spaces with every step up in the dNLS hierarchy ($j \to j+1$) is consistent with what can be observed for the similar situations of the NLS [2] or mKdV hierarchy [38].

As is unsurprising, considering the ill-posedness results already for the NLS hierarchy on the torus [2], the situation for low-regularity well-posedness theory of dNLS hierarchy equations on the torus is dire. One must hope that renormalisation/Wick-ordering or methods of complete integrability can be used in order to achieve any kind of result in this setting.

Regarding ill-posedness for the nonperiodic setting it has turned out to be surprisingly more difficult to achieve a general C^0_{unif} ill-posedness result for the dNLS hierarchy compared with either the NLS or mKdV hierarchy. Explicit soliton solutions for dNLS, which were used in [9] to show the failure of uniform continuity of the flow, are already very delicately constructed functions (evident from the complex choice of coefficients involved). Searching the literature for soliton solutions of higher-order dNLS hierarchy equations yielded only [96], which due to their evidently even more complex structure and little resemblance to the solitons of dNLS suggest that this is a difficult problem to solve in full generality.

We end this subsection by mentioning that, to the author's best knowledge, we are also the first to achieve insight into the structure of coefficients in nonlinear terms in hierarchy equations stemming from completely integrable systems, beyond knowledge of a finite number of hierarchy equations. In particular referring to Proposition 9.7, where we derived a closed form expression for the coefficients of certain nonlinear terms appearing in the dNLS hierarchy equations. Extending such results to the rest of the nonlinear terms, or more generally other hierarchies, is of great interest. This would enable more delicate analysis regarding if the complete integrability structure of the equations has significant influence on the optimal well-posedness results that can be achieved with fixed point methods.

11. Known Estimates

In order to derive our well-posedness theorems, see Section 10, we rely on proving multilinear estimates within the framework of Bourgain spaces that lead to well-posedness. To aid us in proving these multilinear estimates we will make heavy use of linear and bilinear smoothing estimates that were derived by the author in the context of the NLS hierarchy equations [2]. The multilinear $\hat{X}^r_{s,b}$ and $X^p_{s,b}$ estimates that lead to well-posedness in [2] will also be of use.

We cite the necessary estimates in the following subsection for the reader's convenience and to keep this work more self-contained.

11.1. **Smoothing and multilinear estimates.** To keep in line with how the estimates are stated in [2] we introduce the following notational convenience in this subsection: u, v and w will refer to functions in appropriate variants of Bourgain spaces adapted to a particular (linear part of an) equation and data spaces at hand so that the right hand side of the respective estimates remain finite. Keeping with the variable choice of the preceding sections $2j$, for $j \in \mathbb{N}$, will be the power in the phase function of the linear equation with which the estimates are associated.

We begin by stating linear estimates based on Kato smoothing and a maximal-function estimate.

Proposition 11.1 ([2, Proposition 4.1])**.** *Let $b > \frac{1}{2}$, then the following inequalities hold*

(1) for $2 \leq q \leq \infty$ and $\sigma > \frac{1}{2} - \frac{2j}{q}$

$$\|u\|_{L^\infty_x L^q_t} \lesssim \|u\|_{X_{\sigma,b}} \tag{11.1}$$

(2) for $2 \leq p \leq \infty$ and $\sigma = -\frac{2j-1}{2}(1 - \frac{2}{p})$

$$\|u\|_{L^p_x L^2_t} \lesssim \|u\|_{X_{\sigma,b}} \tag{11.2}$$

(3) for $4 \leq p \leq \infty$ and $\sigma > \frac{1}{2} - \frac{1}{p}$

$$\|u\|_{L^p_x L^\infty_t} \lesssim \|u\|_{X_{\sigma,b}} \tag{11.3}$$

In addition we will be making use of a Strichartz-type estimate that is more adapted (and thus more useful) to our Fourier-Lebesgue space setting, referred to most often in the literature as a Fefferman-Stein estimate. In the L^2-based setting it reduces to the well-known L^6-Strichartz estimate for (higher-order) Schrödinger equations.

Proposition 11.2 ([2, Corollary 4.6])**.** *Let $j \geq 1$, $0 \leq \frac{1}{r} < \frac{3}{4}$ and $b > \frac{1}{r}$, then one has the estimate*

$$\|I^{\frac{2(j-1)}{3r}} u\|_{L^{3r}_{xt}} \lesssim \|u\|_{\hat{X}^r_{0,b}}. \tag{11.4}$$

Moving on, we may now recall the pair of bilinear operators introduced in [2]: For $j \in \mathbb{N}$ and $1 \leq p \leq \infty$ define $I^{\pm}_{p,j}$ by its Fourier transform

$$\mathcal{F}_x I^{\pm}_{p,j}(f,g)(\xi) = c \int_* k^{\pm}_j(\xi_1,\xi_2)^{\frac{1}{p}} \hat{f}(\xi_1)\hat{g}(\xi_2)\, \mathrm{d}\xi_1 \tag{11.5}$$

where the symbol is given by

$$k^{\pm}_j(\xi_1,\xi_2) = |\xi_1 \pm \xi_2|(|\xi_1|^{2j-2} + |\xi_2|^{2j-2}). \tag{11.6}$$

We may now state the $X_{s,b}$ variant of a bilinear estimate involving our bilinear operator(s). To state the proposition we make use of the Fourier-Lebesgue space norms $\|f\|_{\widehat{L^p}} = \|\hat{f}\|_{L^{p'}}$.

Proposition 11.3 ([2, Corollary 4.3])**.** *Let $1 \leq q \leq r_{1,2} \leq p < \infty$ and $b_i > \frac{1}{r_i}$. Then we have*

$$\|I^{\pm}_{p,j}(u, v_\pm)\|_{\widehat{L^q_x}\widehat{L^p_t}} \lesssim \|u\|_{\hat{X}^{r_1}_{0,b_1}} \|v\|_{\hat{X}^{r_2}_{0,b_2}} \tag{11.7}$$

where $v_+ = \overline{v}$ and $v_- = v$.

Interpreting this bilinear operator as a multiplication operator we may find its adjoint (see [2, Section 4.2] for details) and gain an additional set of bilinear estimates associated with the adjoint. Let $I^{\pm,*}_{p,j}$ denote this adjoint. It has the symbol

$$k^{+,*}_j(\xi_1,\xi_2) = |\xi_1|(|\xi_1|^{2j-2} + |\xi_2|^{2j-2}), \quad \text{or} \tag{11.8}$$

$$k^{-,*}_j(\xi_1,\xi_2) = |\xi_1 + 2\xi_2|(|\xi_1|^{2j-2} + |\xi_2|^{2j-2}). \tag{11.9}$$

We have the following $X_{s,b}$ estimates regarding $I^{\pm,*}_{p,j}$:

Proposition 11.4 ([2, Corollary 4.4]). *Let $1 < q \le r_{1,2} \le p < \infty$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r_1} + \frac{1}{r_2}$ and $b_i > \frac{1}{r_i}$. Then the estimate*

$$\|I^{\pm,*}_{p,j}(u, v_\mp)\|_{\hat{X}^{r_1'}_{0,-b_1}} \lesssim \|u\|_{\widehat{L^{q'}_x}\widehat{L^{p'}_t}} \|v\|_{\hat{X}^{r_2}_{0,b_2}} \tag{11.10}$$

holds. If alternatively $0 \le \frac{1}{\rho'} \le \frac{1}{r'}$ and $\beta < -\frac{1}{\rho'}$ we have

$$\|I^{\pm,*}_{\rho',j}(u, v_\mp)\|_{\hat{X}^{r}_{0,\beta}} \lesssim \|u\|_{\widehat{L^{r}_{xt}}} \|v\|_{\hat{X}^{\rho'}_{0,-\beta}}. \tag{11.11}$$

In both cases $v_+ = \overline{v}$ and $v_- = v$.

Finally we will later also make use of the trilinear $X_{s,b}$ estimates that leads to well-posedness in Fourier-Lebesgue and/or modulation spaces. Recall:

Proposition 11.5 ([2, Proposition 5.1]). *Let $1 < r \le 2$, $s = \frac{j-1}{r'}$, $\alpha \in \mathbb{N}_0^3$ with $|\alpha| = 2(j-1)$. Then there exist $b' < 0$ and $b' + 1 > b > \frac{1}{r}$ such that one has*

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}. \tag{11.12}$$

Proposition 11.6 ([2, Proposition 5.6]). *Let $j \ge 2$, $2 \le p < \infty$, $s = \frac{j-1}{2}$, $\alpha \in \mathbb{N}_0^3$ with $|\alpha| = 2(j-1)$. Then there exist $b' < 0$ and $b' + 1 > b > \frac{1}{2}$ such that one has*

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{X^p_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{X^p_{s,b}}. \tag{11.13}$$

11.2. **Basic estimate on the resonance relation.** As mentioned in the introduction, the additional derivative in the nonlinear terms of dNLS hierarchy equations adds difficulty (over the NLS hierarchy equations) in their analysis. Additional arguments are necessary to overcome this difficulty. The first step in this direction was the introduction and use of the gauge-transformation in order to simplify, or more precisely, remove ill-behaved terms from, the equations. See Section 9.2.

The second step we take in tackling well-posedness estimates for the dNLS hierarchy equations is exploiting the resonance relation, the effectiveness of which was already demonstrated in [37]. In the absence of an analogue of the exact factorisation for the resonance relation for higher-order dNLS hierarchy equations one may still recover the essence:

Lemma 11.7 ([31, Lemma 2.3]). *Let $\alpha > 1$, $\xi_1, \xi_2, \xi_3 \in \mathbb{R}$ and set $\xi = \xi_1 + \xi_2 + \xi_3$. Then one has*

$$||\xi|^\alpha - |\xi_1|^\alpha + |\xi_2|^\alpha - |\xi_3|^\alpha| \gtrsim |\xi_1 + \xi_2||\xi_2 + \xi_3||\xi_{max}|^{\alpha-2}, \tag{11.14}$$

where $\xi_{max} = \max\{|\xi_1|, |\xi_2|, |\xi_3|, |\xi|\}$.

Using this estimate, in combination with the flexibility $X_{s,b}$ spaces offer, will suffice in order to derive the multilinear estimates that lead to well-posedness we are after.

