

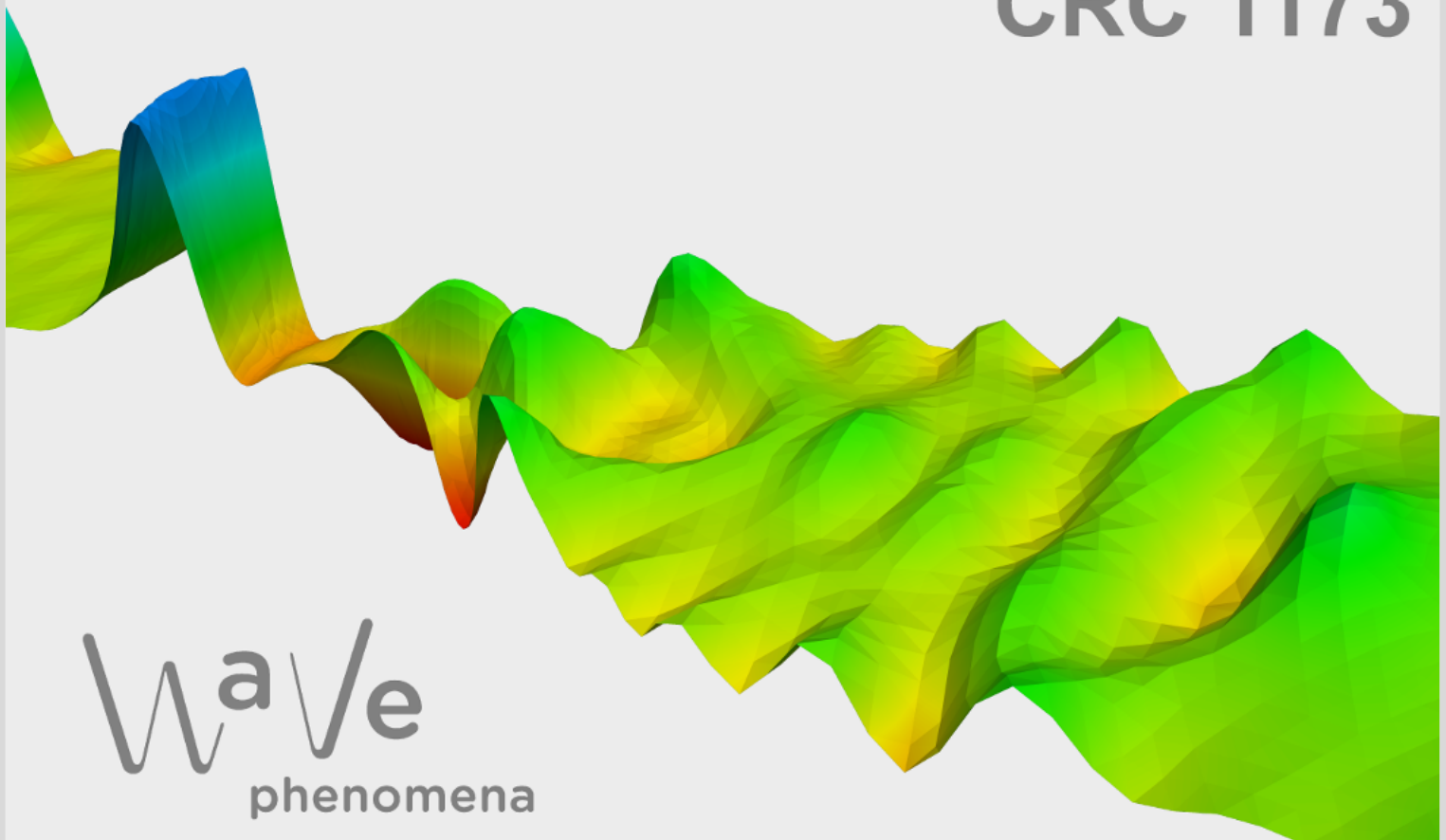
Long-wave KdV hierarchy approximation of the NLS hierarchy with nonzero boundary conditions

Robert Wegner

CRC Preprint 2026/36, July 2026

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave
phenomena

Participating universities



Funded by



Long-wave KdV Hierarchy Approximation of the NLS Hierarchy with Nonzero Boundary Conditions

Robert Wegner

Abstract

We study the approximation of certain renormalized conserved quantities for the NLS hierarchy with nonzero boundary conditions, in the long-wave regime, by the energies of the KdV hierarchy. We extend this to all $n \in \mathbb{N}$ by proving an approximation result for the transmission coefficient of the Lax operator of the NLS hierarchy, which is a Dirac operator in the nonrelativistic regime, by the transmission coefficient of a Schrödinger operator, which is the Lax operator of the KdV hierarchy. This yields a formal approximation result between the hierarchies, which we quantify using energy methods and previously established well-posedness results.

Keywords: NLS hierarchy, KdV hierarchy, Gross-Pitaevskii equation, Korteweg-de Vries equation, transmission coefficient, Lax operator, Dirac operator, long-wave approximation

AMS Subject Classification (2020): 35Q51, 35Q53, 35Q55, 37K10

Contents

1	Introduction	2
1.1	From NLS to GP and $\widetilde{\text{GP}}$	3
1.2	From GP to the hydrodynamic formulation hGP and the long-wave rescaling εGP	6
1.3	Approximation of εGP by transport and KdV equations from [12, 13]	8
1.4	Approximation of Hamiltonians and transmission coefficients	9
1.5	Formal approximation of hierarchies	12
1.6	Review of well-posedness results for the hierarchies	15
1.7	Quantitative approximation of hierarchies	17
1.8	Examples	20
1.9	Notations and definitions	22
2	Definition and approximation of the Hamiltonians $\mathcal{H}_n^{\varepsilon\text{GP}}$ by $\mathcal{E}_n^{\text{KdV}}$	23
2.1	Lax pairs	23
2.2	Jost solutions and their asymptotic expansion	25
2.3	Transmission coefficients and their asymptotic expansion	26
2.4	Recurrence relations for the densities of the Hamiltonians	30
2.5	Proofs of Theorem 1.1 and Corollary 1.5	33
3	Approximation of solutions to the $\varepsilon\widetilde{\text{GP}}$ hierarchy by solutions to the KdV hierarchy	39
3.1	Proof of Proposition 1.10	39
3.2	Proof of Proposition 1.11	40
3.3	Proof of Theorem 1.12	42
A	Approximation of the perturbation determinant $\alpha^{\varepsilon\text{GP}}$ by α^{KdV}	45
A.1	Definitions	45
A.2	Approximation of perturbation determinants	48

B Overview of Hamiltonian structures	57
B.1 GP	57
B.2 hGP	57
B.3 ε GP	58
B.4 KdV	59

1 Introduction

We fix $N \in \mathbb{N}$ and consider the “ N -truncated” defocusing NLS hierarchy ¹

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\text{NLS}}(q)}{\delta \bar{q}} \quad \forall n \in \{0, \dots, N\} \quad (\text{NLS}_n)$$

for a function $q = q(\mathbf{t}, x) = q(t_0, \dots, t_N, x) : \mathbb{R}^{N+1} \times \mathbb{R} \rightarrow \mathbb{C}$. Here $\mathcal{H}_n^{\text{NLS}}$ are the Hamiltonians of the NLS hierarchy, the first of which are

$$\mathcal{H}_0^{\text{NLS}}(q) = \int_{\mathbb{R}} q \bar{q} \, dx \quad (\text{Mass})$$

$$\mathcal{H}_1^{\text{NLS}}(q) = -i \int_{\mathbb{R}} q \bar{q}_x \, dx \quad (\text{Momentum})$$

$$\mathcal{H}_2^{\text{NLS}}(q) = \int_{\mathbb{R}} q_x \bar{q}_x + q^2 \bar{q}^2 \, dx \quad (\text{Energy})$$

$$\mathcal{H}_3^{\text{NLS}}(q) = i \int_{\mathbb{R}} q \bar{q}_{xxx} - 4q^2 \bar{q} \bar{q}_x - q_x q \bar{q}^2 \, dx$$

$$\mathcal{H}_4^{\text{NLS}}(q) = \int_{\mathbb{R}} q_{xx} \bar{q}_{xx} - 6q^2 \bar{q} \bar{q}_{xx} - 5q^2 \bar{q}_x^2 - 6q q_x \bar{q} \bar{q}_x - q q_{xx} \bar{q}^2 + 2q^3 \bar{q}^3 \, dx.$$

We are interested in long-wave solutions q of the form

$$q(\mathbf{t}, x) = \left(1 + \frac{\varepsilon^2}{2}(W_+ - W_-)(\mathbf{T}, X)\right) \exp\left(i \frac{\varepsilon}{\sqrt{2}} \int_0^X (W_+ + W_-)(\mathbf{T}, Y) \, dY\right) e^{i\varphi(\mathbf{t})} \quad (1.1)$$

which satisfy the nonzero boundary conditions

$$\lim_{x \rightarrow \pm\infty} q(\mathbf{t}, x) e^{-i\psi(\mathbf{t})} = q_{\pm} \text{ where } |q_{\pm}| = 1. \quad (\text{NZBC})$$

We use the rescaled space and time variables

$$X = \sqrt{2}\varepsilon x \quad \mathbf{T} = (T_n)_{0 \leq n \leq N} = \left(\frac{dT_n}{dt_n} t_n\right)_{0 \leq n \leq N}$$

with

$$\frac{dT_{2n}}{dt_{2n}} = 2(\sqrt{2}\varepsilon)^{2n-1} \quad \frac{dT_{2n-1}}{dt_{2n-1}} = (\sqrt{2}\varepsilon)^{2n-1}, \quad (1.2)$$

where $\varepsilon > 0$ is a small parameter. The functions $\varphi, \psi : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$ are assumed to depend only on $\mathbf{t}$, representing uniform-in- x phase rotations. We assume $W_+, W_- : \mathbb{R}^{N+1} \times \mathbb{R} \rightarrow \mathbb{R}$ and use the variable $W = (W_+, W_-)$ to denote the pair. In order to match NZBC for q , we choose for W the zero boundary condition

$$\lim_{|X| \rightarrow \infty} W(\mathbf{T}, X) = 0.$$

¹Some authors define the NLS hierarchy to consist of only the even flows and refer to the odd flows as the mKdV hierarchy.

We are interested only in the dynamics of W . Define $M = \lfloor \frac{N-1}{2} \rfloor$. Our main result is Theorem 1.12 below, which states that the behavior of the N -truncated NLS hierarchy with NZBC in the long-wave regime can be described by the M -truncated KdV hierarchy

$$\partial_{T_n} U = \frac{1}{2} \partial_X \frac{\delta \mathcal{E}_n^{\text{KdV}}(U)}{\delta U} \quad \forall n \in \{0, \dots, M\} \quad (\text{KdV}_n)$$

for a function $U = U(\mathbf{T}, X) = U(T_0, \dots, T_M, X) : \mathbb{R}^{M+1} \times \mathbb{R} \rightarrow \mathbb{R}$. Here $\mathcal{E}_n^{\text{KdV}}$ are the Hamiltonians of the KdV hierarchy, the first of which are

$$\mathcal{E}_{-1}^{\text{KdV}}(U) = \int_{\mathbb{R}} U \, dX \quad (\text{Mass})$$

$$\mathcal{E}_0^{\text{KdV}}(U) = \int_{\mathbb{R}} U^2 \, dX \quad (\text{Momentum})$$

$$\mathcal{E}_1^{\text{KdV}}(U) = \int_{\mathbb{R}} U_X^2 + 2U^3 \, dX \quad (\text{Energy})$$

$$\mathcal{E}_2^{\text{KdV}}(U) = \int_{\mathbb{R}} U_{XX}^2 + 10U_X^2 U + 5U^4 \, dX$$

$$\mathcal{E}_3^{\text{KdV}}(U) = \int_{\mathbb{R}} U_{XXX}^2 + 14U_{XX}^2 U + 70U_X^2 U^2 + 14U^5 \, dX.$$

Theorem 1.12 is stated below after further preliminary discussion. In Examples 1.13–1.20 at the end of the introduction several explicit demonstrations of our main result may be found.

1.1 From NLS to GP and GP

Let us begin by recalling the defocusing nonlinear Schrödinger equation

$$iq_t + q_{xx} = 2q|q|^2 \quad (\text{NLS})$$

for a wavefunction $q = q(t, x) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$ that satisfies the zero boundary condition

$$\lim_{|x| \rightarrow \infty} q(t, x) = 0. \quad (\text{ZBC})$$

It is closely related to the defocusing Gross-Pitaevskii equation

$$iq_t + q_{xx} = 2q(|q|^2 - 1), \quad (\text{GP})$$

which we associate with the nonzero boundary conditions

$$\lim_{x \rightarrow \pm \infty} q(t, x) = q_{\pm} \text{ where } |q_{\pm}| = 1. \quad (\text{NZBC})$$

Both equations are formally equivalent, as the transformation $q \mapsto e^{2it} q$ maps solutions of NLS to solutions of GP. It is preferable to work with the system GP–NZBC instead of NLS–NZBC because solutions to the former system are stationary at infinity, which is not the case for the latter system. When writing NLS and GP, we generally mean the systems NLS–ZBC and GP–NZBC.

The Hamiltonians $\mathcal{H}_n^{\text{NLS}}$ are formally conserved quantities for both NLS and GP, but ill-defined in the setting of NZBC. We use instead an alternative sequence of Hamiltonians $(\mathcal{H}_n^{\text{GP}})_{n \geq -1}$ which is

compatible with NZBC. The initial ones are

$$\mathcal{H}_0^{\text{GP}}(q) = \int_{\mathbb{R}} q\bar{q} - 1 \, dx \quad (\text{Mass})$$

$$\mathcal{H}_1^{\text{GP}}(q) = -i \int_{\mathbb{R}} q\bar{q}_x \, dx \quad (\text{Momentum})$$

$$\mathcal{H}_2^{\text{GP}}(q) = \int_{\mathbb{R}} q_x \bar{q}_x + (q\bar{q} - 1)^2 \, dx \quad (\text{Energy})$$

$$\mathcal{H}_3^{\text{GP}}(q) = i \int_{\mathbb{R}} q\bar{q}_{xxx} - 4q^2 \bar{q}q_x - q_x q\bar{q}^2 + 4q\bar{q}_x \, dx$$

$$\begin{aligned} \mathcal{H}_4^{\text{GP}}(q) = \int_{\mathbb{R}} & q_{xx} \bar{q}_{xx} - 6q^2 \bar{q}q_{xx} - 5q^2 \bar{q}_x^2 - 6qq_x \bar{q}q_x - qq_{xx} \bar{q}^2 \\ & - 6q_x \bar{q}_x + 2q^3 \bar{q}^3 - 6q^2 \bar{q}^2 + 6q\bar{q} - 2 \, dx. \end{aligned}$$

In this setting there exists an additional non-trivial conserved quantity: the asymptotic phase change

$$\mathcal{H}_{-1}^{\text{GP}}(q) = -i \log_{\text{pr}} \left(\frac{q_-}{q_+} \right).$$

Here $\log_{\text{pr}}$ denotes the principal logarithm.

We consider also another sequence of Hamiltonians $(\mathcal{H}_n^{\widetilde{\text{GP}}})_{n \geq -1}$, which is defined by $\mathcal{H}_{2n}^{\widetilde{\text{GP}}} = \mathcal{H}_{2n}^{\text{GP}}$ and the equivalent relations

$$\mathcal{H}_{2n-1}^{\widetilde{\text{GP}}} = \sum_{m=0}^n \binom{-\frac{1}{2}}{n-m} 4^{n-m} 2^{-\mathbb{1}_{\{m=0\}}} \mathcal{H}_{2m-1}^{\text{GP}} \quad \text{and} \quad \mathcal{H}_{2n-1}^{\text{GP}} = 2^{\mathbb{1}_{\{n=0\}}} \sum_{m=0}^n \binom{\frac{1}{2}}{n-m} 4^{n-m} \mathcal{H}_{2m-1}^{\widetilde{\text{GP}}}. \quad (1.3)$$

Here the initial odd Hamiltonians are

$$\begin{aligned} \mathcal{H}_{-1}^{\widetilde{\text{GP}}}(q) &= \frac{1}{2} \mathcal{H}_{-1}^{\text{GP}}(q) \\ \mathcal{H}_1^{\widetilde{\text{GP}}}(q) &= -i \int_{\mathbb{R}} q\bar{q}_x \, dx - \mathcal{H}_{-1}^{\text{GP}}(q) \\ \mathcal{H}_3^{\widetilde{\text{GP}}}(q) &= i \int_{\mathbb{R}} q\bar{q}_{xxx} - 4q^2 \bar{q}q_x - q_x q\bar{q}^2 + 6q\bar{q}_x \, dx + 3\mathcal{H}_{-1}^{\text{GP}}(q). \end{aligned}$$

These Hamiltonians play a crucial role in the preceding work [67], which concerns the well-posedness of the NLS hierarchy with NZBC. In principle, the Hamiltonians $\mathcal{H}_0^{\text{GP}}, \dots, \mathcal{H}_8^{\text{GP}}$ have previously been used² in [12, 13], but to the best of our knowledge³ the authors were not aware of the “correct” definition for all $n \in \mathbb{N}$. We elaborate on this in Section 1.3. They have also appeared in [60, End of §1.3.4].

Using the invertible coefficient matrices $\mathbf{c}_{\text{GP} \rightarrow \widetilde{\text{GP}}}, \mathbf{c}_{\widetilde{\text{GP}} \rightarrow \text{GP}} \in \mathbb{R}^{(N+1) \times (N+1)}$ given by

$$\begin{aligned} \mathbf{c}_{\text{GP} \rightarrow \text{GP}}^{2n, 2m} &= \mathbb{1}_{\{n=m\}} & \mathbf{c}_{\text{GP} \rightarrow \text{GP}}^{2n, 2m-1} &= 0 \\ \mathbf{c}_{\text{GP} \rightarrow \text{GP}}^{2n-1, 2m} &= 0 & \mathbf{c}_{\widetilde{\text{GP}} \rightarrow \text{GP}}^{2n-1, 2m-1} &= 2^{\mathbb{1}_{\{n=0\}}} \binom{\frac{1}{2}}{n-m} 4^{n-m} \end{aligned}$$

²They use the variables $\partial_x \Theta_\varepsilon = 6(W_+ + W_-)$ and $N_\varepsilon = -6(W_+ - W_- + \frac{\varepsilon^2}{4}(W_+ - W_-)^2)$, and their choice of energies and momenta corresponds to ours with $\mathcal{E}_n(N_\varepsilon, \Theta_\varepsilon) = \frac{18}{(\sqrt{2\varepsilon})^{2n+1}} \mathcal{H}_{2n}^{\widetilde{\text{GP}}}(W)$ and $\mathcal{P}_n(N_\varepsilon, \Theta_\varepsilon) = \sqrt{2} \frac{18}{(\sqrt{2\varepsilon})^{2n+1}} \mathcal{H}_{2n-1}^{\widetilde{\text{GP}}}(W)$. We have checked this only for $n \in \{1, 2, 3\}$ due to computational restrictions. Note that (101) in the reference contains a term twice, which we understand to be unintentional.

³We base this on [13, Remark 4.7].

and $\mathbf{c}_{\text{GP} \mapsto \widetilde{\text{GP}}} = (\mathbf{c}_{\widetilde{\text{GP}} \mapsto \text{GP}})^{-1}$, we can write

$$\left(\frac{\delta \mathcal{H}_n^{\widetilde{\text{GP}}}}{\delta \bar{q}} \right)_{0 \leq n \leq N} = \mathbf{c}_{\text{GP} \mapsto \widetilde{\text{GP}}} \left(\frac{\delta \mathcal{H}_n^{\text{GP}}}{\delta \bar{q}} \right)_{0 \leq n \leq N} \quad \left(\frac{\delta \mathcal{H}_n^{\text{GP}}}{\delta \bar{q}} \right)_{0 \leq n \leq N} = \mathbf{c}_{\widetilde{\text{GP}} \mapsto \text{GP}} \left(\frac{\delta \mathcal{H}_n^{\widetilde{\text{GP}}}}{\delta \bar{q}} \right)_{0 \leq n \leq N}.$$

Let us also state at this point a relation between $\mathcal{H}_n^{\text{NLS}}$ and $\mathcal{H}_n^{\text{GP}}$, which is a consequence of (2.37)–(2.38) below (see also [30, Chapter 1, (10.25)]):

$$\frac{\delta \mathcal{H}_{2n}^{\text{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^n \binom{n - \frac{1}{2}}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m}^{\text{GP}}(q)}{\delta \bar{q}} \quad (1.4)$$

$$\frac{\delta \mathcal{H}_{2n-1}^{\text{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^n \binom{n-1}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m-1}^{\text{GP}}(q)}{\delta \bar{q}}. \quad (1.5)$$

Similarly, we have

$$\frac{\delta \mathcal{H}_{2n}^{\text{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^n \binom{n - \frac{1}{2}}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m}^{\widetilde{\text{GP}}}(q)}{\delta \bar{q}} \quad (1.6)$$

$$\frac{\delta \mathcal{H}_{2n-1}^{\text{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^n \binom{n - \frac{1}{2}}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m-1}^{\widetilde{\text{GP}}}(q)}{\delta \bar{q}}. \quad (1.7)$$

We do not state these relations for the Hamiltonians themselves because $\mathcal{H}_n^{\text{NLS}}(q)$ can only be well-defined if q satisfies ZBC, while $\mathcal{H}_n^{\text{GP}}(q)$ and $\mathcal{H}_n^{\widetilde{\text{GP}}}(q)$ require NZBC. We define again invertible coefficient matrices $\mathbf{c}_{\text{NLS} \mapsto \text{GP}}, \mathbf{c}_{\text{GP} \mapsto \text{NLS}}, \mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}}, \mathbf{c}_{\widetilde{\text{GP}} \mapsto \text{NLS}} \in \mathbb{R}^{(N+1) \times (N+1)}$ by

$$\begin{aligned} \mathbf{c}_{\text{NLS} \mapsto \text{GP}}^{2n, 2m} &= \binom{n - \frac{1}{2}}{n-m} (-4)^{n-m} & \mathbf{c}_{\text{NLS} \mapsto \text{GP}}^{2n, 2m-1} &= 0 \\ \mathbf{c}_{\text{NLS} \mapsto \text{GP}}^{2n-1, 2m} &= 0 & \mathbf{c}_{\text{NLS} \mapsto \text{GP}}^{2n-1, 2m-1} &= \binom{n-1}{n-m} (-4)^{n-m} \\ \mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}}^{2n, 2m} &= \binom{n - \frac{1}{2}}{n-m} (-4)^{n-m} & \mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}}^{2n, 2m-1} &= 0 \\ \mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}}^{2n-1, 2m} &= 0 & \mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}}^{2n-1, 2m-1} &= \binom{n - \frac{1}{2}}{n-m} (-4)^{n-m} \end{aligned}$$

and $\mathbf{c}_{\text{GP} \mapsto \text{NLS}} = (\mathbf{c}_{\text{NLS} \mapsto \text{GP}})^{-1}$, $\mathbf{c}_{\widetilde{\text{GP}} \mapsto \text{NLS}} = (\mathbf{c}_{\text{NLS} \mapsto \widetilde{\text{GP}}})^{-1}$.

In analogy to the definition of the NLS hierarchy, we consider the N -truncated GP and $\widetilde{\text{GP}}$ hierarchies, which consist, respectively, of the Hamiltonian flows

$$i \partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\text{GP}}(q)}{\delta \bar{q}} \quad (\text{GP}_n)$$

$$i \partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\widetilde{\text{GP}}}(q)}{\delta \bar{q}} \quad (\widetilde{\text{GP}}_n)$$

for $n \in \{0, \dots, N\}$. While the GP hierarchy⁴ is a more immediate renormalization of the NLS hierarchy for working with NZBC, the $\widetilde{\text{GP}}$ hierarchy is the “better” hierarchy to work with in both [67] and this work. Note, however, that by Proposition 1.11 below the theories of the NLS, GP, and $\widetilde{\text{GP}}$ hierarchies are equivalent, up to the application of explicit transformations.

⁴Not to be confused with the “Gross–Pitaevskii hierarchy” from kinetic theory (see e. g. [29]).

Let it be said that $\frac{\delta \mathcal{H}_{-1}^{\text{GP}}}{\delta \bar{q}} = 0$ and hence the Hamiltonian flow generated by this conserved quantity is trivial, meaning the asymptotic phase change $\mathcal{H}_{-1}^{\text{GP}}$ is not of particular importance in this work. Nevertheless, Theorem 1.1 below *does* require the involvement of $\mathcal{H}_{-1}^{\text{GP}}$, and its proof motivates our choice of normalization for this conserved quantity, which is different from $\mathcal{H}_{-1}^{\text{GP}}$ that appears in [60].

1.2 From GP to the hydrodynamic formulation hGP and the long-wave rescaling εGP

The solutions we are interested in have no vacuum, meaning that there exists a constant $\delta > 0$ for which

$$\inf_{x \in \mathbb{R}} |q(t, x)| > \delta.$$

As a result, it is possible to apply the so-called Madelung transform, which maps q to the pair (a, v) with

$$a = |q| \quad v = \text{Im} \left(\frac{q_x}{q} \right) \quad q = a \exp \left(i \int_0^x v(y) dy \right) e^{i\varphi}.$$

Here φ is a uniform-in- x choice of phase, which is information that has to be provided in order to invert the Madelung transform. It is well-known that if q solves GP, then (a, v) solve a hydrodynamic system, which we call the hydrodynamic Gross-Pitaevskii equations:

$$\begin{aligned} \partial_t a + a \partial_x v + 2v \partial_x a &= 0 \\ \partial_t v + \partial_x(v^2) + 4a \partial_x a &= \partial_x \left(\frac{\partial_{xx} a}{a} \right). \end{aligned}$$

We prefer to work with the slightly different pair of variables $w = (w_+, w_-)$ where

$$w_{\pm} = \frac{v}{2} \pm (a - 1).$$

We still refer to the system of PDEs that w solves as the hydrodynamic Gross-Pitaevskii equations:

$$\partial_t w_- = -(3w_- + w_+ - 2) \partial_x w_- + \frac{1}{2} \partial_x \left(\frac{\partial_{xx}(w_+ - w_-)}{2a} \right) \quad (\text{hGP-1})$$

$$\partial_t w_+ = -(w_- + 3w_+ + 2) \partial_x w_+ + \frac{1}{2} \partial_x \left(\frac{\partial_{xx}(w_+ - w_-)}{2a} \right). \quad (\text{hGP-2})$$

Of course, this system has an infinite family of conserved energies given by

$$\mathcal{H}_n^{\text{hGP}}(w) = \mathcal{H}_n^{\text{GP}}(q) \quad \mathcal{H}_n^{\widetilde{\text{hGP}}}(w) = \mathcal{H}_n^{\widetilde{\text{GP}}}(q)$$

for $n \geq -1$. We define the N -truncated hGP and $\widetilde{\text{hGP}}$ hierarchies as the hierarchies of Hamiltonian flows

$$\partial_{t_n} w = \frac{1}{4} \begin{pmatrix} -\partial_x a^{-1} - a^{-1} \partial_x & \partial_x a^{-1} - a^{-1} \partial_x \\ -\partial_x a^{-1} + a^{-1} \partial_x & \partial_x a^{-1} + a^{-1} \partial_x \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta w_+} \mathcal{H}_n^{\text{hGP}}(w) \\ \frac{\delta}{\delta w_-} \mathcal{H}_n^{\text{hGP}}(w) \end{pmatrix} \quad (\text{hGP}_n)$$

$$\partial_{t_n} w = \frac{1}{4} \begin{pmatrix} -\partial_x a^{-1} - a^{-1} \partial_x & \partial_x a^{-1} - a^{-1} \partial_x \\ -\partial_x a^{-1} + a^{-1} \partial_x & \partial_x a^{-1} + a^{-1} \partial_x \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta w_+} \mathcal{H}_n^{\widetilde{\text{hGP}}}(w) \\ \frac{\delta}{\delta w_-} \mathcal{H}_n^{\widetilde{\text{hGP}}}(w) \end{pmatrix} \quad (\widetilde{\text{hGP}}_n)$$

for $n \in \{0, \dots, N\}$. Note that here $a = 1 + \frac{1}{2}(w_+ - w_-)$, $|a| > \delta$, and a^{-1} should be understood as the operator of multiplication by a^{-1} . In Appendix B.2 we detail the complete Hamiltonian structure of

(hGP_n), which is chosen so that $q \leftrightarrow w$ is a symplecto-/Poisson morphism. It is indeed the case that (hGP_1), ($\widetilde{\text{hGP}}_1$), and (hGP-1)–(hGP-2) are all identical.

For a small $\varepsilon > 0$ we now set

$$X = \sqrt{2}\varepsilon x \quad T_2 = 2\sqrt{2}\varepsilon t$$

and consider the rescaled variables

$$\begin{aligned} W_{\pm}(T_2, X) &= \varepsilon^{-2} w_{\pm}(t, x) \\ A(T_2, X) &= 1 + \frac{\varepsilon^2}{2} (W_+(T_2, X) - W_-(T_2, X)) = a(t, x). \end{aligned}$$

Transforming back to q yields the ansatz (1.1) above, up to the use of T_2 instead of the time vector $\mathbf{T}$, and the necessary reconstruction of the uniform-in- x phase function $\varphi(\mathbf{t})$. We refer to the system that $W = (W_+, W_-)$ solves as the ε -rescaled hydrodynamic Gross-Pitaevskii equations:

$$\partial_{T_2} W_+ = -\partial_X W_+ + \frac{\varepsilon^2}{4} \left(-(2W_- + 6W_+) \partial_X W_+ + 2\partial_X \left(\frac{\partial_{XX}(W_+ - W_-)}{2A} \right) \right) \quad (\varepsilon\text{GP-1})$$

$$\partial_{T_2} W_- = \partial_X W_- + \frac{\varepsilon^2}{4} \left(-(6W_- + 2W_+) \partial_X W_- + 2\partial_X \left(\frac{\partial_{XX}(W_+ - W_-)}{2A} \right) \right). \quad (\varepsilon\text{GP-2})$$

They have the infinite family of conserved energies

$$\mathcal{H}_n^{\varepsilon\text{GP}}(W) = \mathcal{H}_n^{\text{hGP}}(w) \quad \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) = \mathcal{H}_n^{\text{h}\widetilde{\text{GP}}}(w)$$

for $n \geq -1$, the first of which are

$$\mathcal{H}_{-1}^{\varepsilon\text{GP}}(W) = -\frac{\varepsilon}{\sqrt{2}} \int_{\mathbb{R}} W_+ + W_- \, dX \quad (\text{Phase Change})$$

$$\mathcal{H}_0^{\varepsilon\text{GP}}(W) = \frac{1}{\sqrt{2}\varepsilon} \int_{\mathbb{R}} \left(1 + \varepsilon^2 \frac{W_+ - W_-}{2} \right)^2 - 1 \, dX \quad (\text{Mass})$$

$$\mathcal{H}_1^{\varepsilon\text{GP}}(W) = -\frac{\varepsilon}{\sqrt{2}} \int_{\mathbb{R}} \left(1 + \frac{\varepsilon^2}{2} (W_+ - W_-) \right)^2 (W_+ + W_-) \, dX \quad (\text{Momentum})$$

$$\begin{aligned} \mathcal{H}_2^{\varepsilon\text{GP}}(W) &= \frac{1}{\sqrt{2}\varepsilon} \int \frac{\varepsilon^6}{2} (\partial_X (W_+ - W_-))^2 + \varepsilon^4 \left(1 + \varepsilon^2 \frac{W_+ - W_-}{2} \right)^2 (W_+ + W_-)^2 \\ &\quad + \left(\left(1 + \varepsilon^2 \frac{W_+ - W_-}{2} \right)^2 - 1 \right)^2 \, dX. \end{aligned} \quad (\text{Energy})$$

We define the N -truncated εGP and $\varepsilon\widetilde{\text{GP}}$ hierarchies as the hierarchies of Hamiltonian flows

$$\partial_{T_n} W = \frac{dt_n}{dT_n} \frac{1}{2\varepsilon^2} \begin{pmatrix} -\partial_X A^{-1} - A^{-1} \partial_X & \partial_X A^{-1} - A^{-1} \partial_X \\ -\partial_X A^{-1} + A^{-1} \partial_X & \partial_X A^{-1} + A^{-1} \partial_X \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{H}_n^{\varepsilon\text{GP}}(W) \\ \frac{\delta}{\delta W_-} \mathcal{H}_n^{\varepsilon\text{GP}}(W) \end{pmatrix} \quad (\varepsilon\text{GP}_n)$$

$$\partial_{T_n} W = \frac{dt_n}{dT_n} \frac{1}{2\varepsilon^2} \begin{pmatrix} -\partial_X A^{-1} - A^{-1} \partial_X & \partial_X A^{-1} - A^{-1} \partial_X \\ -\partial_X A^{-1} + A^{-1} \partial_X & \partial_X A^{-1} + A^{-1} \partial_X \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \\ \frac{\delta}{\delta W_-} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \end{pmatrix} \quad (\varepsilon\widetilde{\text{GP}}_n)$$

for $n \in \{0, \dots, N\}$. Here

$$A(\mathbf{T}, X) = 1 + \frac{\varepsilon^2}{2} (W_+ - W_-)(\mathbf{T}, X)$$

is the rescaled amplitude. In Appendix B.3 we detail the complete Hamiltonian structure of (εGP_n), which is chosen so that $w \leftrightarrow W$ is a symplecto-/Poisson morphism. Recall that the hGP , $\text{h}\widetilde{\text{GP}}$, εGP , and $\varepsilon\widetilde{\text{GP}}$ hierarchies are equipped with the zero boundary conditions

$$\lim_{|x| \rightarrow \infty} w(\mathbf{t}, x) = 0 \quad \lim_{|X| \rightarrow \infty} W(\mathbf{T}, X) = 0.$$

1.3 Approximation of $\varepsilon\mathbf{GP}$ by transport and $\mathbf{KdV}$ equations from [12, 13]

Observe that the leading order behavior of $(\varepsilon\mathbf{GP-1})$ – $(\varepsilon\mathbf{GP-2})$ is given by two transport equations with constant velocities ± 1 . As such, we can think of W_+ and W_- as two largely decoupled waves moving in opposite directions. This was observed and rigorously studied in [10]. The next-to-leading order of the system is reminiscent of the $\mathbf{KdV}$ equation

$$\partial_T U = 6UU_X - U_{XXX}, \quad (\mathbf{KdV})$$

although non-trivial interaction terms such as $W_+ \partial_X W_-$ formally persist in the limit $\varepsilon \rightarrow 0$. In [12, 13] F. B  thuel, P. Gravejat, J.-C. Saut, and D. Smets studied $\mathbf{hGP}$ in the long-wave regime described above and proved that the system $(\varepsilon\mathbf{GP-1})$ – $(\varepsilon\mathbf{GP-2})$ indeed behaves at leading order like two transport equations, and at next-to-leading order like two $\mathbf{KdV}$ equations. Specifically, they proved [12, Theorem 1], which states that for any solution q of $\mathbf{GP}$ in a Zhidkov space of sufficient regularity, for which $W(0)$ satisfies additional estimates, there exist two solutions $\pm U_\pm$ of $\mathbf{KdV}$ so that

$$\left\| 6W_\pm \left(\frac{8}{\varepsilon^2} \tau, X \mp \frac{8}{\varepsilon^2} \tau \right) - U_\pm(\tau, \pm X) \right\|_{H^s} \leq \varepsilon^2 C(W(0)) e^{C(W(0))|\tau|} \quad (1.8)$$

for all $\tau \in \mathbb{R}$ and $s \in \mathbb{N}$. Both the transport equation $(\mathbf{KdV}_0)$ and of course $\mathbf{KdV} = (\mathbf{KdV}_1)$ are part of the $\mathbf{KdV}$ hierarchy. Note that after switching to a co-moving frame to remove the leading-order transport behavior, we need to accelerate the timescale by a factor of order ε^2 in order to see the $\mathbf{KdV}$ approximation.