12. Estimates leading to well-posedness

With all necessary smoothing estimates that we will need recalled, as well as previous $X_{s,b}$ estimates that we will want to make use of, we are ready to prove the propositions that serve as proof of our Theorem 10.5 and 10.6. The discussion of the gauge-transformation in Section 9.2 combined with these Theorems then also suffice to argue the validity of Theorem 10.1 and 10.3.

As is the case for the NLS hierarchy equations, cubic nonlinear terms are more difficult to deal with than their quintic and higher-order counterparts. So we will be dealing with cubic and higher-order terms separately.

12.1. Multilinear estimates in $\hat{X}^r_{s,b}$ spaces.

Proposition 12.1. *Let $j \geq 2$, $1 < r \leq 2$, $s \geq \frac{1}{2} + \frac{j-1}{r'}$, and*

$$\alpha \in \left\{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{N}_0^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = 2j-1,\, \alpha_2 \neq 0\right\},$$

then there exist $b' < 0 < \frac{1}{r} < b < b'+1$ such that the following estimate holds:

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}. \tag{12.1}$$

Proof. It suffices to prove this estimate for $s = \frac{1}{2} + \frac{j-1}{r'}$ fixed. For the proof we want to rely, for the most difficult frequency constellations, on the cubic estimate in Proposition 11.5, which was proven in the author's previous work on the NLS hierarchy [2]. Relying on the 'equivalent' NLS estimate to prove well-posedness for dNLS was already a successful technique employed in [37]. Though in addition to the arguments presented there we have to utilise the full gain of the modulation in order to close the estimate.

In particular, one is able to re-use the NLS estimate, if the frequencies ξ, ξ_1, ξ_2, ξ_3 allow for the following inequality:

$$\langle\xi\rangle^s |\xi_1|^{\alpha_1} |\xi_2|^{\alpha_2} |\xi_3|^{\alpha_3} \lesssim \langle\xi\rangle^{s-\frac{1}{2}} \langle\xi_1\rangle^{\alpha_1+\frac{1}{2}} \langle\xi_2\rangle^{\alpha_2-\frac{1}{2}} \langle\xi_3\rangle^{\alpha_3+\frac{1}{2}}. \tag{12.2}$$

This is also where it becomes relevant that we assume that at least a single derivative falls on $\overline{u_2}$. Otherwise $\alpha_2 - \frac{1}{2}$ may become negative which would in turn require far more detailed analysis, since the NLS estimate could not be as easily applied.

Furthermore, from here on we will assume, by symmetry, that u_1 has larger frequency than u_3 and the largest frequency of $|\xi_1|$, $|\xi_2|$, and $|\xi_3|$ shall synonymously be known as $|\xi_{max}|$.

(1) (12.2) holds. In this case we may use the inequality (12.2) and 'reinterpret' the cubic nonlinearity as one how it would appear in an NLS hierarchy equation:

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \|\partial_x^{\alpha_1}(J^{\frac{1}{2}} u_1) \partial_x^{\alpha_2-1}(J^{\frac{1}{2}} \overline{u_2}) \partial_x^{\alpha_3}(J^{\frac{1}{2}} u_3)\|_{\hat{X}^r_{s-\frac{1}{2},b'}} \tag{12.3}$$

$$\lesssim \prod_{i=1}^{3} \|J^{\frac{1}{2}} u_i\|_{\hat{X}^r_{s-\frac{1}{2},b}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}, \tag{12.4}$$

where we were then immediately able to apply (11.12) and arrive at our desired upper bound.

But when is (12.2) true, i.e. which other cases do we still have to deal with?

- It certainly holds if $|\xi_2|\langle\xi\rangle \lesssim \langle\xi_1\rangle\langle\xi_3\rangle$, as from this (12.2) is quite immediate.
- When $|\xi_2| \lesssim 1$ or $|\xi| \lesssim 1$ then (12.2) must also hold. This is because either the frequency $|\xi_2|$ is negligible and can easily be traded against ξ_1 or ξ_3, or because there exist at least two high-frequency factors between which derivatives can be traded painlessly.

So in all other cases that follow this one we may assume, without loss of generality, that $|\xi| \sim \langle \xi \rangle$, $|\xi_2| \sim \langle \xi_2 \rangle$, and $\langle \xi \rangle \langle \xi_2 \rangle \gg \langle \xi_1 \rangle \langle \xi_3 \rangle$, and we will do so without further mention. Further we will also be showcasing the estimate on condition that the modulation of the product is maximal $\langle \sigma_0 \rangle = \langle \sigma_{max} \rangle$. Cases where the modulations of individual factors are maximal can be proven analogously since the remaining cases are non- or at most semi-resonant.

(2) $|\xi_1| = |\xi_{max}|$. In this case, because of $\langle \xi \rangle \langle \xi_2 \rangle \gg \langle \xi_1 \rangle \langle \xi_3 \rangle$, it must hold that $\langle \xi \rangle \gg \langle \xi_3 \rangle$ which in turn implies $\langle \xi \rangle \lesssim \langle \xi_1 + \xi_2 \rangle$. From this we may also derive

$$|\xi \xi_2| \lesssim |(\xi_1 + \xi_2)\xi_2| \lesssim |\xi_{max}||\xi_1 + \xi_2| = |\xi_1(\xi_1 + \xi_2)|. \tag{12.5}$$

Further, using our general estimate for the modulation Lemma 11.7 we have

$$\langle \sigma_0 \rangle \gtrsim |\xi_1 + \xi_2||\xi_2 + \xi_3||\xi_{max}|^{2j-2} \gtrsim |\xi \xi_2||\xi_1|^{2j-2} \tag{12.6}$$

at our disposal. With all preparations done we may focus on proving the estimate.

As our first step we shift all derivatives of the product, except for one guaranteed to lie on u_2, to u_1 and use our estimate for the modulation (12.6).

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \|\Lambda^{b'} J^{s+\frac{1}{2}}((I^{|\alpha|-\frac{3}{2}} u_1)(I^{\frac{3}{2}} \overline{u_2})(I^{-\frac{1}{2}} u_3))\|_{\widehat{L^r_{xt}}} \tag{12.7}$$

$$\lesssim \|J^{s+\frac{1}{2}-\frac{1}{r'}+}((I^{|\alpha|-\frac{3}{2}-\frac{2j-2}{r'}+} u_1)(I^{\frac{3}{2}-\frac{1}{r'}+} \overline{u_2})(I^{-\frac{1}{2}} u_3))\|_{\widehat{L^r_{xt}}} \tag{12.8}$$

Now using (12.5) we may shift derivatives again to arrive at:

$$\lesssim \|J^{s+\frac{1}{2}-\frac{1}{r'}-\frac{1}{r}+}(I^{\frac{1}{r}}((I^{|\alpha|-\frac{3}{2}-\frac{2j-2}{r'}+\frac{1}{r}+} u_1)(I^{\frac{3}{2}-\frac{1}{r'}-\frac{1}{r}+} \overline{u_2}))(I^{-\frac{1}{2}} u_3))\|_{\widehat{L^r_{xt}}}, \tag{12.9}$$

where we are now ready to upgrade $I^{\frac{1}{r}}$ to our well-known bilinear operator $I^+_{r,j}$ and then apply its corresponding estimate, after dealing with u_3 by Hölder's inequality.

$$\lesssim \|J^{s-\frac{1}{2}+}(I^{\frac{1}{r}}((I^{|\alpha|-\frac{3}{2}-\frac{2j-2}{r'}+\frac{1}{r}+} u_1)(I^{\frac{1}{2}+} \overline{u_2}))(I^{-\frac{1}{2}} u_3))\|_{\widehat{L^r_{xt}}} \tag{12.10}$$

$$\lesssim \|I^+_{r,j}(I^{|\alpha|+s-2-(2j-2)+\frac{1}{r}+} u_1, I^{\frac{1}{2}+} \overline{u_2})\|_{\widehat{L^r_{xt}}} \|I^{-\frac{1}{2}} u_3\|_{\widehat{L^\infty_{xt}}} \tag{12.11}$$

$$\lesssim \|J^{s-1+\frac{1}{r}+} u_1\|_{\hat{X}^r_{0,b}} \|I^{\frac{1}{2}+} u_2\|_{\hat{X}^r_{0,b}} \|I^{\frac{1}{r}-\frac{1}{2}+} u_3\|_{\hat{X}^r_{0,b}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}} \tag{12.12}$$

The last inequality holds, so long as $\frac{1}{r} - 1 < 0$ and $s \geq \frac{1}{2} + \frac{j-1}{r'} \geq \max(\frac{1}{2}+, \frac{1}{r} - \frac{1}{2}+)$, which is the case for $r > 1$.

(3) $|\xi_2| = |\xi_{max}|$. The argument in this case is quite similar to the preceding case, only that now we do not have to account for the guaranteed derivative on u_2 as this is the high-frequency factor to which we shift all derivatives anyway.

When $|\xi_2|$ is maximal it follows that $|\xi_2| \sim |\xi| \sim |\xi_1 + \xi_2| \sim |\xi_2 + \xi_3|$ and for the modulation, again using Lemma 11.7, we can estimate $\langle \sigma_0 \rangle \gtrsim |\xi_1 + \xi_2||\xi_2 + \xi_3||\xi_{max}|^{2j-2} \gtrsim |\xi_2|^{2j}$. For proving our estimate this leads us to

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{\hat{X}^r_{s,b'}} \lesssim \|\Lambda^{b'}(u_1(I^{|\alpha|+s} \overline{u_2}) u_3)\|_{\widehat{L^r_{xt}}} \tag{12.13}$$

$$\lesssim \|(I^{-\frac{1}{2}} u_1) I^{\frac{1}{r}}((I^{|\alpha|+s+\frac{1}{2}-\frac{1}{r}-\frac{2j}{r'}+} \overline{u_2}) u_3)\|_{\widehat{L^r_{xt}}} \tag{12.14}$$

$$\lesssim \|I^{-\frac{1}{2}} u_1\|_{\widehat{L^\infty_{xt}}} \|I^{\frac{1}{r}}((I^{|\alpha|+s+\frac{1}{2}-\frac{1}{r}-\frac{2j}{r'}+} \overline{u_2}) u_3)\|_{\widehat{L^r_{xt}}}$$

$$\lesssim \|I^{\frac{1}{r}-\frac{1}{2}+} u_1\|_{\hat{X}^r_{0,b}} \|I^+_{r,j}(I^{|\alpha|+s+\frac{1}{2}-\frac{1}{r}-\frac{2j}{r'}-\frac{2j}{r}+} \overline{u_2}, u_3)\|_{\widehat{L^r_{xt}}} \tag{12.15}$$

$$\lesssim \|I^{\frac{1}{r}-\frac{1}{2}+} u_1\|_{\hat{X}^r_{0,b}} \|I^{s-\frac{1}{2}-\frac{1}{r}+} \overline{u_2}\|_{\hat{X}^r_{0,b}} \|u_3\|_{\hat{X}^r_{0,b}} \lesssim \prod_{i=1}^{3} \|u_i\|_{\hat{X}^r_{s,b}}, \tag{12.16}$$

which is the desired upper bound, if $r > 1$, so the proof is complete.

□

Since in the proof of the necessary quintilinear (and higher-order) estimate to argue our well-posedness Theorems we rely on the fact $s < \frac{1}{r}$, we will argue the estimate for the full range of parameter $1 < r \leq 2$ in two parts. First we will prove Proposition 12.2 below, which for a comparatively higher level of regularity establishes the multilinear estimate near the endpoint $r \to 1$. This we can then in turn interpolate with the L^2-based estimate that is part of Proposition 12.4 in order to cover the full parameter range.

Proposition 12.2. *Let $j \geq 2$, $2 \leq k \leq 2j$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2j - k$. Then there exists an $1 < r_0 \ll 2$ such that for all $1 < r < r_0$ and $s > \frac{1}{2} + \frac{j-k}{kr'}$ there exist $b' < 0 < \frac{1}{r} < b < b' + 1$ such that the following estimate holds*

$$\| \prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i \|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{\hat{X}^r_{s,b}}, \tag{12.17}$$

where exactly k of the factors $v_1, v_2, \ldots, v_{2k+1}$ are equal to $\overline{u_i}$ and otherwise just equal to u_i.

Proof. Without loss of generality we may assume that the frequencies of the $2k+1$ factors are ordered decreasingly, i.e. $|\xi_1| \geq |\xi_2| \geq \ldots |\xi_{2k+1}|$. We will analyse the product based on the number of high-frequency factors present.

Throughout the proof we will need to make use of the fact $s - \frac{1}{r} < 0$, which we may achieve by choosing $1 < r_0 \ll 2$ appropriately small.

(1) $|\xi_4| \gtrsim |\xi_1|$, so we have at least 4 high-frequency factors. In this case every factor of the product passes through a norm that is invariant with respect to complex conjugation, so we may ignore its distribution among the factors in this case.

The idea of the proof in this case is to use the Fefferman-Stein estimate for the high-frequency factors (of which we need 4 in order to ensure its applicability) and a Sobolev-type embedding for the rest. We start by distributing the derivatives of the norm and those in the product on the high-frequency factors in addition to leaving a little leeway for embeddings on the remaining factors. Then we use the Hausdorff-Young inequality to 'remove the hat' from the space.

$$\| \prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i \|_{\hat{X}^r_{s,b'}} \lesssim \|(J^\sigma u_1)(J^\sigma u_2)(J^\sigma u_3)(J^\sigma u_4) \prod_{i=5}^{2k+1} J^{s-\frac{1}{r}-} u_i \|_{L^r_{xt}} \tag{12.18}$$

This requires $\sigma \geq 0$, $4\sigma + (2k-3)(s - \frac{1}{r}) \geq s + 2j - k$ as well as $s - \frac{1}{r} - < 0$ the latter of which is ensured by our choice of r_0 at the beginning of this proof. By now using Hölder's inequality, Hausdorff-Young again (to put the hat back on L^∞) and a Sobolev-type embedding we arrive at:

$$\lesssim \prod_{i=1}^{4} \|J^\sigma u_i\|_{L^{4r}_{xt}} \prod_{i=5}^{2k+1} \|J^{s-\frac{1}{r}-} u_i\|_{\widehat{L^\infty_{xt}}} \lesssim \prod_{i=1}^{4} \|J^\sigma u_i\|_{L^{4r}_{xt}} \prod_{i=5}^{2k+1} \|u_i\|_{\hat{X}^r_{s,b}} \tag{12.19}$$

As announced before, we may now use the Fefferman-Stein estimate (11.4) which grants us a gain of $\frac{2(j-1)}{4r}$ derivatives on each of the high-frequency factors, but leaves us in the wrong $\hat{X}^{\frac{4r}{3}}_{0,\frac{3}{4r}+}$ space. To remedy this we may use a Sobolev-type inequality for which we have to spend $\frac{1}{4r}+$ derivatives.

$$\lesssim \prod_{i=1}^{4} \|J^{\sigma - \frac{2(j-1)}{4r} + \frac{1}{4r}+} u_i\|_{\hat{X}^r_{0,b}} \prod_{i=5}^{2k+1} \|u_i\|_{\hat{X}^r_{s,b}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{\hat{X}^r_{s,b}} \tag{12.20}$$

The reader my verify that for $s > \frac{1}{2} + \frac{j-k}{kr'}$ and the choice $\sigma = s + \frac{2(j-1)}{4r} - \frac{1}{4r} -$ the requirements gathered involving σ can be fulfilled and we may justify the final inequality to arrive at the desired upper bound.

(2) $|\xi| \sim |\xi_1| \gg |\xi_2|$, so we have only a single high-frequency factor.

One needs to take care as to what the distribution of complex conjugates in the product is. We will showcase a proof of the estimate in the instance that the product we are dealing with is equal to $\overline{u_1}(\prod_{i=2}^{2k-3} v_i) u_{2k-2} u_{2k-1} u_{2k} \overline{u_{2k+1}}$ (ignoring derivatives). This aligns with the requirement, that k of the factors are complex conjugates. The other cases, for different distributions of complex conjugates can be dealt with in a similar fashion and we omit the details.

With only a single large frequency we have immediate control over the symbols of the bilinear operators $I^{+}_{r,j}$ and $I^{-,*}_{\rho',j}$. We proceed by shifting all derivatives of the norm and in the product onto the high-frequency factor in addition to some extra derivatives we will later need for Sobolev-type embeddings.

$$\| \prod_{i=1}^{2k+1} \partial_x^{\alpha_i} v_i \|_{\hat{X}^r_{s,b'}} \lesssim \| J^{\sigma} \overline{u_1} (\prod_{i=2}^{2k-1} J^{s-\frac{1}{r}-} v_i) J^s u_{2k} J^{s-\frac{1}{r}+\frac{1}{\rho'}-} \overline{u_{2k+1}} \|_{\hat{X}^r_{0,b'}} \tag{12.21}$$

Here we have introduced $\sigma \geq 0$ and this inequality holds so long as $\sigma + (2k-2)(s - \frac{1}{r}) + s + (s - \frac{1}{r} + \frac{1}{\rho'}) > s + 2j - k$. Furthermore $s - \frac{1}{r} - < 0$ is ensured by our choice of r_0 at the beginning of this proof.

Next we introduce the bilinear operator $I^{+}_{r,j}$ which grants us a gain of $\frac{2j-1}{r}$ derivatives on the high-frequency factor:

$$\lesssim \| I^{+}_{r,j} (J^{\sigma - \frac{2j-1}{r}} \overline{u_1}, J^s u_{2k}) (\prod_{i=2}^{2k-1} J^{s-\frac{1}{r}-} v_i) J^{s-\frac{1}{r}+\frac{1}{\rho'}-} \overline{u_{2k+1}} \|_{\hat{X}^r_{0,b'}} \tag{12.22}$$

$$\lesssim \| I^{-,*}_{\rho',j} (I^{+}_{r,j} (J^{\sigma - \frac{2j-1}{r} - \frac{2j-1}{\rho'}} \overline{u_1}, J^s u_{2k}) \prod_{i=2}^{2k-1} J^{s-\frac{1}{r}-} v_i, J^{s-\frac{1}{r}+\frac{1}{\rho'}-} \overline{u_{2k+1}}) \|_{\hat{X}^r_{0,b'}} \tag{12.23}$$

Choosing $\frac{1}{\rho'} \leq \frac{1}{r}$ with $\frac{2(j-k)}{r} < \frac{2j}{\rho'}$ we now also introduce its dual $I^{-,*}_{\rho',j}$, which grants us $\frac{2j-1}{\rho'}$ on the high-frequency factor. Now applying the continuity of the dual bilinear operator (11.10) and Hölder's inequality we may derive

$$\lesssim \| I^{+}_{r,j} (J^{\sigma - \frac{2j-1}{r} - \frac{2j-1}{\rho'}} \overline{u_1}, J^s u_{2k}) \prod_{i=2}^{2k-1} J^{s-\frac{1}{r}-} v_i \|_{\widehat{L^r_{xt}}} \| J^{s-\frac{1}{r}+\frac{1}{\rho'}-} u_{2k+1} \|_{\hat{X}^{\rho'}_{0,-b'}} \tag{12.24}$$

$$\lesssim \| I^{+}_{r,j} (J^{\sigma - \frac{2j-1}{r} - \frac{2j-1}{\rho'}} \overline{u_1}, J^s u_{2k}) \|_{\widehat{L^r_{xt}}} \prod_{i=2}^{2k-1} \| J^{s-\frac{1}{r}-} u_i \|_{\widehat{L^{\infty}_{xt}}} \| J^{s-\frac{1}{r}+\frac{1}{\rho'}-} u_{2k+1} \|_{\hat{X}^{\rho'}_{0,-b'}}$$

For the first factor we apply the continutiy of the bilinear operator (11.7), for the factors in the product we use a Sobolev-type embedding and for the final factor Young's inequality:

$$\lesssim \| J^{\sigma - \frac{2j-1}{r} - \frac{2j-1}{\rho'}} u_1 \|_{\hat{X}^r_{0,b}} \| J^s u_{2k} \|_{\hat{X}^r_{0,b}} \prod_{i=2}^{2k-1} \| J^s u_i \|_{\hat{X}^r_{0,b}} \| J^s u_{2k+1} \|_{\hat{X}^r_{0,b}} \tag{12.25}$$

By choosing $\sigma = \frac{2k-1}{r} + 2j - k - (2k-1)s - \frac{1}{\rho'} + > 0$ and our choice of ρ' the reader may verify that $\sigma - \frac{2j-1}{r} - \frac{2j-1}{\rho'} < s$ and that the other requirements with respect to σ are fulfilled. Hence we have accomplished the proof of the estimate in this case.