It turns out to be the case that for the higher equations in the $\mathbf{NLS}$, $\mathbf{GP}$ and $\widetilde{\mathbf{GP}}$ hierarchies, or more precisely the $\varepsilon\mathbf{GP}$ and $\varepsilon\widetilde{\mathbf{GP}}$ hierarchies, the higher equations in the $\mathbf{KdV}$ hierarchy show up at leading order as approximating equations. We can demonstrate this by formally expanding

$$A^{-1} = \sum_{k=0}^{\infty} \left(-\frac{\varepsilon^2}{2} (W_+ - W_-) \right)^k$$

and computing the initial orders in powers of ε of the right-hand sides of $(\varepsilon\mathbf{GP}_n)$ and $(\varepsilon\widetilde{\mathbf{GP}}_n)$. For the flows with even n there is no difference between the $\varepsilon\mathbf{GP}$ and $\varepsilon\widetilde{\mathbf{GP}}$ hierarchies, and we obtain

$$\partial_{T_0} W = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (\varepsilon\widetilde{\mathbf{GP}}_0)$$

$$\partial_{T_2} W = \begin{pmatrix} -(W_+)_X \\ (W_-)_X \end{pmatrix} + \frac{\varepsilon^2}{4} \begin{pmatrix} (W_+)_{XXX} - (W_-)_{XXX} - 6(W_+)_X W_+ - 2(W_+)_X W_- \\ (W_+)_{XXX} - (W_-)_{XXX} - 2(W_-)_X W_+ - 6(W_-)_X W_- \end{pmatrix} + O(\varepsilon^4) \quad (\varepsilon\widetilde{\mathbf{GP}}_2)$$

$$\begin{aligned} \partial_{T_4} W &= \begin{pmatrix} (W_+)_{XXX} - 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} \quad (\varepsilon\widetilde{\mathbf{GP}}_4) \\ &+ \frac{\varepsilon^2}{4} \begin{pmatrix} -(W_+)_{XXXXX} + (W_-)_{XXXXX} + 13(W_+)_{XXX} W_+ + 3(W_+)_{XXX} W_- + 29(W_+)_{XX} (W_+)_X \\ -(W_+)_{XXXXX} + (W_-)_{XXXXX} + 13(W_-)_{XXX} W_- + 3(W_-)_{XXX} W_+ + 29(W_-)_{XX} (W_-)_X \end{pmatrix} \\ &+ \frac{\varepsilon^2}{4} \begin{pmatrix} 9(W_-)_{XX} (W_-)_X + 3(W_+)_{XX} (W_-)_X - (W_+)_X (W_-)_{XX} + 3(W_-)_{XXX} W_- - 3(W_-)_{XXX} W_+ \\ 9(W_+)_{XX} (W_+)_X + 3(W_+)_X (W_-)_{XX} - (W_+)_X (W_-)_{XX} + 3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- \end{pmatrix} \\ &+ \frac{\varepsilon^2}{4} \begin{pmatrix} -45(W_+)_X W_+^2 - 6(W_+)_X W_+ W_- + 3(W_+)_X W_-^2 \\ 45(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- - 3(W_-)_X W_+^2 \end{pmatrix} + O(\varepsilon^4). \end{aligned}$$

Observe that the leading orders correspond precisely to the initial equations $(\mathbf{KdV}_n)$ in the $\mathbf{KdV}$ hierarchy, and this happens on a purely algebraic level. For the flows with odd n we first look at the $\varepsilon\mathbf{GP}$

hierarchy:

$$\partial_{T_1} W = \begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} \quad (\varepsilon \text{GP}_1)$$

$$\partial_{T_3} W = \frac{1}{\varepsilon^2} \begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} + \begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} + O(\varepsilon^2) \quad (\varepsilon \text{GP}_3)$$

$$\begin{aligned} \partial_{T_5} W &= \frac{1}{2\varepsilon^4} \begin{pmatrix} -(W_+)_X \\ -(W_-)_X \end{pmatrix} + \frac{1}{\varepsilon^2} \begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} \\ &\quad + \begin{pmatrix} (W_+)_{XXXXX} - 10(W_+)_{XXX} W_+ - 20(W_+)_{XX} (W_+)_X + 30(W_+)_X W_+^2 \\ (W_-)_{XXXXX} + 10(W_-)_{XXX} W_- + 20(W_-)_{XX} (W_-)_X + 30(W_-)_X W_-^2 \end{pmatrix} + O(\varepsilon^2). \end{aligned} \quad (\varepsilon \text{GP}_5)$$

Here the dynamics at leading order are mixed up. The $\varepsilon \widetilde{\text{GP}}$ hierarchy is a renormalization which separates these dynamics:

$$\partial_{T_1} W = \begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} \quad (\varepsilon \widetilde{\text{GP}}_1)$$

$$\begin{aligned} \partial_{T_3} W &= \begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} + \frac{\varepsilon^2}{4} \begin{pmatrix} -3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- \\ 3(W_-)_{XXX} W_+ + 3(W_-)_{XXX} W_- \end{pmatrix} \\ &\quad + \frac{\varepsilon^2}{4} \begin{pmatrix} -9(W_+)_{XX} (W_+)_X - 3(W_+)_{XX} (W_-)_X + 3(W_-)_{XXX} W_- + 3(W_-)_{XXX} W_+ + 3(W_+)_X W_-^2 \\ 9(W_-)_{XX} (W_-)_X + 3(W_+)_X (W_-)_{XX} - 3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- + 3(W_-)_X W_+^2 \end{pmatrix} \\ &\quad + \frac{\varepsilon^2}{4} \begin{pmatrix} 3(W_+)_X (W_-)_{XX} + 9(W_-)_{XX} (W_-)_X + 15(W_+)_X W_+^2 + 6(W_+)_X W_+ W_- \\ -3(W_+)_{XX} (W_-)_X - 9(W_+)_{XX} (W_+)_X + 15(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- \end{pmatrix} + O(\varepsilon^4) \end{aligned} \quad (\varepsilon \widetilde{\text{GP}}_3)$$

$$\partial_{T_5} W = \begin{pmatrix} (W_+)_{XXXXX} - 10(W_+)_{XXX} W_+ - 20(W_+)_{XX} (W_+)_X + 30(W_+)_X W_+^2 \\ (W_-)_{XXXXX} + 10(W_-)_{XXX} W_- + 20(W_-)_{XX} (W_-)_X + 30(W_-)_X W_-^2 \end{pmatrix} + O(\varepsilon^2). \quad (\varepsilon \widetilde{\text{GP}}_5)$$

The correspondence between the leading order of the equations in the $\varepsilon \widetilde{\text{GP}}$ hierarchy and the equations in the KdV hierarchy, for all $n \in \mathbb{N}$, is the content of Theorems 1.1 and 1.8 below, and also our main result Theorem 1.12. Let us state clearly at this point that for the case of $\text{GP} = (\text{GP}_2)$ these theorems *do not* generalize the next-to-leading order KdV approximation result, i. e. (1.8). They only generalize the leading order transport approximation result. Note also that the leading order contains no interaction terms, which simplifies the analysis.

This work was inspired by [13, Remarks 1.4, 1.5, 4.7], where the authors comment on their discovery of “[...] a strong and somewhat striking relationship between the (GP) invariants and the (KdV) invariants”. Specifically, they noticed that certain linear combinations of the conserved quantities $(\mathcal{H}_k^{\varepsilon \widetilde{\text{GP}}}(W))_{-1 \leq k \leq 2n}$ approximate the energy $\mathcal{E}_{n-1}^{\text{KdV}}(\pm W_{\pm})$ up to an error of size $O(\varepsilon^2)$ for $0 \leq n \leq 4$. They subsequently remark that “It would be of interest to investigate further the relationships between (GP) and (KdV), in particular at the level of the spectral problems associated to the corresponding inverse scattering methods.”, and that “[the relationship mentioned above] might be extended to higher order (GP) and (KdV) invariants, provided one was first able to compute some expressions for them.”.

We carry this out, extending the relationship between $(\mathcal{H}_k^{\varepsilon \widetilde{\text{GP}}}(W))_{-1 \leq k \leq 2n}$ and $\mathcal{E}_{n-1}^{\text{KdV}}(\pm W_{\pm})$ to all $n \in \mathbb{N}$.

1.4 Approximation of Hamiltonians and transmission coefficients

Theorem 1.1. *All identities below hold up to formal integration by parts. For all $n \in \mathbb{N}$ there exist polynomials $\mathbf{P}_n$, depending on ε , W_+ , W_- , and up to $\lfloor \frac{n}{2} \rfloor$ derivatives thereof, such that*

$$\mathcal{H}_{2n}^{\varepsilon \widetilde{\text{GP}}}(W) \mp 2\mathcal{H}_{2n-1}^{\varepsilon \widetilde{\text{GP}}}(W) = (\sqrt{2}\varepsilon)^{2n+1} \left(\mathcal{E}_{n-1}^{\text{KdV}}(\pm W_{\pm}) + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{P}_{2n} \mp \mathbf{P}_{2n-1} dX \right) \quad (1.9)$$

In particular

$$\mathcal{H}_{2n}^{\varepsilon \widetilde{GP}}(W) = \frac{(\sqrt{2}\varepsilon)^{2n+1}}{2} \left(\mathcal{E}_{n-1}^{KdV}(-W_-) + \mathcal{E}_{n-1}^{KdV}(W_+) + \varepsilon^2 \int_{\mathbb{R}} \mathbf{P}_{2n} dX \right) \quad (1.10)$$

$$\mathcal{H}_{2n-1}^{\varepsilon \widetilde{GP}}(W) = \frac{(\sqrt{2}\varepsilon)^{2n+1}}{4} \left(\mathcal{E}_{n-1}^{KdV}(-W_-) - \mathcal{E}_{n-1}^{KdV}(W_+) + \varepsilon^2 \int_{\mathbb{R}} \mathbf{P}_{2n-1} dX \right). \quad (1.11)$$

Furthermore, there are no constant or linear monomials in $\mathbf{P}_n$. Lastly, there exist polynomials $\mathbf{Q}_{2n}$, depending on ε , W_+ , W_- , and up to $n-1$ derivatives thereof, such that

$$\int_{\mathbb{R}} \mathbf{P}_{2n} dX = \frac{1}{4} \int_{\mathbb{R}} (\partial_X^n (W_+ - W_-))^2 dX + \int_{\mathbb{R}} \mathbf{Q}_{2n} dX. \quad (1.12)$$

Remark 1.2. The initial values of the error polynomials $\mathbf{P}_n$ are as follows:

$$\begin{aligned} \mathbf{P}_{-1} &= 0 & \mathbf{P}_0 &= \frac{1}{4}(W_+ - W_-)^2 & \mathbf{P}_1 &= \frac{1}{4}(-W_+^3 + W_+^2 W_- + W_+ W_-^2 - W_-^3) \\ \mathbf{P}_2 &= \frac{1}{4}(W_+ - W_-)_X^2 + \frac{1}{4}(3W_+^3 - W_+^2 W_- + W_+ W_-^2 - 3W_-^3) \\ &\quad + \frac{\varepsilon^2}{32}(5W_+^4 - 4W_+^3 W_- - 2W_+^2 W_-^2 - 4W_+ W_-^3 + 5W_-^4) \\ \mathbf{P}_3 &= \frac{1}{4}(-(W_+)_X^2(5W_+ + W_-) + 6(W_+)_X(W_-)_X(W_+ + W_-) - (W_-)_X^2(5W_- + W_+)) \\ &\quad + \frac{1}{4}(-5W_+^4 + 2W_+^3 W_- - 2W_+ W_-^3 + 5W_-^4) \\ &\quad + \frac{\varepsilon^2}{32}(-7W_+^5 + 5W_+^4 W_- + 2W_+^3 W_-^2 + 2W_+^2 W_-^3 + 5W_+ W_-^4 - 7W_-^5). \end{aligned}$$

In order to explain the origin of this result, we need to consider the Lax operators

$$\begin{aligned} L^{\varepsilon GP}(W) &= \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon \partial_X \\ i\sqrt{2}\varepsilon \partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix} & L^{KdV}(\pm W_{\pm}) &= -\partial_X^2 \pm W_{\pm} \\ L_0^{\varepsilon GP} &= \begin{pmatrix} 1 & i\sqrt{2}\varepsilon \partial_X \\ i\sqrt{2}\varepsilon \partial_X & -1 \end{pmatrix} & L_0^{KdV} &= -\partial_X^2. \end{aligned}$$

Their relevance is explained in Section 2.1, where we give a brief overview of the Lax pair formalisms for the equations under consideration. In the following, we denote by $\mathcal{Q}_j$, $j \in \{1, 2, 3, 4\}$ the four open quadrants of the complex plane. We take two spectral parameters $\lambda \in \overline{\mathcal{Q}_1} \cup \overline{\mathcal{Q}_3}$ and $k \in \overline{\mathcal{Q}_1}$ such that $\lambda = \pm\sqrt{1 + 2\varepsilon^2 k^2}$, and consider the scattering problems associated to $L^{\varepsilon GP} - \lambda$ and $L^{KdV} - k^2$. This is discussed in detail in Section 2, where we define in (2.14) the corresponding transmission coefficients $a^{\varepsilon GP}(\lambda, k; W)$ and $a^{KdV}(k; \pm W_{\pm})$. They are conserved quantities for every flow in their respective hierarchies and admit for $\lambda \in \overline{\mathcal{Q}_1}$ the following asymptotic expansions in powers of $2k$ at infinity:

$$\log a^{\varepsilon GP}(\lambda, k; W) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^{\varepsilon GP}(W)}{(2k)^{n+1}} \quad \log a^{KdV}(k; \pm W_{\pm}) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{E}_n^{KdV}(\pm W_{\pm})}{(2k)^{2n+3}}. \quad (1.13)$$

The precise sense of expansion is given in Definition 2.2 below. The following corollary presents a different perspective on Theorem 1.1. Here we define for functions $f = f(\lambda) : \mathbb{C} \rightarrow \mathbb{C}$ the even and odd parts

$$\text{Ev}_{\lambda} f(\lambda) = \frac{f(\lambda) + f(-\lambda)}{2} \quad \text{Od}_{\lambda} f(\lambda) = \frac{f(\lambda) - f(-\lambda)}{2}. \quad (1.14)$$

Proposition 1.3. Let $W = (W_+, W_-) \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$, $\varepsilon \in (0, 1)$, and $\lambda, k \in \overline{\mathcal{Q}_1}$ with $|k| > c(W)$ so that Lemma 2.3 below is applicable and hence the asymptotic expansions (1.13) exist. Then

$$\text{Ev}_\lambda[\log a^{\varepsilon GP}(\lambda, k; W)] \mp \frac{1}{\lambda} \text{Od}_\lambda[\log a^{\varepsilon GP}(\lambda, k; W)] \sim \log a^{KdV}(k; (\pm W_\pm)) + \sum_{n=-1}^{\infty} \frac{O(\varepsilon^2)}{(2k)^{n+1}} \quad (1.15)$$

in the sense of asymptotic expansions as $2ik \rightarrow \infty$. Here $O(\varepsilon^2)$ must be understood as in Theorem 1.1, i. e. in terms of polynomials in ε , W_+ , W_- , and derivatives thereof.

If instead $\lambda \in (0, 1)$, $k \in i\mathbb{R}_+$, and $\kappa = -ik < \frac{1}{\sqrt{2}\varepsilon}$ satisfies

$$\max\left\{\frac{1}{\kappa}, \frac{1}{\kappa^3}\right\} \|W\|_{L^2}^2 \leq \frac{1}{30},$$

then

$$\left| \text{Ev}_\lambda[\log a^{\varepsilon GP}(\lambda, k; W)] \mp \frac{1}{\lambda} \text{Od}_\lambda[\log a^{\varepsilon GP}(\lambda, k; W)] - \log a^{KdV}(k; \pm W_\pm) \right| \leq 8\varepsilon^2 \kappa^{-1} \|W\|_{L^2}^2. \quad (1.16)$$

Proof. The asymptotic expansion relation (1.15) is a reformulation of Lemma 2.7, or alternatively Theorem 1.1, involving the identities (1.3), (2.18), (2.19), and (2.49). The second approximation result (1.16) is taken from Lemma A.6. $\square$

Remark 1.4 (Splitting of the Dirac operator in the nonrelativistic limit). Setting $c = \frac{1}{\varepsilon^2}$, we have

$$\frac{L^{\varepsilon GP}}{\varepsilon^2} = c \begin{pmatrix} 0 & \sqrt{2}i\partial_X \\ \sqrt{2}i\partial_X & 0 \end{pmatrix} + c^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -W_- & 0 \\ 0 & -W_+ \end{pmatrix},$$

which is a one-dimensional Dirac operator with speed of light c and mass $m = 1$. On the other hand, $L^{KdV}(\pm W_\pm)$ is the Schrödinger operator with potential $\pm W_\pm$. The limit $\varepsilon \rightarrow 0$ that we are interested in corresponds to the nonrelativistic limit $c \rightarrow \infty$, where one expects to recover from the Dirac operator a combination of two Schrödinger operators. The problem of taking this limit is well-studied in mathematical physics. We refer to [82, Chapter 6] for an introduction in a general setting. As an example of such a result, in [34, Corollary 4.1] (see also [84, Theorem 4.1]) it is proven that

$$\left(\frac{L^{\varepsilon GP} \mp 1 - \lambda}{\varepsilon^2} \right)^{-1} \xrightarrow{\varepsilon \rightarrow 0} \mp (L^{KdV}(\pm W_\pm) + \kappa^2)^{-1} \begin{pmatrix} \frac{1 \pm 1}{2} & 0 \\ 0 & \frac{1 \mp 1}{2} \end{pmatrix}$$

uniformly strongly for any $\lambda \in \mathbb{C} \setminus \mathbb{R}$. More generally, one may study the asymptotic behavior of the Dirac operator in the nonrelativistic regime by considering various quantities as holomorphic functions of $\frac{1}{c^2}$ around $c = \infty$ and expanding them in a power series. For example, in [34, Theorem 2.1] such a holomorphic representation of the Dirac operator is used to prove the aforementioned result. Most relevant for us is [16, Theorem 4.2], where such a holomorphic representation of the scattering matrix associated to $\frac{L^{\varepsilon GP} \mp 1}{\varepsilon^2}$ is given, and the initial expansion coefficients are explicitly calculated. In particular, the scattering matrix associated to $L^{KdV}(\pm W_\pm)$ is recovered in the limit. Since the scattering matrices contain the transmission coefficients $a^{\varepsilon GP}(\lambda; W)$ and $a^{KdV}(\kappa, \pm W_\pm)$, this result is in correspondence with (1.15), although it is not identical. Admitting that we cannot speak with confidence about this broad field of mathematical physics, we hope that Proposition 1.3 may present a novel perspective on the nonrelativistic limit of the one-dimensional Dirac operator.

A consequence of Proposition 1.3 is that the energies $\mathcal{E}_n^{KdV}$ on the right-hand side of (1.9) can be used to control the Sobolev norms of $\pm W_\pm$ individually, up to an error of size ε^2 . This is the content of the following corollary, which is proven in Section 2.5 by combining Theorem 1.1 with [61, Theorem 9.1]. Here we use the functionals

$$\mathcal{E}_n^{\varepsilon GP, \pm}(W) = \sum_{l=0}^n \binom{n}{l} (\sqrt{2}\varepsilon)^{-2l-3} (\mathcal{H}_{2l+2}^{\varepsilon GP}(W) \mp 2\mathcal{H}_{2l+1}^{\varepsilon GP}(W)),$$

which are of course conserved under every flow in the $\varepsilon\widetilde{\text{GP}}$ and $\varepsilon\widetilde{\text{GP}}$ hierarchies.

Corollary 1.5. *Let $n \in \mathbb{N}$ with $n \geq 1$ and $\varepsilon \in (0, 1)$. There exists some $\delta = \delta(n) \in (0, 1)$ and $C = C(n) > 0$ such that for all $W \in H^{n+1}$ with $\|W\|_{H^{-1}} < \delta$ we have $\mathcal{E}_n^{\varepsilon\widetilde{\text{GP}}, \pm}(W) < \infty$ and*

$$\left| \mathcal{E}_n^{\varepsilon\widetilde{\text{GP}}, \pm}(W) - \|W_{\pm}\|_{H^n}^2 - \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)\|_{H^n}^2 \right| \leq C \|W_{\pm}\|_{H^n}^2 \|W_{\pm}\|_{H^{-1}} + \varepsilon^2 C \|W\|_{H^n}^2 (1 + \|W\|_{H^n})^{n+1}.$$

Remark 1.6. *We use Corollary 1.5 to prove the crucial estimates (1.23) and (1.24) below, which control the size of W globally in time. One may also attempt such estimates from (1.22), or more specifically the conserved energies $\sum_{l=0}^{h-1} \tau^{2(h-1-l)} \binom{h-1}{l} \mathcal{H}_{2l+2}^{\text{GP}}$ constructed in [59, 60] for $\tau \geq 2$. They control the size of the solution as measured by the semiclassical Sobolev norm*

$$\|(|q|^2 - 1, \partial_x q)\|_{H_\tau^h} = \left(\int_{\mathbb{R}} (\tau^2 + \xi^2)^h (|\widehat{|q|^2 - 1}|^2 + |\widehat{\partial_x q}|^2) d\xi \right)^{\frac{1}{2}}.$$

After obtaining an estimate, it is necessary to pass it through a rescaling and then the Madelung transform. The first step turns out to be problematic due to the inhomogeneity causing negative powers of ε to appear. This could be resolved by lifting the restriction $\tau \geq 2$.

1.5 Formal approximation of hierarchies

In the previous section we have established an approximative relationship between the Hamiltonians of the $\widetilde{\text{GP}}$ and KdV hierarchies. This suggests a method of proof for the pattern we have observed in Section 1.3 to hold for the leading orders of the Hamiltonian PDEs in the $\varepsilon\widetilde{\text{GP}}$ hierarchy. Such an argument requires the symplectic structures of GP and KdV to have a similar approximative relationship, which turns out to be the case. In Appendix B the reader may find a detailed description of the symplectic/Poisson structures of the hierarchies in question. Let us write $(\widetilde{\text{GP}}_n)$ in the Hamiltonian formulation

$$i\partial_{t_n} \begin{pmatrix} q \\ \bar{q} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta \bar{q}} \mathcal{H}_n^{\widetilde{\text{GP}}}(q) \\ \frac{\delta}{\delta q} \mathcal{H}_n^{\widetilde{\text{GP}}}(q) \end{pmatrix}.$$

Conjugation of the Hamiltonian operator with the map $(q, \bar{q}) \mapsto (w_+, w_-)$ and a careful change of variables for the functional derivative shows that the Hamiltonian formulation of $(\text{h}\widetilde{\text{GP}}_n)$ is

$$\partial_{t_n} \begin{pmatrix} w_+ \\ w_- \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -\partial_x a^{-1} - a^{-1} \partial_x & \partial_x a^{-1} - a^{-1} \partial_x \\ -\partial_x a^{-1} + a^{-1} \partial_x & \partial_x a^{-1} + a^{-1} \partial_x \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta w_+} \mathcal{H}_n^{\text{h}\widetilde{\text{GP}}}(w) \\ \frac{\delta}{\delta w_-} \mathcal{H}_n^{\text{h}\widetilde{\text{GP}}}(w) \end{pmatrix}.$$

Note that here $\partial_x a^{-1}$ and $a^{-1} \partial_x$ are compositions of operators. Then the Hamiltonian formulation of $(\varepsilon\widetilde{\text{GP}}_n)$ is

$$\begin{aligned} \partial_{T_n} \begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{dt_n}{dT_n} \frac{1}{2\varepsilon^2} \begin{pmatrix} -\partial_X A^{-1} - A^{-1} \partial_X & \partial_X A^{-1} - A^{-1} \partial_X \\ -\partial_X A^{-1} + A^{-1} \partial_X & \partial_X A^{-1} + A^{-1} \partial_X \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \\ \frac{\delta}{\delta W_-} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \end{pmatrix} \\ &= \frac{dt_n}{dT_n} \frac{2}{\varepsilon^2} \left(\begin{pmatrix} -\frac{1}{2} \partial_X & 0 \\ 0 & \frac{1}{2} \partial_X \end{pmatrix} + \varepsilon^2 \mathbf{E}_n \right) \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \\ \frac{\delta}{\delta W_-} \mathcal{H}_n^{\varepsilon\widetilde{\text{GP}}}(W) \end{pmatrix}, \end{aligned}$$

where the scaling factor $\frac{dt_n}{dT_n}$ was defined in (1.2) and

$$\begin{aligned} \mathbf{E}_n &= \frac{A^{-1} - 1}{2\varepsilon^2} \begin{pmatrix} -\partial_X & 0 \\ 0 & \partial_X \end{pmatrix} + \frac{A_X A^{-2}}{4\varepsilon^2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \\ &= -\frac{1}{4} \frac{W_+ - W_-}{1 + \frac{\varepsilon^2}{2}(W_+ - W_-)} \begin{pmatrix} -\partial_X & 0 \\ 0 & \partial_X \end{pmatrix} + \frac{1}{8} \frac{(W_+ - W_-)_X}{(1 + \frac{\varepsilon^2}{2}(W_+ - W_-))^2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}. \end{aligned}$$

Applying Theorem 1.1 yields

$$\begin{aligned}
\partial_{T_{2n}} \begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{dt_{2n}}{dT_{2n}} \frac{2}{\varepsilon^2} \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2 \mathbf{E}_{2n} \right) \frac{(\sqrt{2}\varepsilon)^{2n+1}}{2} \\
&\quad \times \left(\begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2 \begin{pmatrix} \frac{\delta}{\delta W_+} \int_{\mathbb{R}} \mathbf{P}_{2n} dX \\ \frac{\delta}{\delta W_-} \int_{\mathbb{R}} \mathbf{P}_{2n} dX \end{pmatrix} \right) \\
&= \begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2 \mathbf{R}_{2n} \\
\partial_{T_{2n-1}} \begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{dt_{2n-1}}{dT_{2n-1}} \frac{2}{\varepsilon^2} \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2 \mathbf{E}_{2n-1} \right) \frac{(\sqrt{2}\varepsilon)^{2n+1}}{4} \\
&\quad \times \left(\begin{pmatrix} -\frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2 \begin{pmatrix} \frac{\delta}{\delta W_+} \int_{\mathbb{R}} \mathbf{P}_{2n-1} dX \\ \frac{\delta}{\delta W_-} \int_{\mathbb{R}} \mathbf{P}_{2n-1} dX \end{pmatrix} \right) \\
&= \begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} \begin{pmatrix} -\frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2 \mathbf{R}_{2n-1}
\end{aligned}$$

with residual terms

$$\begin{aligned}
\mathbf{R}_{2n} &= \mathbf{E}_{2n} \begin{pmatrix} \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2 \mathbf{E}_{2n} \right) \begin{pmatrix} \frac{\delta}{\delta W_+} \int_{\mathbb{R}} \mathbf{P}_{2n} dX \\ \frac{\delta}{\delta W_-} \int_{\mathbb{R}} \mathbf{P}_{2n} dX \end{pmatrix} \\
\mathbf{R}_{2n-1} &= \mathbf{E}_{2n-1} \begin{pmatrix} -\frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \\ \frac{\delta}{\delta W_-} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \end{pmatrix} + \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2 \mathbf{E}_{2n-1} \right) \begin{pmatrix} \frac{\delta}{\delta W_+} \int_{\mathbb{R}} \mathbf{P}_{2n-1} dX \\ \frac{\delta}{\delta W_-} \int_{\mathbb{R}} \mathbf{P}_{2n-1} dX \end{pmatrix}.
\end{aligned}$$

Remark 1.7. The initial values of the residual terms $\mathbf{R}_n$ are as follows:

$$\begin{aligned}
\mathbf{R}_0 &= \mathbf{R}_1 = 0 \\
\mathbf{R}_2 &= \frac{1}{4} \begin{pmatrix} (W_+)_{XXX} - 6(W_+)_X W_+ - (W_-)_{XXX} - 2(W_+)_X W_- \\ -(W_-)_{XXX} - 6(W_-)_X W_- + (W_+)_{XXX} - 2(W_-)_X W_+ \end{pmatrix} + O(\varepsilon^2) \\
\mathbf{R}_3 &= \frac{1}{4} \begin{pmatrix} -3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- - 9(W_+)_{XX} (W_+)_X - 3(W_+)_{XX} (W_-)_X \\ -3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- - 9(W_+)_{XX} (W_+)_X - 3(W_+)_{XX} (W_-)_X \end{pmatrix} \\
&\quad + \frac{1}{4} \begin{pmatrix} 3(W_+)_X (W_-)_{XX} + 15(W_+)_X W_+^2 + 6(W_+)_X W_+ W_- + 3(W_+)_X W_-^2 \\ 3(W_+)_X (W_-)_{XX} + 15(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- + 3(W_-)_X W_+^2 \end{pmatrix} \\
&\quad + \frac{1}{4} \begin{pmatrix} 3(W_-)_{XXX} W_+ + 3(W_-)_{XXX} W_- + 9(W_-)_{XX} (W_-)_X \\ 3(W_-)_{XXX} W_+ + 3(W_-)_{XXX} W_- + 9(W_-)_{XX} (W_-)_X \end{pmatrix} + O(\varepsilon^2) \\
\mathbf{R}_4 &= \frac{1}{4} \begin{pmatrix} -(W_+)_{XXXXX} + 13(W_+)_{XXX} W_+ + 3(W_+)_{XXX} W_- + 29(W_+)_{XX} (W_+)_X + 3(W_+)_{XX} (W_-)_X \\ -(W_+)_{XXXXX} + 13(W_-)_{XXX} W_- + 3(W_+)_{XXX} W_+ + 29(W_-)_{XX} (W_-)_X + 3(W_+)_{XX} (W_-)_{XX} \end{pmatrix} \\
&\quad + \frac{1}{4} \begin{pmatrix} -(W_+)_{XX} (W_-)_{XX} - 45(W_+)_{XX} W_+^2 - 6(W_+)_{XX} W_+ W_- + 3(W_+)_{XX} W_-^2 + (W_-)_{XXXXX} \\ -(W_+)_{XX} (W_-)_{XX} + 45(W_-)_{XX} W_-^2 + 6(W_-)_{XX} W_+ W_- - 3(W_-)_{XX} W_+^2 + (W_-)_{XXXXX} \end{pmatrix} \\
&\quad + \frac{1}{4} \begin{pmatrix} -3(W_-)_{XXX} W_+ + 3(W_-)_{XXX} W_- + 9(W_-)_{XX} (W_-)_X \\ -3(W_+)_{XXX} W_- + 3(W_-)_{XXX} W_+ + 9(W_+)_{XX} (W_+)_X \end{pmatrix} + O(\varepsilon^2).
\end{aligned}$$

It is the KdV-like structure of $\mathbf{R}_2$ that is proven in [12, 13] to yield a KdV approximation at next-to-leading order.

We define invertible coefficient matrices $\mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}, \mathbf{c}_{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} \in \mathbb{R}^{(N+1) \times (N+1)}$ by

$$\begin{aligned} \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}^{2n, 2m} &= \binom{n - \frac{1}{2}}{n - m} \left(-\frac{2}{\varepsilon^2} \right)^{n-m} & \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}^{2n, 2m-1} &= 0 \\ \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}^{2n-1, 2m} &= 0 & \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}^{2n-1, 2m-1} &= \binom{n - \frac{1}{2}}{n - m} \left(-\frac{2}{\varepsilon^2} \right)^{n-m} \end{aligned}$$

and $\mathbf{c}_{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} = (\mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}})^{-1}$. For any of the coefficient matrices $\mathbf{c}_{\dots \rightarrow \dots}$ that we use, we define for their transposes the notation

$$\mathbf{c}^{\dots \rightarrow \dots} = (\mathbf{c}_{\dots \rightarrow \dots})^T \quad \mathbf{c}_{m,n}^{\dots \rightarrow \dots} = \mathbf{c}_{\dots \rightarrow \dots}^{n,m}.$$

Then with

$$(\nabla_{\mathbf{t}} \mathbf{T}) = \text{diag} \left(\left(\frac{dT_n}{dt_n} \right)_{0 \leq n \leq N} \right) \quad (\nabla_{\mathbf{T}} \mathbf{t}) = \text{diag} \left(\left(\frac{dt_n}{dT_n} \right)_{0 \leq n \leq N} \right)$$

we can write

$$\begin{aligned} \mathbf{c}_{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} &= (\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}_{\widetilde{\text{GP}} \rightarrow \text{NLS}} (\nabla_{\mathbf{T}} \mathbf{t}) & \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} &= (\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}_{\text{NLS} \rightarrow \widetilde{\text{GP}}} (\nabla_{\mathbf{T}} \mathbf{t}) \\ \mathbf{c}_{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} &= (\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}_{\widetilde{\text{GP}} \rightarrow \text{NLS}} (\nabla_{\mathbf{T}} \mathbf{t}) & \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} &= (\nabla_{\mathbf{T}} \mathbf{t}) \mathbf{c}_{\text{NLS} \rightarrow \widetilde{\text{GP}}} (\nabla_{\mathbf{t}} \mathbf{T}). \end{aligned}$$

For later, we define also

$$\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \text{GP}} = (\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\text{NLS} \rightarrow \text{GP}} (\nabla_{\mathbf{T}} \mathbf{t}) \quad \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \rightarrow \varepsilon \text{GP}} = \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} \mathbf{c}^{\text{NLS} \rightarrow \varepsilon \text{GP}}.$$

For the convenience of the reader, we note the following facts: $\mathbf{c}_{\dots \rightarrow \dots}$ is always lower triangular and $\mathbf{c}^{\dots \rightarrow \dots}$ always upper triangular. Furthermore, the diagonal values are all 1, these matrices are always invertible, and the row and column indices of any nonzero entry have matching parity.

Theorem 1.8 (Formal approximation of the $\varepsilon \widetilde{\text{GP}}$ hierarchy by the KdV hierarchy). *Suppose that q is formally a solution of the N -truncated NLS hierarchy. If $W(\mathbf{T})$ is defined by (1.1), then $W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} \mathbf{T})$ solves the N -truncated $\varepsilon \widetilde{\text{GP}}$ hierarchy and*

$$\frac{d}{dT_n} W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} \mathbf{T}) = \left(\begin{aligned} & -(-1)^{n-\frac{1}{2}} \partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{\lfloor \frac{n-1}{2} \rfloor}^{\text{KdV}} (W_+(\mathbf{T})) \\ & -\frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{\lfloor \frac{n-1}{2} \rfloor}^{\text{KdV}} (-W_-(\mathbf{T})) \end{aligned} \right) + \varepsilon^2 \mathbf{R}_n(\mathbf{T}).$$

Correspondingly, $W(\mathbf{T})$ solves

$$\begin{aligned} \partial_{T_{2n}} W(\mathbf{T}) &= \sum_{m=0}^n \binom{n - \frac{1}{2}}{n - m} \left(\frac{2}{\varepsilon^2} \right)^{n-m} \left(\begin{aligned} & -\frac{1}{2} \partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{m-1}^{\text{KdV}} (W_+(\mathbf{T})) \\ & -\frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{m-1}^{\text{KdV}} (-W_-(\mathbf{T})) \end{aligned} \right) + \varepsilon^2 \mathbf{R}_{2m}(\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} \mathbf{T}) \\ \partial_{T_{2n-1}} W(\mathbf{T}) &= \sum_{m=0}^n \binom{n - \frac{1}{2}}{n - m} \left(\frac{2}{\varepsilon^2} \right)^{n-m} \left(\begin{aligned} & \frac{1}{2} \partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{m-1}^{\text{KdV}} (W_+(\mathbf{T})) \\ & -\frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{m-1}^{\text{KdV}} (-W_-(\mathbf{T})) \end{aligned} \right) + \varepsilon^2 \mathbf{R}_{2m-1}(\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \rightarrow \text{NLS}} \mathbf{T}). \end{aligned}$$

Proof. By the chain rule we have

$$\nabla_{\mathbf{T}} (W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} \mathbf{T})) = \mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} (\nabla_{\mathbf{T}} W) (\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} \mathbf{T}).$$

Together with (1.6)–(1.7) and the definition of $\mathbf{c}_{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}}$, this implies that $W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widetilde{\text{GP}}} \mathbf{T})$ solves the N -truncated $\varepsilon \widetilde{\text{GP}}$ hierarchy. Then the claimed formulas follow from the elaborations above the theorem. $\square$

Remark 1.9 (Conjecture on the next-to-leading order of the $\varepsilon\widehat{\text{GP}}$ hierarchy). *We have verified the following formula to hold for $n \in \{1, 2, 3, 4\}$:*

$$\begin{aligned} 4(\mathbf{R}_{2n} + \mathbf{R}_{2n-1}) &= \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}_n^{\text{KdV}}(W_+) + \partial_{XX} \left(\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-) \right) + \text{interaction terms} + O(\varepsilon^2) \right) \\ &\quad O(\varepsilon^0) \\ 4(\mathbf{R}_{2n} - \mathbf{R}_{2n-1}) &= \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta W_-} \mathcal{E}_n^{\text{KdV}}(W_-) + \partial_{XX} \left(\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+) \right) + \text{interaction terms} + O(\varepsilon^2) \right) \\ &\quad O(\varepsilon^0) \end{aligned}$$

Here “interaction terms” are those which contain as factors both W_+ and W_- , or, respectively, some derivatives thereof. The significance of the above formulas is the following: suppose $W(\mathbf{T})$ is a solution to the $\varepsilon\widehat{\text{GP}}$ hierarchy. Then Theorem 1.8 implies

$$\begin{aligned} \frac{1}{2} \left(\frac{d}{dT_{2n}} + \frac{d}{dT_{2n-1}} \right) W(\mathbf{T}) &= \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-(\mathbf{T})) \right) \\ &\quad + \frac{\varepsilon^2}{8} \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}_n^{\text{KdV}}(W_+(\mathbf{T})) + \partial_{XX} \left(\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_{n-1}^{\text{KdV}}(-W_-(\mathbf{T})) \right) + \text{interaction terms} + O(\varepsilon^2) \right) \\ &\quad O(\varepsilon^0) \\ \frac{1}{2} \left(\frac{d}{dT_{2n}} - \frac{d}{dT_{2n-1}} \right) W(\mathbf{T}) &= \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+(\mathbf{T})) \right) \\ &\quad + \frac{\varepsilon^2}{8} \left(-\frac{1}{2}\partial_X \frac{\delta}{\delta W_-} \mathcal{E}_n^{\text{KdV}}(W_-(\mathbf{T})) + \partial_{XX} \left(\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}_{n-1}^{\text{KdV}}(W_+(\mathbf{T})) \right) + \text{interaction terms} + O(\varepsilon^2) \right) \\ &\quad O(\varepsilon^0) \end{aligned}$$

This suggests that an extension of Theorem 1.8 and Theorem 1.12 below to the next-to-leading order is plausible: while $(\varepsilon\widehat{\text{GP}}_{2n})$ has its leading order behavior given by (KdV_{n-1}) , switching to a “co- $(\varepsilon\widehat{\text{GP}}_{2n-1})$ -moving” frame allows us to see, on a longer/slower timescale, the dynamics of (KdV_n) in one component of our solution W . In the case $n = 1$, i. e. that of $(\varepsilon\widehat{\text{GP}}_2)$, this corresponds to the utilization of right- or left-moving frames in order to see the dynamics of (KdV_1) in one component of the solution. This is the content of [12]. Interestingly, this suggests the co-moving frame used in this result should be thought of as a renormalization using $(\varepsilon\widehat{\text{GP}}_1)$, instead of (KdV_0) , although they are both transport equations.

In the general case, supposing the above formulas can be proven to hold, the term wrapped in two derivatives and the interaction terms, which are a priori not negligible, may still turn out to be of order $O(\varepsilon^2)$ in the energy estimates by the same method that is used in [12]: one component of the solution travels rapidly to infinity due to the dynamics of (KdV_{n-1}) happening on a faster timescale.