(3) $|\xi_2| \gtrsim |\xi_1| \gg |\xi_3|$ or $|\xi_3| \gtrsim |\xi_1| \gg |\xi_4|$, so we have two or three high-frequency factors. In this case, also depending on which factors are complex conjugates, we may parenthesise differently in use of the bilinear operator to the preceding case. Different distributions of complex conjugates may be dealt with by using either $I^+_{r,j}$ or $I^-_{r,j}$ (and their duals) appropriately. The arguments are similar to case we have already dealt with, so we choose to omit the details.

□

12.2. **Multilinear estimates in $X^p_{s,b}$ spaces.** The proof of the cubic estimate in modulation space-based $X_{s,b}$ spaces is very similar to the proof of Proposition 12.1 (in the $r = 2$ case), where the equivalent estimate for Fourier-Lebesgue-based spaces is showcased. We choose to omit the details that are analogous and only show the necessary additional arguments.

Proposition 12.3. *Let $j \geq 2$, $2 \leq p < \infty$, $s \geq \frac{j}{2}$, and*

$$\alpha \in \left\{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{N}_0^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = 2j-1,\ \alpha_2 \neq 0\right\},$$

then there exist $b' < 0 < \frac{1}{r} < b < b' + 1$ such that one has the estimate

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{X^p_{s,b'}} \lesssim \prod_{i=1}^{3} \|u_i\|_{X^p_{s,b}}. \tag{12.26}$$

Proof. Main idea of the proof is again to reuse the corresponding NLS estimate for cubic terms (that is Proposition 11.6 in this case) in the most difficult resonant cases. We argue along the lines of the first case in the proof of Proposition 12.1, replacing any mention of an $\hat{X}^r_{s,b}$ space with the appropriate $X^p_{s,b}$ space.

What is left is to argue the remaining two cases where either ξ_1 or ξ_2 is the maximal frequency (here we have also adopted the convention that the frequency of u_1 is greater than that of u_3, without loss).

For both cases we begin by using the trivial embedding $X^p_{s,b'} \supset X_{s,b'}$, so that we may reuse what was argued in the $r = 2$ case in Proposition 12.1. Following along the lines of the proof one arrives at a bound

$$\|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{X^p_{s,b'}} \lesssim \|\partial_x^{\alpha_1} u_1 \partial_x^{\alpha_2} \overline{u_2} \partial_x^{\alpha_3} u_3\|_{X_{s,b'}} \tag{12.27}$$

$$\lesssim \|J^{s-\frac{1}{2}+} u_1\|_{\hat{X}_{0,b}} \|u_2\|_{\hat{X}_{0+,b}} \|u_3\|_{\hat{X}_{\frac{1}{2}+,b}}, \tag{12.28}$$

where possibly the roles of u_1, u_2 and u_3 are interchanged depending on the exact case (i.e. $|\xi_1| = |\xi_{max}|$ or $|\xi_2| = |\xi_{max}|$). Now, using the Sobolev-type embedding for modulation spaces (8.6), we may bound this by our desired right-hand side so long as $s - \frac{1}{2} + \frac{1}{2} - \frac{1}{p} < s$ and $\frac{1}{2} + \frac{1}{2} - \frac{1}{p} < s$, which can be achieved for $p < \infty$, $j \geq 2$ and $s \geq \frac{j}{2}$ as claimed. □

Proposition 12.4. *Let $j \geq 2$, $2 \leq p \leq \infty$, $2 \leq k \leq 2j$, $s > \frac{1}{2} + \frac{1}{4k} - \frac{2k+1}{2kp}$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2j - k$, then there exist $b' < 0 < \frac{1}{2} < b < b' + 1$ such that the following estimate holds:*

$$\|\prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i\|_{X^p_{s,b'}} \lesssim \prod_{i=1}^{2k+1} \|u_i\|_{X^p_{s,b}}. \tag{12.29}$$

Additionally, the distribution of complex conjugates on the factors u_i may be chosen arbitrarily.

Proof. We will assume, without loss of generality by symmetry, that the frequencies of the factors in the product are order decreasingly, i.e. $|\xi_1| \geq |\xi_2| \geq \ldots \geq |\xi_{2k+1}|$.

Depending on if the largest frequency of one of the factors is comparable (or not) to the frequency of the product we differentiate between two cases.

The reader may note that in both cases each factor passes through a mixed $L_x^p L_t^q$ which is invariant with respect to complex conjugation. This justifies the addition to the theorem, that the distribution of complex conjugates may be chosen arbitrarily.

Idea of the proof is to use reduce the proof to the L^2 case, where Kato smoothing for two of the 'factors' and the maximal function estimate for the rest is used, and to then use a Sobolev-type embedding to get back to the correct modulation space. The latter is what leads to the restriction on s, i.e. in the L^2 case we reach scaling up to an epsilon.

(1) $|\xi| \sim |\xi_1|$. Since u_1 is the factor with the largest frequency, comparable with the product itself, we may redistribute all derivatives in the product accordingly. In the same step we use the trivial embedding $X^p_{s,b'} \supset X^2_{s,b'}$, for $p \geq 2$, and introduce $\sigma \geq 0$ to be choosen later as well as $r = \infty -$.

$$\| \prod_{i=1}^{2k+1} \partial_x^{\alpha_i} u_i \|_{X^p_{s,b'}} \lesssim \| (J^{s+2j-k-\frac{2j-1}{2}(1-\frac{2}{r})+\sigma} u_1) \prod_{i=2}^{2k+1} J^{-\frac{\sigma}{2k}} u_i \|_{X_{\frac{2j-1}{2}(1-\frac{2}{r}),b'}} \tag{12.30}$$

Now we may use the modulation with exponent $b' = -\frac{1}{2}+$ by applying the dual version of Kato's smoothing estimate (11.2), followed by an application of Hölder's inequality.

$$\lesssim \| (J^{s+2j-k-\frac{2j-1}{2}(1-\frac{2}{r})+\sigma} u_1) \prod_{i=2}^{2k+1} J^{-\frac{\sigma}{2k}} u_i \|_{L_x^{r'} L_t^2} \tag{12.31}$$

$$\lesssim \| J^{s+2j-k-\frac{2j-1}{2}(1-\frac{2}{r})+\sigma} u_1 \|_{L_x^\infty L_t^2} \prod_{i=2}^{2k+1} \| J^{-\frac{\sigma}{2k}} u_i \|_{L_x^{2kr'} L_t^\infty} \tag{12.32}$$

Another application of Kato's smoothing inequality (11.2) for the first factor and the maximal function estimate (11.3) leads us to:

$$\lesssim \| J^{s+2j-k-\frac{2j-1}{2}(2-\frac{2}{r})+\sigma} u_1 \|_{X_{0,b}} \prod_{i=2}^{2k+1} \| J^{-\frac{\sigma}{2k}+\frac{1}{2}-\frac{1}{2kr'}+} u_i \|_{X_{0,b}} \tag{12.33}$$

$$\lesssim \| J^{s+2j-k-\frac{2j-1}{2}(2-\frac{2}{r})+\sigma+\frac{1}{2}-\frac{1}{p}+} u_1 \|_{X^p_{0,b}} \prod_{i=2}^{2k+1} \| J^{-\frac{\sigma}{2k}+\frac{1}{2}-\frac{1}{2kr'}+\frac{1}{2}-\frac{1}{p}+} u_i \|_{X^p_{0,b}} \tag{12.34}$$

where we have applied the Sobolev-type embedding (8.6) to each of the factors. This product as a whole may be bounded by our desired right hand side in (12.29) on condition that

$$2j - k - \frac{2j-1}{2}(2 - \frac{2}{r}) + \sigma + \frac{1}{2} - \frac{1}{p} < 0 \qquad \text{and} \tag{12.35}$$

$$-\frac{\sigma}{2k} + \frac{1}{2} - \frac{1}{2kr'} + \frac{1}{2} - \frac{1}{p} < s. \tag{12.36}$$

We leave it to the reader to verify that, so long as $s > \frac{1}{2} + \frac{1}{4k} - \frac{2k+1}{2kp}$, these inequalities hold, if one chooses $\sigma = k - \frac{3}{2} + \frac{1}{p} -$ which clearly also fulfils $\sigma \geq 0$.

(2) $|\xi| \ll |\xi_1|$ so that we must have $|\xi_1| \sim |\xi_2|$. The proof in this case is similar in spirit to the preceding case, only that, since the frequency of the product is small, it is more beneficial to apply Kato's smoothing inequality to the first two factors.

After redistributing derivatives beneficially and moving to L^2-based Bourgain spaces as above, we use the modulation of the product (with exponent $b' = -\frac{1}{2}+$)

for a Sobolev embedding in time. In the space variable we also sacrifice a total of $\frac{1}{2}-$ derivatives for a Sobolev embedding to L^{1+}.

$$\|\prod_{i=1}^{2k+1}\partial_x^{\alpha_i}u_i\|_{X^p_{s,b'}} \lesssim \|(J^{\frac{s}{2}+\sigma_1+\frac{1}{4}-}u_1)(J^{\frac{s}{2}+\sigma_1+\frac{1}{4}-}u_2)\prod_{i=3}^{2k+1}J^{\sigma_2}u_i\|_{L^{1+}_{xt}} \tag{12.37}$$

Here we have introduced $\sigma_1 \geq 0$ and $\sigma_2 \leq 0$ which are to be chosen later under the constraint $2\sigma_1 + (2k-1)\sigma_2 = 2j-k$. Next we may apply Hölder's inequality in preparation for applications of Kato smoothing (11.1) for the first two factors. For the remaining factors one has to be careful: Either one can apply the maximal function estimate (11.3) if one has enough factors, that is $k > 3$, or one resorts to using a Sobolev embedding which works just as well for $k = 2$ or $k = 3$.

$$\lesssim \|J^{\frac{s}{2}+\sigma_1+\frac{1}{4}-}u_1\|_{L^\infty_x L^{2+}_t}\|J^{\frac{s}{2}+\sigma_1+\frac{1}{4}-}u_2\|_{L^\infty_x L^{2+}_t}\prod_{i=3}^{2k+1}\|J^{\sigma_2}u_i\|_{L^r_x L^\infty_t} \tag{12.38}$$

$$\lesssim \|J^{\frac{s}{2}+\sigma_1+\frac{1}{4}+\frac{1}{2}-\frac{2j}{2+}-}u_1\|_{X_{0,b}}\|J^{\frac{s}{2}+\sigma_1+\frac{1}{4}+\frac{1}{2}-\frac{2j}{2+}-}u_2\|_{X_{0,b}}\prod_{i=3}^{2k+1}\|J^{\sigma_2+\frac{1}{2}-\frac{1}{r}+}u_i\|_{X_{0,b}}$$

Here we have introduced r such that $\frac{2k}{r} = \frac{1}{1+} = 1-$ in an intermediate step. This final product may again be bounded by our desired right hand side in (12.29) after an application of the Sobolev-type embedding for modulation spaces (8.6), if the following conditions are met:

$$\sigma_1 + \frac{1}{4} + \frac{1}{2} - \frac{2j}{2} + \frac{1}{2} - \frac{1}{p} < \frac{s}{2} \qquad \text{and} \qquad \sigma_2 + \frac{1}{2} - \frac{1}{r} + \frac{1}{2} - \frac{1}{p} < s. \tag{12.39}$$

By choosing

$$\sigma_1 = j - \frac{3}{4} - \frac{2k-1}{8k} + \frac{2k-1}{4kp} \qquad \text{and} \qquad \sigma_2 = -\frac{1}{2} + \frac{1}{4k} - \frac{1}{2kp} + \frac{1}{2k-1} \tag{12.40}$$

one may verify that these conditions (and those placed upon σ_1 and σ_2) are met so long as $s > \frac{1}{2} + \frac{1}{4k} - \frac{2k+1}{2kp}$ and thus the proof is complete.

□

From Propositions 12.3 and 12.4, possibly also using the gauge-transformation, the well-posedness Theorems we mentioned at the beginning of this section are now immediate from general theory on $X^p_{s,b}$ spaces. See [2, Section 1.2] for references on this matter.

Moving back to estimates in Fourier-Lebesgue-based spaces, we may now use the L^2-based (that is $p = 2$) estimate that is contained in Proposition 12.4 and interpolate (by the complex multilinear interpolation method) with the near-endpoint estimate from Proposition 12.2 in order to cover the full parameter range $1 < r \leq 2$ that is necessary to argue our well-posedness Theorems in such spaces.

Corollary 12.5. *Let $j \geq 2$, $1 < r \leq 2$, $2 \leq k \leq 2j$, $s > \frac{1}{r} - \frac{1}{2}$ and $\alpha \in \mathbb{N}_0^{2k+1}$ with $|\alpha| = 2j-k$. Then there exist $b' < 0 < \frac{1}{r} < b < b'+1$ such that the following estimate holds*

$$\|\partial_x^{\alpha_1}u_1\prod_{i=1}^{k}\partial_x^{\alpha_{2i}}\overline{u_{2i}}\partial_x^{\alpha_{2i+1}}u_{2i+1}\|_{\hat{X}^r_{s,b'}} \lesssim \prod_{i=1}^{2k+1}\|u_i\|_{\hat{X}^r_{s,b}}. \tag{12.41}$$

The Theorems mentioned at the beginning of this section regarding well-posedness in Fourier-Lebesgue spaces may now be derived from Proposition 12.1 and Corollary 12.5, possibly in combination with use of the gauge-transformation, with standard theory on $\hat{X}^r_{s,b}$ spaces. See [2, Section 1.2] for references on this matter.

13. Proofs of ill-posedness results

With our well-posedness results established, we now proceed to demonstrate that these results are, in a certain sense, optimal. Specifically, the following arguments will prove Theorems 10.8 and 10.9, showing that it is impossible to achieve well-posedness for the equations of interest below the regularity threshold we have already identified using the direct application of the contraction mapping theorem. Additionally, we will show that for periodic initial data, achieving analogous results to those in the nonperiodic case from the previous sections is also unfeasible with the contraction mapping principle.

The argument we use was initially investigated in [14] and then later refined in [92]. By now it has found widespread use to show ill-posedness results for power-type nonlinearities appearing in a wide variety of dispersive equations.

Proof of Theorem 10.8. Let us assume that the flow $S : \hat{H}^s_r(\mathbb{R}) \times (-T, T) \to \hat{H}^s_r(\mathbb{R})$ of the Cauchy problem (10.1) for a general nonlinearity $N(u)$ containing a cubic term with $2j-1$ derivatives is thrice continuously differentiable. (See the discussion in Remark 10.10 for what 'general nonlinearity' means.) We will as a necessary condition on the regularity of the initial data that $s \geq \frac{1}{2} + \frac{j-1}{r'}$.

For initial datum $u_0(x) = \delta\phi(x)$, where $\delta > 0$ and $\phi \in \hat{H}^s_r(\mathbb{R})$ for any $1 \leq r \leq \infty$ and $s \in \mathbb{R}$ are to be chosen later, we calculate the third derivative of the flow at the origin. Let u denote the solution corresponding to u_0 as initial data, then

$$\frac{\partial^3 u}{\partial \delta^3}\bigg|_{\delta=0} \sim \int_0^t U(t-t')N_3(U(t')u_0)\,\mathrm{d}t', \tag{13.1}$$

where we use $U(t)$ to denote the linear propagator of our equation and $N_3(u)$ to refer only to the cubic nonlinear terms in the nonlinearity of our equation. The higher-order nonlinear terms disappear from the third derivative of the flow, because we are evaluating it at $\delta = 0$.

For our choice of initial data we now introduce parameters $N \gg 1$ and $\gamma \ll 1$ that are to be chosen later. With these in hand we may set $\hat{\phi}(\xi) = \gamma^{-\frac{1}{r'}} N^{-s}\chi(\xi)$, where $\chi(\xi)$ is the characteristic function of the interval $[N, N+\gamma]$. The factors in the definition of ϕ are chosen such that we have $\|\phi\|_{\hat{H}^s_r} \sim 1$.

Our next step is inserting our initial datum u_0 into (13.1):

$$\mathcal{F}_x\left(\frac{\partial^3 u}{\partial \delta^3}\bigg|_{\delta=0}\right)(\xi,t) \sim \xi^{2j-1}\int_*\int_0^t e^{it(-\xi^{2j}+\xi_1^{2j}-\xi_2^{2j}+\xi_3^{2j})}\hat{\phi}(\xi_1)\hat{\overline{\phi}}(\xi_2)\hat{\phi}(\xi_3)\,\mathrm{d}t\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2$$

In order to properly bound the inner t-integral we must have control of the resonance relation $\Phi = -\xi^{2j}+\xi_1^{2j}-\xi_2^{2j}+\xi_3^{2j}$ of which Lemma 11.7 tells us that we may bound it by $\Phi \sim \gamma^2 N^{2j-2}$. Hence we see a choice of $\gamma \sim N^{-(j-1)}$ is sensible. We continue working on a lower bound:

$$\begin{aligned}
&\gtrsim tN^{-3s}\gamma^{-\frac{3}{r'}}N^{2j-1}\,\chi * \chi * \chi(\xi)\\
&\gtrsim tN^{-3s}\gamma^{-\frac{3}{r'}}N^{2j-1}\gamma^2\,\chi(\xi)\\
&\sim tN^{-2s+2j-1}\gamma^{2-\frac{2}{r'}}(\gamma^{-\frac{1}{r'}}N^{-2}\chi(\xi)) = tN^{-2s+\frac{2j-2}{r'}+1}\hat{\phi}(\xi).
\end{aligned}$$

Here we may now take the the $\hat{H}^s_r(\mathbb{R})$ norm of both sides, keeping in mind our choice of ϕ leading to $\|\phi\|_{\hat{H}^s_r} \sim 1$. Thus we have a lower bound on the third derivative of the flow

$$\|\,\frac{\partial^3 u}{\partial \delta^3}\bigg|_{\delta=0}\|_{\hat{H}^s_r} \gtrsim tN^{-2s+\frac{2j-2}{r'}+1}. \tag{13.2}$$

In order for this quantity to stay bounded (a necessity, if the flow shall be thrice continuously differentiable) we must have $-2s + \frac{2j-2}{r'} + 1 \leq 0 \iff s \geq \frac{1}{2} + \frac{j-1}{r'}$, since otherwise we can let $N \to \infty$ and thus produce a contradiction. □

We will omit the proof of Theorem 10.9 as it follows along the same lines as the $r = 2$ case in the preceding proof. The key insight to be had is, because it suffices the look at the high-high-high interaction, with frequencies located on a single interval of length $o(1)$, the exact choice of Hölder exponents p, q in the modulation spaces is irrelevant. This argument was also given in [63]. Hence the C^3 ill-posedness result in modulation spaces parallels the $r = 2$ case in Fourier-Lebesgue spaces in terms of regularity ($s < \frac{j}{2}$), but with arbitrary exponents p, q.

Having addressed the non-periodic setting, we now present two propositions that establish our ill-posedness results for gauged dNLS equations on the torus. Their proofs follow arguments well-known to the relevant literature and correspond to Theorems 10.11 and 10.12, respectively.

Proposition 13.1. *The flow $S : \hat{H}^s_r(\mathbb{T}) \times (-T, T) \to \hat{H}^s_r(\mathbb{T})$ of the Cauchy problem for the fourth-order dNLS hierarchy equation (which corresponds to $j = 2$)*

$$i\partial_t u - \partial_x^4 u = \partial_x(-iu^2\overline{u}_{xx} - 4i|u|^2 u_{xx} - 2i|u_x|^2 u - 3iu_x^2\overline{u} - \frac{15}{2}|u|^4 u_x + \frac{5i}{2}|u|^6 u)$$

cannot be thrice continuously differentiable for any $1 \leq r \leq \infty$ and $s \in \mathbb{R}$.

Proof. The proof of this proposition works similarly to the one given by the author in [2, Proposition 6.3], which in turn was based on an argument by Bourgain [14].