1.6 Review of well-posedness results for the hierarchies

The introduction of [67] contains an overview of well-posedness results for NLS, GP, hGP, KdV, and the associated hierarchies. Since this work concerns the hierarchies, we focus our review of the literature on well-posedness results specifically for hierarchies. We start with the KdV hierarchy, where initial existence and uniqueness results were given by J.-C. Saut in [75] and M. Schwarz in [44]. Later, C. E. Kenig, G. Ponce and L. Vega produced a number of works [47, 48, 49, 50, 51, 52, 53], developing dispersive estimates and using them to prove the local well-posedness of various generalized KdV-type hierarchies in weighted Sobolev spaces. Another local well-posedness result in weighted Sobolev spaces for small initial data is given in [72]. This work also contains an ill-posedness result for all of the higher equations in the KdV hierarchy, in the sense of failure of the flow map to be C^2 in the Sobolev spaces $H^s(\mathbb{R})$. Nevertheless, by employing subtle energy estimates and a parabolic regularization argument, C. E. Kenig and D. Pilod [46] proved local well-posedness in $H^s(\mathbb{R})$ at high regularity. We combine this with the conserved quantities that H. Koch and D. Tataru constructed in [61], which prevent the Sobolev norm of a local solution from blowing up in finite time, to formulate the following global well-posedness result. Here and thereafter we use the notation $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$.

Proposition 1.10 ([46, Theorem 1.1] and [61, Theorem 9.1]). *Let $N \in \mathbb{N}$, $s \geq 4N - \frac{9}{2}$. For every $U_0 \in H^s$ there exists a unique $U \in C(\mathbb{R}^{N+1}; H^s)$ with initial data $U(0) = U_0$ that satisfies (KdV $_n$) in the sense of distributions for all $n \in \{0, \dots, N\}$. For every $s' \in [0, s]$ we have*

$$\|U\|_{L_T^\infty H_X^{s'}} \leq C(s')(1 + \|U_0\|_{H^{-1}}^{6s'}) \|U_0\|_{H^{s'}}. \quad (1.17)$$

Since we shall work with high regularity this formulation is all that we need, but it is far from the state of the art in the literature. In the low regularity regime A. Rybkin [74] and also R. Killip, M. Viřan, and X. Zhang [55, 56] achieved a breakthrough by using a perturbation determinant approach to construct conserved quantities that control the H^{-1} -norm of the solution. This allowed them to prove the global well-posedness of KdV in H^{-1} on the line and the torus, and their “method of commuting flows” has since been applied to several other integrable PDEs (see e. g. [15, 39, 54]). We use the uniform estimates they give for solutions to the KdV hierarchy in the proof of (1.17). Recently, the method of commuting flows has been used to prove the global well-posedness of the KdV hierarchy in $H^{-1}(\mathbb{R})$ by F. Klaus, H. Koch, and B. Liu [58].

Moving on to the NLS hierarchy, we mention the work [38] by A. Grünrock where the Fourier restriction norm method is used to prove global well-posedness and ill-posedness results for the flows in the NLS hierarchy, and also the KdV hierarchy, in Fourier-Lebesgue spaces at low regularity. This was extended by J. Adams in [4] to the even flows in the NLS hierarchy, and recently in [3] to the dNLS hierarchy. Note also [71], where orbital stability of the dark soliton solutions to the odd flows (NLS $_{2n+1}$) of the NLS hierarchy, with real-valued q , is proven.

Of direct relevance is the preceding work [67], where the global well-posedness of the NLS hierarchy with NZBC is proven, i. e. precisely the hierarchy we study here. While the local well-posedness is based on the aforementioned methods from C. E. Kenig, G. Ponce, and L. Vega, the globalization argument uses certain conserved energies constructed in [59, 60] by H. Koch and X. Liao. This can be understood as the analogue of [61] in the setting of NZBC. Since this is essential for Proposition 1.11 below, we recall now some of their definitions.

We define for $s, s' \in \mathbb{N}$ and $\tau > 0$ the (semi-classical/weighted) Sobolev spaces

$$(H^s, \|\cdot\|_{H^s}) = (\{u \in L^2 : (1 + |\xi|^2)^{\frac{s}{2}} \widehat{u} \in L^2\}, \|(1 + |\xi|^2)^{\frac{s}{2}} \widehat{\cdot}\|_{L^2}) \quad (1.18)$$

$$(H_\tau^s, \|\cdot\|_{H_\tau^s}) = (\{u \in L^2 : (\tau^2 + |\xi|^2)^{\frac{s}{2}} \widehat{u} \in L^2\}, \|(\tau^2 + |\xi|^2)^{\frac{s}{2}} \widehat{\cdot}\|_{L^2}) \quad (1.19)$$

$$(H^{s',1}, \|\cdot\|_{H^{s',1}}) = (\{u \in H^{s'} : xu \in H^{s'}\}, \|x \cdot \|_{H^{s'}} + \|\cdot\|_{H^{s'}}) \quad (1.20)$$

and the energy functionals

$$E^s(q) = \| |q|^2 - 1 \|_{H^{s-1}} + \|q_x\|_{H^{s-1}} \quad E^{s',1}(q) = \| |q|^2 - 1 \|_{H^{s'-1,1}} + \|q_x\|_{H^{s'-1,1}}.$$

We recall the complete metric space (X^s, d^s) whose definition is

$$X^s = \{q \in H_{loc}^s : E^{s-1}(q) < \infty\} / \mathbb{S}^1 \quad d^s(q, p) = \left(\int_{\mathbb{R}} \inf_{\lambda \in \mathbb{S}^1} \|\text{sech}(\cdot - y)(\lambda p - q)\|_{H^s}^2 dy \right)^{\frac{1}{2}}. \quad (1.21)$$

This space was used by H. Koch and X. Liao in [59, 60] to prove the global well-posedness of GP for $s > 0$. Their globalization argument relies on the aforementioned conserved energies, as they can be used to control the energy functionals $E^s(q)$.

We state now the well-posedness result for the NLS, GP, $\widetilde{\text{GP}}$ and $\varepsilon\widetilde{\text{GP}}$ hierarchies that we use. It is a consequence of [67, Theorem 1.2] and the theory used in its proof.

Proposition 1.11 (Well-posedness of the NLS, GP, and $\widetilde{\text{GP}}$ hierarchies). *Let $N \geq 2$ and define $\omega = \frac{N-1}{2}$. Let $s, s' \in \mathbb{N}$ such that $s \geq 3N - 1 + \omega$ and $\frac{s+3\omega}{2} \leq s' \leq s - N$. For every $q_0 : \mathbb{R} \rightarrow \mathbb{C}$ of the form $q_0 = q_* + p_0$, where $q_* \in H_{loc}^{s+1-\lceil\omega\rceil+N}$ with $E^{s+1-\lceil\omega\rceil+N,1}(q_*) < \infty$ and $p_0 \in H^s \cap H^{s',1}$, there exists a unique function $q : \mathbb{R}^{N+1} \times \mathbb{R} \rightarrow \mathbb{C}$ with $q(0) = q_0$ satisfying the following:*

- (i) For every $n \in \{0, \dots, N\}$ the functions $q(\mathbf{t}, x)$, $q(\mathbf{c}^{NLS \rightarrow GP} \mathbf{t}, x)$, and $q(\mathbf{c}^{NLS \rightarrow \widetilde{GP}} \mathbf{t}, x)$ solve respectively (NLS_n) , (GP_n) , and $(\widetilde{GP}_n)$ in the sense of distributions.
- (ii) We have $q \in C_b(\mathbb{R}^{N+1}; X^s)$ and

$$\sup_{\mathbf{t} \in \mathbb{R}^{N+1}} E^s(q(\mathbf{t})) \leq C(N, s, E^s(q_0)). \quad (1.22)$$

- (iii) The function $p = q - e^{-i(\mathbf{c}^{GP \rightarrow NLS} \mathbf{t})_0} q_*$ satisfies $p \in C(\mathbb{R}^{N+1}; H^s \cap H^{s', 1})$. Furthermore, $p(\mathbf{t})$, $p(\mathbf{c}^{NLS \rightarrow GP} \mathbf{t})$, and $p(\mathbf{c}^{NLS \rightarrow \widetilde{GP}} \mathbf{t})$ solve the corresponding perturbative formulations of respectively (NLS_n) , (GP_n) , and $(\widetilde{GP}_n)$, for the respective fronts $e^{-i(\mathbf{c}^{GP \rightarrow NLS} \mathbf{t})_0} q_*$, $e^{-it_0} q_*$, and $e^{-it_0} q_*$, strongly in $H^{s-n} \cap H^{s'-n, 1}$.

- (iv) If $\mathcal{H}_2^{GP}(q_0) < b < \frac{8}{3}$ for some $b > 0$, then there exists $\delta = \delta(b) > 0$ such that

$$\inf_{\mathbf{t} \in \mathbb{R}^{N+1}, x \in \mathbb{R}} |q(\mathbf{t}, x)| > \delta.$$

In this case we can define for some $\varepsilon > 0$ the function W via the ansatz (1.1), and find that $W(\mathbf{c}^{NLS \rightarrow \varepsilon \widetilde{GP}} \mathbf{T}) \in C(\mathbb{R}^{N+1}; H^{s-1})$ solves $(\varepsilon \widetilde{GP}_n)$ strongly in $H^{s-2[\frac{n}{2}]-2}$. Furthermore, for all $h \in \{1, \dots, s-2\}$ there exists some $\varepsilon_0 = \varepsilon_0(h, \|W(\mathbf{0})\|_{H^{h+1}}) > 0$, $c = c(h) > 0$, and $C = C(h) > 0$ such that the following holds. Define

$$\tau = \max\{1, c\|W(\mathbf{0})\|_{H^{h+1}}^{\frac{2}{3}}\}.$$

Then $\varepsilon \in (0, \varepsilon_0)$ implies

$$\begin{aligned} \|W_{\pm}(\mathbf{T})\|_{H_{\tau}^h}^2 + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)(\mathbf{T})\|_{H_{\tau}^h}^2 &\leq C \left(\|W_{\pm}(\mathbf{0})\|_{H_{\tau}^h}^2 + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)(\mathbf{0})\|_{H_{\tau}^h}^2 \right) \\ &\quad + \varepsilon^2 \tau^2 C \left(\|W(\mathbf{0})\|_{H_{\tau}^h}^2 + \frac{\varepsilon^2}{4} \|\partial_X(W_+ - W_-)(\mathbf{0})\|_{H_{\tau}^h}^2 \right) \end{aligned} \quad (1.23)$$

for all $\mathbf{T} \in \mathbb{R}^{N+1}$.

1.7 Quantitative approximation of hierarchies

Let us give some justification for why the study of integrable hierarchies is interesting. Due to Noether's theorem, the higher flows in an integrable hierarchy are symmetries of each equation in the hierarchy. Any statement about the NLS/KdV/GP/... hierarchy is therefore also a statement about the NLS/KdV/GP/... equation. These equations are all central to the study of nonlinear PDEs and ubiquitous in the physical modelling of nonlinear waves (see e. g. [8, 9, 40] for applications of NLS and [31, 32, 33, 37, 64, 69] for applications of GP). Besides, integrable hierarchies directly appear in mathematical physics in connection with algebraic geometry [5, 26, 27]. For the reader interested in more algebraic approaches to these integrable hierarchies, we refer to [25, 35, 36]. We are particularly interested in the relevance of these hierarchies in the theory of amplitude and modulation equations.

The transport equation (KdV_0) and the Korteweg-de Vries equation (KdV_1) play an important role as long-wave amplitude equations in the study of one-dimensional nonlinear waves. For example, they have been used in the description of long-wave solutions to the water wave problem [21, 22, 79], a Boussinesq equation [77], a capillary-gravity wave system [80], the NLS equation [20, 76, 19], the GP equation [12, 13], and the Euler-Poisson system [68]. For particular classes of dispersive nonlinear equations, KdV acts as a *universal* long-wave amplitude equation [14] (see also [78, §12.2]). Other examples of long-wave amplitude and modulation equations are NLS, the Ginzburg-Landau equation, and the Whitham modulation equations. For a broad perspective on the theory of such amplitude and

modulation equations, we refer to the books [78, 85, 86] and the survey [28]. Many of these amplitude equations are completely integrable, an aspect which has been studied in [45] by L. A. Kalyakin. Another perspective on this phenomenon was given in [17] by F. Calogero, who argued that for certain large classes of PDEs the process of passing to an effective limit equation preserves integrability, and hence the resulting limit equations ought to be integrable if the original class contained any integrable equations. This idea was further developed toward the classification of integrable equations.

Returning specifically to (KdV_1) as a long-wave amplitude equation, we note that it is common for the (trivial transport equation) (KdV_0) to predict the leading order behavior, and (KdV_1) the next-to-leading order behavior. One is therefore tempted to conjecture that the higher equations (KdV_n) in the KdV hierarchy may be used in the description of the higher order asymptotic behavior of long-wave solutions to dispersive nonlinear equations. Yet, such results, and in general direct physical applications of the KdV hierarchy in the modelling of nonlinear waves, are few and far between. In this direction, there exist results by Y. Kodama, T. Taniuti, and Y. Hiraoka [41, 62]. They were used by D. Bambusi [7] to show that, indeed, (KdV_2) shows up after (KdV_0) and (KdV_1) in the aforementioned long-wave approximation of water waves. Furthermore, using the reductive perturbation method of T. Taniuti [81], R. A. Kraenkel, M. A. Manna, and J. G. Pereira [63] discovered that the full KdV hierarchy appears in the study of the asymptotic behavior of long, shallow water waves. Other results in a similar direction are [23, 43]. In [42, Remark 4.6.10] S. Hofbauer noted that a result of her dissertation, using the inverse scattering transform method to prove an approximation result for KdV by a linear Schrödinger equation, can be extended to the higher equations in the KdV hierarchy.

We state now our main result, which shows rigorously that the equations in the KdV hierarchy act as long-wave amplitude equations for the $\widehat{\text{GP}}$ hierarchy. Recall our notations $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$ and $M = \lfloor \frac{N-1}{2} \rfloor$. Define matrices $\mathbf{b}^+, \mathbf{b}^- \in \mathbb{R}^{(M+1) \times (N+1)}$ by

$$\mathbf{b}_{m,n}^+ = -(-1)^n \mathbb{1}_{\{m = \lfloor \frac{n-1}{2} \rfloor\}} \quad \mathbf{b}_{m,n}^- = \mathbb{1}_{\{m = \lfloor \frac{n-1}{2} \rfloor\}}.$$

The purpose of this matrix is to match the correct flows in the NLS hierarchy with the corresponding flows in the KdV hierarchy, as described in Theorem 1.8.

Theorem 1.12 (Quantitative approximation of the $\varepsilon \widehat{\text{GP}}$ hierarchy by the KdV hierarchy). *Let $N \geq 2$, $s \in \mathbb{N}$, and $q_0 : \mathbb{R} \rightarrow \mathbb{C}$ satisfy the assumptions of Proposition 1.11 for some $s' \in \mathbb{N}$ and $q_*, p_0 : \mathbb{R} \rightarrow \mathbb{C}$. Assume in addition, as in (iv) of Proposition 1.11, that $\mathcal{H}_2^{\text{GP}}(q_0) < \frac{8}{3}$ so that $|q(\mathbf{t}, x)| > \delta > 0$ and hence we can define for some $\varepsilon > 0$ the function W according to the ansatz (1.1). We have $W \in C(\mathbb{R}^{N+1}; H^{s-1})$, and if $\varepsilon < \varepsilon_0(s, \|W(\mathbf{0})\|_{H^{s-1}})$ then*

$$\|W\|_{L_T^\infty H_X^{h-1}} \leq C(h, \|W(\mathbf{0})\|_{H^h}) \|W(\mathbf{0})\|_{H^h} \quad \forall h \in \{2, \dots, s-1\}. \quad (1.24)$$

Furthermore, $W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widehat{\text{GP}}} \mathbf{T}) \in C(\mathbb{R}^{N+1}; H^{s-1})$ solves $(\varepsilon \widehat{\text{GP}}_n)$ in the sense of distributions.

Let $\pm U_\pm \in C(\mathbb{R}^{M+1}; H^{s-1})$ be the solutions to the M -truncated KdV hierarchy, given by Proposition 1.10, with initial data $U_\pm(\mathbf{0}) = W_\pm(\mathbf{0})$. They satisfy

$$\|U_\pm\|_{L_T^\infty H_X^h} \leq C(h, \|U_\pm(\mathbf{0})\|_{H^{-1}}) \|U_\pm(\mathbf{0})\|_{H^h} \quad \forall h \in \{0, \dots, s-1\}. \quad (1.25)$$

Define $[\mathbf{T}] = |T_2| + \dots + |T_N|$. For all $h \in \{0, \dots, \lfloor \frac{s-N}{2} \rfloor\}$ there exists $C = C(N, h, \delta) > 0$ such that

$$\begin{aligned} \left\| \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{T}) \end{pmatrix} - W(\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widehat{\text{GP}}} \mathbf{T}) \right\|_{H^h} &\leq \varepsilon^2 [\mathbf{T}]^{\frac{1}{2}} \exp\left([\mathbf{T}] C(1 + \|(W, U)\|_{L_T^\infty H_X^{2M}}^M)\right) \\ &\times C(1 + \|W\|_{L_T^\infty H_X^{N+1+h}}^{N+4}) (1 + [\mathbf{T}] \|(W, U)\|_{L_T^\infty H_X^{2(M+h)}}^{N-1})^{\frac{h}{2}}. \end{aligned} \quad (1.26)$$

This result represents a collection of approximations, and the timescales involved have several edge cases. Note, for example, that $\mathbf{c}^{\text{NLS} \rightarrow \varepsilon \widehat{\text{GP}}}$ contains negative powers of ε . We shall therefore take some

time now to explain this result. It is convenient to introduce for a *time-independent* function $q : \mathbb{R} \rightarrow \mathbb{C}$ with $|q| > \delta > 0$ the notation

$$[q \mapsto W](q)(X) = \frac{1}{\varepsilon^2} \left(\frac{1}{2} \operatorname{Im} \left[\frac{q_x}{q} \right] + (|q| - 1), \frac{1}{2} \operatorname{Im} \left[\frac{q_x}{q} \right] - (|q| - 1) \right)(x).$$

This means that if $q : \mathbb{R}^{N+1} \times \mathbb{R} \rightarrow \mathbb{C}$ is a solution to the N -truncated **NLS** hierarchy, and we define $W : \mathbb{R}^{N+1} \times \mathbb{R} \rightarrow \mathbb{R}^2$ according to (1.1), then the following hold:

$$\begin{aligned} q(\mathbf{t}) \text{ solves the } N\text{-truncated NLS hierarchy} & \quad \text{and} \quad [q \mapsto W](q(\mathbf{t})) = W(\mathbf{T}) \\ q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t}) \text{ solves the } N\text{-truncated GP hierarchy} & \quad \text{and} \quad [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) = W(\mathbf{c}^{\text{NLS} \mapsto \varepsilon \text{GP}} \mathbf{T}) \\ q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t}) \text{ solves the } N\text{-truncated } \widetilde{\text{GP}} \text{ hierarchy} & \quad \text{and} \quad [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) = W(\mathbf{c}^{\text{NLS} \mapsto \varepsilon \widetilde{\text{GP}}} \mathbf{T}). \end{aligned}$$

We see now that Theorem 1.12 is an approximation result written in terms of the $\widetilde{\text{GP}}/\varepsilon \widetilde{\text{GP}}$ hierarchy. We want to sketch the consequences of Theorem 1.12 for the equations (NLS_n) , (GP_n) , and $(\widetilde{\text{GP}}_n)$ for $n \in \{0, \dots, 7\}$. To keep the notation simple, we shall use the formal statement

$$W(\mathbf{c}^{\text{NLS} \mapsto \varepsilon \widetilde{\text{GP}}} \mathbf{T}) = \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{T}) \end{pmatrix} + \begin{cases} O(\varepsilon^2) & \text{as long as } [\mathbf{T}] \lesssim 1, [\mathbf{T}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [\mathbf{T}] = 0 \end{cases}$$

to represent the precise statement (1.26), which as explained above is an approximation result for the $\widetilde{\text{GP}}$ hierarchy. A corresponding approximation result for the **NLS** hierarchy reads

$$W(\mathbf{T}) = \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{T}) \end{pmatrix} + \begin{cases} O(\varepsilon^2) & \text{as long as } [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{T}] \lesssim 1, [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{T}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{T}] = 0 \end{cases},$$

and for the **GP** hierarchy it is

$$W(\mathbf{c}^{\text{NLS} \mapsto \varepsilon \text{GP}} \mathbf{T}) = \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \varepsilon \text{GP}} \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \varepsilon \text{GP}} \mathbf{T}) \end{pmatrix} + \begin{cases} O(\varepsilon^2) & \text{as long as } [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \varepsilon \text{GP}} \mathbf{T}] \lesssim 1, [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \varepsilon \text{GP}} \mathbf{T}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [\mathbf{c}^{\varepsilon \widetilde{\text{GP}} \mapsto \varepsilon \text{GP}} \mathbf{T}] = 0 \end{cases}.$$

Here the dependence of various time-related quantities on ε is not clearly visible, so let us write it in a more explicit form. We assume that N is even. The approximation results can be expressed as

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+ \left(\left((\sqrt{2}\varepsilon)^{2m+1} ((\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t})_{2m+1} - 2(\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t})_{2m+2}) \right)_{0 \leq m \leq M} \right) \\ U_- \left(\left((\sqrt{2}\varepsilon)^{2m+1} ((\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t})_{2m+1} + 2(\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t})_{2m+2}) \right)_{0 \leq m \leq M} \right) \end{pmatrix} \\ &+ \begin{cases} O(\varepsilon^2) & \text{as long as } [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t}] \lesssim 1, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{NLS}} \mathbf{t}] = 0 \end{cases} \end{aligned}$$

for the **NLS** hierarchy,

$$\begin{aligned} [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+ \left(\left((\sqrt{2}\varepsilon)^{2m+1} ((\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t})_{2m+1} - 2(\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t})_{2m+2}) \right)_{0 \leq m \leq M} \right) \\ U_- \left(\left((\sqrt{2}\varepsilon)^{2m+1} ((\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t})_{2m+1} + 2(\mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t})_{2m+2}) \right)_{0 \leq m \leq M} \right) \end{pmatrix} \\ &+ \begin{cases} O(\varepsilon^2) & \text{as long as } [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t}] \lesssim 1, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{c}^{\widetilde{\text{GP}} \mapsto \text{GP}} \mathbf{t}] = 0 \end{cases} \end{aligned}$$

for the GP hierarchy, and

$$[q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) = \begin{pmatrix} U_+ \left(((\sqrt{2}\varepsilon)^{2m+1} (t_{2m+1} - 2t_{2m+2}))_{0 \leq m \leq M} \right) \\ U_- \left(((\sqrt{2}\varepsilon)^{2m+1} (t_{2m+1} + 2t_{2m+2}))_{0 \leq m \leq M} \right) \end{pmatrix} \\ + \begin{cases} O(\varepsilon^2) & \text{as long as } [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{t}] \lesssim 1, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{t}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1}, [(\nabla_{\mathbf{t}} \mathbf{T}) \mathbf{t}] = 0 \end{cases}$$

for the $\widetilde{\text{GP}}$ hierarchy.

1.8 Examples

We use the above formulations to give explicit approximation results along one-dimensional time-rays in $\mathbb{R}^{N+1}$ for the cases $N = 0, \dots, 7$.

Example 1.13 $((\text{NLS}_0) = (\text{GP}_0) = (\widetilde{\text{GP}}_0) = \text{“phase rotation”})$. Take

$$N = 0 \quad M = -1 \quad \mathbf{T} = (2(\sqrt{2}\varepsilon)^{-1}t_0).$$

Then

$$[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) = \begin{pmatrix} U_+(0) \\ U_-(0) \end{pmatrix} \quad \text{for all } t_0 \in \mathbb{R}.$$

Here we use the convention $\mathbb{R}^{M+1} = \mathbb{R}^0 = \{0\}$.

Example 1.14 $((\text{NLS}_1) = (\text{GP}_1) = (\widetilde{\text{GP}}_1) = \text{“transport”})$. Take

$$N = 1 \quad M = 0 \quad \mathbf{T} = (0, \sqrt{2}\varepsilon t_1).$$

Then

$$[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) = \begin{pmatrix} U_+(\sqrt{2}\varepsilon t_1) \\ U_-(\sqrt{2}\varepsilon t_1) \end{pmatrix} \quad \text{for all } t_1 \in \mathbb{R}.$$

Example 1.15 $((\text{NLS}_2) = \text{NLS} \text{ and } (\text{GP}_2) = (\widetilde{\text{GP}}_2) = \text{GP})$. Take

$$N = 2 \quad M = 0 \quad \mathbf{T} = (0, 0, 2\sqrt{2}\varepsilon t_2).$$

Then

$$[q \mapsto W](q(\mathbf{t})) = \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(-2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) = \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(-2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) = \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(-2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1}.$$

In fact, we have $[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t}))$. This is a consequence of the fact that the difference between these flows for q is one of phase rotation, which does not show up in the variable W .

This result corresponds to [10, Theorem 2], i. e. it is an approximation of GP by two transport equations. As discussed above, it is known from [13, 12] that in the subsequent order KdV = (KdV₁) shows up as the approximating equation. It is natural to conjecture that this may be generalized for the cases $N \geq 3$, i. e. to characterize the effect of the terms $O(\varepsilon^2)$ up to $O(\varepsilon^4)$. Remark 1.9 above provides some evidence towards this.

Example 1.16 $((\text{NLS}_3) = \text{mKdV}, (\text{GP}_3), \text{ and } (\widetilde{\text{GP}}_3))$. Take

$$N = 3 \quad M = 1 \quad \mathbf{T} = (0, 0, 0, (\sqrt{2}\varepsilon)^3 t_3).$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) &= \begin{pmatrix} U_+(0, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(0, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3}. \end{aligned}$$

This is the first new result that Theorem 1.12 yields. Here mKdV and renormalized versions of mKdV are approximated by two transport equations in leading order, and then by two KdV equations at next-to-leading order. Observe that for the NLS and GP hierarchies the timescale is ε^{-3} , which is better than the expected ε^{-1} . This is because the approximation of transport equations by transport equations has zero error.

Example 1.17 $((\text{NLS}_4) = \text{“4-th order NLS”}, (\text{GP}_4) = (\widetilde{\text{GP}}_4) = \text{“4-th order GP”})$. Take

$$N = 4 \quad M = 1 \quad \mathbf{T} = (0, 0, 0, 0, 2(\sqrt{2}\varepsilon)^3 t_4).$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon \mathbf{c}_{2,4}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_4, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(-2\sqrt{2}\varepsilon \mathbf{c}_{2,4}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_4, -2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+(0, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(0, 2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) &= \begin{pmatrix} U_+(0, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(0, 2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-3}. \end{aligned}$$

For the NLS hierarchy the case $n = 2$ “dominates”, leading to the shorter timescale ε^{-1} . For the $\widetilde{\text{GP}}$ hierarchy the dynamics are cleanly separated, leading to a better timescale. Remark 1.9 suggests that studying $(\varepsilon \widetilde{\text{GP}}_4)$ in a “co- $(\varepsilon \widetilde{\text{GP}}_3)$ -moving” frame allows us to see, on a longer timescale, the dynamics of (KdV_2) in one component of our solution W .

Example 1.18 $((\text{NLS}_5) = \text{“5-th order mKdV”}, (\text{GP}_5), \text{ and } (\widetilde{\text{GP}}_5))$. Take

$$N = 5 \quad M = 2 \quad \mathbf{T} = (0, 0, 0, 0, 0, (\sqrt{2}\varepsilon)^5 t_5).$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) &= \begin{pmatrix} U_+(0, 0, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(0, 0, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-5}. \end{aligned}$$

Here (KdV₂) shows up for the first time. Note that once again the approximation of transport equations by transport equations has no error, yielding the better-than-expected timescale ε^{-3} for the NLS and GP hierarchies.

Example 1.19 ((NLS₆) = “6-th order NLS”, (GP₆) = ($\widetilde{\text{GP}}_6$) = “6-th order GP”). Take

$$N = 6 \quad M = 2 \quad \mathbf{T} = (0, 0, 0, 0, 0, 0, 2(\sqrt{2}\varepsilon)^5 t_6).$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon \mathbf{c}_{2,6}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_6, -2(\sqrt{2}\varepsilon)^3 \mathbf{c}_{4,6}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_6, -2(\sqrt{2}\varepsilon)^5 t_6) \\ U_- (2\sqrt{2}\varepsilon \mathbf{c}_{2,6}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_6, 2(\sqrt{2}\varepsilon)^3 \mathbf{c}_{4,6}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_6, 2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+(0, 0, -2(\sqrt{2}\varepsilon)^5 t_6) \\ U_-(0, 0, 2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-5} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) &= \begin{pmatrix} U_+(0, 0, -2(\sqrt{2}\varepsilon)^5 t_6) \\ U_-(0, 0, 2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-5}. \end{aligned}$$

Example 1.20 ((NLS₇) = “7-th order mKdV”, (GP₇), and ($\widetilde{\text{GP}}_7$)). Take

$$N = 7 \quad M = 3 \quad \mathbf{T} = (0, 0, 0, 0, 0, 0, 0, (\sqrt{2}\varepsilon)^7 t_7).$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^5 \mathbf{c}_{5,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^5 \mathbf{c}_{5,7}^{\widetilde{\text{GP}} \mapsto \text{NLS}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \text{GP}} \mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^5 \mathbf{c}_{5,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^5 \mathbf{c}_{5,7}^{\widetilde{\text{GP}} \mapsto \text{GP}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\text{NLS} \mapsto \widetilde{\text{GP}}} \mathbf{t})) &= \begin{pmatrix} U_+(0, 0, 0, (\sqrt{2}\varepsilon)^7 t_7) \\ U_-(0, 0, 0, (\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-7}. \end{aligned}$$

At this point we stop because all the edge cases have been exhibited and the pattern continues as expected.

1.9 Notations and definitions

We use $\mathbb{N} = \{n \in \mathbb{Z} : n \geq 0\}$ but note that in some cases where we write $n \in \mathbb{N}$, and when the intended interpretation is clear from the context, $n = 0$ will be an inadmissible value.

We often write $A \leq C(\text{things})B$ to indicate that there exists a constant, depending only on the given parameters, for which the estimate holds. Such constants may change from one line to the next without being given a new name.

Whenever we consider a function space such as H^s without explicit domain, the domain is implicitly assumed to be $\mathbb{R}$. The target space may be $\mathbb{R}$, $\mathbb{C}$, or corresponding spaces of vectors or matrices, and must be determined from context.

We denote by $\mathbf{t} = (t_0, \dots, t_N)$, $\mathbf{T} = (T_0, \dots, T_N)$, $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$ certain vectors of times. For matrices $\mathbf{c} \in \mathbb{C}^{(N+1) \times (N+1)}$ we denote the entry in the n -th row and m -th column by $\mathbf{c}_{n,m}$. The only exception is the matrix of Jost solutions $\Phi^{\dagger, \pm} = (\Phi_1^{\dagger, \pm} | \Phi_2^{\dagger, \pm})$.

We define the Pauli matrices

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For real numbers $x, y \in \mathbb{R}$, we define $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$, as well as the Japanese bracket $\langle x \rangle = \sqrt{x^2 + 1}$ and the floor and ceiling functions $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$ and $\lceil x \rceil = \min\{n \in \mathbb{Z} : n \geq x\}$.

For functions $f : \mathbb{R} \rightarrow \mathbb{C}^n$ we use the following convention for the Fourier transform:

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x) dx.$$

We extend the Fourier transform to the tempered distributions $\mathcal{S}'(\mathbb{R}; \mathbb{C}^n)$ as usual.

We denote by $\mathcal{Q}_j = i^{j-1}\{z \in \mathbb{C} : \Re z > 0, \Im z > 0\}$, $j \in \{1, 2, 3, 4\}$ the four open quadrants of the complex plane.

Acknowledgements. I thank my supervisor Xian Liao for her support and guidance over the past few years, and for many enjoyable and insightful discussions.

2 Definition and approximation of the Hamiltonians $\mathcal{H}_n^{\varepsilon \text{GP}}$ by $\mathcal{E}_n^{\text{KdV}}$

2.1 Lax pairs

Formally, a function $q : \mathbb{R}^2 \rightarrow \mathbb{C}$ is a solution to GP if and only if the time-dependent operators

$$L^{\text{GP}} = \begin{pmatrix} i\partial_x & -iq \\ i\bar{q} & -i\partial_x \end{pmatrix}$$

and

$$P^{\text{GP}} = i \begin{pmatrix} 2\partial_x^2 - q\bar{q} + 1 & -2q\partial_x - q_x \\ 2\bar{q}\partial_x + \bar{q}_x & -2\partial_x^2 + q\bar{q} - 1 \end{pmatrix}$$

solve the so-called Lax equation

$$\partial_t L^{\text{GP}} = [P^{\text{GP}}, L^{\text{GP}}].$$

Here L^{GP} is called the Lax operator of GP and $(L^{\text{GP}}, P^{\text{GP}})$ the Lax pair. Setting $L^{\text{NLS}} = L^{\text{GP}}$ and

$$P^{\text{NLS}} = i \begin{pmatrix} 2\partial_x^2 - q\bar{q} & -2q\partial_x - q_x \\ 2\bar{q}\partial_x + \bar{q}_x & -2\partial_x^2 + q\bar{q} \end{pmatrix}$$

yields a Lax pair for NLS. A direct consequence of the Lax pair formalism is that eigenvalues of the Lax operator are conserved quantities. A deeper consequence is that the potential q can be reconstructed from the scattering data associated to the Lax operator L^{GP} , and this scattering data undergoes a linear evolution in time. As such, one can formally linearize NLS or GP by conjugating the flow with the direct and inverse scattering transforms. This is called the inverse scattering transform (IST) method, and it is

the reason why NLS and GP are considered to be completely integrable PDEs. Note that understanding the mapping properties of the IST on standard function spaces is non-trivial, especially for nonzero boundary conditions such as in the case of GP. For NLS and GP the scattering data is obtained from the Zakharov-Shabat scattering problem, which is given as the cases $\dagger \in \{\text{NLS}, \text{GP}\}$ in Definition 2.1 below. Since the IST method is not the focus of this work, we shall merely highlight here the seminal papers [2, 87] and refer to the recent review [73] regarding the development of the IST method, with a focus on GP. We emphasize the connection to [67, Section 2] in a preceding work, where the theory that we use here is set up, with matching notation, and more background on complete integrability and the IST method is given.

We shall now give Lax pairs for hGP, εGP and KdV. Set $q = ae^{i\varphi}$ and $\theta = \frac{i}{2}(\varphi - \frac{\pi}{2})$. In [66] X. Liao and M. Plum note that the explicit unitary matrix

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-\theta} & e^{\theta} \\ e^{-\theta} & -e^{\theta} \end{pmatrix} \quad M^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{\theta} & e^{-\theta} \\ e^{\theta} & -e^{-\theta} \end{pmatrix}$$

yields the unitary equivalence

$$L^{\text{GP}} = M^{-1} L^{\text{hGP}} M,$$

where

$$L^{\text{hGP}} = \begin{pmatrix} -w_- + 1 & i\partial_x \\ i\partial_x & -w_+ - 1 \end{pmatrix}.$$

Together with

$$P^{\text{hGP}} = \begin{pmatrix} -\frac{1}{2}(3w_- + w_+)_x - (3w_- + w_+ - 2)\partial_x & 2i\partial_x^2 - \frac{i}{2} \frac{(w_+ - w_-)_{xx}}{w_+ - w_- + 2} \\ 2i\partial_x^2 - \frac{i}{2} \frac{(w_+ - w_-)_{xx}}{w_+ - w_- + 2} & -\frac{1}{2}(w_- + 3w_+)_x - (w_- + 3w_+ + 2)\partial_x \end{pmatrix}$$

this forms a Lax pair for the hydrodynamic formulation hGP of GP.

By rescaling, we can construct from $(L^{\text{hGP}}, P^{\text{hGP}})$ the Lax pair

$$L^{\varepsilon\text{GP}} = \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix}$$

and

$$P^{\varepsilon\text{GP}} = \begin{pmatrix} P_{1,1}^{\varepsilon\text{GP}} & P_{1,2}^{\varepsilon\text{GP}} \\ P_{2,1}^{\varepsilon\text{GP}} & P_{2,2}^{\varepsilon\text{GP}} \end{pmatrix} \text{ where } \begin{pmatrix} P_{1,1}^{\varepsilon\text{GP}} \\ P_{1,2}^{\varepsilon\text{GP}} \\ P_{2,1}^{\varepsilon\text{GP}} \\ P_{2,2}^{\varepsilon\text{GP}} \end{pmatrix} = \begin{pmatrix} -\frac{\varepsilon}{\sqrt{2}}(3\varepsilon^2 W_- + \varepsilon^2 W_+)_X - (3\varepsilon^2 W_- + \varepsilon^2 W_+ - 2)\sqrt{2}\varepsilon\partial_X \\ 4i\varepsilon^2\partial_X^2 - i\varepsilon^2 \frac{(\varepsilon^2 W_+ - \varepsilon^2 W_-)_{XX}}{\varepsilon^2 W_+ - \varepsilon^2 W_- + 2} \\ 4i\varepsilon^2\partial_X^2 - i\varepsilon^2 \frac{(\varepsilon^2 W_+ - \varepsilon^2 W_-)_{XX}}{\varepsilon^2 W_+ - \varepsilon^2 W_- + 2} \\ -\frac{\varepsilon}{\sqrt{2}}(\varepsilon^2 W_- + 3\varepsilon^2 W_+)_X - (\varepsilon^2 W_- + 3\varepsilon^2 W_+ + 2)\sqrt{2}\varepsilon\partial_X \end{pmatrix}$$

for εGP . Lastly, we recall the well-known Lax pair

$$\tilde{L}^{\text{KdV}} = -\partial_X^2 + U \quad P^{\text{KdV}} = -4\partial_X^3 + 3U_X + 6U\partial_X$$

for KdV. For the sake of uniformity we treat the scattering problem for KdV as a first-order system, i. e. we use instead of $\tilde{L}^{\text{KdV}}$ the Lax operator

$$L^{\text{KdV}} = \begin{pmatrix} -\partial_X^2 + U & 0 \\ U_X & -\partial_X^2 + U \end{pmatrix}.$$

We are interested in these Lax pairs because the conservation of the spectral data of the Lax operator implies the existence of infinitely many conserved quantities. More specifically, we shall construct in the subsequent section the Hamiltonians $\mathcal{H}_n^{\text{NLS}}, \mathcal{H}_n^{\text{GP}}, \mathcal{H}_n^{\text{hGP}}, \mathcal{H}_n^{\varepsilon\text{GP}}$, and $\mathcal{E}_n^{\text{KdV}}$ from the transmission coefficients associated to each scattering problem.

2.2 Jost solutions and their asymptotic expansion

The Jost solutions for each scattering problem depend on a complex parameter $\lambda \in \mathbb{C}$. We prefer to use for GP and hGP the alternative variable $z = \sqrt{\lambda^2 - 1}$, and for ε GP and KdV the variable $k = \frac{z}{\sqrt{2\varepsilon}}$. Due to the multivaluedness of the relation $\lambda \leftrightarrow z$, it is common to consider pairs $(\lambda, z) \in \mathbb{C}^2$ as lying on a Riemann surface. Since we also consider the additional parameter k , we define

$$\mathbb{K} = \{(\lambda, z, k) \in \mathbb{C}^3 : \lambda^2 - z^2 = 1 \text{ and } z = \sqrt{2\varepsilon}k\}. \quad (2.1)$$

We decompose $\mathbb{K}$ into two sheets $\mathbb{K} = \overline{\mathbb{K}_+} \cup \overline{\mathbb{K}_-}$ where

$$\mathbb{K}_\pm = \{(\lambda, z, k) \in \mathbb{K} : \lambda \in \mathbb{C} \setminus ((-\infty, -1] \cup [1, \infty)), \pm \operatorname{Im} z > 0\}.$$

By an abuse of notation, we may write $\lambda \in \overline{\mathbb{K}_\pm}$, $z \in \overline{\mathbb{K}_\pm}$, or $k \in \overline{\mathbb{K}_\pm}$ to mean $(\lambda, z, k) \in \overline{\mathbb{K}_\pm}$. Similarly, for functions f with domain $\overline{\mathbb{K}_\pm}$ we write $f(\lambda) = f(z) = f(k) = f(\lambda, z, k)$.

In order to give a uniform treatment of all the scattering problems under consideration, we define

$$\text{Eqns} = \{\text{NLS}, \text{GP}, \text{hGP}, \varepsilon\text{GP}, \text{KdV}\}$$

and write $\dagger \in \text{Eqns}$ when we wish for $\dagger$ to refer to any of the equations under consideration. Using this notation, we shall characterize all five scattering problems under consideration using a single set of variables:

$$(q^\dagger, q_\pm^\dagger, x^\dagger, \lambda^\dagger, z^\dagger) = \begin{cases} (q, 0, x, \lambda, \lambda) & , \dagger = \text{NLS} \\ (q, q_\pm, x, \lambda, z) & , \dagger = \text{GP} \\ (w, 0, x, \lambda, z) & , \dagger = \text{hGP} \\ (W, 0, X, \lambda, k) & , \dagger = \varepsilon\text{GP} \\ (U, 0, X, k^2, k) & , \dagger = \text{KdV} \end{cases}.$$

In the following we always take $j \in \{1, 2\}$ and set $\gamma_1 = 1$ and $\gamma_2 = -1$.

Definition 2.1 (Jost solutions of $\dagger$). Given a potential $q^\dagger$ with $\lim_{x \rightarrow \pm\infty} q^\dagger = q_\pm^\dagger$, we define the Jost solution $\Phi_j^{\dagger, \pm}$ as the solution to the scattering problem

$$\begin{aligned} \Phi_j^{\dagger, \pm} : \mathbb{R} \times \overline{\mathbb{K}_{\mp\gamma_j}} &\longrightarrow \mathbb{C}^2 \\ L^\dagger \Phi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) &= \lambda^\dagger \Phi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) \end{aligned} \quad (2.2)$$

$$\lim_{x \rightarrow \pm\infty} \Phi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) e^{x^\dagger i z^\dagger \gamma_j} = E_j^{\dagger, \pm}(z^\dagger). \quad (2.3)$$

We define in addition the modified Jost solutions as $\Psi_j^{\dagger, \pm} = \Phi_j^{\dagger, \pm} e^{x^\dagger i z^\dagger \gamma_j}$. The eigenvalue problem (2.2) is equivalent to

$$\partial_{x^\dagger} \Phi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) = A^\dagger(x^\dagger, z^\dagger; q^\dagger) \Phi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) \quad (2.4)$$

$$\partial_{x^\dagger} \Psi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) = (A^\dagger(x^\dagger, z^\dagger; q^\dagger) + i z^\dagger \gamma_j) \Psi_j^{\dagger, \pm}(x^\dagger, z^\dagger; q^\dagger) \quad (2.5)$$

The matrices $A^\dagger(x^\dagger, z^\dagger; q^\dagger), E^{\dagger, \pm}(z^\dagger) \in \mathbb{C}^{2 \times 2}$ are defined in the following table:

$\dagger$	$A^\dagger$	$E^{\dagger, \pm}$
NLS	$\begin{pmatrix} -i\lambda & q \\ \bar{q} & i\lambda \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
GP	$\begin{pmatrix} -i\lambda & q \\ \bar{q} & i\lambda \end{pmatrix}$	$\begin{pmatrix} q_\pm & i(z - \lambda) \\ i(\lambda - z) & \bar{q}_\pm \end{pmatrix}$
hGP	$\begin{pmatrix} 0 & -i(\lambda + 1 + w_+) \\ -i(\lambda - 1 + w_-) & 0 \end{pmatrix}$	$\frac{i}{\sqrt{2}} \begin{pmatrix} \lambda - z + 1 & -\lambda + z - 1 \\ -\lambda + z + 1 & -\lambda + z + 1 \end{pmatrix} \begin{pmatrix} e^{\theta_\pm} & 0 \\ 0 & e^{-\theta_\pm} \end{pmatrix}$
ε GP	$\begin{pmatrix} 0 & -\frac{i\varepsilon}{\sqrt{2}} \left(\frac{\lambda+1}{\varepsilon^2} + W_+ \right) \\ -\frac{i\varepsilon}{\sqrt{2}} \left(\frac{\lambda-1}{\varepsilon^2} + W_- \right) & 0 \end{pmatrix}$	$\frac{i}{\sqrt{2}} \begin{pmatrix} \lambda - \sqrt{2}\varepsilon k + 1 & -\lambda + \sqrt{2}\varepsilon k - 1 \\ -\lambda + \sqrt{2}\varepsilon k + 1 & -\lambda + \sqrt{2}\varepsilon k + 1 \end{pmatrix} \begin{pmatrix} e^{\theta_\pm} & 0 \\ 0 & e^{-\theta_\pm} \end{pmatrix}$
KdV	$\begin{pmatrix} 0 & 1 \\ U - k^2 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ -ik & ik \end{pmatrix}$

Here $\theta_\pm = \frac{i}{2}(\varphi_\pm - \frac{\pi}{2})$ where $e^{i\varphi_\pm} = q_\pm$. We also define $A^{\dagger, \pm}(z^\dagger; q^\dagger) = \lim_{x^\dagger \rightarrow \pm\infty} A^\dagger(x^\dagger, z^\dagger; q^\dagger)$. Observe that $E^{\dagger, \pm}$ is indeed a solution to (2.5) at $x^\dagger = \pm\infty$.

Our choice of normalization for these eigenvalue problems yields the identity

$$\Phi_j^{\text{GP}, \pm}(x, z) = M(x) \Phi_j^{\text{hGP}, \pm}(x, z) = M\left(\frac{X}{\sqrt{2\varepsilon}}\right) \Phi_j^{\varepsilon\text{GP}, \pm}(X, k). \quad (2.6)$$

2.3 Transmission coefficients and their asymptotic expansion

We recall [67, Definition 2.4].

Definition 2.2. Let $D \subset \mathbb{C}$, $d \in \mathbb{N}$ and $1 \leq p \leq \infty$.

- (i) We say that a function $f = f(z) \in C^\infty(D; \mathbb{C}^d)$ has an **asymptotic expansion in powers of $2iz$ at infinity on D** if there exist $(f_n)_{n \in \mathbb{N}} \subset \mathbb{C}^d$ such that

$$\forall N \in \mathbb{N} \quad \lim_{|z| \rightarrow \infty} (2iz)^N \left(f(z) - \sum_{n=0}^N \frac{f_n}{(2iz)^n} \right) = 0.$$

We call f_n the **expansion coefficients** of f and note that they are unique. If f and g have such an expansion then $2izf$ and fg do as well. If $g(z) \neq 0$ for $|z|$ sufficiently large, then fg^{-1} also has such an expansion.

- (ii) We say that a function $f = f(x, z) \in C^\infty(\mathbb{R} \times D; \mathbb{C}^d)$ has an **L^p -smooth asymptotic expansion in powers of $2iz$ at infinity on D** if there exist $(f_n)_{n \in \mathbb{N}} \subset (C^\infty \cap L^p)(\mathbb{R}; \mathbb{C}^d)$ such that

$$\forall k, N \in \mathbb{N} \quad \lim_{|z| \rightarrow \infty} \left\| (2iz)^N \left(\partial_x^k f(x, z) - \sum_{n=0}^N \frac{\partial_x^k f_n(x)}{(2iz)^n} \right) \right\|_{L_x^p(\mathbb{R})} = 0.$$

In this case $\partial_x f$ has such an expansion as well. If g has such an L^∞ -smooth expansion and $\inf_{x \in \mathbb{R}} |g(x, z)| > 0$ for $|z|$ sufficiently large, then fg^{-1} also has an L^p -smooth asymptotic expansion. If $p = 1$ then $\int_{-\infty}^\infty f(x, z) dx$ has an asymptotic expansion in powers of $2iz$ at infinity on D with expansion coefficients $\int_{-\infty}^\infty f_n(x) dx$.

When f has an asymptotic expansion of some kind with expansion coefficients $(f_n)_{n \in \mathbb{N}}$, we write

$$f \sim \sum_{n=0}^\infty \frac{f_n}{(2iz)^n}.$$

We collect now a number of results on the Jost solutions which are known in the literature, although rarely arranged in this form. For the rest of the section we focus only on $\mathbb{K}_+$.

Lemma 2.3. Let $q^\dagger \in C^\infty(\mathbb{R})$ with $q^\dagger - q_\pm^\dagger \in \mathcal{S}(\mathbb{R}_\pm)$.

(i) The Jost solutions $\Phi_j^{\dagger, \pm}$ exist and are smooth in $x^\dagger$ and analytic in $z^\dagger$ up to a finite set of choices for $z^\dagger$.

(ii) There exists a constant $c(q^\dagger) > 0$ such that $\text{Im } z^\dagger > c(q^\dagger)$ implies

$$\text{the limits } \lim_{x^\dagger \rightarrow \infty} \Psi_1^{\dagger, -}(x^\dagger, z^\dagger), \lim_{x^\dagger \rightarrow -\infty} \Psi_2^{\dagger, +}(x^\dagger, z^\dagger) \text{ exist} \quad (2.7)$$

$$\lim_{x^\dagger \rightarrow \infty} \partial_{x^\dagger} \Psi_1^{\dagger, -}(x^\dagger, z^\dagger) = \lim_{x^\dagger \rightarrow -\infty} \partial_{x^\dagger} \Psi_2^{\dagger, +}(x^\dagger, z^\dagger) = 0 \quad (2.8)$$

$$\sup_{x^\dagger \in \mathbb{R}} \left| \Psi_1^{\dagger, -}(x^\dagger, z^\dagger) - E_1^{\dagger, -}(z^\dagger) \right| < \frac{|E_1^{\dagger, -}(z^\dagger)|}{c(q^\dagger)|z^\dagger|} \quad (2.9)$$

$$\sup_{x^\dagger \in \mathbb{R}} \left| \Psi_2^{\dagger, +}(x^\dagger, z^\dagger) - E_2^{\dagger, +}(z^\dagger) \right| < \frac{|E_2^{\dagger, +}(z^\dagger)|}{c(q^\dagger)|z^\dagger|}. \quad (2.10)$$

(iii) Define on $\overline{\mathbb{K}_+} \cap \{\text{Im } z^\dagger > c, \pm \text{Im } \lambda > 0\}$ the quantities

$$\sigma^\dagger = \frac{\partial_{x^\dagger} \Psi_{1,1}^{\dagger, -}}{\Psi_{1,1}^{\dagger, -}} \quad \tilde{\sigma}^\dagger = -\frac{\partial_{x^\dagger} \Psi_{2,2}^{\dagger, +}}{\Psi_{2,2}^{\dagger, +}} \quad (2.11)$$

$$r^\dagger = \frac{\Psi_{1,2}^{\dagger, -}}{\Psi_{1,1}^{\dagger, -}} \quad \tilde{r}^\dagger = \frac{\Psi_{2,1}^{\dagger, +}}{\Psi_{2,2}^{\dagger, +}}. \quad (2.12)$$

They are bounded and have L^∞ -smooth asymptotic expansions in powers of $2iz^\dagger$ at infinity. The densities $\sigma^\dagger$ and $\tilde{\sigma}^\dagger$ are in addition integrable, and their asymptotic expansions are also L^1 -smooth.

Proof. Let us first discuss the case of GP. Here (i) and (ii) can be proven by the well-known method of writing the boundary value problem as a Volterra integral equation and using a Neumann series ansatz. For a demonstration we refer to [24, Proposition 3], and also the theory in [30, Chapter 1]. We refer also to [67, Section 2] for a treatment of these matters with a different approach, and specifically [67, Lemma 2.5] as a reference for (iii). These same methods are strictly easier to apply for NLS and KdV because the potential under consideration has zero boundary data and the structure of the scattering problem is similar. For more information on the direct scattering theory of KdV we refer to [1]. Using (2.6), the cases hGP and ε GP can be obtained directly from GP. $\square$

If $\text{Im } z^\dagger = 0$ then $\Phi^{\dagger, -} = (\Phi_1^{\dagger, -}, \Phi_2^{\dagger, -})$ and $\Phi^{\dagger, +} = (\Phi_1^{\dagger, +}, \Phi_2^{\dagger, +})$ are two fundamental solution matrices to (2.4), hence there exist $a^\dagger(z^\dagger; q^\dagger), b^\dagger(z^\dagger; q^\dagger) \in \mathbb{C}$ such that

$$\Phi_1^{\dagger, -}(x^\dagger, z^\dagger) = a^\dagger(z^\dagger; q^\dagger) \Phi_1^{\dagger, +}(x^\dagger, z^\dagger) + b^\dagger(z^\dagger; q^\dagger) \Phi_2^{\dagger, +}(x^\dagger, z^\dagger). \quad (2.13)$$

We call $a^\dagger(z^\dagger; q^\dagger)$ the transmission coefficient of $\dagger$ with potential $q^\dagger$. It satisfies

$$a^\dagger = \frac{\det(\Phi_1^{\dagger, -} | \Phi_2^{\dagger, +})}{\det E^{\dagger, +}} = \frac{\Phi_{1,1}^{\dagger, -} \Phi_{2,2}^{\dagger, +} - \Phi_{1,2}^{\dagger, -} \Phi_{2,1}^{\dagger, +}}{\Phi_{1,1}^{\dagger, +} \Phi_{2,2}^{\dagger, +} - \Phi_{1,2}^{\dagger, +} \Phi_{2,1}^{\dagger, +}} = \frac{\Psi_{1,1}^{\dagger, -} \Psi_{2,2}^{\dagger, +} - \Psi_{1,2}^{\dagger, -} \Psi_{2,1}^{\dagger, +}}{\Psi_{1,1}^{\dagger, +} \Psi_{2,2}^{\dagger, +} - \Psi_{1,2}^{\dagger, +} \Psi_{2,1}^{\dagger, +}}. \quad (2.14)$$

Using this formula, we extend $a^\dagger(z^\dagger)$ analytically to $z^\dagger \in \overline{\mathbb{K}_+}$ up to a finite set of points. Since the determinant is multiplicative, and due to (2.6), we have

$$a^{\text{GP}}(z; q) = a^{\text{hGP}}(z; w) = a^{\varepsilon\text{GP}}(k; W).$$

Note that

$$\det E^{\dagger, \pm} = \begin{cases} 1 & , \dagger = \text{NLS} \\ 2z(\lambda - z) & , \dagger = \text{GP} \\ -2z(\lambda - z) & , \dagger = \text{hGP} \\ -2\sqrt{2}\varepsilon k(\lambda - \sqrt{2}\varepsilon k) & , \dagger = \varepsilon\text{GP} \\ 2ik & , \dagger = \text{KdV} \end{cases}.$$

Using (2.5) to substitute $\Phi_{1,2}^{\dagger, -}$ in (2.14) and taking the limit $x \rightarrow \infty$, or, respectively, by substituting $\Phi_{2,1}^+$ and taking the limit $x \rightarrow -\infty$, we find that

$$\lim_{x^\dagger \rightarrow \infty} \Psi_{1,1}^{\dagger, -} = a^\dagger \beta_1^\dagger \quad \lim_{x^\dagger \rightarrow -\infty} \Psi_{2,2}^{\dagger, +} = a^\dagger \beta_2^\dagger$$

where

$$\beta_1^\dagger = \frac{\det E^{\dagger, +}}{E_{2,2}^{\dagger, +} + (A_{2,2}^{\dagger, +} + iz^\dagger)^{-1} A_{1,2}^{\dagger, +} E_{2,1}^{\dagger, +}} = \begin{cases} 1 & , \dagger = \text{NLS} \\ q_+ & , \dagger = \text{GP} \\ e^{\theta_+} \frac{i}{\sqrt{2}} (1 + \lambda - z) & , \dagger = \text{hGP} \\ e^{\theta_+} \frac{i}{\sqrt{2}} (1 + \lambda - \sqrt{2}\varepsilon k) & , \dagger = \varepsilon\text{GP} \\ 1 & , \dagger = \text{KdV} \end{cases}$$

$$\beta_2^\dagger = \frac{\det E^{\dagger, -}}{E_{1,1}^{\dagger, -} + (A_{1,1}^{\dagger, -} - iz^\dagger)^{-1} A_{2,1}^{\dagger, -} E_{1,2}^{\dagger, -}} = \begin{cases} 1 & , \dagger = \text{NLS} \\ \bar{q}_- & , \dagger = \text{GP} \\ e^{-\theta_-} \frac{i}{\sqrt{2}} (1 + z - \lambda) & , \dagger = \text{hGP} \\ e^{-\theta_-} \frac{i}{\sqrt{2}} (1 + \sqrt{2}\varepsilon k - \lambda) & , \dagger = \varepsilon\text{GP} \\ ik & , \dagger = \text{KdV} \end{cases}.$$

On the other hand, (2.3) implies

$$\lim_{x^\dagger \rightarrow -\infty} \Psi_{1,1}^{\dagger, -} = E_{1,1}^{\dagger, -} = \begin{cases} 1 & , \dagger = \text{NLS} \\ q_- & , \dagger = \text{GP} \\ e^{\theta_-} \frac{i}{\sqrt{2}} (1 + \lambda - z) & , \dagger = \text{hGP} \\ e^{\theta_-} \frac{i}{\sqrt{2}} (1 + \lambda - \sqrt{2}\varepsilon k) & , \dagger = \varepsilon\text{GP} \\ 1 & , \dagger = \text{KdV} \end{cases}.$$

$$\lim_{x^\dagger \rightarrow \infty} \Psi_{2,2}^{\dagger, +} = E_{2,2}^{\dagger, +} = \begin{cases} 1 & , \dagger = \text{NLS} \\ \bar{q}_+ & , \dagger = \text{GP} \\ e^{-\theta_+} \frac{i}{\sqrt{2}} (1 + z - \lambda) & , \dagger = \text{hGP} \\ e^{-\theta_+} \frac{i}{\sqrt{2}} (1 + \sqrt{2}\varepsilon k - \lambda) & , \dagger = \varepsilon\text{GP} \\ ik & , \dagger = \text{KdV} \end{cases}.$$

Due to (2.9)–(2.10), there exists for sufficiently large $\text{Im } z > c = c(q)$ a branch $\log^\dagger$ of the logarithm, only depending on $q_\pm^\dagger$, for which $\log(\Psi_{1,1}^{\dagger, -}(\beta_1^\dagger)^{-1})$ and $\log(\Psi_{2,2}^{\dagger, +}(\beta_2^\dagger)^{-1})$ are continuously differentiable in $x^\dagger$ on all of $\mathbb{R}$. Specifically, writing $\log_{\text{pr}}$ for the principal logarithm, we choose

$$\log^\dagger(\omega) = \begin{cases} \log_{\text{pr}}(\omega) & , \dagger \in \{\text{NLS}, \text{KdV}\} \\ \log_{\text{pr}}(\omega e^{i(\varphi_+ - \varphi_-)}) - \log_{\text{pr}}(e^{i(\varphi_+ - \varphi_-)}) & , \dagger = \text{GP} \\ \log_{\text{pr}}(\omega e^{\frac{i}{2}(\varphi_+ - \varphi_-)}) - \log_{\text{pr}}(e^{\frac{i}{2}(\varphi_+ - \varphi_-)}) & , \dagger \in \{\text{hGP}, \varepsilon\text{GP}\} \end{cases} \quad \forall \omega \in \mathbb{C}.$$

Then

$$\log^\dagger a^\dagger - \begin{cases} 0 & , \dagger = \text{NLS} \\ i(\varphi_- - \varphi_+) & , \dagger = \text{GP} \\ \frac{i}{2}(\varphi_- - \varphi_+) & , \dagger \in \{\text{hGP}, \varepsilon\text{GP}\} \\ 0 & , \dagger = \text{KdV} \end{cases} = \int_{\mathbb{R}} \sigma^\dagger dx^\dagger = \int_{\mathbb{R}} \tilde{\sigma}^\dagger dx^\dagger. \quad (2.15)$$

We are interested in the asymptotic expansions

$$\sigma^\dagger(x^\dagger, z^\dagger; q^\dagger) \sim \sum_{n=0}^{\infty} \frac{\sigma_n^\dagger(x^\dagger; q^\dagger)}{(2iz^\dagger)^n} \quad r^\dagger(x^\dagger, z^\dagger; q^\dagger) \sim \sum_{n=0}^{\infty} \frac{r_n^\dagger(x^\dagger; q^\dagger)}{(2iz^\dagger)^n}. \quad (2.16)$$

By (iii) from Lemma 2.3 these expansions indeed exist, and the expansion coefficients are bounded and integrable in $x^\dagger$. Furthermore, the asymptotic expansion of $\sigma^\dagger$ is L^1 -smooth, meaning that we can take the integral and thereby obtain an asymptotic expansion in powers of $2iz^\dagger$ for the transmission coefficient of the form

$$\log^\dagger a^\dagger - \begin{cases} 0 & , \dagger = \text{NLS} \\ i(\varphi_- - \varphi_+) & , \dagger = \text{GP} \\ \frac{i}{2}(\varphi_- - \varphi_+) & , \dagger \in \{\text{hGP}, \varepsilon\text{GP}\} \\ 0 & , \dagger = \text{KdV} \end{cases} \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} \sigma_n^\dagger dx^\dagger}{(2iz^\dagger)^n} \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} \tilde{\sigma}_n^\dagger dx^\dagger}{(2iz^\dagger)^n}. \quad (2.17)$$

For the cases $\dagger \in \{\text{GP}, \text{hGP}, \varepsilon\text{GP}\}$ clarification is necessary, as the expansion coefficients both in (2.16) and (2.17) are not the same on all of $\overline{\mathbb{K}_+}$. Instead, they depend on the sign of $\text{Im } \lambda$, or alternatively a choice of bijection $\lambda \leftrightarrow (z, k)$. We shall choose the expansion coefficients so that (2.16)–(2.17) hold true for $\text{Im } \lambda > 0$.

Using these expansion coefficients, we define for $\dagger \neq \text{KdV}$ the Hamiltonians

$$\mathcal{H}_n^\dagger(q^\dagger) = -(-i)^n ((\sqrt{2\varepsilon})^{(n+1)})^{\mathbb{1}_{\{\dagger=\varepsilon\text{GP}\}}} \int_{\mathbb{R}} \sigma_{n+1}^\dagger(x^\dagger; q^\dagger) dx^\dagger + \mathbb{1}_{\{n=-1\}} \begin{cases} 0 & , \dagger = \text{NLS} \\ \varphi_- - \varphi_+ & , \dagger = \text{GP} \\ \frac{\varphi_- - \varphi_+}{2} & , \dagger \in \{\text{hGP}, \varepsilon\text{GP}\} \end{cases}, \quad (2.18)$$

and for $\dagger = \text{KdV}$ the energies

$$\mathcal{E}_n^{\text{KdV}}(U) = (-1)^n \int_{\mathbb{R}} \sigma_{2n+3}^{\text{KdV}}(X; U) dX. \quad (2.19)$$

We do not define conserved quantities corresponding to the even orders because they vanish (see Lemma 2.6 below). These Hamiltonians act as expansion coefficients for the logarithm of the transmission coefficient. Specifically, for $(\lambda, z, k) \in \mathbb{K}_+ \cap \overline{\mathcal{Q}_1}$ we have as $\text{Im } z \rightarrow \infty$ the asymptotic expansions

$$\begin{aligned} \log a^\dagger(z^\dagger; q^\dagger) &\sim i \sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^\dagger(q^\dagger)}{(2z)^{n+1}} & \forall \dagger \in \{\text{NLS}, \text{GP}, \text{hGP}, \varepsilon\text{GP}\} \\ \log a^{\text{KdV}}(k; U) &\sim i \sum_{n=-1}^{\infty} \frac{\mathcal{E}_n^{\text{KdV}}(U)}{(2k)^{2n+3}}. \end{aligned}$$

In particular

$$\mathcal{H}_n^{\text{GP}}(q) = \mathcal{H}_n^{\text{hGP}}(w) = \mathcal{H}_n^{\varepsilon\text{GP}}(W). \quad (2.20)$$

Note that we can reformulate (1.3) as

$$\log a^{\text{GP}}(z; q) \sim i \sum_{n=0}^{\infty} \frac{\mathcal{H}_{2n}^{\widetilde{\text{GP}}}(q)}{(2z)^{2n+1}} + i \frac{\lambda}{z} \sum_{n=0}^{\infty} \frac{2^{\mathbf{1}_{\{n=0\}}} \mathcal{H}_{2n-1}^{\widetilde{\text{GP}}}(q)}{(2z)^{2n}},$$

which gives an alternative view on the definition of the $\widetilde{\text{GP}}$ hierarchy and its relation to the GP hierarchy.

As mentioned above, for the cases $\dagger \in \{\text{GP}, \text{hGP}, \varepsilon\text{GP}\}$ these expansion coefficients are specific to our choice $\text{Im } \lambda > 0$. A crucial property of the Hamiltonians $\mathcal{H}_n^\dagger$ is that they are pairwise Poisson commuting. This is a well-known result and it implies that each Hamiltonian is a conserved quantity for every Hamiltonian flow.

Lemma 2.4. *Let $\dagger \in \text{Eqns}$ and $q^\dagger \in C^\infty(\mathbb{R})$ with $q^\dagger - q_\pm^\dagger \in \mathcal{S}(\mathbb{R}_\pm)$. Below, we write $\mathcal{H}_n^{\text{KdV}} = \mathcal{E}_n^{\text{KdV}}$. For all $n, m \geq 0$ and $z_1^\dagger, z_2^\dagger \in \mathbb{K}_+^\dagger$ with $\text{Im } z_1^\dagger, \text{Im } z_2^\dagger > c(q^\dagger)$ we have*

$$\{\mathcal{H}_n^\dagger, \mathcal{H}_m^\dagger\}^\dagger(q^\dagger) = 0 \quad \{\log a^\dagger(z_1^\dagger), \log a^\dagger(z_2^\dagger)\}(q^\dagger) = 0.$$

The Poisson brackets $\{\dots, \dots\}^\dagger$ are given in Appendix B.

Proof. For NLS, GP and KdV, which are all part of the AKNS hierarchy, we refer to [58, Theorem B.7]. Specifically for GP we note also a different proof given in [30, III.§2]. With (2.20), we can derive from the case GP the cases hGP and εGP . $\square$

2.4 Recurrence relations for the densities of the Hamiltonians

From (2.5) we derive Riccati equations for $\sigma^\dagger$ and $\tilde{\sigma}^\dagger$ with $\dagger \in \{\text{NLS}, \text{GP}\}$:

$$\partial_x \sigma^{\text{NLS}} = (2i\lambda) \sigma^{\text{NLS}} + \frac{q_x}{q} \sigma^{\text{NLS}} - (\sigma^{\text{NLS}})^2 + q\bar{q} \quad (2.21)$$

$$-\partial_x \tilde{\sigma}^{\text{NLS}} = (2i\lambda) \tilde{\sigma}^{\text{NLS}} - \frac{\bar{q}_x}{\bar{q}} \tilde{\sigma}^{\text{NLS}} - (\tilde{\sigma}^{\text{NLS}})^2 + q\bar{q} \quad (2.22)$$

$$\partial_x \sigma^{\text{GP}} = (2iz) \sigma^{\text{GP}} - (\sigma^{\text{GP}})^2 + q\bar{q} - 1 + q_x \frac{\sigma^{\text{GP}} + i(\lambda - z)}{q} \quad (2.23)$$

$$-\partial_x \tilde{\sigma}^{\text{GP}} = (2iz) \tilde{\sigma}^{\text{GP}} - (\tilde{\sigma}^{\text{GP}})^2 + q\bar{q} - 1 - \bar{q}_x \frac{\tilde{\sigma}^{\text{GP}} + i(\lambda - z)}{\bar{q}} \quad (2.24)$$

They were first derived in [87]. For $\dagger \in \{\text{hGP}, \varepsilon\text{GP}\}$ we prefer to work with the systems for the pairs of variables $(\sigma^\dagger, r^\dagger)$ and $(\tilde{\sigma}^\dagger, \tilde{r}^\dagger)$, which are specifically

$$\partial_x \sigma^{\text{hGP}} = (2iz) \sigma^{\text{hGP}} - (\sigma^{\text{hGP}})^2 - w_- w_+ - w_- (1 + \lambda) + w_+ (1 - \lambda) - i(w_+)_x r^{\text{hGP}} \quad (2.25)$$

$$\partial_x \sigma^{\text{hGP}} = (2iz) \sigma^{\text{hGP}} - (\sigma^{\text{hGP}})^2 - w_- w_+ - w_- (1 + \lambda) + w_+ (1 - \lambda) - i(w_+)_x r^{\text{hGP}} \quad (2.26)$$

$$-\partial_x \tilde{\sigma}^{\text{hGP}} = (2iz) \tilde{\sigma}^{\text{hGP}} - (\tilde{\sigma}^{\text{hGP}})^2 - w_- w_+ - w_- (1 + \lambda) + w_+ (1 - \lambda) - i(w_-)_x \tilde{r}^{\text{hGP}} \quad (2.27)$$

$$\partial_x r^{\text{hGP}} = i(1 + \lambda + w_+) (r^{\text{hGP}})^2 + i(1 - \lambda - w_-) \quad (2.28)$$

$$-\partial_x \tilde{r}^{\text{hGP}} = i(1 - \lambda - w_-) (\tilde{r}^{\text{hGP}})^2 + i(1 + \lambda + w_+) \quad (2.29)$$

$$\partial_X \sigma^{\varepsilon\text{GP}} = (2ik) \sigma^{\varepsilon\text{GP}} - (\sigma^{\varepsilon\text{GP}})^2 + \frac{1}{2} \left(-\varepsilon^2 W_- W_+ - W_- (1 + \lambda) + W_+ (1 - \lambda) - i\sqrt{2}\varepsilon (W_+)_X r^{\varepsilon\text{GP}} \right) \quad (2.30)$$

$$-\partial_X \tilde{\sigma}^{\varepsilon\text{GP}} = (2ik) \tilde{\sigma}^{\varepsilon\text{GP}} - (\tilde{\sigma}^{\varepsilon\text{GP}})^2 + \frac{1}{2} \left(-\varepsilon^2 W_- W_+ - W_- (1 + \lambda) + W_+ (1 - \lambda) - i\sqrt{2}\varepsilon (W_-)_X \tilde{r}^{\varepsilon\text{GP}} \right) \quad (2.31)$$

$$\partial_X r^{\varepsilon\text{GP}} = \frac{i\varepsilon}{\sqrt{2}} (\varepsilon^{-2} (1 + \lambda) + W_+) (r^{\varepsilon\text{GP}})^2 + \frac{i\varepsilon}{\sqrt{2}} (\varepsilon^{-2} (1 - \lambda) - W_-) \quad (2.32)$$

$$-\partial_X \tilde{r}^{\varepsilon\text{GP}} = \frac{i\varepsilon}{\sqrt{2}} (\varepsilon^{-2} (1 - \lambda) - W_-) (\tilde{r}^{\varepsilon\text{GP}})^2 + \frac{i\varepsilon}{\sqrt{2}} (\varepsilon^{-2} (1 + \lambda) + W_+). \quad (2.33)$$

The case KdV is distinct in that we consider a single equation for σ^{KdV} , but also the system $(\tilde{\sigma}^{\text{KdV}}, \tilde{r}^{\text{KdV}})$:

$$\partial_X \sigma^{\text{KdV}} = (2ik)\sigma^{\text{KdV}} - (\sigma^{\text{KdV}})^2 + U \quad (2.34)$$

$$-\partial_X \tilde{\sigma}^{\text{KdV}} = (2ik)\tilde{\sigma}^{\text{KdV}} - (\tilde{\sigma}^{\text{KdV}})^2 + U + U_X \tilde{r}^{\text{KdV}} \quad (2.35)$$

$$-\partial_X \tilde{r}^{\text{KdV}} = \left(U + \frac{1}{4}(2ik)^2 \right) (\tilde{r}^{\text{KdV}})^2 - 1. \quad (2.36)$$

These equations prescribe recurrence relations for the expansion coefficients $\sigma_n^\dagger$. In order to avoid an abuse of notation, we shall use bold variables $\boldsymbol{\sigma}^\dagger, \tilde{\boldsymbol{\sigma}}^\dagger, \mathbf{r}^\dagger, \tilde{\mathbf{r}}^\dagger$, and so on to denote the formal power series corresponding to each function. By substituting the formal power series $\boldsymbol{\sigma}^{\text{NLS}}, \tilde{\boldsymbol{\sigma}}^{\text{NLS}}, \boldsymbol{\sigma}^{\text{GP}}$ and $\tilde{\boldsymbol{\sigma}}^{\text{GP}}$ into the Riccati equations (2.21)–(2.24), we can derive recurrence relations for the expansion coefficients $\sigma_n^{\text{NLS}}, \tilde{\sigma}_n^{\text{NLS}}, \sigma_n^{\text{GP}}$ and $\tilde{\sigma}_n^{\text{GP}}$. For the case NLS this yields

$$\sigma_0^{\text{NLS}} = 0 \quad \sigma_1^{\text{NLS}} = -q\bar{q} \quad \sigma_{n+1}^{\text{NLS}} = \partial_x \sigma_n^{\text{NLS}} - \frac{q_x}{q} \sigma_n^{\text{NLS}} + \sum_{k=0}^n \sigma_k^{\text{NLS}} \sigma_{n-k}^{\text{NLS}}.$$

Due to the presence of λ on the right-hand side of (2.23), the recurrence relation for σ_n^{GP} is awkward, so we use instead the relation

$$\boldsymbol{\sigma}^{\text{NLS}}(x, \lambda) = \boldsymbol{\sigma}^{\text{GP}}(x, z) + i(\lambda - z),$$

which is proven by using it to substitute $\boldsymbol{\sigma}^{\text{NLS}}$ in (2.21) and observing that this yields (2.23). We can then derive the expansion coefficients σ_n^{GP} from σ_n^{NLS} and vice versa. We assume now that $\lambda, z \in \overline{\mathcal{Q}_1}$. Then we can write $z = \sqrt{\lambda^2 - 1}$ and $\lambda = \sqrt{z^2 + 1}$ using the principal square root, and obtain the relations

$$\sigma_{2m}^{\text{NLS}} = \sum_{k=0}^m \binom{m-1}{m-k} (-4)^{m-k} \sigma_{2k}^{\text{GP}} \quad \sigma_{2m+1}^{\text{NLS}} = \sum_{k=0}^m \binom{m-\frac{1}{2}}{m-k} (-4)^{m-k} \sigma_{2k+1}^{\text{GP}} + (-1)^{m+1} C_m \quad (2.37)$$

$$\sigma_{2m}^{\text{GP}} = \sum_{k=0}^m \binom{m-1}{m-k} 4^{m-k} \sigma_{2k}^{\text{NLS}} \quad \sigma_{2m+1}^{\text{GP}} = \sum_{k=0}^m \binom{m-\frac{1}{2}}{m-k} 4^{m-k} \sigma_{2k+1}^{\text{NLS}} + C_m. \quad (2.38)$$

Here C_n are the Catalan numbers. They can be defined either by the recurrence relation they solve, or by their generating function:

$$C_0 = 1 \quad C_{n+1} = \sum_{k=0}^n C_k C_{n-k} \quad \sum_{n=0}^{\infty} X^n C_n = \frac{1 - \sqrt{1 - 4X}}{2X}. \quad (2.39)$$

If $\lambda \notin \overline{\mathcal{Q}_1}$ or $z \notin \overline{\mathcal{Q}_1}$ then the relations (2.37)–(2.38) may need to be corrected by sign changes.