In the present setting we may observe, that the symbol of the cubic nonlinearity in the fourth-order hierarchy equation can be written as

$$n_3(k_1, k_2, k_3) = (k_1 + k_2 + k_3)(2k_1^2 + k_2^2 + 2k_3^2 + k_1k_2 + k_2k_3 + 3k_1k_3). \tag{13.3}$$

Following along the details of [2, Proposition 6.3], i.e. differentiating the flow thrice (with respect to δ) with initial data $\delta\phi(x)$, where $\delta > 0$ and $\hat{\phi}(k) = k^{-s}(\delta_{k,N} + \delta_{k,N_0})$ and looking for a lower bound on the $H^s(\mathbb{T})$ norm of this third derivative, one arrives at the same conclusion. Only for (N, N_0, N_0) and $(N, -N, -N)$ (or appropriate permutations thereof) an overall frequency of N is achieved. Inserting these constellations into (13.3) one may derive a lower bound of $N^s tN^{3-s}(1 + N^{-2s}) \gtrsim tN^3$ for the $H^s(\mathbb{T})$ norm of the derivative. For $N \to \infty$ this diverges, so we know the flow cannot be thrice continuously differentiable. We leave working out further details to the reader. □

If one lowers the assumption on the regularity of the initial data, one is able to strengthen the form of ill-posedness that is derived to failure of uniform continuity using an argument originally developed in [60].

Proposition 13.2. *Let $j \in \mathbb{N}$, $1 \leq r \leq \infty$ and $s < j - \frac{1}{2}$. The flow $S : \hat{H}^s_r(\mathbb{T}) \times (-T, T) \to \hat{H}^s_r(\mathbb{T})$ of the Cauchy problem*

$$i\partial_t u + (-1)^{j+1}\partial_x^{2j} u = iu^2\partial_x^{2j-1}\overline{u} \tag{13.4}$$

cannot be uniformly continuous on bounded sets.

Proof. The proof of this proposition follows the same argument already given by the author for [2, Proposition 6.4], so we will not repeat the details here. The only difference is that one must choose a different particular solution of (13.4), which is a slightly modified (i.e. adapted to the dNLS hierarchy setting) version of [2, eq. (6.5)]. In particular it suffices to use the family of solutions

$$u_{N,a}(x, t) = N^{-s} a \exp(i(Nx - N^{2j}t + N^{2j-1-2s}|a|^2 t)) \tag{13.5}$$

in this case.

Note that as all derivatives in this equation fall on $\overline{u}$ this is in fact a gauged dNLS equation. Changing the sign in front of the nonlinearity allows one to solve (using the same family $u_{N,a}$) the equation where all derivatives fall on u instead. $\square$

Appendix A. The first few dNLS hierarchy equations

For future reference and the interested reader we would like to list the first few equations of the dNLS hierarchy and the resulting equations after they have been gauge transformed for the Schrödinger-like ones. A similar listing concerning the NLS hierarchy equations may be found in [2, Appendix A].

We will give the equations in terms of the potentials q and r as in the description of the hierarchy with a specific choice of α_n left to the reader (except for all other α_n being zero), as in Section 9. The usual identification $r = \pm\overline{q}$ leads to the well-known equations found elsewhere in the literature. For the non-gauge transformed equations we give them once with the nonlinearity as a total derivative, as in the representation (9.7), and again but with the derivative applied.

For the equations which have been adjusted with the gauge transform we use the convention $\alpha_{2j-1} = \alpha_n = 2^{2j-1}$ which has been in use throughout the rest of the text as well.

(1) $n = 0$. transport equation $q_t = \alpha_0 q_x$
(2) $n = 1$. classic dNLS equation

$$q_t = \frac{i\alpha_1}{2}(q_{xx} + \partial_x(-iq^2 r)) = \frac{i\alpha_1}{2}(q_{xx} - 2iqq_x r - iq^2 r_x)$$

After gauge transformation:

$$iq_t + q_{xx} = -iq^2 r_x - \frac{1}{2}q^3 r^2$$

(3) $n = 2$.

$$\begin{aligned} q_t &= -\frac{\alpha_2}{4}(q_{xxx} + \partial_x(-3iqq_x r - \frac{3}{2}q^3 r^2)) \\ &= -\frac{\alpha_2}{4}(q_{xxx} - 3iq_x^2 r - 3iqq_x r_x - 3iqq_{xx} r - 3q^3 r r_x - \frac{9}{2}q^2 q_x r^2) \end{aligned}$$

(4) $n = 3$. fourth order dNLS equation

$$\begin{aligned} q_t &= -\frac{i\alpha_3}{8}(q_{xxxx} + \partial_x(-iq^2 r_{xx} - 4iqq_{xx} r - 2iqq_x r_x - 3iq_x^2 r \\ &- \frac{15}{2}q^2 q_x r^2 + \frac{5i}{2}q^4 r^3)) \\ &= -\frac{i\alpha_3}{8}(q_{xxxx} - iq^2 r_{xxx} - 4iqq_{xxx} r - 4iqq_x r_{xx} - 6iqq_{xx} r_x \\ &- 10iq_x q_{xx} r - 5iq_x^2 r_x - \frac{15}{2}q^2 q_{xx} r^2 - 15q^2 q_x r r_x \\ &- 15qq_x^2 r^2 + \frac{15i}{2}q^4 r2r_x + 10iq^3 q_x r^3) \end{aligned}$$

After gauge transformation:

$$\begin{aligned} iq_t - q_{xxxx} &= iq^2 r_{xxx} + 2iqq_x r_{xx} + 4iqq_{xx} r_x + 3iq_x^2 r_x + q^3 r r_{xx} \\ &+ \frac{5}{2}q^2 q_{xx} r^2 - \frac{1}{2}q^3 r_x^2 + 4q^2 q_x r r_x + \frac{5}{2}qq_x^2 r^2 + \frac{3i}{2}q^4 r^2 r_x + \frac{3}{8}q^5 r^4 \end{aligned}$$

(5) $n = 4$.

$$
\begin{aligned}
q_t = &\frac{\alpha_4}{16}(q_{xxxxx} + \partial_x(-5iqq_{xxx}r - 5iqq_x r_{xx} - 5iqq_{xx}r_x - 10iq_x q_{xx}r \\
&- 5iq_x^2 r_x - 5q^3 rr_{xx} - \frac{25}{2}q^2 q_{xx}r^2 - \frac{5}{2}q^3 r_x^2 - 15q^2 q_x rr_x - \frac{35}{2}qq_x^2 r^2 \\
&+ \frac{35i}{2}q^3 q_x r^3 + \frac{35}{8}q^5 r^4)) \\
= &\frac{\alpha_4}{16}(q_{xxxxx} - 5iqq_{xxxx}r - 5iqq_x r_{xxx} - 10iqq_{xxx}r_x - 10iqq_{xx}r_{xx} - 10iq_x^2 r_{xx} \\
&- 10iq_{xx}^2 r - 25iq_x q_{xx}r_x - 15q_x q_{xxx}r - 5q^3 rr_{xxx} - \frac{25}{2}q^2 q_{xxx}r^2 - 10q^3 r_x r_{xx} \\
&- 30q^2 q_x rr_{xx} - 40q^2 q_{xx}rr_x - 60qq_x q_{xx}r^2 - \frac{45}{2}q^2 q_x r_x^2 - 65qq_x^2 rr_x - \frac{35}{2}q_x^3 r^2 \\
&+ \frac{35i}{2}q^3 q_{xx}r^3 + \frac{105i}{2}q^3 q_x r^2 r_x + \frac{35}{2}q^5 r^3 r_x + \frac{105i}{2}q^2 q_x^2 r^3 + \frac{175}{8}q^4 q_x r^4)
\end{aligned}
$$

(6) $n = 5$. sixth order dNLS equation

$$
\begin{aligned}
q_t = &\frac{i\alpha_5}{32}(q_{xxxxxx} + \partial_x(-iq^2 r_{xxxx} - 6iqq_{xxxx}r - 4iqq_x r_{xxx} - 9iqq_{xxx}r_x \\
&- 15iq_x q_{xxx}r - 11iqq_{xx}r_{xx} - 10iq_x^2 r_{xx} - 10iq_{xx}^2 r - 25q_x q_x x r_x - \frac{35}{2}q^2 q_{xxx}r^2 \\
&- 35q^2 q_x rr_{xx} - 35q^2 q_{xx}rr_x - 70qq_x q_{xx}r^2 - \frac{35}{2}q^2 q_x r_x^2 - 70qq_x^2 rr_x - \frac{35}{2}q_x^3 r^2 \\
&+ \frac{35i}{2}q^4 r^2 r_{xx} + 35iq^3 q_{xx}r^3 + \frac{35i}{2}q^4 rr_x^2 + 70iq^3 q_x r^2 r_x + 70iq^2 q_x^2 r^3 \frac{315}{8}q^4 q_x r^4 \\
&- \frac{63i}{8}q^6 r^5)) \\
= &\frac{i\alpha_5}{32}(q_{xxxxxx} - iq^2 r_{xxxxx} - 6iqq_{xxxxx}r - 6iqq_x r_{xxxx} - 15iqq_{xxxx}r_x \\
&- 21iq_x q_{xxxx}r - 15iqq_{xx}r_{xxx} - 14iq_x^2 r_{xxx} - 20iqq_{xxx}r_{xx} - 35iq_{xx}q_{xxx}r \\
&- 49iq_x q_{xxx}r_x - 56iq_x q_{xx}r_{xx} - 35iq_{xx}^2 r_x - \frac{35}{2}q^2 q_{xxxx}r^2 - 35q^2 q_x r_{xxx}r \\
&- 70q^2 q_{xxx}rr_x - 105qq_x q_{xxx}r^2 - 70q^2 q_{xx}rr_{xx} - 70q^2 q_x r_x r_{xx} - 140qq_x^2 rr_{xx} \\
&- 70qq_{xx}^2 r^2 - \frac{105}{2}q^2 q_{xx}r_x^2 - 350qq_x q_{xx}rr_x - \frac{245}{2}q_x^2 q_{xx}r^2 - 105qq_x^2 r_x^2 \\
&- 105q_x^3 rr_x + \frac{35i}{2}q^4 r^2 r_{xxx} + 35iq^3 q_{xxx}r^3 + 70iq^4 rr_x r_{xx} + 140iq^3 q_x r^2 r_{xx} \\
&+ 175iq^3 q_{xx}r^2 r_x + 245iq^2 q_x q_{xx}r^3 + \frac{35i}{2}q^4 r_x^3 + 210iq^3 q_x rr_x^2 + 420iq^2 q_x^2 r^2 r_x \\
&+ 140iqq_x^3 r^3 + \frac{315}{8}q^4 q_{xx}r^4 + \frac{315}{2}q^4 q_x r^3 r_x \frac{315}{2}q^3 q_x^2 r^4 - \frac{315i}{8}q^6 r^4 r_x \\
&- \frac{189i}{4}q^5 q_x r^5)
\end{aligned}
$$