Defining $q^\sharp(x) = q(-x)$, we can observe that $\boldsymbol{\sigma}^{\text{NLS}}(x, \lambda; q)$ and $\tilde{\boldsymbol{\sigma}}^{\text{NLS}}(-x, \lambda; \bar{q}^\sharp)$ both solve (2.21), and $\boldsymbol{\sigma}^{\text{GP}}(x, z; q)$ and $\tilde{\boldsymbol{\sigma}}^{\text{GP}}(-x, z; \bar{q}^\sharp)$ both solve (2.23). A consequence of the recurrence relations discussed above is that σ_n^{NLS} and σ_n^{GP} consist of monomials whose total number of derivatives has the opposite parity of n . Then $(\partial_x^k q^\sharp)(x) = (-1)^k (\partial_x^k q)(-x)$ together with (2.17) yields

$$\int \sigma_n^\dagger(x, z; q) dx = (-1)^{n+1} \int \overline{\sigma_n^\dagger(x, z; q)} dx \quad (2.40)$$

for $\dagger \in \{\text{NLS}, \text{GP}\}$, implying that the functionals $\mathcal{H}_n^{\text{NLS}}, \mathcal{H}_n^{\text{GP}}, \mathcal{H}_n^{\text{hGP}}$ and $\mathcal{H}_n^{\varepsilon\text{GP}}$ are real-valued. We assume again that $\lambda, k \in \overline{\mathcal{Q}_1}$ are in the first quadrant of the complex plane in order to have an explicit expansion

$$\lambda = \sqrt{1 + 2\varepsilon^2 k^2} = \sum_{n=-1}^{\infty} \frac{c_n}{(2ik)^n} \quad \text{where} \quad c_n = \begin{cases} -\frac{i\varepsilon}{\sqrt{2}} & , n = -1 \\ \frac{i}{(\sqrt{2}\varepsilon)^n} \frac{1}{n} \binom{n-1}{\frac{n-1}{2}} & , n \geq 0 \text{ odd} \\ 0 & , \text{else} \end{cases} \quad (2.41)$$

Then we can derive from (2.30)–(2.33) the recurrence relations

$$\sigma_0^{\varepsilon\text{GP}} = \frac{c_{-1}}{2}(W_+ + W_-) \quad (2.42)$$

$$\sigma_1^{\varepsilon\text{GP}} = \frac{c_{-1}}{2}\partial_X(W_+ + W_-) + \frac{c_{-1}^2}{4}(W_+ + W_-)^2 + \frac{\varepsilon^2}{2}W_-W_+ + \frac{1}{2}(W_- - W_+) + \frac{i\varepsilon}{\sqrt{2}}\partial_X W_+ r_0^{\varepsilon\text{GP}} \quad (2.43)$$

$$\sigma_{n+1}^{\varepsilon\text{GP}} = \partial_X \sigma_n^{\varepsilon\text{GP}} + \sum_{k=0}^n \sigma_{n-k}^{\varepsilon\text{GP}} \sigma_k^{\varepsilon\text{GP}} + \frac{1}{2}c_n(W_- + W_+) + \frac{i\varepsilon}{\sqrt{2}}\partial_X W_+ r_n^{\varepsilon\text{GP}}. \quad (2.44)$$

and

$$r_0^{\varepsilon\text{GP}} = 1 \quad (2.45)$$

$$r_1^{\varepsilon\text{GP}} = -\frac{i\varepsilon}{\sqrt{2}}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right) \quad (2.46)$$

$$\begin{aligned} r_{n+1}^{\varepsilon\text{GP}} &= \partial_X r_n^{\varepsilon\text{GP}} + \frac{ic_n}{\sqrt{2}\varepsilon} + \frac{i\varepsilon}{\sqrt{2}}\left(-W_+ - \frac{1}{\varepsilon^2}\right) \sum_{k=0}^n r_{n-k}^{\varepsilon\text{GP}} r_k^{\varepsilon\text{GP}} \\ &\quad - \frac{i}{\sqrt{2}\varepsilon} \sum_{m=0}^n \sum_{k=0}^m c_{n-m} r_{m-k}^{\varepsilon\text{GP}} r_k^{\varepsilon\text{GP}} - \frac{ic_{-1}}{\sqrt{2}\varepsilon} \sum_{k=1}^n r_{n+1-k}^{\varepsilon\text{GP}} r_k^{\varepsilon\text{GP}}. \end{aligned} \quad (2.47)$$

In the same way recurrence relations can be derived for $\tilde{\sigma}^{\varepsilon\text{GP}}$, $\tilde{r}^{\varepsilon\text{GP}}$, σ^{hGP} , $\tilde{\sigma}^{\text{hGP}}$, r^{hGP} , and $\tilde{r}^{\text{hGP}}$, which we omit here for brevity.

Lastly, we derive from (2.34)–(2.36) the recurrence relations

$$\sigma_0^{\text{KdV}} = 0 \quad \sigma_1^{\text{KdV}} = -U \quad \sigma_{n+1}^{\text{KdV}} = \partial_X \sigma_n^{\text{KdV}} + \sum_{k=0}^n \sigma_k^{\text{KdV}} \sigma_{n-k}^{\text{KdV}}. \quad (2.48)$$

and

$$\tilde{\sigma}_0^{\text{KdV}} = 0 \quad \tilde{\sigma}_1^{\text{KdV}} = -U \quad \tilde{\sigma}_{n+1}^{\text{KdV}} = -\partial_X \tilde{\sigma}_n^{\text{KdV}} + \sum_{k=0}^n \tilde{\sigma}_k^{\text{KdV}} \tilde{\sigma}_{n-k}^{\text{KdV}} - U_X \tilde{r}_n^{\text{KdV}}$$

as well as

$$\tilde{r}_0^{\text{KdV}} = 0 \quad \tilde{r}_1^{\text{KdV}} = 2 \quad \tilde{r}_{n+1}^{\text{KdV}} = -\partial_X \tilde{r}_n^{\text{KdV}} - U \sum_{k=0}^n \tilde{r}_k^{\text{KdV}} \tilde{r}_{n-k}^{\text{KdV}} - \frac{1}{4} \sum_{k=2}^n \tilde{r}_k^{\text{KdV}} \tilde{r}_{n+2-k}^{\text{KdV}}.$$

Note that something subtle has happened in (2.45): the equation (2.30) only implies $(r_0^{\varepsilon\text{GP}})^2 = 1$, so we need to use the additional information

$$r_0^{\varepsilon\text{GP}} = \lim_{k \rightarrow i\infty} \frac{E_{1,2}^{\varepsilon\text{GP},-}(k)}{E_{1,1}^{\varepsilon\text{GP},-}(k)} = \lim_{k, \lambda \rightarrow i\infty} \frac{1 + \sqrt{2}\varepsilon k - \lambda}{1 - \sqrt{2}\varepsilon k + \lambda} = 1.$$

On the other hand, we have $\tilde{r}_0^{\varepsilon\text{GP}} = -1$. These statements are true only in the case $\text{Im } \lambda > 0$, which we have assumed by considering the asymptotic expansion for $\lambda, k \in \overline{\mathcal{Q}_1}$. On other regions of $\overline{\mathbb{K}_+}$, where instead $\text{Im } \lambda < 0$ while of course $\text{Im } k > 0$, we have $r_0^{\varepsilon\text{GP}} = -1$, $\tilde{r}_0^{\varepsilon\text{GP}} = 1$. In addition we must also replace c_n by $-c_n$, hence the expansion coefficients change significantly. This is relevant for the following lemma, which uses the operators Ev_λ and Od_λ defined in (1.14).

Lemma 2.5. *For $\lambda, k \in \overline{\mathcal{Q}_1}$ we have the following asymptotic expansions in powers of $2ik$ at infinity:*

$$\text{Ev}_\lambda[\log a^{\varepsilon\text{GP}}(k)] \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} \text{Re } \sigma_{2n+1}^{\varepsilon\text{GP}} dX}{(2ik)^{2n+1}} \quad \text{Od}_\lambda[\log a^{\varepsilon\text{GP}}(k)] \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} i \text{Im } \sigma_{2n}^{\varepsilon\text{GP}} dX}{(2ik)^{2n}}. \quad (2.49)$$

Proof. Due to (2.40), the real and imaginary parts on the right-hand side are superfluous. Define now $W_{\pm}^{\sharp}(X) = W_{\pm}(-X)$. An analysis of (2.30)–(2.33) reveals that

$$\begin{aligned} r^{\varepsilon\text{GP}}(X, (-\lambda, k); (W_+, W_-)) &= \tilde{r}^{\varepsilon\text{GP}}(-X, (\lambda, k); (-W_-^{\sharp}, -W_+^{\sharp})) \\ \sigma^{\varepsilon\text{GP}}(X, (-\lambda, k); (W_+, W_-)) &= \tilde{\sigma}^{\varepsilon\text{GP}}(-X, (\lambda, k); (-W_-^{\sharp}, -W_+^{\sharp})). \end{aligned}$$

Observe that $(W_+, W_-) \mapsto (-W_-^{\sharp}, -W_+^{\sharp})$ corresponds to $q \mapsto q^{\sharp}$. Then one can proceed as in the proof of (2.40) to obtain

$$\int_{\mathbb{R}} \tilde{\sigma}_n^{\varepsilon\text{GP}}(-X, (\lambda, k); (-W_-^{\sharp}, -W_+^{\sharp})) \, dX = (-1)^{n+1} \int_{\mathbb{R}} \sigma_n^{\varepsilon\text{GP}}(X, (\lambda, k); (W_+, W_-)) \, dX,$$

which yields the claimed expansions. $\square$

2.5 Proofs of Theorem 1.1 and Corollary 1.5

Let us emphasize that these recurrence relations are convenient for the explicit computation of the expansion coefficients, but difficult to work with in proofs by induction. Instead, it is preferable to work with their generating functions, which solve the Riccati equations derived above, as operations on these generating functions elegantly represent complicated steps of a proof by induction. As an example, consider the following lemma.

Lemma 2.6. *For all $U \in \mathcal{S}(\mathbb{R}; \mathbb{R})$ and $n \geq 0$*

$$\int_{\mathbb{R}} \sigma_{2n}^{\text{KdV}}(X; U) \, dX = 0.$$

Proof. We decompose $\sigma^{\text{KdV}} = E + O$ into the even and odd orders of the formal power series. Then (2.34) implies

$$O_X = (2ik)E - 2EO.$$

In particular

$$E = \frac{O_X}{2ik - 2O} = -\frac{(2ik - 2O)_X}{2ik - 2O} = -\frac{1}{2}(\log(2ik - 2O))_X.$$

Since the right-hand side is a formal power series where each coefficient is a derivative of a polynomial in U and its derivatives, its integral vanishes. $\square$

This is essentially a proof by induction, but all the work of the induction step is hidden in the calculus of formal power series. In order to transform it into an actual proof by induction one needs to compare the coefficients at each order, which yields recurrence relations. The calculations with generating functions then represent the inductive step for these recurrence relations. We generally skip the base case, which sometimes also falls out of the generating function calculations, but otherwise can always be checked manually.

Using the generating-function approach, we prove our approximation result for the densities of the Hamiltonians. In the following one should understand $k \in i\mathbb{R}_+$ to be sufficiently large so that $\lambda \in i\mathbb{R}$. **Lemma 2.7.** *Define $W_{\pm}^{\sharp}(X) = W_{\pm}(-X)$. We assume that $k \in i\mathbb{R}$ for the definition of the real and imaginary part of formal power series in the variable $2ik$. Let λ be a purely imaginary formal power*

series in the variable $2ik$ which satisfies $\lambda^2 = 1 - \frac{\varepsilon^2}{2}(2ik)^2$. Then

$$\operatorname{Re}[\sigma^{\varepsilon GP}](X, k; W) - \frac{1}{i\lambda} \operatorname{Im}[\sigma^{\varepsilon GP}](X, k; W) = \sigma^{KdV}(X, k; -W_-) + O(\varepsilon^2) \quad (2.50)$$

$$\operatorname{Re}[\sigma^{\varepsilon GP}](X, k; W) + \frac{1}{i\lambda} \operatorname{Im}[\sigma^{\varepsilon GP}](X, k; W) = \tilde{\sigma}^{KdV}(-X, k; W_+^\sharp) + O(\varepsilon^2) \quad (2.51)$$

$$\operatorname{Re}[\tilde{\sigma}^{\varepsilon GP}](X, k; W) - \frac{1}{i\lambda} \operatorname{Im}[\tilde{\sigma}^{\varepsilon GP}](X, k; W) = \tilde{\sigma}^{KdV}(X, k; -W_-) + O(\varepsilon^2) \quad (2.52)$$

$$\operatorname{Re}[\tilde{\sigma}^{\varepsilon GP}](X, k; W) + \frac{1}{i\lambda} \operatorname{Im}[\tilde{\sigma}^{\varepsilon GP}](X, k; W) = \sigma^{KdV}(-X, k; W_+^\sharp) + O(\varepsilon^2) \quad (2.53)$$

and

$$\begin{aligned} \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right](X, k; W) - \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right](X, k; W) &= O(\varepsilon^2) \\ \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right](X, k; W) + \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right](X, k; W) &= \tilde{\mathbf{r}}^{KdV}(-X, k; W_+^\sharp) + O(\varepsilon^2) \\ \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right](X, k; W) - \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right](X, k; W) &= \tilde{\mathbf{r}}^{KdV}(X, k; -W_-) + O(\varepsilon^2) \\ \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right](X, k; W) + \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right](X, k; W) &= O(\varepsilon^2). \end{aligned}$$

Proof. We define

$$\begin{aligned} \mu &= \operatorname{Re}[\sigma^{\varepsilon GP}] + \frac{1}{i\lambda} \operatorname{Im}[\sigma^{\varepsilon GP}] & \nu &= \operatorname{Re}[\sigma^{\varepsilon GP}] - \frac{1}{i\lambda} \operatorname{Im}[\sigma^{\varepsilon GP}] \\ \tilde{\mu} &= \operatorname{Re}[\tilde{\sigma}^{\varepsilon GP}] + \frac{1}{i\lambda} \operatorname{Im}[\tilde{\sigma}^{\varepsilon GP}] & \tilde{\nu} &= \operatorname{Re}[\tilde{\sigma}^{\varepsilon GP}] - \frac{1}{i\lambda} \operatorname{Im}[\tilde{\sigma}^{\varepsilon GP}] \\ \alpha &= \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right] + \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right] & \beta &= \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right] - \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon GP}\right] \\ \tilde{\alpha} &= \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right] + \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right] & \tilde{\beta} &= \operatorname{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right] - \frac{1}{i\lambda} \operatorname{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\tilde{\mathbf{r}}^{\varepsilon GP}\right] \end{aligned}$$

From (2.32)–(2.33) we derive the system

$$\begin{aligned} \alpha_X - W_+ \alpha^2 + (2ik)^2 \left(\frac{\varepsilon^2}{2} W_+ \left(\frac{\alpha - \beta}{2} \right)^2 - \left(\frac{\alpha + \beta}{2} \right) \left(\frac{\alpha - \beta}{2} \right) \right) + 1 - \frac{\varepsilon^2}{2} W_- &= 0 \\ \beta_X - (W_+ + 2\varepsilon^{-2}) \beta^2 + (2ik)^2 \left(\frac{\varepsilon^2}{2} W_+ \left(\frac{\alpha - \beta}{2} \right)^2 - \beta \left(\frac{\alpha - \beta}{2} \right) \right) - \frac{\varepsilon^2}{2} W_- &= 0 \\ -\tilde{\alpha}_X + (W_- - 2\varepsilon^{-2}) \tilde{\alpha}^2 - (2ik)^2 \left(\frac{\varepsilon^2}{2} W_- \left(\frac{\tilde{\alpha} - \tilde{\beta}}{2} \right)^2 - \tilde{\alpha} \left(\frac{\tilde{\alpha} - \tilde{\beta}}{2} \right) \right) + \frac{\varepsilon^2}{2} W_+ &= 0 \\ -\tilde{\beta}_X + W_- \tilde{\alpha}^2 - (2ik)^2 \left(\frac{\varepsilon^2}{2} W_- \left(\frac{\tilde{\alpha} - \tilde{\beta}}{2} \right)^2 - \left(\frac{\tilde{\alpha} + \tilde{\beta}}{2} \right) \left(\frac{\tilde{\alpha} - \tilde{\beta}}{2} \right) \right) + 1 + \frac{\varepsilon^2}{2} W_+ &= 0. \end{aligned}$$

This implies

$$\alpha, \tilde{\beta} = O(\varepsilon^0) \quad \tilde{\alpha}, \beta = O(\varepsilon^2) \quad (2.54)$$

by induction. More precisely, we have

$$\begin{aligned}\alpha_X - \left(W_+ + \frac{1}{4}(2ik)^2\right)\alpha^2 + 1 &= O(\varepsilon^2) \\ -\tilde{\beta}_X - \left(-W_- + \frac{1}{4}(2ik)^2\right)\tilde{\beta}^2 + 1 &= O(\varepsilon^2).\end{aligned}$$

As a result, the claimed approximations for $\mathbf{r}^{\varepsilon\text{GP}}$ and $\tilde{\mathbf{r}}^{\varepsilon\text{GP}}$ hold. On the other hand, we obtain from (2.30)–(2.31) the system

$$\begin{aligned}\mu_X - (2ik)\mu + \mu^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\mu - \nu}{2}\right)^2 - W_+ + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_+ \alpha &= 0 \\ \nu_X - (2ik)\nu + \nu^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\mu - \nu}{2}\right)^2 + W_- + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_+ \beta &= 0 \\ -\tilde{\mu}_X - (2ik)\tilde{\mu} + \tilde{\mu}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\tilde{\mu} - \tilde{\nu}}{2}\right)^2 - W_+ + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_- \tilde{\alpha} &= 0 \\ -\tilde{\nu}_X - (2ik)\tilde{\nu} + \tilde{\nu}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\tilde{\mu} - \tilde{\nu}}{2}\right)^2 + W_- + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_- \tilde{\beta} &= 0.\end{aligned}$$

Again with induction, we find that

$$\begin{aligned}\mu_X - (2ik)\mu + \mu^2 - W_+ + (W_+)_X \alpha &= O(\varepsilon^2) \\ \nu_X - (2ik)\nu + \nu^2 + W_- &= O(\varepsilon^2) \\ -\tilde{\mu}_X - (2ik)\tilde{\mu} + \tilde{\mu}^2 - W_+ &= O(\varepsilon^2) \\ -\tilde{\nu}_X - (2ik)\tilde{\nu} + \tilde{\nu}^2 + W_- + (W_-)_X \tilde{\beta} &= O(\varepsilon^2).\end{aligned}$$

This concludes the proof of the approximations for $\sigma^{\varepsilon\text{GP}}$ and $\tilde{\sigma}^{\varepsilon\text{GP}}$. $\square$

Proof of Theorem 1.1. We assume again a setting where $\lambda, k \in \overline{\mathcal{Q}_1}$. From the definition (2.18) and the fact that the Hamiltonians are real, we obtain

$$\begin{aligned}\mathcal{H}_{2n}^{\varepsilon\text{GP}}(W) &= (-1)^{n+1}(\sqrt{2}\varepsilon)^{2n+1} \int_{\mathbb{R}} \text{Re } \sigma_{2n+1}^{\varepsilon\text{GP}}(X; W) \, dX \\ \mathcal{H}_{2n-1}^{\varepsilon\text{GP}}(W) &= (-1)^n(\sqrt{2}\varepsilon)^{2n} \int_{\mathbb{R}} \text{Im } \sigma_{2n}^{\varepsilon\text{GP}}(X; W) \, dX.\end{aligned}$$

Now combining (1.3) with (2.19) and Lemma 2.7 yields (1.9). Note that $\sigma_n^{\text{KdV}}(\pm W_{\pm})$ is a polynomial in U and its derivatives with an even number of derivatives present in each monomial. Therefore $\int_{\mathbb{R}} \sigma_n^{\text{KdV}}(\mp X; \pm W_{\pm}^{\sharp}) \, dX = \int_{\mathbb{R}} \sigma_n^{\text{KdV}}(X; \pm W_{\pm}) \, dX$.

The main work is done, but we would like to gain a precise understanding of the number of derivatives present in $\mathbf{P}_n$ and prove (1.12). Below we prove results about $\mathcal{H}^{\varepsilon\text{GP}}$, which then directly imply the claims about $\mathcal{H}^{\varepsilon\widetilde{\text{GP}}}$. For $k \in \mathbb{N}$ and a polynomial P in W_+ , W_- , and their derivatives, we define $\pi_k P$ to be the sum of all monomials in P whose total number of derivatives is k . In the following we identify such polynomials P up to formal integration by parts, as the claims of Theorem 1.1 concern only integrals $\int_{\mathbb{R}} P \, dX$.

We set now $\dagger = \varepsilon \text{GP}$ in order to simplify the notation. From (2.42)–(2.44) we derive

$$\begin{aligned}\text{Re } \sigma_0^\dagger &= 0 \\ \text{Re } \sigma_1^\dagger &= -\frac{\varepsilon^2}{8}(W_+ - W_-)^2 + \frac{1}{2}(W_- - W_+) \\ \text{Re } \sigma_{n+1}^\dagger &= \partial_X \text{Re } \sigma_n^\dagger + \sum_{k=0}^n (\text{Re } \sigma_{n-k}^\dagger \text{Re } \sigma_k^\dagger - \text{Im } \sigma_{n-k}^\dagger \text{Im } \sigma_k^\dagger) - \frac{\varepsilon}{\sqrt{2}} \partial_X W_+ \text{Im } r_n^\dagger\end{aligned}$$

and

$$\begin{aligned}\text{Im } \sigma_0^\dagger &= -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-) \\ \text{Im } \sigma_1^\dagger &= \frac{\varepsilon}{2\sqrt{2}} \partial_X (W_+ - W_-) \\ \text{Im } \sigma_{n+1}^\dagger &= \partial_X \text{Im } \sigma_n^\dagger + \sum_{k=0}^n 2 \text{Re } \sigma_{n-k}^\dagger \text{Im } \sigma_k^\dagger + \frac{1}{2} \text{Im } c_n (W_- + W_+) + \frac{\varepsilon}{\sqrt{2}} \partial_X W_+ \text{Re } r_n^\dagger\end{aligned}$$

as well as

$$\begin{aligned}\text{Im } r_0^\dagger &= 0 \\ \text{Im } r_1^\dagger &= -\frac{\varepsilon}{\sqrt{2}} \left(W_+ - W_- + \frac{2}{\varepsilon^2} \right) \\ \text{Im } r_{n+1}^\dagger &= \partial_X \text{Im } r_n^\dagger + \frac{\varepsilon}{\sqrt{2}} \left(-W_+ - \frac{1}{\varepsilon^2} \right) \sum_{k=0}^n (\text{Re } r_{n-k}^\dagger \text{Re } r_k^\dagger - \text{Im } r_{n-k}^\dagger \text{Im } r_k^\dagger) \\ &\quad + \frac{1}{\sqrt{2}\varepsilon} \sum_{m=0}^n \sum_{k=0}^m \text{Im } c_{n-m} 2 \text{Re } r_{m-k}^\dagger \text{Im } r_k^\dagger + \frac{\text{Im } c_{-1}}{\sqrt{2}\varepsilon} \sum_{k=1}^n 2 \text{Re } r_{n+1-k}^\dagger \text{Im } r_k^\dagger.\end{aligned}$$

We do not write out the recurrence relation for $\text{Re } r_n^\dagger$ since it plays no significant role. The first fact that we read from these recurrence relations is that $\text{Re } \sigma_{2n+1}^\dagger$ and $\text{Im } \sigma_{2n}^\dagger$ are sums of monomials with an even number of derivatives, while all monomials in $\text{Re } \sigma_{2n}^\dagger$ and $\text{Im } \sigma_{2n+1}^\dagger$ have an odd number of derivatives. Studying the recurrence relation for $\text{Im } r_n^\dagger$ explicitly, we see that $\pi_k r_n^\dagger = 0$ for all $k \geq n-1$ and

$$\begin{aligned}\pi_0 \text{Im } r_1^\dagger &= -\frac{\varepsilon}{\sqrt{2}} \left(W_+ - W_- + \frac{2}{\varepsilon^2} \right) \\ \pi_n \text{Im } r_{n+1}^\dagger &= \partial_X \pi_{n-1} \text{Im } r_n^\dagger,\end{aligned}$$

from which we deduce for all $n \geq 1$ that

$$\pi_{n-1} \text{Im } r_n^\dagger = -\frac{\varepsilon}{\sqrt{2}} \partial_X^{n-1} \left(W_+ - W_- + \frac{2}{\varepsilon^2} \right).$$

Similarly, we find that $\pi_k \text{Im } \sigma_n^\dagger = 0$ for all $k \geq n$ and

$$\begin{aligned}\pi_0 \text{Im } \sigma_0^\dagger &= -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-) \\ \pi_1 \text{Im } \sigma_1^\dagger &= \frac{\varepsilon}{2\sqrt{2}} \partial_X (W_+ - W_-) \\ \pi_{n+1} \text{Im } \sigma_{n+1}^\dagger &= \pi_n \partial_X \text{Im } \sigma_n^\dagger,\end{aligned}$$

which implies

$$\pi_n \operatorname{Im} \sigma_n^\dagger = \begin{cases} -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-) & , n = 0 \\ \frac{\varepsilon}{2\sqrt{2}} \partial_X^n (W_+ - W_-) & , n \geq 1 \end{cases}.$$

Lastly, we have $\pi_k \operatorname{Re} \sigma_n^\dagger = 0$ for $k \geq n$ and

$$\begin{aligned} \pi_0 \operatorname{Re} \sigma_1^\dagger &= -\frac{\varepsilon^2}{8}(W_+ - W_-)^2 + \frac{1}{2}(W_- - W_+) \\ \pi_n \operatorname{Re} \sigma_{n+1}^\dagger &= \partial_X \pi_{n-1} \operatorname{Re} \sigma_n^\dagger - \sum_{k=0}^n (\pi_k \operatorname{Im} \sigma_k^\dagger)(\pi_{n-k} \operatorname{Im} \sigma_{n-k}^\dagger) - \frac{\varepsilon}{\sqrt{2}} \partial_X W_+ \pi_{n-1} \operatorname{Im} r_n^\dagger \\ &= \partial_X \pi_{n-1} \operatorname{Re} \sigma_n^\dagger - \sum_{k=1}^{n-1} \frac{\varepsilon}{2\sqrt{2}} \partial_X^k (W_+ - W_-) \frac{\varepsilon}{2\sqrt{2}} \partial_X^{n-k} (W_+ - W_-) \\ &\quad + \frac{\varepsilon}{\sqrt{2}} (W_+ + W_-) \frac{\varepsilon}{2\sqrt{2}} \partial_X^n (W_+ - W_-) + \frac{\varepsilon}{\sqrt{2}} \partial_X W_+ \frac{\varepsilon}{\sqrt{2}} \partial_X^{n-1} \left(W_+ - W_- + \frac{2}{\varepsilon^2} \right) \\ &= \partial_X \pi_{n-1} \operatorname{Re} \sigma_n^\dagger + \frac{\varepsilon^2}{8} \left(- \sum_{k=1}^{n-1} \partial_X^k (W_+ - W_-) \partial_X^{n-k} (W_+ - W_-) \right. \\ &\quad \left. + 2(W_+ + W_-) \partial_X^n (W_+ - W_-) + 4\partial_X W_+ \partial_X^{n-1} (W_+ - W_-) \right) \\ &\quad + \mathbb{1}_{\{n=1\}} \partial_X W_+. \end{aligned}$$

This implies

$$\int_{\mathbb{R}} \pi_{2n} \operatorname{Re} \sigma_{2n+1}^\dagger dX = (-1)^{n+1} \frac{\varepsilon^2}{8} \int_{\mathbb{R}} (\partial_X^n (W_+ - W_-))^2 dX,$$

from which we obtain (1.12). So far we have shown that $\operatorname{Re} \sigma_{2n+1}^\dagger$ has up to $2n$ derivatives present in each monomial. Via repeated integration by parts, we can reduce the number of derivatives on the factor which has the most, and thereby ensure that each monomial in $\mathbf{P}_{2n}$ has no more than n derivatives on any factor. For $\operatorname{Im} \sigma_{2n}^\dagger$ we observe that the terms $\pi_{2n} \operatorname{Im} \sigma_{2n}^\dagger$ with the most derivatives vanish under integration by parts, so we obtain at most $n-1$ derivatives on any factor in $\mathbf{P}_{2n-1}$. As a result of these elaborations, we also know that $\mathbf{P}_n$ has at most $\lfloor \frac{n}{2} \rfloor$ derivatives in each monomial.

It remains to prove the absence of constant or linear monomials in $\mathbf{P}_n$ and $\mathbf{Q}_n$. This follows by direct computation for $n \in \{-1, 0\}$. For $n \geq 1$ we shall obtain it from the absence of constant or linear monomials in $\mathcal{H}_n^{\varepsilon \widetilde{\text{GP}}}$ and $\mathcal{E}_{n-1}^{\text{KdV}}$, which for the energies $\mathcal{E}_{n-1}^{\text{KdV}}$ follows directly from (2.48). If P is a polynomial as above, we denote by $\tau_0 P$ its constant part and by $\tau_1 P$ its linear part. First, we note that (2.42)–(2.44) directly imply $\tau_0 \sigma_n^\dagger = 0$. Since $\pi_k \tau_1 \sigma_n^\dagger$ vanishes under integration by parts for all $k \geq 1$, it suffices to determine $\pi_0 \tau_1 \sigma_n^\dagger$. Here (2.42)–(2.44) imply

$$\pi_0 \tau_1 \sigma_0^\dagger = \frac{c_{-1}}{2} (W_+ + W_-) \quad \pi_0 \tau_1 \sigma_1^\dagger = \frac{1}{2} (W_- - W_+) \quad \pi_0 \tau_1 \sigma_{n+1}^\dagger = \frac{c_n}{2} (W_+ + W_-).$$

Since $c_n = 0$ for even n it remains to prove the claim for the odd Hamiltonians

$$\mathcal{H}_{2n-1}^{\varepsilon \widetilde{\text{GP}}}(W) = \int_{\mathbb{R}} \sum_{m=0}^n \binom{-\frac{1}{2}}{n-m} 4^{n-m} (-1)^m (\sqrt{2}\varepsilon)^{2m} \operatorname{Im} \sigma_{2m}^{\varepsilon \text{GP}}(X; W) dX.$$

Here we have

$$\begin{aligned}\pi_0 \tau_1 \sum_{m=0}^n \binom{-\frac{1}{2}}{n-m} 4^{n-m} (-1)^m (\sqrt{2}\varepsilon)^{2m} \operatorname{Im} \sigma_{2m}^{\varepsilon \text{GP}} &= \sum_{m=0}^n \binom{-\frac{1}{2}}{n-m} 4^{n-m} (-1)^m (\sqrt{2}\varepsilon)^{2m} \frac{1}{2} \operatorname{Im} c_{2m-1}(W_+ + W_-) \\ &= -\mathbb{1}_{\{n=0\}} \frac{\varepsilon}{2\sqrt{2}} (W_+ + W_-).\end{aligned}$$

□

Proof of Corollary 1.5. Define for $n \in \mathbb{N}$ the energies

$$\tilde{E}^n(U) = \sum_{l=0}^n \binom{n}{l} \mathcal{E}_l^{\text{KdV}}(U).$$

According to [61, Theorem 9.1] there exists $\delta \in (0, 1)$ such that for every $n \in \mathbb{N}$ the smallness condition $\|U\|_{H^{-1}} \leq \delta$ implies

$$|\tilde{E}^n(U) - \|U\|_{H^n}^2| \leq c(n) \|U\|_{H^{-1}} \|U\|_{H^n}^2.$$

We calculate

$$\begin{aligned}\mathcal{E}_n^{\varepsilon \text{GP}, \pm}(W) &= \sum_{l=0}^n \binom{n}{l} (\sqrt{2}\varepsilon)^{-2l-3} (\mathcal{H}_{2l+2}^{\varepsilon \text{GP}}(W) \mp 2\mathcal{H}_{2l+1}^{\varepsilon \text{GP}}(W)) \\ &= \sum_{l=0}^n \binom{n}{l} \left(\mathcal{E}_l^{\text{KdV}}(\pm W_{\pm}) + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{P}_{2l+2} \mp \mathbf{P}_{2l+1} \, dX \right) \\ &= \sum_{l=0}^n \binom{n}{l} \left(\mathcal{E}_l^{\text{KdV}}(\pm W_{\pm}) + \frac{\varepsilon^2}{8} \|\partial_X^{l+1}(W_+ - W_-)\|_{L^2}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1} \, dX \right) \\ &= \tilde{E}^n(\pm W_{\pm}) + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)\|_{H^n}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \sum_{l=0}^n \binom{n}{l} (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1}) \, dX \\ &= \|W_{\pm}\|_{H^n}^2 + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)\|_{H^n}^2 \\ &\quad + \tilde{E}^n(\pm W_{\pm}) - \|W_{\pm}\|_{H^n}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \sum_{l=0}^n \binom{n}{l} (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1}) \, dX\end{aligned}$$

An analysis of the recurrence relations (2.42)–(2.44), (2.45)–(2.47) reveals that the polynomials $\mathbf{Q}_{2l+2}$ and $\mathbf{P}_{2l+1}$ for $0 \leq l \leq n$ have degree less than or equal to $n+3$, and at most n derivatives on a single factor. Furthermore, the total number of derivatives in any monomial is strictly less than $n+1$, and we know from Theorem 1.1 that they contain no constant or linear monomials. For $n \geq 1$ this allows us to combine the Cauchy–Schwarz inequality and the Sobolev embedding $H^1 \hookrightarrow L^\infty$ to estimate

$$\left| \int_{\mathbb{R}} \sum_{l=0}^n (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1}) \, dX \right| \leq C \|W\|_{H^n}^2 (1 + \|W\|_{H^n})^{n+1}.$$

By a similar analysis of $\mathcal{E}_n^{\text{KdV}}$ we find in particular that $\mathcal{H}_n^{\varepsilon \text{GP}}(W) < \infty$ if $W \in H^{\lfloor \frac{n}{2} \rfloor}$. □

3 Approximation of solutions to the $\varepsilon \widetilde{\mathbf{GP}}$ hierarchy by solutions to the KdV hierarchy

3.1 Proof of Proposition 1.10

Proposition 1.10 is an assembly of known results. Most importantly, [46, Theorem 1.1] by C. E. Kenig and D. Pilod states local well-posedness of $(\mathbf{KdV}_n)$. For any initial data, the local solution to the n -th flow $(\mathbf{KdV}_n)$ can be globalized by the standard argument of combining a blow-up alternative with the existence of conserved energies that control the Sobolev norm of the solution. Then the solution to the N -truncated KdV hierarchy is constructed by using this well-posedness result to move for a time t_0 along the flow of $(\mathbf{KdV}_0)$, then for a time t_1 along the flow of $(\mathbf{KdV}_1)$, and so on. Due to the commutativity of the flows, this is independent of the choice of ordering. The commutativity of the flows is a well-known result which can be proven, for example, by proceeding as in [67, Section 4], using the weighted Sobolev spaces $H^s \cap H^{s',1}$. Then by density the commutativity of the flows on H^s follows.

It remains to use certain conserved quantities to prove the uniform bound (1.17). While the energies $\mathcal{E}_n^{\mathbf{KdV}}$ are of course conserved, they do not have the desired coercivity property. A first result controlling the Sobolev norm of the solution by certain combinations of $\mathcal{E}_n^{\mathbf{KdV}}$ in the periodic case was given in [65, Theorem 3.1] by P. D. Lax. We shall use the recent results [56, Theorem 1.1] by R. Killip, M. Viřan, and X. Zhang, and [61, Theorem 9.1] by H. Koch and D. Tataru. The former result is only stated for Schwartz solutions, but the continuity of the local flow maps on Sobolev spaces allows us to approximate.

Proof of Proposition 1.10. Since the low regularity conserved quantities used in [56, Theorem 1.1] are conserved for every flow in the KdV hierarchy, we know that

$$\sup_{\mathbf{T} \in \mathbb{R}^{N+1}} \|U(\mathbf{T})\|_{H^{-1}} \leq C(1 + \|U_0\|_{H^{-1}}^2) \|U_0\|_{H^{-1}},$$

so it suffices to prove for every $T > 0$ that

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{s'}} \leq C(s') \sup_{|\mathbf{T}| \leq T} (1 + \|U(\mathbf{T})\|_{H^{-1}}^{2s'}) \|U_0\|_{H^{s'}}.$$

To show this we use [61, Theorem 9.1], which provides a smallness constant $\delta \in (0, 1)$ and for every $s' > -1$ a functional $\tilde{E}^{s'} : H^{s'} \rightarrow \mathbb{R}$ such that $\|U\|_{H^{-1}} \leq \delta$ implies

$$|\tilde{E}^{s'}(U) - \|U\|_{H^{s'}}^2| \leq c(s') \|U\|_{H^{-1}} \|U\|_{H^{s'}}^2.$$

Crucially, the functionals $\tilde{E}^{s'}$ are conserved along every flow of the KdV hierarchy. In particular, if

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{-1}} \leq \delta_0 = \min \left\{ \delta, \frac{1}{2c(s')} \right\},$$

then

$$\|U(\mathbf{T})\|_{H^{s'}}^2 \leq \tilde{E}^{s'}(U(\mathbf{T})) + c(s') \|U(\mathbf{T})\|_{H^{-1}} \|U(\mathbf{T})\|_{H^{s'}}^2 \leq \tilde{E}^{s'}(U(\mathbf{T})) + \frac{1}{2} \|U(\mathbf{T})\|_{H^{s'}}^2,$$

which implies

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{s'}}^2 \leq 2\tilde{E}^{s'}(U_0) \leq 2(\|U_0\|_{H^{s'}}^2 + c(s') \|U_0\|_{H^{-1}} \|U_0\|_{H^{s'}}^2) \leq 3\|U_0\|_{H^{s'}}^2.$$

We therefore assume $\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{-1}} > \delta_0$ and use the scaling invariance of the KdV hierarchy to regain the required smallness. Specifically, we define $u(\mathbf{T}, X) = \omega^2 U(\Omega \mathbf{T}, \omega X)$ where

$$\omega = \frac{\delta_0^2}{\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{-1}}^2} \in (0, 1) \text{ and } \Omega \in \mathbb{R}^{(N+1) \times (N+1)} \text{ with } \Omega_{n,m} = \mathbb{1}_{\{n=m\}} \omega^{2n+1}.$$

By the representation (3.1) below, u is still a solution to the KdV hierarchy in the sense of this theorem. It satisfies

$$\sup_{|\Omega \mathbf{T}| \leq T} \|u(\mathbf{T})\|_{H^{-1}}^2 = \sup_{|\Omega \mathbf{T}| \leq T} \int_{\mathbb{R}} \frac{\omega}{\frac{1}{\omega^2} + \xi^2} |\widehat{U}(\Omega \mathbf{T}, \xi)|^2 d\xi \leq \omega \sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{-1}}^2 = \delta_0^2.$$

The above reasoning now implies

$$\sup_{|\Omega \mathbf{T}| \leq T} \|u(\mathbf{T})\|_{H^{s'}}^2 \leq 3 \|u(\mathbf{0})\|_{H^{s'}}^2,$$

which after scaling yields

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{s'}}^2 \leq 3\omega^{-2s'} \|U_0\|_{H^{s'}}^2.$$

□

3.2 Proof of Proposition 1.11

Lemma 3.1 (Continuity argument). *For every $C_0 > 0$ and $h \in \mathbb{N}$ there exists some $\delta_0, \tilde{\varepsilon}_0, C_1 > 0$ such that the following holds. Let $r^* \in (0, \infty) \cup \{\infty\}$ and consider two functions $g_{\pm} \in C([0, r^*]; [0, \infty))$ that satisfy for some $\tilde{\varepsilon} \in (0, \tilde{\varepsilon}_0)$ and all $r \in [0, r^*)$*

$$|g_{\pm}(r) - g_{\pm}(0)| \leq C_0(g_{\pm}(r) + g_{\pm}(0))(g_+(r) + g_-(r))^{\frac{1}{2}} + \tilde{\varepsilon}^2 C_0(g_+(r) + g_-(r) + g_+(0) + g_-(0))(1 + g_+(r) + g_-(r) + g_+(0) + g_-(0))^{h+1}.$$

Then $g_+(0) + g_-(0) < \delta_0$ implies

$$g_{\pm}(r) \leq \begin{cases} 3g_{\pm}(0) & , g_{\pm}(0) > 0 \\ 3g_{\mp}(0) & , g_{\pm}(0) = 0 \end{cases} + \tilde{\varepsilon}^2 C_1(g_+(0) + g_-(0))$$

for all $r \in [0, r^*)$.

Proof. Set $E = \tilde{\varepsilon}^2 C_1(g_+(0) + g_-(0))$. We prove by contradiction, so we assume that the estimate fails before r^* . More specifically, we may assume without loss of generality that the estimate for g_+ becomes saturated at some time $r_+ \in (0, r^*)$, and holds for both g_+ and g_- on the interval $(0, r_+)$. This means either $g_+(r_+) = 3g_+(0) + E$ and $g_+(0) > 0$, or $g_+(r_+) = 3g_-(0) + E$ and $g_+(0) = 0$, and furthermore for all $r \in (0, r_+)$ we have $g_{\pm}(r) \leq 3g_{\pm}(0) + E$ if $g_{\pm}(0) > 0$, and $g_{\pm}(r) \leq 3g_{\mp}(0) + E$ if $g_{\pm}(0) = 0$. In the case $g_+(0) > 0$ this implies

$$2g_+(0) + E \leq C_0(8g_+(0) + E)(6g_+(0) + 6g_-(0) + 2E)^{\frac{1}{2}} + \tilde{\varepsilon}^2 C_0(8g_+(0) + 8g_-(0) + 2E)(1 + 8g_+(0) + 8g_-(0) + 2E)^{h+1},$$

which leads to a contradiction if δ_0 and $\tilde{\varepsilon}_0$ are sufficiently small and $g_+(0) + g_-(0) < \delta_0$. The next case is $g_+(0) = 0$ and $g_-(0) > 0$. Here we have

$$g_+(r) \leq C_0 g_+(r)(g_+(r) + g_-(r))^{\frac{1}{2}} + \tilde{\varepsilon}^2 C_0(g_+(r) + g_-(r) + g_-(0))(1 + g_+(r) + g_-(r) + g_-(0))^{h+1}.$$

and $\max\{g_+(r), g_-(r)\} \leq 3g_-(0) + E$ for all $r \in (0, r_+)$. In particular

$$3g_-(0) + E = g_+(r_+) = \lim_{r \nearrow r_+} g_+(r) \leq C_0(3g_-(0) + E)(6g_-(0) + 2E)^{\frac{1}{2}} + \tilde{\varepsilon}^2 C_0(8g_-(0) + E)(1 + 8g_-(0) + E)^{h+1},$$

which again leads to a contradiction for sufficiently small δ_0 and $\tilde{\varepsilon}_0$. The final case is $g_+(0) = g_-(0) = 0$, where

$$g_{\pm}(r) \leq C_0 g_{\pm}(r) (g_+(r) + g_-(r))^{\frac{1}{2}} + \tilde{\varepsilon}^2 C_0 (g_+(r) + g_-(r)) (1 + g_+(r) + g_-(r))^{h+1}.$$

If either g_+ or g_- are ever positive on $[0, r^*)$, then they assume arbitrarily small values for some $r \in (0, r^*)$, which is not compatible with the estimate when δ_0 and $\tilde{\varepsilon}_0$ are sufficiently small. $\square$

We can now prove the proposition.

Proof of Proposition 1.11. This proposition is [67, Corollary 1.7] with the addition of (iv), which we now prove. A lower bound for the distance to vacuum using the energy is given, for example, in [11, Lemma 1] and [83, Lemma 1.2]. Then continuity and boundedness of W in H^{s-1} follow from the local bilipschitz property of the Madelung transform $q \mapsto \left(\left(1 + \frac{w_+ - w_-}{2}\right)^2, w_+ + w_- \right)$ established in [83, Corollary 1.7, Lemma 2.4] together with $q \in C(\mathbb{R}^{N+1}; X^s)$. For a detailed proof of the fact that $(\varepsilon \mathbf{GP}_n)$ is solved in the sense of distributions in the case $n = 2$, we refer to [83, Section 3]. The cases $n \geq 3$ are analogous.

In fact the solutions are strong in the sense that $T_n \mapsto W(\mathbf{c}^{\text{NLS} \mapsto \varepsilon \widetilde{\text{GP}}} \mathbf{T}) \in C^1(H^{s-2} \lfloor \frac{n}{2} \rfloor - 2)$ satisfies $(\varepsilon \widetilde{\text{GP}}_n)$ in that space. We obtain this by transferring the differentiability of the solutions from (iii) of Proposition 1.11 through the Madelung transform. Away from vacuum, this boils down to combining product estimates and the boundedness of certain smooth nonlinear mappings on Sobolev spaces, which can be performed as in [83]. Note that the highest order of derivative appearing in $(\varepsilon \widetilde{\text{GP}}_n)$ is $2 \lfloor \frac{n}{2} \rfloor + 1$.

It remains to prove (1.23), so we let $h \in \{1, \dots, s-2\}$. Here we use Corollary 1.5. In order to obtain an estimate for $\|W_{\pm}\|_{H^h}$ from this bound we need smallness of $\|W_{\pm}\|_{H^{-1}}$. We shall obtain smallness for $\|W\|_{H^{h+1}}$ globally in time by slightly relaxing our scaling. Specifically, we define $\tilde{\mathbf{T}}$, $\tilde{X}$, and $\tilde{W}$ in analogy to $\mathbf{T}$, X , and W , with ε replaced by

$$\tilde{\varepsilon} = \varepsilon \max\{1, \delta^{-\frac{1}{3}} \|W(\mathbf{0})\|_{H^{h+1}}^{\frac{2}{3}}\}$$

for some $\delta = \delta(h) > 0$ to be chosen later. Then $\tilde{W}$ has the same properties as described above, except ε is replaced by $\tilde{\varepsilon}$. Since

$$\tilde{W}(\tilde{\mathbf{T}}, \tilde{X}) = \frac{1}{\tilde{\varepsilon}^2} w(\mathbf{t}, x) = \left(\frac{\varepsilon}{\tilde{\varepsilon}}\right)^2 \frac{1}{\varepsilon^2} w(\mathbf{t}, x) = \left(\frac{\varepsilon}{\tilde{\varepsilon}}\right)^2 W(\mathbf{T}, X) = \left(\frac{\varepsilon}{\tilde{\varepsilon}}\right)^2 W\left(\mathbf{T}, \frac{\varepsilon}{\tilde{\varepsilon}} \tilde{X}\right)$$

we have for some constant $C = C(h) > 0$

$$C^{-1} \left(\frac{\varepsilon}{\tilde{\varepsilon}}\right)^{\frac{3}{2}+h} \|W(\mathbf{T})\|_{H^h} \leq \|\tilde{W}(\tilde{\mathbf{T}})\|_{H^h} \leq C \left(\frac{\varepsilon}{\tilde{\varepsilon}}\right)^{\frac{3}{2}} \|W(\mathbf{T})\|_{H^h},$$

in particular $\|\tilde{W}(\mathbf{0})\|_{H^{h+1}}^2 < \delta$. By Corollary 1.5 there exists a choice of δ and $C = C(h) > 0$ so that $\|\tilde{W}\|_{H^{-1}}^2 < 4\delta$ implies

$$\left| \mathcal{E}_h^{\varepsilon \widetilde{\text{GP}}, \pm}(\tilde{W}) - \|\tilde{W}_{\pm}\|_{H^h}^2 - \frac{\tilde{\varepsilon}^2}{8} \|\partial_X(\tilde{W}_+ - \tilde{W}_-)\|_{H^h}^2 \right| \leq C \|\tilde{W}_{\pm}\|_{H^h}^2 \|\tilde{W}_{\pm}\|_{H^{-1}} + \tilde{\varepsilon}^2 C \|\tilde{W}\|_{H^h}^2 (1 + \|\tilde{W}\|_{H^h})^{h+1}.$$

We use this to estimate

$$\begin{aligned} & \left| \|\tilde{W}_{\pm}(\tilde{\mathbf{T}})\|_{H^h}^2 + \frac{\tilde{\varepsilon}^2}{8} \|\partial_X(\tilde{W}_+ - \tilde{W}_-)(\tilde{\mathbf{T}})\|_{H^h}^2 - \|\tilde{W}_{\pm}(\mathbf{0})\|_{H^h}^2 - \frac{\tilde{\varepsilon}^2}{8} \|\partial_X(\tilde{W}_+ - \tilde{W}_-)(\mathbf{0})\|_{H^h}^2 \right| \\ & \leq C \left(\|\tilde{W}_{\pm}(\tilde{\mathbf{T}})\|_{H^h}^2 \|\tilde{W}_{\pm}(\tilde{\mathbf{T}})\|_{H^{-1}} + \|\tilde{W}_{\pm}(\mathbf{0})\|_{H^h}^2 \|\tilde{W}_{\pm}(\mathbf{0})\|_{H^{-1}} \right) \\ & \quad + \tilde{\varepsilon}^2 C \left(\|\tilde{W}(\tilde{\mathbf{T}})\|_{H^h}^2 (1 + \|\tilde{W}(\tilde{\mathbf{T}})\|_{H^h})^{h+1} + \|\tilde{W}(\mathbf{0})\|_{H^h}^2 (1 + \|\tilde{W}(\mathbf{0})\|_{H^h})^{h+1} \right) \end{aligned}$$

as long as $\|\widetilde{W}(\mathbf{0})\|_{H^{-1}}^2, \|\widetilde{W}(\widetilde{\mathbf{T}})\|_{H^{-1}}^2 < 4\delta$. Define

$$g_{\pm}(r) = \sup_{\{|\widetilde{\mathbf{T}}| \leq r\}} \left(\|\widetilde{W}_{\pm}(\widetilde{\mathbf{T}})\|_{H^h}^2 + \frac{\widetilde{\varepsilon}^2}{8} \|\partial_X(\widetilde{W}_+ - \widetilde{W}_-)(\widetilde{\mathbf{T}})\|_{H^h}^2 \right)$$

$$r^* = \inf\{r > 0 : g_+(r) + g_-(r) \geq 4\delta\}$$

and note that $g_{\pm}$ is continuous.

By choosing $\varepsilon < \varepsilon_0$ for some appropriate $\varepsilon_0 = \varepsilon_0(\delta, \|W(\mathbf{0})\|_{H^{h+1}})$, we can ensure that $\widetilde{\varepsilon} \in (\varepsilon, 1)$. Then the above inequality implies for $r \in [0, r^*)$ that for some $C_0 = C_0(h)$ we have

$$|g_{\pm}(r) - g_{\pm}(0)| \leq C_0(g_{\pm}(r) + g_{\pm}(0))(g_+(r) + g_-(r))^{\frac{1}{2}}$$

$$+ \widetilde{\varepsilon}^2 C_0(g_+(r) + g_-(r) + g_+(0) + g_-(0))(1 + g_+(r) + g_-(r) + g_+(0) + g_-(0))^{h+1}.$$

By Lemma 3.1 there exists some $\delta_0 = \delta_0(h) > 0$, $\widetilde{\varepsilon}_0 = \widetilde{\varepsilon}_0(h) > 0$, and $C_1 = C_1(h) > 0$ such that $g_+(0) + g_-(0) < \delta_0$ and $\widetilde{\varepsilon} < \widetilde{\varepsilon}_0$ imply

$$g_{\pm}(r) \leq \begin{cases} 3g_{\pm}(0) & , g_{\pm}(0) > 0 \\ 3g_{\mp}(0) & , g_{\pm}(0) = 0 \end{cases} + \widetilde{\varepsilon}^2 C_1(g_+(0) + g_-(0))$$

for all $r \in [0, r^*)$. Since $g_{\pm}(0) \lesssim \|\widetilde{W}(\mathbf{0})\|_{H^{h+1}}^2$ we can obtain $g_+(0) + g_-(0) < \min\{\delta_0, \delta\}$ and also $\widetilde{\varepsilon} < \widetilde{\varepsilon}_0$ from an appropriately small choice of $\varepsilon_0 = \varepsilon_0(h, \delta, \delta_0, \|W(\mathbf{0})\|_{H^{h+1}})$. Then the estimate just proved implies $g_+(r) + g_-(r) < 4\delta$, so $r^* = \infty$.

If $g_{\pm}(0) = 0$ then $(\widetilde{W}_+ - \widetilde{W}_-)(\mathbf{0})$ is constant in $\widetilde{X}$, hence $W_+(\mathbf{0}) = W_-(\mathbf{0})$. Therefore $g_{\pm}(0) = 0$ implies $W(\mathbf{0}) = 0$, which is a trivial case. If $g_{\pm}(0) > 0$ then we finally obtain (1.23). $\square$

3.3 Proof of Theorem 1.12

We are now ready to give the proof of our main result. To simplify the notation, we set $\mathbf{c} = \mathbf{c}^{\text{NLS} \mapsto \varepsilon \widetilde{\text{GP}}}$.

Proof of Theorem 1.12. Let $\mathbf{T} \in \mathbb{R}^{N+1}$ and recall that $\mathbf{T} = (T_0, \dots, T_N)$. We shall use an energy inequality with Grönwall's lemma. Note, however, that we do not wish to apply Grönwall's lemma several times, i. e. for each flow in the hierarchy, as this would require reusing solutions obtained by our well-posedness theory as initial data and lead to worse estimates. Instead, we shall apply Grönwall's lemma only once.

Continuous path from $\mathbf{0}$ to $\mathbf{T}$. Let $n \in \{0, \dots, N\}$ and define $m = m(n) = \lfloor \frac{n-1}{2} \rfloor \in \{-1, \dots, M\}$ where $M = \lfloor \frac{N-1}{2} \rfloor$. In order to simplify the argument below we assume that $|T_0|, \dots, |T_N| > 0$. The degenerate cases only require minor modifications. Define $\tau_n = |T_0| + \dots + |T_n|$ as well as $\tau_{-1} = 0$ and set

$$(\mathbf{T}^\tau)_n = \begin{cases} 0 & , \tau \in [0, \tau_{n-1}) \\ \frac{\tau - \tau_{n-1}}{\tau_n - \tau_{n-1}} T_n & , \tau \in [\tau_{n-1}, \tau_n) \\ T_n & , \tau \in [\tau_n, \infty) \end{cases}.$$

Then $(\mathbf{T}^\tau)_{\tau \in [0, \infty)}$ is a continuous path with $\mathbf{T}^0 = \mathbf{0}$ and $(\mathbf{T}^\tau)_n = T_n$ if $\tau \geq \tau_n$. In particular $\mathbf{T}^{\tau_N} = \mathbf{T}$.

We use Theorem 1.8 to write

$$\begin{aligned}
W(\mathbf{cT}^\tau) &= W(\mathbf{0}) + \int_0^{\tau_N \wedge \tau} \frac{d}{d\tau'} W(\mathbf{cT}^{\tau'}) d\tau' \\
&= W(\mathbf{0}) + \sum_{n=0}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \frac{d}{d(\mathbf{T}^{\tau'})_n} W(\mathbf{cT}^{\tau'}) d\tau' \\
&= W(\mathbf{0}) + \sum_{n=0}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \left(\begin{aligned} &(-1)^n \frac{1}{2} \partial_X \frac{\delta}{\delta W_+} \mathcal{E}_m^{\text{KdV}}(W_+(\mathbf{cT}^{\tau'})) \\ & - \frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_m^{\text{KdV}}(-W_-(\mathbf{cT}^{\tau'})) \end{aligned} \right) + \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'}) d\tau'.
\end{aligned}$$

Similarly, we can write

$$\begin{aligned}
\begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}^\tau) \\ U_-(\mathbf{b}^- \mathbf{T}^\tau) \end{pmatrix} &= U(\mathbf{0}) + \int_0^{\tau_N \wedge \tau} \frac{d}{d\tau'} \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}^{\tau'}) \\ U_-(\mathbf{b}^- \mathbf{T}^{\tau'}) \end{pmatrix} d\tau' \\
&= U(\mathbf{0}) + \sum_{n=0}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \frac{d}{d(\mathbf{T}^{\tau'})_n} \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}^{\tau'}) \\ U_-(\mathbf{b}^- \mathbf{T}^{\tau'}) \end{pmatrix} d\tau' \\
&= U(\mathbf{0}) + \sum_{n=0}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \left(\begin{aligned} &(-1)^n \frac{1}{2} \partial_X \frac{\delta}{\delta U_+} \mathcal{E}_m^{\text{KdV}}(U_+(\mathbf{b}^+ \mathbf{T}^{\tau'})) \\ & - \frac{1}{2} \partial_X \frac{\delta}{\delta(-U_-)} \mathcal{E}_m^{\text{KdV}}(-U_-(\mathbf{b}^- \mathbf{T}^{\tau'})) \end{aligned} \right) d\tau'.
\end{aligned}$$

Energy estimate. Define

$$V(\mathbf{T}) = \begin{pmatrix} V_+(\mathbf{T}) \\ V_-(\mathbf{T}) \end{pmatrix} = \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{T}) \end{pmatrix} - W(\mathbf{cT}).$$

Since $W(\mathbf{0}) = U(\mathbf{0})$ and $\mathbf{R}_0 = \mathbf{R}_1 = 0$, we know that $V(\mathbf{T}^\tau) = 0$ for $\tau \in [0, \tau_1]$. We therefore have

$$V(\mathbf{T}^\tau) = \sum_{n=2}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} (\mathbf{F}_n(\tau') - \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'})) d\tau'.$$

with

$$\mathbf{F}_n(\tau) = \begin{pmatrix} -(-1)^n \frac{1}{2} \partial_X \frac{\delta}{\delta W_+} (\mathcal{E}_m^{\text{KdV}}(W_+(\mathbf{cT}^\tau) + V_+(\mathbf{T}^\tau)) - \mathcal{E}_m^{\text{KdV}}(W_+(\mathbf{cT}^\tau))) \\ -\frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} (\mathcal{E}_m^{\text{KdV}}(-W_-(\mathbf{cT}^\tau) - V_-(\mathbf{T}^\tau)) - \mathcal{E}_m^{\text{KdV}}(-W_-(\mathbf{cT}^\tau))) \end{pmatrix} \text{ for } \tau \in [\tau_{n-1}, \tau_n].$$

In particular, for all $h \in \{0, \dots, \lfloor \frac{s-N}{2} \rfloor\}$ we have

$$\int_{\mathbb{R}} (\partial_X^h V(\mathbf{T}^\tau))^2 dX = \sum_{n=2}^N \text{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \int_{\mathbb{R}} 2\partial_X^h V(\mathbf{T}^{\tau'}) \partial_X^h (\mathbf{F}_n(\tau') - \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'})) dX d\tau'.$$

For notational convenience, let us define for a set of functions F in the variable X and parameters $K, L \in \mathbb{N}$ the sets

$$\mathcal{O}^{K,L}(F) = \left\{ \sum_{k=0}^K \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l| \leq L}} \sum_{f=(f_1, \dots, f_k) \in F^k} c_f^{k,l} \prod_{j=1}^k \partial_X^{l_j} f_j : c_f^{k,l} \in \mathbb{C}[\varepsilon] \right\}.$$

Studying the recurrence relations (2.42)–(2.44), (2.45)–(2.47), as we do in the proof of Theorem 1.1 at the end of Section 2, reveals that $\frac{\delta}{\delta W_\pm} \mathcal{H}_n^{\varepsilon \widetilde{\text{GP}}}(W) \in \mathcal{O}^{n+1, 2\lfloor \frac{n}{2} \rfloor}(W_+, W_-)$. A similar analysis of (2.48) yields $\frac{\delta}{\delta U} \mathcal{E}_m^{\text{KdV}}(U) \in \mathcal{O}^{m+1, 2m}(U)$. Then

$$\mathbf{R}_n \in \mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor+1}(W_+, W_-) + \frac{\mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor+1}(W_+, W_-)}{1 + \frac{\varepsilon^2}{2}(W_+ - W_-)} + \frac{\mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor+1}(W_+, W_-)}{(1 + \frac{\varepsilon^2}{2}(W_+ - W_-))^2}.$$

We use Sobolev embedding $H^1 \hookrightarrow L^\infty$ and the H^h -algebra property for $h \geq 1$ as well as [6, Theorem 2.87] to find a constant $C(n, h, \delta) > 0$ such that

$$\|\mathbf{R}_n\|_{H^h} \leq C(n, h, \delta)(1 + \|W\|_{H^{n+1+h}}^{n+4}).$$

Then

$$\int_{\mathbb{R}} \partial_X^h V \varepsilon^2 \partial_X^h \mathbf{R}_n \, dX \leq \mathbb{1}_{\{n \geq 2\}} (\|V\|_{H^h}^2 + \varepsilon^4 C(N, h, \delta)^2 (1 + \|W\|_{H^{n+1+h}}^{2n+8})).$$

We are able to include the indicator function because $\mathbf{R}_0 = \mathbf{R}_1 = 0$. A more detailed analysis of (2.48) yields the representation

$$\frac{1}{2} \partial_X \frac{\delta}{\delta U} \mathcal{E}_m^{\text{KdV}} = (-1)^m \partial_X^{2m+1} U + \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l|+2k=2m+3}} c^{k,l} \prod_{j=1}^k \partial_X^{l_j} U \quad (3.1)$$

with coefficients $c^{k,l} \in \mathbb{C}$. Then

$$\begin{aligned} & \partial_X \frac{\delta}{\delta(\pm W_{\pm})} (\mathcal{E}_m^{\text{KdV}}(\pm W_{\pm} \pm V_{\pm}) - \mathcal{E}_m^{\text{KdV}}(\pm W_{\pm})) \\ &= \pm (-1)^m \partial_X^{2m+1} V_{\pm} + \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l| \leq 2m-1}} c^{k,l} \sum_{i=1}^k \partial_X^{l_i}(\pm V_{\pm}) \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_{\pm}) \prod_{j=i+1}^k \partial_X^{l_j}(\pm W_{\pm} \pm V_{\pm}). \end{aligned}$$

Using integration by parts and $\partial_X^{m+h} V_{\pm} \in H^1(\mathbb{R})$, we find that

$$\int_{\mathbb{R}} \partial_X^h V_{\pm} \partial_X^h \partial_X^{2m+1} V_{\pm} \, dX = \int_{\mathbb{R}} (-1)^m \partial_X \left(\frac{1}{2} (\partial_X^{m+h} V_{\pm})^2 \right) \, dX = 0.$$

Similarly, using repeated integration by parts and the Sobolev embedding $H^1 \hookrightarrow L^\infty$, we estimate

$$\begin{aligned} & \left| \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l| \leq 2m-1}} c^{k,l} \sum_{i=1}^k \int_{\mathbb{R}} \partial_X^h V_{\pm} \partial_X^h \left(\partial_X^{l_i} V_{\pm} \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_{\pm}) \prod_{j=i+1}^k \partial_X^{l_j}(\pm U_{\pm}) \right) \, dX \right| \\ & \leq C^{m+1} \sum_{r=0}^h \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l| \leq 2m-1+2(h-r)}} \sum_{i=1}^k \int_{\mathbb{R}} (\partial_X^r V_{\pm})^2 \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_{\pm}) \prod_{j=i+1}^k \partial_X^{l_j}(\pm U_{\pm}) \, dX \\ & \leq C^{m+1} \sum_{r=0}^h \int_{\mathbb{R}} (\partial_X^r V_{\pm})^2 \, dX \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1, \dots, l_k) \in \mathbb{N}^k \\ |l| \leq 2m-1+2(h-r)}} \sum_{i=1}^k \prod_{\substack{j=1 \\ j \neq i}}^k (\|\partial_X^{l_j} W_{\pm}\|_{L^\infty} + \|\partial_X^{l_j} U_{\pm}\|_{L^\infty}) \\ & \leq C^{m+1} \sum_{r=0}^h \int_{\mathbb{R}} (\partial_X^r V_{\pm})^2 \, dX (1 + \|(W_{\pm}, U_{\pm})\|_{H^{2(m+h-r)}}^m). \end{aligned}$$

Grönwall's lemma for linear systems. In summary, for all $l \in \{0, \dots, h\}$ and $\tau \in [0, \tau_N]$ we have

$$\int_{\mathbb{R}} (\partial_X^l V(\mathbf{T}^\tau))^2 \, dX \leq \mathbb{1}_{\{\tau \geq \tau_1\}} C(N, h, \delta) \left(a_l + \int_0^\tau \sum_{r=0}^h H_{l,r} \int_{\mathbb{R}} (\partial_X^r V(\mathbf{T}^{\tau'}))^2 \, dX \, d\tau' \right),$$

where we consider the real numbers

$$a_l = (\tau_N - \tau_1) \varepsilon^4 (1 + \|W\|_{L_{\mathbf{T}}^\infty H_X^{N+1+l}}^{2N+8})$$

$$H_{l,r} = \mathbb{1}_{r \leq l} (1 + \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2(M+l-r)}}^M)$$

as the entries of a vector $a \in \mathbb{R}^{h+1}$ and a matrix $H \in \mathbb{R}^{(h+1) \times (h+1)}$. Grönwall's lemma for linear systems (see [18, Theorem]) yields

$$\sup_{\tau \in [0, \tau_N]} \int_{\mathbb{R}} (\partial_X^h V(\mathbf{T}^\tau))^2 dX \leq C \left(\left(1 + \int_0^{\tau_N - \tau_1} e^{C(\tau_N - \tau_1 - \tau)H} H d\tau \right) a \right)_h$$

$$= C(e^{C(\tau_N - \tau_1)H} a)_h$$

for some constant $C = C(N, h, \delta) > 0$. It remains to estimate $(e^{C(\tau_N - \tau_1)H} a)_h$. We define some notation for the part $\tilde{H}$ of H which does not lie on the diagonal:

$$\tilde{H} = H - \left(1 + \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2M}}^M \right)$$

$$\tilde{H}_{l,r} = \mathbb{1}_{\{0 \leq r < l\}} g_{l-r} \quad \text{where} \quad g_j = 1 + \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2(M+j)}}^M.$$

Then since $\tilde{H}$ and $H - \tilde{H}$ commute, we have

$$|(e^{(\tau_N - \tau_1)H} a)_h| \leq |(e^{(\tau_N - \tau_1)\tilde{H}} a)_h| e^{(\tau_N - \tau_1)(1 + \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2M}}^M)}.$$

Next, we estimate for $l \geq 1$

$$|(\tilde{H}^l)_{h,r}| = \sum_{h=j_0 > \dots > j_l=r} g_{j_0-j_1} \cdots g_{j_{l-1}-j_l} \leq \binom{h-r-1}{l-1} g_{h-r}^l$$

and then

$$(e^{(\tau_N - \tau_1)\tilde{H}} a)_h = \sum_{l=0}^{\infty} \frac{(\tau_N - \tau_1)^l}{l!} (\tilde{H}^l a)_h = \sum_{l=0}^{\infty} \frac{(\tau_N - \tau_1)^l}{l!} \sum_{r=0}^h (\tilde{H}^l)_{h,r} a_r \leq (h+1) a_h (1 + (\tau_N - \tau_1) g_h)^h.$$

This yields

$$|(e^{(\tau_N - \tau_1)H} a)_h| \leq \varepsilon^4 (\tau_N - \tau_1) C (1 + \|W\|_{L_{\mathbf{T}}^\infty H_X^{N+1+h}}^{2N+8}) (1 + (\tau_N - \tau_1) \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2(M+h)}}^M)^h$$

$$\times \exp \left(C(\tau_N - \tau_1) (1 + \|(W, U)\|_{L_{\mathbf{T}}^\infty H_X^{2M}}^M) \right).$$

□

A Approximation of the perturbation determinant $\alpha^{\varepsilon \mathbf{GP}}$ by $\alpha^{\mathbf{KdV}}$

A.1 Definitions

Let $n \in \{1, 2\}$. We use the convention of identifying functions $f \in L^2(\mathbb{R}; \mathbb{C}^n)$ with the operator of multiplication by f . We write $(L(L^2(\mathbb{R}; \mathbb{C}^n)), \|\cdot\|_{L^2 \rightarrow L^2})$ for the space of bounded linear operators on

$L^2(\mathbb{R}; \mathbb{C}^n)$ equipped with the operator norm, and say that an operator $A \in L(L^2(\mathbb{R}; \mathbb{C}^n))$ is Hilbert-Schmidt if there exists a kernel $K \in L^2(\mathbb{R} \times \mathbb{R}; \mathbb{C}^{n \times n})$ such that

$$Af(X) = \int_{\mathbb{R}} K(X, Y) f(Y) dY.$$

In this case we define the Hilbert-Schmidt norm as

$$\|A\|_{\text{HS}}^2 = \int_{\mathbb{R}^2} |K(X, Y)|^2 dX dY.$$

We say that an operator $A \in L(L^2(\mathbb{R}; \mathbb{C}^n))$ is trace-class if $A = A_1 A_2$ for two Hilbert-Schmidt operators A_1 and A_2 . Writing K_1 and K_2 for their kernels, we define the trace

$$\text{tr}[A] = \text{tr}[A_1 A_2] = \int_{\mathbb{R}} \int_{\mathbb{R}} \text{tr}[K_1(X, Y) K_2(Y, X)] dX dY.$$

In particular

$$\|A\|_{\text{HS}}^2 = \text{tr}[A^* A].$$

If the trace $\text{tr}[K]$ of the kernel K of A is continuous, then

$$\text{tr}[A] = \int_{\mathbb{R}} \text{tr}[K(X, X)] dX.$$

The following lemma collects some results regarding the trace and the Hilbert-Schmidt norm.

Lemma A.1 ([56, Lemmas 1.4, 1.5]). *Let $A, A_j : L^2(\mathbb{R}; \mathbb{C}^n) \rightarrow L^2(\mathbb{R}; \mathbb{C}^n)$, $1 \leq j \leq l$, $2 \leq l \in \mathbb{N}$ be linear operators with finite Hilbert-Schmidt norm. Then*

$$\|A\|_{L^2 \rightarrow L^2} \leq \|A\|_{\text{HS}} \tag{A.1}$$

$$|\text{tr}[A_1 \dots A_l]| \leq \|A_1\|_{\text{HS}} \dots \|A_l\|_{\text{HS}}. \tag{A.2}$$

If $A = A(s) \in C_b^1(\mathbb{R}; \text{HS})$ with $\sup_{s \in \mathbb{R}} \|A(s)\|_{\text{HS}} < \frac{1}{3}$, then there exists an open interval $I \ni 0$ on which the series

$$\alpha(s) = \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[A(s)^l]$$

is absolutely convergent and defines a function $\alpha \in C^1(I; \mathbb{C})$ with derivative

$$\partial_s \alpha(s) = \sum_{l=2}^{\infty} (-1)^l \text{tr}[A(s)^{l-1} \partial_s A(s)]. \tag{A.3}$$

We have the bound

$$\sup_{s \in I} |\alpha(s)| \leq \frac{2}{3} \sup_{s \in I} \|A(s)\|_{\text{HS}}^2.$$

Recall from (2.1) the spectral parameters $(\lambda, z, k) \in \mathbb{K} \subset \mathbb{C}^3$ which satisfy $z^2 = (\sqrt{2\varepsilon}k)^2 = \lambda^2 - 1$. We assume for the rest of this section that $z, k \in i\mathbb{R}_+$ are large and choose $\lambda = \sqrt{1 + 2\varepsilon^2 k^2}$ as the principal square root. We define

$$\kappa = \frac{z}{\sqrt{2\varepsilon}i} = \frac{k}{i} \in \mathbb{R}_+ \qquad \kappa^2 = \frac{1 - \lambda^2}{2\varepsilon^2}.$$

For notational convenience, we use the κ -dependent Sobolev norms

$$\begin{aligned}\|f\|_{H_\kappa^{-1}} &= \left(\int_{\mathbb{R}} \frac{1}{\xi^2 + 4\kappa^2} |\widehat{f}(\xi)|^2 d\xi \right)^{\frac{1}{2}} \leq \min \left\{ \frac{1}{\min\{1, 2\kappa\}} \|f\|_{H^{-1}}, \frac{1}{2\kappa} \|f\|_{L^2} \right\} \\ \|f\|_{L_\kappa^2} &= \left(\int_{\mathbb{R}} \frac{\xi^2 + 2\kappa^2}{\xi^2 + 4\kappa^2} |\widehat{f}(\xi)|^2 d\xi \right)^{\frac{1}{2}} \leq \|f\|_{L^2}.\end{aligned}$$

Here the Fourier transform is defined by

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x) dx.$$

We assume $W = (W_+, W_-) \in L^2(\mathbb{R}; \mathbb{R}^2)$ or $W \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$ depending on the context. We are interested in the Lax operators

$$\begin{aligned}L^{\varepsilon\text{GP}}(W) &= \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon \partial_X \\ i\sqrt{2}\varepsilon \partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix} & L^{\text{KdV}}(\pm W_\pm) &= -\partial_X^2 \pm W_\pm \\ L_0^{\varepsilon\text{GP}} &= \begin{pmatrix} 1 & i\sqrt{2}\varepsilon \partial_X \\ i\sqrt{2}\varepsilon \partial_X & -1 \end{pmatrix} & L_0^{\text{KdV}} &= -\partial_X^2\end{aligned}$$

and their resolvents

$$\begin{aligned}R^{\varepsilon\text{GP}}(W) &= \left(\frac{L^{\varepsilon\text{GP}}(W) - \lambda}{\varepsilon^2} \right)^{-1} & R^{\text{KdV}}(\pm W_\pm) &= (L^{\text{KdV}}(\pm W_\pm) + \kappa^2)^{-1} \\ R_0^{\varepsilon\text{GP}} &= \left(\frac{L_0^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2} \right)^{-1} & R_0^{\text{KdV}} &= (L_0^{\text{KdV}} + \kappa^2)^{-1}.\end{aligned}$$

Define

$$\mathcal{W} = \begin{pmatrix} -W_- & 0 \\ 0 & -W_+ \end{pmatrix}$$

and note that $L^{\varepsilon\text{GP}}(W) = L_0^{\varepsilon\text{GP}} + \varepsilon^2 \mathcal{W}$ and $L^{\text{KdV}}(\pm W_\pm) = L_0^{\text{KdV}} \pm W_\pm$. In the following, we often drop the dependence of the Lax operators and resolvents on W and $(\pm W_\pm)$ from our notation. With

$$\mathcal{U} = \begin{pmatrix} \frac{\lambda+1}{2} & \frac{i\varepsilon}{\sqrt{2}} \partial_X \\ \frac{i\varepsilon}{\sqrt{2}} \partial_X & \frac{\lambda-1}{2} \end{pmatrix}$$

we can write

$$R_0^{\varepsilon\text{GP}} = R_0^{\text{KdV}} \mathcal{U} = \mathcal{U} R_0^{\text{KdV}} = \sqrt{R_0^{\text{KdV}}} \mathcal{U} \sqrt{R_0^{\text{KdV}}}.$$

Next, we use the Green's functions

$$\begin{aligned}G_0^{\text{KdV}} &= \frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1} \left[R_0^{\text{KdV}} \Big|_{\partial_X = i\xi} \right] = \frac{1}{2\kappa} e^{-\kappa|X|} \\ G_0^{\varepsilon\text{GP}} &= \frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1} \left[R_0^{\varepsilon\text{GP}} \Big|_{\partial_X = i\xi} \right] = \frac{1}{2\kappa} e^{-\kappa|X|} \begin{pmatrix} \frac{\lambda+1}{2} & -\frac{i\varepsilon}{\sqrt{2}} \kappa \operatorname{sign}(X) \\ -\frac{i\varepsilon}{\sqrt{2}} \kappa \operatorname{sign}(X) & \frac{\lambda-1}{2} \end{pmatrix}\end{aligned}$$

to define the kernels

$$K_0^{\text{KdV}}(X, Y) = G_0^{\text{KdV}}(X - Y) \quad K_0^{\varepsilon\text{GP}}(X, Y) = G_0^{\varepsilon\text{GP}}(X - Y).$$

Then $K_0^{\text{KdV}} \in L^\infty(\mathbb{R} \times \mathbb{R})$ and $K_0^{\varepsilon\text{GP}} \in (L^\infty(\mathbb{R} \times \mathbb{R}))^{2 \times 2}$ are the integral kernels of R_0^{KdV} and $R_0^{\varepsilon\text{GP}}$ respectively. In order to give a formula for the Green's function $G^{\varepsilon\text{GP}}$, for which $K^{\varepsilon\text{GP}}(X, Y) = G^{\varepsilon\text{GP}}(X, Y)$ is the kernel of $R^{\varepsilon\text{GP}}$, we use the Jost solutions $\Phi_j^{\varepsilon\text{GP}, \pm}$, their boundary data $E^{\varepsilon\text{GP}, \pm}$, and the transmission coefficient $a^{\varepsilon\text{GP}}$. Their definitions may be found in Section 2. For simplicity, we suppress the dependence of these scattering quantities on k from our notation in the following lemma.