After gauge transformation:

$$
\begin{aligned}
iq_t + \partial_x^6 q = &-iq^2 r_{xxxxx} - 4iqq_x r_{xxxx} - 6iqq_{xxxx} r_x - 11iqq_{xx} r_{xxx} - 10iq_x^2 r_{xxx} \\
&- 9iqq_{xxx} r_{xx} - 15iq_x q_{xxx} r_x - 25iq_x q_{xx} r_{xx} - 10iq_{xx}^2 r_x - q^3 r r_{xxxx} \\
&- \frac{7}{2} q^2 q_{xxxx} r^2 + q^3 r_x r_{xxx} - 8q^2 q_x r r_{xxx} - 13q^2 q_{xxx} r r_x - 14qq_x q_{xxx} r^2 \\
&- \frac{1}{2} q^3 r_{xx}^2 - 17q^2 q_{xx} r r_{xx} - 9q^2 q_x r_x r_{xx} - 22qq_x^2 r r_{xx} - \frac{21}{2} qq_{xx}^2 r^2 \\
&- \frac{9}{2} q^2 q_{xx} r_x^2 - 59qq_x q_{xx} r r_x - \frac{35}{2} q_x^2 q_{xx} r^2 - \frac{9}{2} qq_x^2 r_x^2 - 20q_x^3 r r_x \\
&- \frac{5i}{2} q^4 r^2 r_{xxx} - 10iq^4 r r_x r_{xx} - 10iq^3 q_x r^2 r_{xx} - 15iq^3 q_{xx} r^2 r_x - \frac{5i}{2} q^4 r_x^3 \\
&- 25iq^3 y_x r r_x^2 - 25iq^2 q_x^2 r^2 r_x - \frac{5}{2} q^5 r^3 r_{xx} - \frac{35}{8} q^4 q_{xx} r^4 - \frac{5}{4} q^5 r^2 r_x^2 \\
&- 15q^4 q_x r^3 r_x - \frac{35}{4} q^3 q_x^2 r^4 - \frac{5}{16} q^7 r^6 - \frac{15i}{8} q^6 r^4 r_x
\end{aligned}
$$

Note the sign difference of the term $+q^3 r_x r_{xxx}$ to all others with four derivatives lying upon them in the gauge transformed equation. This does not seem to be a mistake originating from the derivation of the equation.

References

[1] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, *The inverse scattering transform-Fourier analysis for nonlinear problems*, Studies in Appl. Math. **53** (1974), no. 4, 249–315. MR450815

[2] J. Adams, *Well-posedness for the NLS hierarchy*, J. Evol. Equ. **24** (2024), no. 4, Paper No. 88, 52. MR4808327

[3] ———, *Wellposedness for the dNLS hierarchy* (2025), available at `arXiv:2502.02181`.

[4] G. P Agrawal, *Nonlinear fiber optics*, Nonlinear Science at the Dawn of the 21st Century, 2000, pp. 195–211.

[5] J. M. Alberty, T. Koikawa, and R. Sasaki, *Canonical structure of soliton equations. I*, Phys. D **5** (1982), no. 1, 43–65. MR666528

[6] D. Anderson and M. Lisak, *Nonlinear asymmetric self-phase modulation and self-steepening of pulses in long optical waveguides*, Phys. Rev. A **27** (1983), 1393–1398.

[7] V. Banica, R. Lucà, N. Tzvetkov, and L. Vega, *Blow-Up for the 1D Cubic NLS*, Comm. Math. Phys. **405** (2024), no. 1, 11. MR4693911

[8] Á. Bényi and K. A. Okoudjou, *Modulation spaces*, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York, 2020. MR4286055

[9] H. A. Biagioni and F. Linares, *Ill-posedness for the derivative Schrödinger and generalized Benjamin-Ono equations*, Trans. Amer. Math. Soc. **353** (2001), no. 9, 3649–3659. MR1837253

[10] J. L. Bona and R. Smith, *The initial-value problem for the Korteweg-de Vries equation*, Philos. Trans. Roy. Soc. London Ser. A **278** (1975), no. 1287, 555–601. MR385355

[11] J. Bona and R. Smith, *Existence of solutions to the Korteweg-de Vries initial value problem*, Nonlinear wave motion (Proc. AMS-SIAM Summer Sem., Clarkson Coll. Tech., Potsdam, N.Y., 1972), 1974, pp. 179–180. MR352736

[12] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations*, Geom. Funct. Anal. **3** (1993), no. 2, 107–156. MR1209299

[13] ———, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation*, Geom. Funct. Anal. **3** (1993), no. 3, 209–262. MR1215780

[14] ———, *Periodic Korteweg de Vries equation with measures as initial data*, Selecta Math. (N.S.) **3** (1997), no. 2, 115–159. MR1466164

[15] E. Brun, G. Li, R. Liu, and Y. Zine, *Global well-posedness of one-dimensional cubic fractional nonlinear Schrödinger equations in negative Sobolev spaces* (2023), available at `arXiv:2311.13370`.

[16] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics, vol. 10, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003. MR2002047

[17] T. Cazenave, L. Vega, and M. C. Vilela, *A note on the nonlinear Schrödinger equation in weak L^p spaces*, Commun. Contemp. Math. **3** (2001), no. 1, 153–162. MR1820017

[18] L. Chaichenets, *Modulation spaces and nonlinear Schrödinger equations*, Ph.D. Thesis, 2018.

[19] M. Chen and B. Guo, *Local well and ill posedness for the modified KdV equations in subcritical modulation spaces*, Commun. Math. Sci. **18** (2020), no. 4, 909–946. MR4129658

[20] Z. Chen, Z. Guo, and C. Huang, *Complex-valued solutions of the mKdV equations in generalized Fourier-Lebesgue spaces* (2024), available at `arXiv:2410.11266`.

[21] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, *Global well-posedness for Schrödinger equations with derivative*, SIAM J. Math. Anal. **33** (2001), no. 3, 649–669. MR1871414

[22] ———, *A refined global well-posedness result for Schrödinger equations with derivative*, SIAM J. Math. Anal. **34** (2002), no. 1, 64–86. MR1950826

[23] ———, *Sharp global well-posedness for KdV and modified KdV on $\mathbb{R}$ and $\mathbb{T}$*, J. Amer. Math. Soc. **16** (2003), no. 3, 705–749. MR1969209

[24] Y. Deng, A. R. Nahmod, and H. Yue, *Optimal local well-posedness for the periodic derivative nonlinear Schrödinger equation*, Comm. Math. Phys. **384** (2021), no. 2, 1061–1107. MR4259382

[25] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian methods in the theory of solitons*, Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1987. Translated from the Russian by A. G. Reyman [A. G. Reĭman]. MR905674

[26] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36. MR257819

[27] H. G. Feichtinger, *Modulation Spaces on Locally Compact Abelian Groups*, University of Vienna, 1983.

[28] X. S. Feng, *The global Cauchy problem for a higher-order nonlinear Schrödinger equation*, Indian J. Pure Appl. Math. **25** (1994), no. 6, 583–606. MR1285221

[29] ———, *The global solutions for a fourth order nonlinear Schrödinger equation*, Acta Math. Sci. (English Ed.) **15** (1995), no. 2, 196–206. MR1350362

[30] A. Fiorenza, M. R. Formica, T. G. Roskovec, and F. Soudský, *Detailed proof of classical Gagliardo-Nirenberg interpolation inequality with historical remarks*, Z. Anal. Anwend. **40** (2021), no. 2, 217–236. MR4237368

[31] J. Forlano and W. J. Trenberth, *On the transport of Gaussian measures under the one-dimensional fractional nonlinear Schrödinger equations*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **36** (2019), no. 7, 1987–2025. MR4020531

[32] J. Ginibre, Y. Tsutsumi, and G. Velo, *On the Cauchy problem for the Zakharov system*, J. Funct. Anal. **151** (1997), no. 2, 384–436. MR1491547

[33] J. Ginibre, *Le problème de Cauchy pour des EDP semi-linéaires périodiques en variables d'espace (d'après Bourgain)*, 1996, pp. Exp. No. 796, 4, 163–187. Séminaire Bourbaki, Vol. 1994/95. MR1423623

[34] L. Grafakos, *Classical Fourier analysis*, Third, Graduate Texts in Mathematics, vol. 249, Springer, New York, 2014. MR3243734

[35] A. Grünrock, *New applications of the Fourier restriction norm method to wellposedness problems for nonlinear evolution equations*, Ph.D. Thesis, 2002.