Lemma A.2. *The Green's function $G^{\varepsilon\text{GP}}$ of the operator*

$$\frac{L^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2} = \begin{pmatrix} -W_- - \frac{\lambda-1}{\varepsilon^2} & \frac{i\sqrt{2}}{\varepsilon} \partial_X \\ \frac{i\sqrt{2}}{\varepsilon} \partial_X & -W_+ - \frac{\lambda+1}{\varepsilon^2} \end{pmatrix},$$

defined by the property that $\left(\frac{L^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2}\right) G^{\varepsilon\text{GP}}(X, Y) = \delta_0(X - Y)$, is given by

$$G^{\varepsilon\text{GP}}(X, Y) = \frac{\varepsilon}{i\sqrt{2} a^{\varepsilon\text{GP}}(k) \det E^{\varepsilon\text{GP}, +}} \begin{cases} \begin{pmatrix} -\Phi_{1,1}^{\varepsilon\text{GP}, -}(X) \Phi_{2,1}^{\varepsilon\text{GP}, +}(Y) & \Phi_{1,1}^{\varepsilon\text{GP}, -}(X) \Phi_{2,2}^{\varepsilon\text{GP}, +}(Y) \\ -\Phi_{1,2}^{\varepsilon\text{GP}, -}(X) \Phi_{2,1}^{\varepsilon\text{GP}, +}(Y) & \Phi_{1,2}^{\varepsilon\text{GP}, -}(X) \Phi_{2,2}^{\varepsilon\text{GP}, +}(Y) \end{pmatrix} & , X > Y \\ \begin{pmatrix} -\Phi_{2,1}^{\varepsilon\text{GP}, +}(X) \Phi_{1,1}^{\varepsilon\text{GP}, -}(Y) & \Phi_{2,1}^{\varepsilon\text{GP}, +}(X) \Phi_{1,2}^{\varepsilon\text{GP}, -}(Y) \\ -\Phi_{2,2}^{\varepsilon\text{GP}, +}(X) \Phi_{1,1}^{\varepsilon\text{GP}, -}(Y) & \Phi_{2,2}^{\varepsilon\text{GP}, +}(X) \Phi_{1,2}^{\varepsilon\text{GP}, -}(Y) \end{pmatrix} & , X < Y \end{cases}.$$

Proof. For every $Y \in \mathbb{R}$ and $X \in \mathbb{R} \setminus \{Y\}$ each column of $G^{\varepsilon\text{GP}}(X, Y)$ is a solution to the scattering problem $\left(\frac{L^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2}\right) G^{\varepsilon\text{GP}}(X, Y) = 0$. Furthermore, the jump condition

$$\begin{aligned} & \lim_{X \searrow Y} G^{\varepsilon\text{GP}}(X, Y) - \lim_{X \nearrow Y} G^{\varepsilon\text{GP}}(X, Y) \\ &= \frac{\varepsilon}{i\sqrt{2} a^{\varepsilon\text{GP}}(k) \det E^{\varepsilon\text{GP}, +}} \begin{pmatrix} \Phi_{2,1}^{\varepsilon\text{GP}, +} \Phi_{1,1}^{\varepsilon\text{GP}, -} - \Phi_{1,1}^{\varepsilon\text{GP}, -} \Phi_{2,1}^{\varepsilon\text{GP}, +} & \Phi_{1,1}^{\varepsilon\text{GP}, -} \Phi_{2,2}^{\varepsilon\text{GP}, +} - \Phi_{2,1}^{\varepsilon\text{GP}, +} \Phi_{1,2}^{\varepsilon\text{GP}, -} \\ \Phi_{2,2}^{\varepsilon\text{GP}, +} \Phi_{1,1}^{\varepsilon\text{GP}, -} - \Phi_{1,2}^{\varepsilon\text{GP}, -} \Phi_{2,1}^{\varepsilon\text{GP}, +} & \Phi_{1,2}^{\varepsilon\text{GP}, -} \Phi_{2,2}^{\varepsilon\text{GP}, +} - \Phi_{2,2}^{\varepsilon\text{GP}, +} \Phi_{1,2}^{\varepsilon\text{GP}, -} \end{pmatrix} (Y) \\ &= \frac{\varepsilon}{i\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

holds, which implies that $\left(\frac{L^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2}\right) G^{\varepsilon\text{GP}}(X, Y) = \delta_0(X - Y)$ in the sense of distributions. $\square$

An analogous representation for the Green's function of L^{KdV} exists, but is not necessary for our purposes.

A.2 Approximation of perturbation determinants

In this section we develop the theory of perturbation determinants of the Lax operators $L^{\varepsilon\text{GP}}$ and L^{KdV} , following in structure the seminal papers [39, 55, 56] which concern L^{KdV} and L^{NLS} . These techniques have recently been used effectively to construct low-regularity conserved quantities and prove well-posedness of various completely integrable PDEs (see also [15, 54]). Our contribution is a proof that an analogue of the approximation property observed in Theorem 1.1 and Lemma 2.7 also holds for the perturbation determinants. This is the content of Lemma A.6 below. The perturbation determinants in question are the quantities

$$\begin{aligned} \log a^{\varepsilon\text{GP}}(k; W) &= \alpha^{\varepsilon\text{GP}}(\kappa; W) - \text{tr}[R_0^{\varepsilon\text{GP}} W] = -\log \det \left(\left(\frac{L_0^{\varepsilon\text{GP}} - \lambda}{\varepsilon^2} \right)^{-1} \frac{L^{\varepsilon\text{GP}}(W) - \lambda}{\varepsilon^2} \right) \\ \log a^{\text{KdV}}(k; \pm W_\pm) &= \alpha^{\text{KdV}}(\kappa; \pm W_\pm) - \text{tr}[R_0^{\text{KdV}}(\pm W_\pm)] = -\log \det \left((L_0^{\text{KdV}} + \kappa^2)^{-1} (L^{\text{KdV}}(\pm W_\pm) + \kappa^2) \right). \end{aligned}$$

Here the formulation on the right-hand side should only be understood in a formal sense. This can be rigorously defined via the absolutely convergent series

$$\alpha^{\varepsilon\text{GP}}(\kappa; W) - \text{tr}[R_0^{\varepsilon\text{GP}}\mathcal{W}] = \sum_{l=1}^{\infty} \frac{(-1)^l}{l} \text{tr}[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]$$

$$\alpha^{\text{KdV}}(\kappa; \pm W_{\pm}) - \text{tr}[R_0^{\text{KdV}}(\pm W_{\pm})] = \sum_{l=1}^{\infty} \frac{(-1)^l}{l} \text{tr}[(R_0^{\text{KdV}}(\pm W_{\pm}))^l],$$

which is the content of Lemma A.4 below. Subsequently, in Lemma A.5, we prove the correspondence between these series and the transmission coefficient that we have stated above. We shall use the smallness of certain Hilbert-Schmidt norms to conclude the absolute convergence of these series as geometric series.

Lemma A.3. *We explicitly compute the following Hilbert-Schmidt norms:*

$$\left\| \sqrt{R_0^{\text{KdV}}} W_{\pm} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 = \frac{1}{\kappa} \|W_{\pm}\|_{H_{\kappa}^{-1}}^2 \quad (\text{A.4})$$

$$\left\| \sqrt{R_0^{\text{KdV}}} i\partial_X W_{\pm} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 = \frac{1}{2\kappa} \|W_{\pm}\|_{L_{\kappa}^2}^2 \quad (\text{A.5})$$

$$\left\| \sqrt{R_0^{\text{KdV}}} \mathcal{U}\mathcal{W} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 = \left| \frac{\lambda+1}{2} \right|^2 \frac{1}{\kappa} \|W_{-}\|_{H_{\kappa}^{-1}}^2 + \left| \frac{\lambda-1}{2} \right|^2 \frac{1}{\kappa} \|W_{+}\|_{H_{\kappa}^{-1}}^2 \quad (\text{A.6})$$

$$+ \frac{\varepsilon^2}{2} \frac{1}{2\kappa} \|W_{-}\|_{L_{\kappa}^2}^2 + \frac{\varepsilon^2}{2} \frac{1}{2\kappa} \|W_{+}\|_{L_{\kappa}^2}^2. \quad (\text{A.7})$$

Proof. For (A.4), we refer to [56, Proposition 2.1]. We shall reduce (A.5) and (A.6) to applications of (A.4). Towards this, we decompose

$$\begin{aligned} \left\| \sqrt{R_0^{\text{KdV}}} \mathcal{U}\mathcal{W} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 &= \text{tr} \left[\sqrt{R_0^{\text{KdV}}} \mathcal{W} \mathcal{U}^* \sqrt{R_0^{\text{KdV}}} \sqrt{R_0^{\text{KdV}}} \mathcal{U}\mathcal{W} \sqrt{R_0^{\text{KdV}}} \right] \\ &= \text{tr} \left[R_0^{\text{KdV}} \begin{pmatrix} W_{-} \frac{\bar{\lambda}+1}{2} & W_{-} \frac{i\varepsilon}{\sqrt{2}} \partial_X \\ W_{+} \frac{i\varepsilon}{\sqrt{2}} \partial_X & W_{+} \frac{\bar{\lambda}-1}{2} \end{pmatrix} R_0^{\text{KdV}} \begin{pmatrix} \frac{\lambda+1}{2} W_{-} & \frac{i\varepsilon}{\sqrt{2}} \partial_X W_{+} \\ \frac{i\varepsilon}{\sqrt{2}} \partial_X W_{-} & \frac{\lambda-1}{2} W_{+} \end{pmatrix} \right] \\ &= \text{tr} \left[R_0^{\text{KdV}} W_{-} \frac{\bar{\lambda}+1}{2} R_0^{\text{KdV}} \frac{\lambda+1}{2} W_{-} \right] + \text{tr} \left[R_0^{\text{KdV}} W_{-} \frac{i\varepsilon}{\sqrt{2}} \partial_X R_0^{\text{KdV}} \frac{i\varepsilon}{\sqrt{2}} \partial_X W_{-} \right] \\ &\quad + \text{tr} \left[R_0^{\text{KdV}} W_{+} \frac{i\varepsilon}{\sqrt{2}} \partial_X R_0^{\text{KdV}} \frac{i\varepsilon}{\sqrt{2}} \partial_X W_{+} \right] + \text{tr} \left[R_0^{\text{KdV}} W_{+} \frac{\bar{\lambda}-1}{2} R_0^{\text{KdV}} \frac{\lambda-1}{2} W_{+} \right] \\ &= \left| \frac{\lambda+1}{2} \right|^2 \left\| \sqrt{R_0^{\text{KdV}}} W_{-} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 + \left| \frac{\lambda-1}{2} \right|^2 \left\| \sqrt{R_0^{\text{KdV}}} W_{+} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 \\ &\quad + \frac{\varepsilon^2}{2} \left\| \sqrt{R_0^{\text{KdV}}} W_{-} i\partial_X \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 + \frac{\varepsilon^2}{2} \left\| \sqrt{R_0^{\text{KdV}}} W_{+} i\partial_X \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2. \end{aligned}$$

It remains to compute

$$\left\| \sqrt{R_0^{\text{KdV}}} i\partial_X W_{\pm} \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 = \left\| \sqrt{R_0^{\text{KdV}}} W_{\pm} i\partial_X \sqrt{R_0^{\text{KdV}}} \right\|_{\text{HS}}^2 = \text{tr}[R_0^{\text{KdV}} W_{\pm} i\partial_X R_0^{\text{KdV}} i\partial_X W_{\pm}],$$

which can be written, using the kernels

$$R_0^{\text{KdV}} W_{\pm} i\partial_X f(X) = \int_{\mathbb{R}} \frac{-i}{2\kappa} e^{-\kappa|X-Y|} (\kappa \text{sign}(X-Y) W_{\pm}(Y) + \partial_Y W_{\pm}(Y)) f(Y) dY$$

$$R_0^{\text{KdV}} i\partial_X W_{\pm} f(X) = \int_{\mathbb{R}} \frac{-i}{2\kappa} e^{-\kappa|X-Y|} \kappa \text{sign}(X-Y) W_{\pm}(Y) f(Y) dY,$$

as

$$\begin{aligned}
& \text{tr}[R_0^{\text{KdV}} W_{\pm} i \partial_X R_0^{\text{KdV}} i \partial_X W_{\pm}] \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{-1}{4\kappa^2} e^{-\kappa|X-Y|} (\kappa \text{sign}(X-Y) W_{\pm}(Y) + \partial_Y W_{\pm}(Y)) e^{-\kappa|Y-X|} \kappa \text{sign}(Y-X) W_{\pm}(X) dX dY \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} \kappa^2 W_{\pm}(X) W_{\pm}(Y) dX dY \\
&+ \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} \kappa \text{sign}(X-Y) W_{\pm}(X) \partial_Y W_{\pm}(Y) dX dY \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} \kappa^2 W_{\pm}(X) W_{\pm}(Y) dX dY \\
&- \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} (2\kappa^2 - 2\kappa\delta_0(X-Y)) W_{\pm}(X) W_{\pm}(Y) dX dY \\
&= - \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} \kappa^2 W_{\pm}(X) W_{\pm}(Y) dX dY + 2\kappa \int_{\mathbb{R}} \frac{1}{4\kappa^2} |W_{\pm}(X)|^2 dX.
\end{aligned}$$

We already know from (A.4) that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{4\kappa^2} e^{-2\kappa|X-Y|} \kappa^2 W_{\pm}(X) W_{\pm}(Y) dX dY = \kappa \|W_{\pm}\|_{H_{\kappa}^{-1}}^2,$$

which finishes the proof. $\square$

As a consequence of this lemma we have $\sqrt{R_0^{\text{KdV}}} \mathcal{W} \sqrt{R_0^{\text{KdV}}}, \sqrt{R_0^{\text{KdV}}} \mathcal{U} \mathcal{W} \sqrt{R_0^{\text{KdV}}} \in L(L^2(\mathbb{R}; \mathbb{C}^2))$. In the following, we sometimes use the operators $R_0^{\text{KdV}} \mathcal{W}$ and $R_0^{\varepsilon \text{GP}} \mathcal{W}$ instead for the sake of clarity in formal calculations, but always in a context where eventually a trace or Hilbert-Schmidt norm is taken, returning us to the well-behaved versions of these operators.

Lemma A.4. *Let $\varepsilon \in (0, 1)$. For all $0 < \kappa < \frac{1}{\sqrt{2}\varepsilon}$ such that*

$$\frac{1}{\kappa} \|W\|_{H_{\kappa}^{-1}}^2 \leq \frac{1}{30} \qquad \frac{1}{\kappa} \|W\|_{L_{\kappa}^2}^2 \leq \frac{1}{15}$$

the series

$$\begin{aligned}
\alpha^{\text{KdV}}(\kappa; \pm W_{\pm}) &= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[(R_0^{\text{KdV}}(\pm W_{\pm}))^l] \\
\alpha^{\varepsilon \text{GP}}(\kappa; W) &= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[(R_0^{\varepsilon \text{GP}} \mathcal{W})^l]
\end{aligned}$$

are absolutely convergent and satisfy

$$\begin{aligned}
\frac{3}{2} |\alpha^{\text{KdV}}(\kappa; \pm W_{\pm})| &\leq \frac{1}{\kappa} \|W_{\pm}\|_{H_{\kappa}^{-1}}^2 \\
\frac{3}{2} |\alpha^{\varepsilon \text{GP}}(\kappa; W)| &\leq \left| \frac{\lambda+1}{2} \right|^2 \frac{1}{\kappa} \|W_{-}\|_{H_{\kappa}^{-1}}^2 + \left| \frac{\lambda-1}{2} \right|^2 \frac{1}{\kappa} \|W_{+}\|_{H_{\kappa}^{-1}}^2 + \frac{\varepsilon^2}{2} \frac{1}{2\kappa} \|W_{-}\|_{L_{\kappa}^2}^2 + \frac{\varepsilon^2}{2} \frac{1}{2\kappa} \|W_{+}\|_{L_{\kappa}^2}^2.
\end{aligned}$$

Moreover, the resolvent

$$R^{\varepsilon \text{GP}} = \left(\frac{L^{\varepsilon \text{GP}} - \lambda}{\varepsilon^2} \right)^{-1} \in L(L^2(\mathbb{R}; \mathbb{C}^2))$$

is represented by the series

$$R^{\varepsilon \text{GP}} = \sum_{l=0}^{\infty} (-1)^l (R_0^{\varepsilon \text{GP}} \mathcal{W})^l R_0^{\varepsilon \text{GP}} = \sum_{l=0}^{\infty} (-1)^l \sqrt{R_0^{\text{KdV}}} \left(\sqrt{R_0^{\text{KdV}}} \mathcal{U} \mathcal{W} \sqrt{R_0^{\text{KdV}}} \right)^l \sqrt{R_0^{\text{KdV}}} \mathcal{U}, \quad (\text{A.8})$$

which is absolutely convergent with respect to the operator norm $\|\cdot\|_{L^2 \rightarrow L^2}$.

Proof. This follows from Lemma A.1 together with the Hilbert-Schmidt norms computed in Lemma A.3. $\square$

We have adapted the method of proof of the following lemma from [57, Theorem 2.6]. We refer to the paragraph above this theorem for some context on the history of this argument, and other approaches to proving that the perturbation determinant equals the logarithm of the transmission coefficient.

Lemma A.5. *Assume the setting of Lemma A.4, and in addition that $W \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$ so that we are in a setting where we have defined the transmission coefficients $a^{\varepsilon\text{GP}}(k; W)$ and $a^{\text{KdV}}(k; \pm W_{\pm})$ (see (2.14)). We have*

$$\alpha^{\varepsilon\text{GP}}(\kappa; W) - \log a^{\varepsilon\text{GP}}(k; W) = \text{tr}[R_0^{\varepsilon\text{GP}}\mathcal{W}] = \frac{1}{2\kappa} \int_{\mathbb{R}} \frac{1}{2} (W_+ - W_-) - \frac{\lambda}{2} (W_+ + W_-) dX \quad (\text{A.9})$$

and similarly

$$\alpha^{\text{KdV}}(\kappa; \pm W_{\pm}) - \log a^{\text{KdV}}(k; \pm W_{\pm}) = \text{tr}[R_0^{\text{KdV}}(\pm W_{\pm})] = \frac{1}{2\kappa} \int_{\mathbb{R}} \pm W_{\pm} dX. \quad (\text{A.10})$$

Proof. We compute

$$\begin{aligned} \text{tr}[R_0^{\varepsilon\text{GP}}\mathcal{W}] &= \text{tr}\left[R_0^{\text{KdV}}\begin{pmatrix} \frac{\lambda+1}{2} & \frac{i\varepsilon}{\sqrt{2}}\partial_X \\ \frac{i\varepsilon}{\sqrt{2}}\partial_X & \frac{\lambda-1}{2} \end{pmatrix}\mathcal{W}\right] \\ &= \text{tr}\left[R_0^{\text{KdV}}\frac{\lambda+1}{2}(-W_-)\right] + \text{tr}\left[R_0^{\text{KdV}}\frac{\lambda-1}{2}(-W_+)\right] \\ &= \frac{\lambda+1}{2} \int_{\mathbb{R}} \frac{1}{2\kappa} e^{-\kappa|X-X|}(-W_-(X)) dX + \frac{\lambda-1}{2} \int_{\mathbb{R}} \frac{1}{2\kappa} e^{-\kappa|X-X|}(-W_+(X)) dX \\ &= \frac{1}{2\kappa} \int_{\mathbb{R}} \frac{1}{2} (W_+ - W_-) - \frac{\lambda}{2} (W_+ + W_-) dX \end{aligned}$$

and

$$\text{tr}[R_0^{\text{KdV}}(\pm W_{\pm})] = \frac{1}{2\kappa} \int_{\mathbb{R}} \pm W_{\pm} dX.$$

For (A.10) we refer to [55, Proposition 2.4], so it remains to prove (A.9).

By explicitly computing the Jost solution $\Psi_{1,1}^{\varepsilon\text{GP},-}(X; 0)$, which is constant, we find that $\log a^{\varepsilon\text{GP}}(k; 0) = 0$, which matches $\alpha^{\varepsilon\text{GP}}(\kappa; 0) = 0$. It therefore suffices to prove for any $W, V \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$ with

$$V = (V_+, V_-) \quad \mathcal{V} = \begin{pmatrix} -V_- & 0 \\ 0 & -V_+ \end{pmatrix}$$

that the following limits exist and are identical:

$$\left. \frac{d}{ds} \right|_{s=0} \log a^{\varepsilon\text{GP}}(k; W + sV) = \left. \frac{d}{ds} \right|_{s=0} (\alpha^{\varepsilon\text{GP}}(\kappa; W + sV) - \text{tr}[R_0^{\varepsilon\text{GP}}(\mathcal{W} + s\mathcal{V})]).$$

Set $\Phi^{\varepsilon\text{GP}} = (\Phi_1^{\varepsilon\text{GP},-} | \Phi_2^{\varepsilon\text{GP},+})$. We start with the formal calculation

$$\begin{aligned} 0 &= \left. \frac{d}{ds} \right|_{s=0} \left(\frac{L^{\varepsilon\text{GP}}(W + sV) - \lambda}{\varepsilon^2} \Phi_j^{\varepsilon\text{GP},\pm}(k; W + sV) \right) \\ &= \mathcal{V} \Phi_j^{\varepsilon\text{GP},\pm}(k; W) + \frac{L^{\varepsilon\text{GP}}(W) - \lambda}{\varepsilon^2} \left. \frac{d}{ds} \right|_{s=0} \Phi_j^{\varepsilon\text{GP},\pm}(k; W + sV), \end{aligned}$$

which implies

$$\frac{d}{ds} \Big|_{s=0} \Phi^{\varepsilon\text{GP}}(X, k; W + sV) = \int_{\mathbb{R}} G^{\varepsilon\text{GP}}(X, Y)(-\mathcal{V}(Y)) \Phi^{\varepsilon\text{GP}}(Y, k; W) dY.$$

For the derivative of the transmission coefficient this yields

$$\begin{aligned} & \frac{d}{ds} \Big|_{s=0} \log a^{\varepsilon\text{GP}}(k; W + sV) \\ &= \frac{d}{ds} \Big|_{s=0} \log \left(\frac{\det \Phi^{\varepsilon\text{GP}}(X, k; W + sV)}{\det E^{\varepsilon\text{GP},+}(k)} \right) \\ &= \text{tr} \left[\Phi^{\varepsilon\text{GP}}(X, k; W)^{-1} \int_{\mathbb{R}} G^{\varepsilon\text{GP}}(X, Y)(-\mathcal{V}(Y)) \Phi^{\varepsilon\text{GP}}(Y, k; W) dY \right] \\ &= \int_{\mathbb{R}} \text{tr} [\Phi^{\varepsilon\text{GP}}(Y, k; W) \Phi^{\varepsilon\text{GP}}(X, k; W)^{-1} G^{\varepsilon\text{GP}}(X, Y)(-\mathcal{V}(Y))] dY \\ &= \int_{\mathbb{R}} \text{tr} \left[\left(\mathbb{1}_{\{X < Y\}} \lim_{X' \nearrow Y} G^{\varepsilon\text{GP}}(X', Y) + \mathbb{1}_{\{X > Y\}} \lim_{X' \searrow Y} G^{\varepsilon\text{GP}}(X', Y) \right) (-\mathcal{V}(Y)) \right] dY \\ &= \int_{\mathbb{R}} G_{1,1}^{\varepsilon\text{GP}}(X, X) V_-(X) + G_{2,2}^{\varepsilon\text{GP}}(X, X) V_+(X) dX. \end{aligned}$$

Here we have used that $G_{1,1}^{\varepsilon\text{GP}}$ and $G_{2,2}^{\varepsilon\text{GP}}$ are continuous. Since the Green's function is bounded and V is Schwartz, the integrals are absolutely convergent and the derivative actually exists. On the other hand, we know from Lemma A.1 that $\alpha^{\varepsilon\text{GP}}(\kappa; W + sV)$ is continuously differentiable. Specifically (A.3) yields

$$\begin{aligned} \frac{d}{ds} \Big|_{s=0} \alpha^{\varepsilon\text{GP}}(\kappa; W + sV) &= \sum_{l=2}^{\infty} (-1)^l \text{tr} [(R_0^{\varepsilon\text{GP}} \mathcal{W})^{l-1} R_0^{\varepsilon\text{GP}} \mathcal{V}] \\ &= \text{tr} [(R_0^{\varepsilon\text{GP}} - R^{\varepsilon\text{GP}}) \mathcal{V}] \\ &= \int_{\mathbb{R}} \text{tr} [(G^{\varepsilon\text{GP}} - G_0^{\varepsilon\text{GP}})(X, X)(-\mathcal{V})(X)] dX \\ &= \int_{\mathbb{R}} (G^{\varepsilon\text{GP}} - G_0^{\varepsilon\text{GP}})_{1,1}(X, X) V_-(X) + (G^{\varepsilon\text{GP}} - G_0^{\varepsilon\text{GP}})_{2,2}(X, X) V_+(X) dX. \end{aligned}$$

Here we have used that $G^{\varepsilon\text{GP}}$ and $G_0^{\varepsilon\text{GP}}$ have continuous diagonal parts. The proof is finished with the additional observation that

$$-\frac{d}{ds} \Big|_{s=0} \text{tr} [R_0^{\varepsilon\text{GP}}(\mathcal{W} + s\mathcal{V})] = \int_{\mathbb{R}} (G_0^{\varepsilon\text{GP}})_{1,1}(X, X) V_-(X) + (G_0^{\varepsilon\text{GP}})_{2,2}(X, X) V_+(X) dX.$$

□

Recall from (1.14) the definition of the operators Ev_λ and Od_λ . In the following, we suppress the dependence of various quantities on λ from the notation, but for the application of these operators this dependence is of course relevant.

Lemma A.6. *Let $W = (W_+, W_-) \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$, $\varepsilon \in (0, 1)$, and $0 < \kappa < \frac{1}{\sqrt{2}\varepsilon}$ with $\lambda \in (0, 1)$. Assume that*

$$\max \left\{ \frac{1}{\kappa}, \frac{1}{\kappa^3} \right\} \|W\|_{L^2}^2 \leq \frac{1}{30}.$$

Then the series $\alpha^{\varepsilon\text{GP}}(\kappa; W)$ and $\alpha^{KdV}(\kappa; \pm W_\pm)$ converge. Furthermore, we have

$$\text{Ev}_\lambda [\text{tr} [R_0^{\varepsilon\text{GP}} \mathcal{W}]] \mp \frac{1}{\lambda} \text{Od}_\lambda [\text{tr} [R_0^{\varepsilon\text{GP}} \mathcal{W}]] = \text{tr} [R_0^{KdV}(\pm W_\pm)] \quad (\text{A.11})$$

$$\text{Ev}_\lambda [\alpha^{\varepsilon\text{GP}}(\kappa; W)] \mp \frac{1}{\lambda} \text{Od}_\lambda [\alpha^{\varepsilon\text{GP}}(\kappa; W)] = \alpha^{KdV}(\kappa; (\pm W_\pm)) + \mathbf{E}(\kappa; W) \quad (\text{A.12})$$

with

$$|\mathbf{E}(\kappa; W)| \leq 8 \frac{\varepsilon^2}{\kappa} \|W\|_{L^2}^2.$$

Proof. Combining (A.9) and (A.10) directly yields (A.11). To prove (A.12) we need to understand

$$\text{Ev}_\lambda[\text{tr}[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]] \mp \frac{1}{\lambda} \text{Od}_\lambda[\text{tr}[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]].$$

Using the Pauli matrices

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

we write

$$\begin{aligned} \text{Ev}_\lambda[R_0^{\varepsilon\text{GP}}\mathcal{W}] &= \frac{1}{2} R_0^{\text{KdV}}(\sigma_3 + i\sqrt{2}\varepsilon\sigma_1\partial_X)\mathcal{W} \\ \text{Od}_\lambda[R_0^{\varepsilon\text{GP}}\mathcal{W}] &= \frac{1}{2} R_0^{\text{KdV}}\lambda\mathcal{W}. \end{aligned}$$

Then

$$\text{Ev}_\lambda[\text{tr}[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]] \mp \frac{1}{\lambda} \text{Od}_\lambda[\text{tr}[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]] = \text{tr}\left[\text{Ev}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l] \mp \frac{1}{\lambda} \text{Od}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l]\right].$$

For $l \in \mathbb{N}$ and $\omega \in \{0, 1, 2\}^l$ we define

$$\bar{\omega} = \sum_{j=1}^l \mathbb{1}_{\{\omega_j=1\}}.$$

One can show directly by induction that

$$\begin{aligned} \text{Ev}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l] &= \sum_{\substack{\omega \in \{0,1\}^l \\ \bar{\omega} \text{ even}}} 2^{-l} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}}(\sigma_3 + \sqrt{2}\varepsilon\sigma_1 i\partial_X)\mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}}\lambda\mathcal{W} & , \omega_j = 1 \end{cases} \\ &= \sum_{\substack{\omega \in \{0,1,2\}^l \\ \bar{\omega} \text{ even}}} 2^{-l} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}}\sigma_3\mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}}\lambda\mathcal{W} & , \omega_j = 1 \\ R_0^{\text{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & , \omega_j = 2 \end{cases} \end{aligned}$$

and

$$\begin{aligned} \text{Od}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l] &= \sum_{\substack{\omega \in \{0,1\}^l \\ \bar{\omega} \text{ odd}}} 2^{-l} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}}(\sigma_3 + \sqrt{2}\varepsilon\sigma_1 i\partial_X)\mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}}\lambda\mathcal{W} & , \omega_j = 1 \end{cases} \\ &= \sum_{\substack{\omega \in \{0,1,2\}^l \\ \bar{\omega} \text{ odd}}} 2^{-l} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}}\sigma_3\mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}}\lambda\mathcal{W} & , \omega_j = 1 \\ R_0^{\text{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & , \omega_j = 2 \end{cases}. \end{aligned}$$

Here the ordering of the product of possibly non-commuting operators does not matter, as long as it is consistent for all $\omega \in \{0, 1\}^l$. We shall fix as a convention that the first term in the product, i. e. the

term with $j = 1$, is multiplied on the left, while the last term is multiplied on the right. Then

$$\begin{aligned}
\text{Ev}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l] \mp \frac{1}{\lambda} \text{Od}_\lambda[(R_0^{\varepsilon\text{GP}}\mathcal{W})^l] &= \sum_{\omega \in \{0,1,2\}^l} 2^{-l} \left(\frac{\mp 1}{\lambda} \right)^{\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}} \sigma_3 \mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}} \lambda \mathcal{W} & , \omega_j = 1 \\ R_0^{\text{KdV}} \sqrt{2\varepsilon} \sigma_1 i \partial_X \mathcal{W} & , \omega_j = 2 \end{cases} \\
&= \sum_{\substack{\omega \in \{0,1,2\}^l \\ \exists j: \omega_j = 2}} 2^{-l} \left(\frac{\mp 1}{\lambda} \right)^{\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}} \sigma_3 \mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}} \lambda \mathcal{W} & , \omega_j = 1 \\ R_0^{\text{KdV}} \sqrt{2\varepsilon} \sigma_1 i \partial_X \mathcal{W} & , \omega_j = 2 \end{cases} \\
&+ \sum_{\omega \in \{0,1\}^l} 2^{-l} \left(\frac{\mp 1}{\lambda} \right)^{\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} \prod_{j=1}^l \begin{cases} R_0^{\text{KdV}} \sigma_3 \mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}} \lambda \mathcal{W} & , \omega_j = 1 \end{cases} \\
&= I + II.
\end{aligned}$$

Using that the trace of off-diagonal matrices vanishes, we estimate

$$\begin{aligned}
\left| \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[I] \right| &\leq \sum_{l=2}^{\infty} \sum_{\substack{\omega \in \{0,1,2\}^l \\ |\{j: \omega_j=2\}| \geq 2 \text{ even}}} |\lambda|^{-\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} 2^{-l} \prod_{j=1}^l \begin{cases} \|\sqrt{R_0^{\text{KdV}}} \sigma_3 \mathcal{W} \sqrt{R_0^{\text{KdV}}}\|_{\text{HS}} & , \omega_j = 0 \\ \|\sqrt{R_0^{\text{KdV}}} \lambda \mathcal{W} \sqrt{R_0^{\text{KdV}}}\|_{\text{HS}} |\lambda| & , \omega_j = 1 \\ \|\sqrt{R_0^{\text{KdV}}} \sqrt{2\varepsilon} \sigma_1 i \partial_X \mathcal{W} \sqrt{R_0^{\text{KdV}}}\|_{\text{HS}} & , \omega_j = 2 \end{cases} \\
&\leq \sum_{l=2}^{\infty} \sum_{\substack{\omega \in \{0,1,2\}^l \\ |\{j: \omega_j=2\}| \geq 2 \text{ even}}} |\lambda|^{-\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} 2^{-l} \kappa^{-\frac{l}{2}} \prod_{j=1}^l \begin{cases} \|W\|_{H_\kappa^{-1}} & , \omega_j = 0 \\ \|W\|_{H_\kappa^{-1}} |\lambda| & , \omega_j = 1 \\ \|W\|_{L_\kappa^2} \varepsilon & , \omega_j = 2 \end{cases}
\end{aligned}$$

Here we have used (A.2), (A.4), and (A.5). Noting that $\|W\|_{H_\kappa^{-1}} \leq \frac{1}{\sqrt{2\kappa}} \|W\|_{L_\kappa^2}$ and $|\lambda| \leq 1$, we finish our estimate:

$$\begin{aligned}
\left| \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[I] \right| &\leq \sum_{l=2}^{\infty} \sum_{\substack{\omega \in \{0,1,2\}^l \\ |\{j: \omega_j=2\}| \geq 2 \text{ even}}} \varepsilon^{|\{\omega_j=2\}|} 2^{-l} \kappa^{-\frac{l}{2}} |\lambda|^{|\{\omega_j=1\}| - \mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} \|W\|_{H_\kappa^{-1}}^{|\{\omega_j=0\}| + |\{\omega_j=1\}|} \|W\|_{L_\kappa^2}^{|\{\omega_j=2\}|} \\
&\leq \sum_{l=2}^{\infty} \|W\|_{L_\kappa^2}^l 2^{-l} \sum_{\substack{\omega \in \{0,1,2\}^l \\ |\{j: \omega_j=2\}| \geq 2 \text{ even}}} \varepsilon^{|\{\omega_j=2\}|} (\sqrt{2\kappa})^{-|\{\omega_j=0\}| - |\{\omega_j=1\}| - \frac{l}{2}} \\
&\leq 2\varepsilon^2 \kappa^2 \sum_{l=2}^{\infty} \kappa^{-\frac{l}{2}} \|W\|_{L_\kappa^2}^l 2^{-l} \sum_{\substack{n=2 \\ n \text{ even}}}^l \sum_{j=0}^l \binom{l}{n, j} (\sqrt{2\varepsilon\kappa})^{n-2} (\sqrt{2\kappa})^{-\frac{3}{2}l} \\
&\leq 2\varepsilon^2 \kappa^2 \sum_{l=2}^{\infty} \kappa^{-\frac{3}{2}l} \|W\|_{L_\kappa^2}^l 2^{-l} 3^l \sqrt{2}^{-\frac{3}{2}l} \leq \frac{\varepsilon^2 \kappa^{-1} \|W\|_{L^2}^2}{1 - \kappa^{-\frac{3}{2}} \|W\|_{L^2}}.
\end{aligned}$$

In order to compute II , we decompose

$$\begin{aligned}
\{0,1\}^l &= \{\omega \in \{0,1\}^l : \omega_1 = 0 \text{ and } \bar{\omega} \text{ even}\} \cup \{\omega \in \{0,1\}^l : \omega_1 = 1 \text{ and } \bar{\omega} \text{ odd}\} \\
&\cup \{\omega \in \{0,1\}^l : \omega_1 = 0 \text{ and } \bar{\omega} \text{ odd}\} \cup \{\omega \in \{0,1\}^l : \omega_1 = 1 \text{ and } \bar{\omega} \text{ even}\}
\end{aligned}$$

and observe that

$$\omega \mapsto (0, \omega_2, \dots, \omega_l) \qquad \omega \mapsto (1, \omega_2, \dots, \omega_l)$$

are bijections between the first two and last two sets. We may sum only over the sets on the left and correspondingly adjust each term as follows:

$$\begin{aligned}
II &= \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0 \\ \bar{\omega} \text{ even}}} 2^{-l} R_0^{\text{KdV}}(\sigma_3 \mp 1) \mathcal{W} \prod_{j=2}^l \begin{cases} R_0^{\text{KdV}} \sigma_3 \mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}} \lambda \mathcal{W} & , \omega_j = 1 \end{cases} \\
&+ \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0 \\ \bar{\omega} \text{ odd}}} 2^{-l} R_0^{\text{KdV}} \left(\frac{\mp \sigma_3}{\lambda} + \lambda \right) \mathcal{W} \prod_{j=2}^l \begin{cases} R_0^{\text{KdV}} \sigma_3 \mathcal{W} & , \omega_j = 0 \\ R_0^{\text{KdV}} \lambda \mathcal{W} & , \omega_j = 1 \end{cases} \\
&= \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0 \\ \bar{\omega} \text{ even}}} 2^{-l} (\sigma_3 \mp 1) \lambda^{\bar{\omega}} \sigma_3^{l-1-\bar{\omega}} (R_0^{\text{KdV}} \mathcal{W})^l + \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0 \\ \bar{\omega} \text{ odd}}} 2^{-l} (\sigma_3 \mp 1 \pm 2\varepsilon^2 \kappa^2) \left(\frac{\mp 1}{\lambda} \right) \lambda^{\bar{\omega}} \sigma_3^{l-1-\bar{\omega}} (R_0^{\text{KdV}} \mathcal{W})^l \\
&= (\sigma_3 \mp 1) (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l \sigma_3 \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0}} 2^{-l} \left(\frac{\mp 1}{\lambda} \right)^{\mathbb{1}_{\{\bar{\omega} \text{ odd}\}}} \lambda^{\bar{\omega}} \sigma_3^{\bar{\omega}} + 2\varepsilon^2 \kappa^2 (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l \sigma_3 \sum_{\substack{\omega \in \{0,1\}^l \\ \omega_1=0 \\ \bar{\omega} \text{ odd}}} 2^{-l} \left(\frac{-1}{\lambda} \right) \lambda^{\bar{\omega}} \sigma_3^{\bar{\omega}} \\
&= III + IV.
\end{aligned}$$

The further treatment of *III* and *IV* is a matter of computation. We have

$$\begin{aligned}
III &= (\sigma_3 \mp 1) (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sigma_3 \left(\sum_{\substack{\bar{\omega}=0 \\ \bar{\omega} \text{ even}}}^{l-1} \binom{l-1}{\bar{\omega}} \lambda^{\bar{\omega}} \mp \sigma_3 \sum_{\substack{\bar{\omega}=0 \\ \bar{\omega} \text{ odd}}}^{l-1} \binom{l-1}{\bar{\omega}} \lambda^{\bar{\omega}-1} \right) \\
&= \frac{\sigma_3 \mp 1}{2} (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{1-l} \left(\sigma_3 \sum_{j=0}^{\lfloor \frac{l-1}{2} \rfloor} \binom{l-1}{2j} (1 - 2\varepsilon^2 \kappa^2)^j \mp \sum_{j=0}^{\lfloor \frac{l}{2} \rfloor - 1} \binom{l-1}{2j+1} (1 - 2\varepsilon^2 \kappa^2)^j \right) \\
&= \frac{(\sigma_3 \mp 1)^2}{4} (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l + \frac{\sigma_3 \mp 1}{2} (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{1-l} \sum_{j=0}^{\infty} \left(\sigma_3 \binom{l-1}{2j} \mp \binom{l-1}{2j+1} \right) ((1 - 2\varepsilon^2 \kappa^2)^j - 1) \\
&= V + VI.
\end{aligned}$$

The most important term is *V*, because

$$\sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[V] = \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \left(\frac{(1 \mp 1)^2}{4} \text{tr}[(R_0^{\text{KdV}}(-W_-))^l] + \frac{(-1 \mp 1)^2}{4} \text{tr}[(R_0^{\text{KdV}} W_+)^l] \right) = \alpha^{\text{KdV}}(\kappa; \pm W_{\pm}).$$

We calculate further

$$\begin{aligned}
IV &= -2\varepsilon^2 \kappa^2 (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{\substack{\bar{\omega}=0 \\ \bar{\omega} \text{ odd}}}^{l-1} \binom{l-1}{\bar{\omega}} \lambda^{\bar{\omega}-1} \\
&= -2\varepsilon^2 \kappa^2 (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{j=0}^{\lfloor \frac{l}{2} \rfloor - 1} \binom{l-1}{2j+1} (1 - 2\varepsilon^2 \kappa^2)^j \\
&= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m \sum_{j=0}^{\infty} \binom{l-1}{2j+1} \binom{j}{m-1}
\end{aligned}$$

and

$$\begin{aligned}
VI &= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} (\sigma_3 \mp 1) \sum_{j=0}^{\infty} \left(\sigma_3 \binom{l-1}{2j} \mp \binom{l-1}{2j+1} \right) \sum_{m=1}^{\infty} \binom{j}{m} (-2\varepsilon^2 \kappa^2)^m \\
&= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m \sum_{j=0}^{\infty} (1 \mp \sigma_3) \binom{l}{2j+1} \binom{j}{m}.
\end{aligned}$$

Then a binomial identity, which may be verified by a calculation with generating functions, yields

$$\begin{aligned}
IV + VI &= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m \sum_{j=0}^{\infty} \left((1 \mp \sigma_3) \binom{l}{2j+1} \binom{j}{m} + \binom{l-1}{2j+1} \binom{j}{m-1} \right) \\
&= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m \left((1 \mp \sigma_3) 2^{l-1-2m} \binom{l-1-m}{m} + 2^{l-2m} \binom{l-1-m}{m-1} \right) \\
&= (R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m 4^{-m} \left(\frac{1 \mp \sigma_3}{2} \binom{l-1-m}{m} + \binom{l-1-m}{m-1} \right).
\end{aligned}$$

We find that

$$\begin{aligned}
&\sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[IV + VI] \\
&= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m 4^{-m} \left(\text{tr} \left[(R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l \frac{1 \mp \sigma_3}{2} \right] \binom{l-1-m}{m} + \text{tr}[(R_0^{\text{KdV}} \sigma_3 \mathcal{W})^l] \binom{l-1-m}{m-1} \right) \\
&= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \sum_{m=1}^{\infty} (-2\varepsilon^2 \kappa^2)^m 4^{-m} \left(\text{tr}[(R_0^{\text{KdV}} (\pm W_{\pm}))^l] \binom{l-1-m}{m} \right. \\
&\quad \left. + (\text{tr}[(R_0^{\text{KdV}} (-W_-))^l] + \text{tr}[(R_0^{\text{KdV}} W_+)^l]) \frac{1}{2} \binom{l-1-m}{m-1} \right),
\end{aligned}$$

and finally

$$\begin{aligned}
\left| \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \text{tr}[IV + VI] \right| &\leq \sum_{m=1}^{\infty} \varepsilon^{2m} \kappa^{2m} 2^{-m} \sum_{l=2m}^{\infty} 2^{l-m} \kappa^{-\frac{l}{2}} (\|W_-\|_{H_{\kappa}^{-1}}^l + \|W_+\|_{H_{\kappa}^{-1}}^l) \\
&\leq \sum_{m=1}^{\infty} \varepsilon^{2m} \kappa^m \|W\|_{H_{\kappa}^{-1}}^{2m} \frac{2}{1 - 2\kappa^{-\frac{1}{2}} \|W\|_{H_{\kappa}^{-1}}} \\
&\leq \frac{2\varepsilon^2 \kappa \|W\|_{H_{\kappa}^{-1}}^2}{1 - 2\varepsilon^2 \kappa \|W\|_{H_{\kappa}^{-1}}^2} \frac{2}{1 - 2\kappa^{-\frac{1}{2}} \|W\|_{H_{\kappa}^{-1}}} \\
&\leq \frac{\varepsilon^2 \kappa^{-1} \|W\|_{L^2}^2}{1 - \varepsilon^2 \kappa^{-1} \|W\|_{L^2}^2} \frac{1}{1 - \kappa^{-\frac{3}{2}} \|W\|_{L^2}} \\
&\leq \frac{\varepsilon^2 \kappa^{-1} \|W\|_{L^2}^2}{(1 - \kappa^{-\frac{3}{2}} \|W\|_{L^2})^2}.
\end{aligned}$$

Finally, we estimate

$$\frac{1}{(1 - \kappa^{-\frac{3}{2}} \|W\|_{L^2})^2} + \frac{1}{1 - \kappa^{-\frac{3}{2}} \|W\|_{L^2}} \leq \frac{1}{(1 - \frac{1}{\sqrt{3}})^2} + \frac{1}{1 - \frac{1}{\sqrt{3}}} \leq 8.$$

□

B Overview of Hamiltonian structures

In this section we present concisely and formally the Hamiltonian structures of GP, hGP, ε GP, and KdV.

B.1 GP

We consider formally a manifold where we denote points by $\mathbf{q} = (q, \bar{q})$ and tangent vectors by $\mathbf{p} = (p, \bar{p})$ and $\mathbf{r} = (r, \bar{r})$. We define the Hamiltonian operator J^{GP} , the Riemannian metric g^{GP} and the symplectic form ω^{GP} by

$$\begin{aligned} J_{\mathbf{q}}^{\text{GP}} &= -\frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ g_{\mathbf{q}}^{\text{GP}}(\mathbf{p}, \mathbf{r}) &= \frac{1}{2} \int_{\mathbb{R}} (p\bar{r} + \bar{p}r) \, dx \\ \omega_{\mathbf{q}}^{\text{GP}}(\mathbf{p}, \mathbf{r}) &= -i \int_{\mathbb{R}} (p\bar{r} - \bar{p}r) \, dx \end{aligned}$$

They satisfy

$$\omega_{\mathbf{q}}^{\text{GP}}(\mathbf{p}, J_{\mathbf{q}}^{\text{GP}} \mathbf{r}) = g_{\mathbf{q}}^{\text{GP}}(\mathbf{p}, \mathbf{r}).$$

We define the gradient $\nabla_{\mathbf{q}}^{g^{\text{GP}}}$ with respect to the metric g^{GP} as well as the functional derivative via

$$g_{\mathbf{q}}^{\text{GP}}(\mathbf{p}, \nabla_{\mathbf{q}}^{g^{\text{GP}}} \mathcal{H}) = \int_{\mathbb{R}} \mathbf{p} \left(\frac{\delta \mathcal{H}(\mathbf{q})}{\delta q}, \frac{\delta \mathcal{H}(\mathbf{q})}{\delta \bar{q}} \right)^T \, dx = \left. \frac{d}{ds} \right|_{s=0} \mathcal{H}(\mathbf{q} + s\mathbf{p}).$$

This implies

$$\nabla_{\mathbf{q}}^{g^{\text{GP}}} \mathcal{H} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \left(\frac{\delta \mathcal{H}(\mathbf{q})}{\delta q}, \frac{\delta \mathcal{H}(\mathbf{q})}{\delta \bar{q}} \right)^T.$$

Then (GP_n) can be written as

$$i\partial_{t_n} \mathbf{q} = iJ_{\mathbf{q}}^{\text{GP}} \nabla_{\mathbf{q}}^{g^{\text{GP}}} \mathcal{H}_n^{\text{GP}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \left(\frac{\delta \mathcal{H}_n^{\text{GP}}(\mathbf{q})}{\delta q}, \frac{\delta \mathcal{H}_n^{\text{GP}}(\mathbf{q})}{\delta \bar{q}} \right)^T.$$

The Poisson bracket is

$$\{\mathcal{F}, \mathcal{G}\}^{\text{GP}}(\mathbf{q}) = \int_{\mathbb{R}} \left(\frac{\delta \mathcal{F}}{\delta q} \frac{\delta \mathcal{G}}{\delta \bar{q}} - \frac{\delta \mathcal{F}}{\delta \bar{q}} \frac{\delta \mathcal{G}}{\delta q} \right) \, dx.$$

B.2 hGP

We consider formally a manifold where we denote points by $w = (w_+, w_-)$ and tangent vectors by $u = (u_+, u_-)$ and $v = (v_+, v_-)$. We define the Hamiltonian operator J^{hGP} , the Riemannian metric g^{hGP} and the symplectic form ω^{hGP} by

$$\begin{aligned} J_w^{\text{hGP}} &= \frac{1}{8} \begin{pmatrix} -\partial_x a^{-1} + 4a\partial_x^{-1} & \partial_x a^{-1} + 4a\partial_x^{-1} \\ -\partial_x a^{-1} - 4a\partial_x^{-1} & \partial_x a^{-1} - 4a\partial_x^{-1} \end{pmatrix} \\ g_w^{\text{hGP}}(u, v) &= \int_{\mathbb{R}} \frac{1}{4} (u_+ - u_-)(v_+ - v_-) + a^2 \partial_x^{-1} (u_+ + u_-) \partial_x^{-1} (v_+ + v_-) \, dx \\ \omega_w^{\text{hGP}}(u, v) &= \int_{\mathbb{R}} a(v_+ - v_-) \partial_x^{-1} (u_+ + u_-) - a(u_+ - u_-) \partial_x^{-1} (v_+ + v_-) \, dx. \end{aligned}$$

They satisfy

$$\omega_w^{\text{hGP}}(u, J_w^{\text{hGP}} v) = g_w^{\text{hGP}}(u, v).$$

We define the gradient $\nabla_w^{g^{\text{hGP}}}$ with respect to the metric g^{hGP} as well as the functional derivative via

$$g_w^{\text{hGP}}(v, \nabla_w^{g^{\text{hGP}}} \mathcal{H}) = \int_{\mathbb{R}} (v_+, v_-) \left(\frac{\delta \mathcal{H}(w)}{\delta w_+}, \frac{\delta \mathcal{H}(w)}{\delta w_-} \right)^T dx = \frac{d}{ds} \Big|_{s=0} \mathcal{H}(w + sv).$$

This implies

$$\nabla_w^{g^{\text{hGP}}} \mathcal{H} = \frac{1}{4} \begin{pmatrix} 4 - \partial_x a^{-2} \partial_x & -4 - \partial_x a^{-2} \partial_x \\ -4 - \partial_x a^{-2} \partial_x & 4 - \partial_x a^{-2} \partial_x \end{pmatrix} \left(\frac{\delta \mathcal{H}(w)}{\delta w_+}, \frac{\delta \mathcal{H}(w)}{\delta w_-} \right)^T.$$

Then (hGP_n) can be written as

$$\partial_{t_n} w = J_w^{\text{hGP}} \nabla_w^{g^{\text{hGP}}} \mathcal{H}_n^{\text{hGP}} = \frac{1}{4} \begin{pmatrix} -\partial_x a^{-1} - a^{-1} \partial_x & \partial_x a^{-1} - a^{-1} \partial_x \\ -\partial_x a^{-1} + a^{-1} \partial_x & \partial_x a^{-1} + a^{-1} \partial_x \end{pmatrix} \left(\frac{\delta \mathcal{H}_n^{\text{hGP}}(w)}{\delta w_+}, \frac{\delta \mathcal{H}_n^{\text{hGP}}(w)}{\delta w_-} \right)^T.$$

The Poisson bracket is

$$\{\mathcal{F}, \mathcal{G}\}^{\text{hGP}}(w) = \frac{1}{4} \int_{\mathbb{R}} \begin{pmatrix} \frac{\delta \mathcal{F}}{\delta w_+} & \frac{\delta \mathcal{F}}{\delta w_-} \end{pmatrix} \begin{pmatrix} -\partial_x a^{-1} - a^{-1} \partial_x & \partial_x a^{-1} - a^{-1} \partial_x \\ -\partial_x a^{-1} + a^{-1} \partial_x & \partial_x a^{-1} + a^{-1} \partial_x \end{pmatrix} \begin{pmatrix} \frac{\delta \mathcal{G}}{\delta w_+} \\ \frac{\delta \mathcal{G}}{\delta w_-} \end{pmatrix} dx.$$

B.3 εGP

We consider formally a manifold where we denote points by $W = (W_+, W_-)$ and tangent vectors by $U = (U_+, U_-)$ and $V = (V_+, V_-)$. We define the Hamiltonian operator $J^{\varepsilon\text{GP}}$, the Riemannian metric $g^{\varepsilon\text{GP}}$ and the symplectic form $\omega^{\varepsilon\text{GP}}$ by

$$\begin{aligned} J_W^{\varepsilon\text{GP}} &= \frac{1}{8} \begin{pmatrix} -\sqrt{2}\varepsilon \partial_X A^{-1} + \frac{4}{\sqrt{2}\varepsilon} A \partial_X^{-1} & \sqrt{2}\varepsilon \partial_X A^{-1} + \frac{4}{\sqrt{2}\varepsilon} A \partial_X^{-1} \\ -\sqrt{2}\varepsilon \partial_X A^{-1} - \frac{4}{\sqrt{2}\varepsilon} A \partial_X^{-1} & \sqrt{2}\varepsilon \partial_X A^{-1} - \frac{4}{\sqrt{2}\varepsilon} A \partial_X^{-1} \end{pmatrix} \\ g_W^{\varepsilon\text{GP}}(U, V) &= \int_{\mathbb{R}} \frac{\varepsilon^3}{4\sqrt{2}} (U_+ - U_-)(V_+ - V_-) + \frac{\varepsilon}{2\sqrt{2}} A^2 \partial_X^{-1} (U_+ + U_-) \partial_X^{-1} (V_+ + V_-) dX \\ \omega_W^{\varepsilon\text{GP}}(U, V) &= \frac{\varepsilon^2}{2} \int_{\mathbb{R}} A((V_+ - V_-) \partial_X^{-1} (U_+ + U_-) - (U_+ - U_-) \partial_X^{-1} (V_+ + V_-)) dX. \end{aligned}$$

They satisfy

$$\omega_W^{\varepsilon\text{GP}}(U, J_W^{\varepsilon\text{GP}} V) = g_W^{\varepsilon\text{GP}}(U, V).$$

We define the gradient $\nabla_W^{g^{\varepsilon\text{GP}}}$ with respect to the metric $g^{\varepsilon\text{GP}}$ as well as the functional derivative via

$$g_W^{\varepsilon\text{GP}}(V, \nabla_W^{g^{\varepsilon\text{GP}}} \mathcal{H}) = \int_{\mathbb{R}} (V_+, V_-) \left(\frac{\delta \mathcal{H}(W)}{\delta W_+}, \frac{\delta \mathcal{H}(W)}{\delta W_-} \right)^T dX = \frac{d}{ds} \Big|_{s=0} \mathcal{H}(W + sV).$$

This implies

$$\nabla_W^{g^{\varepsilon\text{GP}}} \mathcal{H} = \frac{1}{\sqrt{2}\varepsilon^3} \begin{pmatrix} 2 - \varepsilon^2 \partial_X A^{-2} \partial_X & -2 - \varepsilon^2 \partial_X A^{-2} \partial_X \\ -2 - \varepsilon^2 \partial_X A^{-2} \partial_X & 2 - \varepsilon^2 \partial_X A^{-2} \partial_X \end{pmatrix} \left(\frac{\delta \mathcal{H}(W)}{\delta W_+}, \frac{\delta \mathcal{H}(W)}{\delta W_-} \right)^T.$$

Then (εGP_n) can be written as

$$\partial_{T_n} W = \frac{dt_n}{dT_n} J_W^{\varepsilon\text{GP}} \nabla_W^{g^{\varepsilon\text{GP}}} \mathcal{H}_n^{\varepsilon\text{GP}} = \frac{dt_n}{dT_n} \frac{1}{2\varepsilon^2} \begin{pmatrix} -\partial_X A^{-1} - A^{-1} \partial_X & \partial_X A^{-1} - A^{-1} \partial_X \\ -\partial_X A^{-1} + A^{-1} \partial_X & \partial_X A^{-1} + A^{-1} \partial_X \end{pmatrix} \left(\frac{\delta \mathcal{H}_n^{\varepsilon\text{GP}}(W)}{\delta W_+}, \frac{\delta \mathcal{H}_n^{\varepsilon\text{GP}}(W)}{\delta W_-} \right)^T.$$

The Poisson bracket is

$$\{\mathcal{F}, \mathcal{G}\}^{\varepsilon\text{GP}}(W) = \frac{1}{2\varepsilon^2} \int_{\mathbb{R}} \begin{pmatrix} \frac{\delta \mathcal{F}}{\delta W_+} & \frac{\delta \mathcal{F}}{\delta W_-} \end{pmatrix} \begin{pmatrix} -\partial_X A^{-1} - A^{-1} \partial_X & \partial_X A^{-1} - A^{-1} \partial_X \\ -\partial_X A^{-1} + A^{-1} \partial_X & \partial_X A^{-1} + A^{-1} \partial_X \end{pmatrix} \begin{pmatrix} \frac{\delta \mathcal{G}}{\delta W_+} \\ \frac{\delta \mathcal{G}}{\delta W_-} \end{pmatrix} dX.$$

B.4 KdV

We consider formally a manifold where we denote points by U and tangent vectors by F and G . We define the Hamiltonian operator J^{KdV} , the Riemannian metric g^{KdV} and the symplectic form ω^{KdV} by

$$\begin{aligned} J_U^{\text{KdV}} &= \frac{1}{2} \partial_X \\ g_U^{\text{KdV}}(F, G) &= \int_{\mathbb{R}} FG \, dX \\ \omega_U^{\text{KdV}} &= \frac{1}{2} \int_{\mathbb{R}} F \partial_X^{-1} G - G \partial_X^{-1} F \, dX. \end{aligned}$$

They satisfy

$$\omega_U^{\text{KdV}}(F, J_U^{\text{KdV}} G) = g_U^{\text{KdV}}(F, G).$$

We define the gradient $\nabla_U^{g^{\text{KdV}}}$ with respect to the metric g^{KdV} as well as the functional derivative via

$$g_U^{\text{KdV}}(F, \nabla_U^{g^{\text{KdV}}} \mathcal{H}) = \int_{\mathbb{R}} F \frac{\delta \mathcal{H}(U)}{\delta U} \, dX = \left. \frac{d}{ds} \right|_{s=0} \mathcal{H}(U + sF).$$

This implies

$$\nabla_U^{g^{\text{KdV}}} \mathcal{H} = \frac{\delta \mathcal{H}(U)}{\delta U}.$$

Then (KdV_n) can be written as

$$\partial_{T_n} U = J_U^{\text{KdV}} \nabla_U^{g^{\text{KdV}}} \mathcal{E}_n^{\text{KdV}} = \frac{1}{2} \partial_X \frac{\delta \mathcal{E}_n^{\text{KdV}}(U)}{\delta U}.$$

The Poisson bracket is

$$\{\mathcal{F}, \mathcal{G}\}^{\text{KdV}}(U) = \frac{1}{2} \int_{\mathbb{R}} \frac{\delta \mathcal{F}}{\delta U} \partial_X \frac{\delta \mathcal{G}}{\delta U} \, dX.$$

Statements and Declarations.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

No datasets were generated or analysed during the current study.

The author has no competing interests to declare that are relevant to the content of this article.

SymPy [70] was used to carry out symbolic computations for the purposes of exploration and verification of various identities involving recurrence relations. Various versions of the OpenAI services ChatGPT and Codex were used for mathematical and grammatical proofreading, literature search, and exploratory discussion.

References

- [1] M. J. Ablowitz and P. A. Clarkson. *Solitons, nonlinear evolution equations and inverse scattering*. Vol. 149. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1991, pp. xii+516.
- [2] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur. “The inverse scattering transform-Fourier analysis for nonlinear problems”. In: *Stud. Appl. Math.* 53.4 (1974), pp. 249–315.

- [3] J. Adams. “Local and global well-posedness for the higher-order NLS & dNLS hierarchy equations”. English. Dissertation (Ph.D.)—Heinrich Heine Universität Düsseldorf; supervisors: A. Grünrock and J. Saal. PhD thesis. Düsseldorf: Heinrich Heine Universität Düsseldorf, 2025.
- [4] J. Adams. “Well-posedness for the NLS hierarchy”. In: *J. Evol. Equ.* 24.4 (2024), p. 88.
- [5] H. Aratyn and A. S. Sorin, eds. *Integrable hierarchies and modern physical theories*. Vol. 18. NATO Science Series II: Mathematics, Physics and Chemistry. Kluwer Academic Publishers, Dordrecht, 2001, pp. x+435.
- [6] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Grundlehren der mathematischen Wissenschaften. Springer, Heidelberg, 2011, pp. xvi+523.
- [7] D. Bambusi. “Hamiltonian studies on counter-propagating water waves”. In: *Water Waves* 3.1 (2021), pp. 49–83.
- [8] W. Bao. “The nonlinear Schrödinger equation and applications in Bose-Einstein condensation and plasma physics”. In: *Dynamics in models of coarsening, coagulation, condensation and quantization*. Vol. 9. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. World Sci. Publ., Hackensack, NJ, 2007, pp. 141–239.
- [9] L. Bergé. “Wave collapse in physics: principles and applications to light and plasma waves”. In: *Phys. Rep.* 303.5-6 (1998), pp. 259–370.
- [10] F. Béthuel, R. Danchin, and D. Smets. “On the linear wave regime of the Gross-Pitaevskii equation”. In: *J. Anal. Math.* 110 (2010), pp. 297–338.
- [11] F. Béthuel, P. Gravejat, and J.-C. Saut. “Existence and properties of travelling waves for the Gross-Pitaevskii equation”. In: *Stationary and time dependent Gross-Pitaevskii equations*. Vol. 473. Contemp. Math. Amer. Math. Soc., Providence, RI, 2008, pp. 55–103.
- [12] F. Béthuel, P. Gravejat, J.-C. Saut, and D. Smets. “On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation II”. In: *Comm. Partial Differ. Equ.* 35.1 (2010), pp. 113–164.
- [13] F. Béthuel, P. Gravejat, J.-C. Saut, and D. Smets. “On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation. I”. In: *Int. Math. Res. Not. IMRN* 14 (2009), pp. 2700–2748.
- [14] T. J. Bridges. “A universal form for the emergence of the Korteweg-de Vries equation”. In: *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 469.2153, 20120707 (2013), p. 15.
- [15] B. Bringmann, R. Killip, and M. Visan. “Global well-posedness for the fifth-order KdV equation in $H^{-1}(\mathbb{R})$ ”. In: *Ann. PDE* 7.2 (2021), p. 21.
- [16] W. Bulla, F. Gesztesy, and K. Unterkofler. “Holomorphy of the scattering matrix with respect to c^{-2} for Dirac operators and an explicit treatment of relativistic corrections”. In: *Comm. Math. Phys.* 144.2 (1992), pp. 391–416.
- [17] D. Calogero. “Why Are Certain Nonlinear PDEs Both Widely Applicable and Integrable?” In: *What Is Integrability?* Ed. by V. E. Zakharov. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 1–62.
- [18] J. Chandra and P. W. Davis. “Linear generalizations of Gronwall’s inequality”. In: *Proc. Amer. Math. Soc.* 60 (1976), pp. 157–160.
- [19] M. Chirilus-Bruckner, W.-P. Düll, and G. Schneider. “Validity of the KdV equation for the modulation of periodic traveling waves in the NLS equation”. In: *J. Math. Anal. Appl.* 414.1 (2014), pp. 166–175.
- [20] D. Chiron and F. Rousset. “The KdV/KP-I limit of the nonlinear Schrödinger equation”. In: *SIAM J. Math. Anal.* 42.1 (2010), pp. 64–96.
- [21] W. Craig. “An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits”. In: *Comm. Partial Differential Equations* 10.8 (1985), pp. 787–1003.
- [22] W. Craig and M. D. Groves. “Hamiltonian long-wave approximations to the water-wave problem”. In: *Wave Motion* 19.4 (1994), pp. 367–389.
- [23] H. Demiray. “Modified reductive perturbation method as applied to long water-waves: the Korteweg-de Vries hierarchy”. In: *Int. J. Nonlinear Sci.* 6.1 (2008), pp. 11–20.

- [24] F. Demontis, B. Prinari, C. van der Mee, and F. Vitale. “The inverse scattering transform for the defocusing nonlinear Schrödinger equations with nonzero boundary conditions”. In: *Stud. Appl. Math.* 131.1 (2013), pp. 1–40.
- [25] L. A. Dickey. *Soliton Equations and Hamiltonian Systems*. Vol. 12. Advanced Series in Mathematical Physics. Singapore: World Scientific, 1991, p. 310.
- [26] R. Donagi and T. Shaska, eds. *Integrable systems and algebraic geometry. Vol. 1*. Vol. 458. London Mathematical Society Lecture Note Series. A celebration of Emma Previato’s 65th birthday. Cambridge University Press, Cambridge, 2020, pp. viii+412.
- [27] R. Donagi and T. Shaska, eds. *Integrable systems and algebraic geometry. Vol. 2*. Vol. 459. London Mathematical Society Lecture Note Series. A celebration of Emma Previato’s 65th birthday. Cambridge University Press, Cambridge, 2020, pp. viii+528.
- [28] G. A. El and M. A. Hoefer. “Dispersive shock waves and modulation theory”. In: *Phys. D* 333 (2016), pp. 11–65.
- [29] L. Erdős, B. Schlein, and H.-T. Yau. “Derivation of the Gross-Pitaevskii hierarchy for the dynamics of Bose-Einstein condensate”. In: *Comm. Pure Appl. Math.* 59.12 (2006), pp. 1659–1741.
- [30] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian methods in the theory of solitons*. English. Classics in Mathematics. Translated from the 1986 Russian original by Alexey G. Reyman. Springer, Berlin, 2007, pp. x+592.
- [31] R.P. Feynman. “Superfluidity and superconductivity”. In: *Rev. Mod. Phys.* 29.2 (1957), pp. 205–212.
- [32] D. Franco, S. Giorgini, L.P. Pitaevskii, and S. Stringari. “Theory of Bose-Einstein condensation in trapped gases”. In: *Rev. Mod. Phys.* 71.3 (1999), pp. 463–512.
- [33] C.L. Gardner. “The quantum hydrodynamic model for semiconductor devices”. In: *SIAM J. Appl. Math.* 54.2 (1994), pp. 409–427.
- [34] F. Gesztesy, H. Grosse, and B. Thaller. “A rigorous approach to relativistic corrections of bound state energies for spin-1/2 particles”. In: *Ann. Inst. Henri Poincaré Phys. Théor.* 40.2 (1984), pp. 159–174.
- [35] F. Gesztesy and H. Holden. *Soliton equations and their algebro-geometric solutions. Vol. I*. Vol. 79. Cambridge Studies in Advanced Mathematics. (1 + 1)-dimensional continuous models. Cambridge University Press, Cambridge, 2003, pp. xii+505.
- [36] F. Gesztesy, H. Holden, J. Michor, and G. Teschl. *Soliton equations and their algebro-geometric solutions. Vol. II*. Vol. 114. Cambridge Studies in Advanced Mathematics. (1 + 1)-dimensional discrete models. Cambridge University Press, Cambridge, 2008, pp. x+438.
- [37] J. Grant. “Pressure and stress tensor expressions in the fluid mechanical formulation of the Bose condensate equations”. In: *J. phys. A: Math. Nucl. Gen.* 6.11 (1973), pp. L151–L153.
- [38] A. Grünrock. “On the hierarchies of higher order mKdV and KdV equations”. In: *Cent. Eur. J. Math.* 8.3 (2010), pp. 500–536.
- [39] B. Harrop-Griffiths, R. Killip, and M. Vişan. “Sharp well-posedness for the cubic NLS and mKdV in $H^s(\mathbb{R})$ ”. In: *Forum Math. Pi* 12 (2024), e6.
- [40] A. Hasegawa. “An historical review of application of optical solitons for high speed communications”. In: vol. 10. 3. Optical solitons—perspectives and applications. 2000, pp. 475–485.
- [41] Y. Hiraoka and Y. Kodama. “Normal form and solitons”. In: *Integrability*. Vol. 767. Lecture Notes in Phys. Springer, Berlin, 2009, pp. 175–214.
- [42] S. Hofbauer. “Long wave approximation over and beyond the natural time scale”. PhD thesis. University of Stuttgart, Apr. 2024.
- [43] R. I. Ivanov. “Hamiltonian models for the propagation of long gravity waves, higher-order KdV-type equations and integrability”. In: *Nonlinear dispersive waves*. Adv. Math. Fluid Mech. Birkhäuser/Springer, Cham, 2024, pp. 81–97.
- [44] M. Schwarz Jr. “The initial value problem for the sequence of generalized Korteweg-de Vries equations”. In: *Adv. Math.* 54.1 (1984), pp. 22–56.
- [45] L. A. Kalyakin. “Long-wave asymptotics. Integrable equations as the asymptotic limit of nonlinear systems”. In: *Uspekhi Mat. Nauk* 44.1(265) (1989), pp. 5–34, 247.

- [46] C. E. Kenig and D. Pilod. “Local well-posedness for the KdV hierarchy at high regularity”. In: *Adv. Differ. Equ.* 21.9-10 (2016), pp. 801–836.
- [47] C. E. Kenig, G. Ponce, and L. Vega. “Higher-order nonlinear dispersive equations”. In: *Proc. Amer. Math. Soc.* 122.1 (1994), pp. 157–166.
- [48] C. E. Kenig, G. Ponce, and L. Vega. “On the generalized Benjamin-Ono equation”. In: *Trans. Amer. Math. Soc.* 342.1 (1994), pp. 155–172.
- [49] C. E. Kenig, G. Ponce, and L. Vega. “On the hierarchy of the generalized KdV equations”. In: *Singular limits of dispersive waves (Lyon, 1991)*. Vol. 320. NATO Adv. Sci. Inst. Ser. B: Phys. Plenum, New York, 1994, pp. 347–356.
- [50] C. E. Kenig, G. Ponce, and L. Vega. “Oscillatory integrals and regularity of dispersive equations”. In: *Indiana Univ. Math. J.* 40.1 (1991), pp. 33–69.
- [51] C. E. Kenig, G. Ponce, and L. Vega. “Small solutions to nonlinear Schrödinger equations”. In: *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* 10.3 (1993), pp. 255–288.
- [52] C. E. Kenig, G. Ponce, and L. Vega. “Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle”. In: *Comm. Pure Appl. Math.* 46.4 (1993), pp. 527–620.
- [53] C. E. Kenig, G. Ponce, and L. Vega. “Well-posedness of the initial value problem for the Korteweg-de Vries equation”. In: *J. Amer. Math. Soc.* 4.2 (1991), pp. 323–347.
- [54] R. Killip, T. Laurens, and M. Viřan. “Sharp well-posedness for the Benjamin-Ono equation”. In: *Invent. Math.* 236.3 (2024), pp. 999–1054.
- [55] R. Killip and M. Viřan. “KdV is well-posed in H^{-1} ”. In: *Ann. of Math. (2)* 190.1 (2019), pp. 249–305.
- [56] R. Killip, M. Viřan, and X. Zhang. “Low regularity conservation laws for integrable PDE”. In: *Geom. Funct. Anal.* 28.4 (2018), pp. 1062–1090.
- [57] F. Klaus. “Nonlinear Schrödinger Equations with Rough Data”. PhD thesis. Karlsruher Institut für Technologie (KIT), 2022. 143 pp.
- [58] F. Klaus, H. Koch, and B. Liu. “Well-posedness for the KdV hierarchy”. In: *arXiv e-prints* (2023), pp. 1–139.
- [59] H. Koch and X. Liao. “Conserved energies for the one dimensional Gross-Pitaevskii equation”. In: *Adv. Math.* 377, 107467 (2021).
- [60] H. Koch and X. Liao. “Conserved energies for the one dimensional Gross-Pitaevskii equation: low regularity case”. In: *Adv. Math.* 420, 108996 (2023).
- [61] H. Koch and D. Tataru. “Conserved energies for the cubic nonlinear Schrödinger equation in one dimension”. In: *Duke Math. J.* 167.17 (2018), pp. 3207–3313.
- [62] Y. Kodama and T. Taniuti. “Higher order approximation in the reductive perturbation method. III. The weakly dissipative system”. In: *J. Phys. Soc. Japan* 47.5 (1979), pp. 1706–1716.
- [63] R. A. Kraenkel, M. A. Manna, and J. G. Pereira. “The Korteweg-de Vries hierarchy and long water-waves”. In: *J. Math. Phys.* 36.1 (1995), pp. 307–320.
- [64] L. Landau. “Theory of the superfluidity of Helium II”. In: *Phys. Rev.* 60.4 (1941), pp. 356–358.
- [65] P. D. Lax. “Periodic solutions of the KdV equation”. In: *Comm. Pure Appl. Math.* 28 (1975), pp. 141–188.
- [66] X. Liao and M. Plum. “Eigenvalue analysis of the lax operator for the one-dimensional cubic nonlinear defocusing Schrödinger equation”. In: *SIAM J. Math. Anal.* 55.6 (2023), pp. 7399–7425.
- [67] X. Liao and R. Wegner. *Global well-posedness of the NLS hierarchy with nonzero boundary condition*. 2025.
- [68] Y. Liu and X. Yang. “On the long-wave approximation for the Euler-Poisson system”. In: *Adv. Math.* 372 (2020), pp. 107300, 41.
- [69] M. I. Loffredo and L. M. Morato. “On the creation of quantized vortex lines in rotating He II”. In: *Il Nuovo Cimento B (1971-1996)* 108.2 (1993), pp. 205–215.
- [70] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, Š. Roučka, A. Saboo,

- I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz. “SymPy: symbolic computing in Python”. In: *PeerJ Comput. Sci.* 3 (Jan. 2017), e103.
- [71] T. Ozawa, S.-F. Tian, and Z.-J. Wu. “Orbital stability of dark solitons for the modified Korteweg–de Vries hierarchy with nonzero boundary conditions”. In: *Math. Ann.* 394.1 (2026), Paper No. 23, 51.
- [72] D. Pilod. “On the Cauchy problem for higher-order nonlinear dispersive equations”. In: *J. Differ. Equ.* 245.8 (2008), pp. 2055–2077.
- [73] B. Prinari. “Inverse scattering transform for nonlinear Schrödinger systems on a nontrivial background: a survey of classical results, new developments and future directions”. In: *J. Nonlinear Math. Phys.* 30.2 (2023), pp. 317–383.
- [74] A. Rybkin. “Regularized perturbation determinants and KdV conservation laws for irregular initial profiles”. In: *Topics in operator theory. Volume 2. Systems and mathematical physics*. Vol. 203. Oper. Theory Adv. Appl. Birkhäuser Verlag, Basel, 2010, pp. 427–444.
- [75] J.-C. Saut. “Quelques généralisations de l’équation de Korteweg-de Vries. II”. In: *J. Differential Equations* 33.3 (1979), pp. 320–335.
- [76] G. Schneider. “Approximation of the Korteweg-de Vries equation by the nonlinear Schrödinger equation”. In: *J. Differential Equations* 147.2 (1998), pp. 333–354.
- [77] G. Schneider. “The long wave limit for a Boussinesq equation”. In: *SIAM J. Appl. Math.* 58.4 (1998), pp. 1237–1245.
- [78] G. Schneider and H. Uecker. *Nonlinear PDEs*. Vol. 182. Graduate Studies in Mathematics. A dynamical systems approach. American Mathematical Society, Providence, RI, 2017, pp. xiii+575.
- [79] G. Schneider and C. E. Wayne. “The long-wave limit for the water wave problem. I. The case of zero surface tension”. In: *Comm. Pure Appl. Math.* 53.12 (2000), pp. 1475–1535.
- [80] G. Schneider and C. E. Wayne. “The rigorous approximation of long-wavelength capillary-gravity waves”. In: *Arch. Ration. Mech