[36] ———, *An improved local well-posedness result for the modified KdV equation*, Int. Math. Res. Not. **61** (2004), 3287–3308. MR2096258

[37] ———, *Bi- and trilinear Schrödinger estimates in one space dimension with applications to cubic NLS and DNLS*, Int. Math. Res. Not. **41** (2005), 2525–2558. MR2181058

[38] ———, *On the hierarchies of higher-order mKdV and KdV equations*, Cent. Eur. J. Math. **8** (2010), no. 3, 500–536. MR2653659

[39] A. Grünrock and S. Herr, *Low regularity local well-posedness of the derivative nonlinear Schrödinger equation with periodic initial data*, SIAM J. Math. Anal. **39** (2008), no. 6, 1890–1920. MR2390318

[40] A. Grünrock and L. Vega, *Local well-posedness for the modified KdV equation in almost critical $\widehat{H^r_s}$-spaces*, Trans. Amer. Math. Soc. **361** (2009), no. 11, 5681–5694. MR2529909

[41] S. Guo, *On the 1D cubic nonlinear Schrödinger equation in an almost critical space*, J. Fourier Anal. Appl. **23** (2017), no. 1, 91–124. MR3602811

[42] S. Guo, X. Ren, and B. Wang, *Local well-posedness for the derivative nonlinear Schrödinger equation with L^2-subcritical data*, Discrete Contin. Dyn. Syst. **41** (2021), no. 9, 4207–4253. MR4256461

[43] Z. Guo and Y. Wu, *Global well-posedness for the derivative nonlinear Schrödinger equation in $H^{\frac{1}{2}}(\mathbb{R})$*, Discrete Contin. Dyn. Syst. **37** (2017), no. 1, 257–264. MR3583477

[44] B. Harrop-Griffiths, R. Killip, and M. Vişan, *Large-data equicontinuity for the derivative NLS*, Int. Math. Res. Not. IMRN **6** (2023), 4601–4642. MR4565673

[45] N. Hayashi and T. Ozawa, *On the derivative nonlinear Schrödinger equation*, Phys. D **55** (1992), no. 1-2, 14–36. MR1152001

[46] ———, *Finite energy solutions of nonlinear Schrödinger equations of derivative type*, SIAM J. Math. Anal. **25** (1994), no. 6, 1488–1503. MR1302158

[47] S. Herr, *On the Cauchy problem for the derivative nonlinear Schrödinger equation with periodic boundary condition*, Int. Math. Res. Not. (2006), Art. ID 96763, 33. MR2219223

[48] H. Hirayama, M. Ikeda, and T. Tanaka, *Well-posedness for the fourth-order Schrödinger equation with third order derivative nonlinearities*, NoDEA Nonlinear Differential Equations Appl. **28** (2021), no. 5, Paper No. 46, 72. MR4274694

[49] Z. Huo and Y. Jia, *The Cauchy problem for the fourth-order nonlinear Schrödinger equation related to the vortex filament*, J. Differential Equations **214** (2005), no. 1, 1–35. MR2143510

[50] ———, *Well-posedness for the fourth-order nonlinear derivative Schrödinger equation in higher dimension*, J. Math. Pures Appl. (9) **96** (2011), no. 2, 190–206. MR2818712

[51] T. Kato, *Unconditional well-posendness for the fourth order nonlinear Schrodinger type equations on the torus* (2025), available at `arXiv:2501.11455`.

[52] C. E. Kenig and D. Pilod, *Local well-posedness for the KdV hierarchy at high regularity*, Adv. Differential Equations **21** (2016), no. 9-10, 801–836. MR3513119

[53] C. E. Kenig, G. Ponce, and L. Vega, *Oscillatory integrals and regularity of dispersive equations*, Indiana Univ. Math. J. **40** (1991), no. 1, 33–69. MR1101221

[54] ———, *Small solutions to nonlinear Schrödinger equations*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **10** (1993), no. 3, 255–288. MR1230709

[55] ———, *Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle*, Comm. Pure Appl. Math. **46** (1993), no. 4, 527–620. MR1211741
[56] ———, *Higher-order nonlinear dispersive equations*, Proc. Amer. Math. Soc. **122** (1994), no. 1, 157–166. MR1195480
[57] ———, *On the hierarchy of the generalized KdV equations*, Singular limits of dispersive waves (Lyon, 1991), 1994, pp. 347–356. MR1321214
[58] ———, *A bilinear estimate with applications to the KdV equation*, J. Amer. Math. Soc. **9** (1996), no. 2, 573–603. MR1329387
[59] ———, *Quadratic forms for the 1-D semilinear Schrödinger equation*, Trans. Amer. Math. Soc. **348** (1996), no. 8, 3323–3353. MR1357398
[60] ———, *On the ill-posedness of some canonical dispersive equations*, Duke Math. J. **106** (2001), no. 3, 617–633. MR1813239
[61] R. Killip, M. Ntekoume, and M. Vişan, *On the well-posedness problem for the derivative nonlinear Schrödinger equation*, Anal. PDE **16** (2023), no. 5, 1245–1270. MR4628747
[62] R. Killip, M. Vişan, and X. Zhang, *Low regularity conservation laws for integrable PDE*, Geom. Funct. Anal. **28** (2018), no. 4, 1062–1090. MR3820439
[63] F. Klaus, *Wellposedness of NLS in modulation spaces*, J. Fourier Anal. Appl. **29** (2023), no. 1, Paper No. 9, 37. MR4534495
[64] F. Klaus, H. Koch, and B. Liu, *Wellposedness for the KdV hierarchy* (2023), available at `arXiv:2309.12773`.
[65] F. Klaus and R. Schippa, *A priori estimates for the derivative nonlinear Schrödinger equation*, Funkcial. Ekvac. **65** (2022), no. 3, 329–346. MR4515528
[66] H. Koch and N. Tzvetkov, *Nonlinear wave interactions for the Benjamin-Ono equation*, Int. Math. Res. Not. **30** (2005), 1833–1847. MR2172940
[67] H. Koch and D. Tataru, *A priori bounds for the 1D cubic NLS in negative Sobolev spaces*, Int. Math. Res. Not. IMRN **16** (2007), Art. ID rnm053, 36. MR2353092
[68] ———, *Energy and local energy bounds for the 1-d cubic NLS equation in $H^{-\frac{1}{4}}$*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **29** (2012), no. 6, 955–988. MR2995102
[69] ———, *Conserved energies for the cubic nonlinear Schrödinger equation in one dimension*, Duke Math. J. **167** (2018), no. 17, 3207–3313. MR3874652
[70] P. D. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math. **21** (1968), 467–490. MR235310
[71] C. Miao, Y. Wu, and G. Xu, *Global well-posedness for Schrödinger equation with derivative in $H^{\frac{1}{2}}(\mathbb{R})$*, J. Differential Equations **251** (2011), no. 8, 2164–2195. MR2823664
[72] K. Mio, T. Ogino, K. Minami, and S. Takeda, *Modified nonlinear Schrödinger equation for Alfvén waves propagating along the magnetic field in cold plasmas*, J. Phys. Soc. Japan **41** (1976), no. 1, 265–271. MR462141
[73] L. Molinet, J. C. Saut, and N. Tzvetkov, *Ill-posedness issues for the Benjamin-Ono and related equations*, SIAM J. Math. Anal. **33** (2001), no. 4, 982–988. MR1885293
[74] T. Oh and C. Sulem, *On the one-dimensional cubic nonlinear Schrödinger equation below L^2*, Kyoto J. Math. **52** (2012), no. 1, 99–115. MR2892769
[75] T. Oh and N. Tzvetkov, *Quasi-invariant Gaussian measures for the cubic fourth order nonlinear Schrödinger equation*, Probab. Theory Related Fields **169** (2017), no. 3-4, 1121–1168. MR3719064
[76] T. Oh and Y. Wang, *Global well-posedness of the one-dimensional cubic nonlinear Schrödinger equation in almost critical spaces*, J. Differential Equations **269** (2020), no. 1, 612–640. MR4081534
[77] ———, *On global well-posedness of the modified KdV equation in modulation spaces*, Discrete Contin. Dyn. Syst. **41** (2021), no. 6, 2971–2992. MR4235636
[78] T. Ozawa, *On the nonlinear Schrödinger equations of derivative type*, Indiana Univ. Math. J. **45** (1996), no. 1, 137–163. MR1406687
[79] R. S. Palais, *The symmetries of solitons*, Bull. Amer. Math. Soc. (N.S.) **34** (1997), no. 4, 339–403. MR1462745
[80] N. Pattakos, *NLS in the modulation space $M_{2,q}(\mathbb{R})$*, J. Fourier Anal. Appl. **25** (2019), no. 4, 1447–1486. MR3977124
[81] D. Pilod, *On the Cauchy problem for higher-order nonlinear dispersive equations*, J. Differential Equations **245** (2008), no. 8, 2055–2077. MR2446185
[82] R. Sasaki, *Canonical structure of soliton equations. II. The Kaup-Newell system*, Phys. D **5** (1982), no. 1, 66–74. MR666529
[83] J.-C. Saut, *Quelques généralisations de l'équation de Korteweg-de Vries. II*, J. Differential Equations **33** (1979), no. 3, 320–335. MR543702

[84] P. Sjölin, *Regularity of solutions to the Schrödinger equation*, Duke Math. J. **55** (1987), no. 3, 699–715. MR904948

[85] C. Sulem and P.-L. Sulem, *The nonlinear Schrödinger equation*, Applied Mathematical Sciences, vol. 139, Springer-Verlag, New York, 1999. Self-focusing and wave collapse. MR1696311

[86] H. Takaoka, *Well-posedness for the one-dimensional nonlinear Schrödinger equation with the derivative nonlinearity*, Adv. Differential Equations **4** (1999), no. 4, 561–580. MR1693278

[87] ———, *Global well-posedness for Schrödinger equations with derivative in a nonlinear term and data in low-order Sobolev spaces*, Electron. J. Differential Equations (2001), No. 42, 23. MR1836810

[88] T. Tao, *Nonlinear dispersive equations*, CBMS Regional Conference Series in Mathematics, vol. 106, Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. Local and global analysis. MR2233925

[89] M. Tsutsumi and I. Fukuda, *On solutions of the derivative nonlinear Schrödinger equation. Existence and uniqueness theorem*, Funkcial. Ekvac. **23** (1980), no. 3, 259–277. MR621533

[90] ———, *On solutions of the derivative nonlinear Schrödinger equation. II*, Funkcial. Ekvac. **24** (1981), no. 1, 85–94. MR634894

[91] N. Tzoar and M. Jain, *Self-phase modulation in long-geometry optical waveguides*, Phys. Rev. A **23** (1981), 1266–1270.

[92] N. Tzvetkov, *Remark on the local ill-posedness for KdV equation*, C. R. Acad. Sci. Paris Sér. I Math. **329** (1999), no. 12, 1043–1047. MR1735881

[93] A. Vargas and L. Vega, *Global wellposedness for 1D non-linear Schrödinger equation for data with an infinite L^2 norm*, J. Math. Pures Appl. (9) **80** (2001), no. 10, 1029–1044. MR1876762

[94] Y. Wu, *Global well-posedness for the nonlinear Schrödinger equation with derivative in energy space*, Anal. PDE **6** (2013), no. 8, 1989–2002. MR3198590

[95] ———, *Global well-posedness on the derivative nonlinear Schrödinger equation*, Anal. PDE **8** (2015), no. 5, 1101–1112. MR3393674

[96] H. Zhou, J. Yu, and J. Han, *Several exact solutions of the reduced fourth-order flow equation of the Kaup-Newell system*, Wave Motion **109** (2022), Paper No. 102840, 12. MR4348344

Affidavit

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist.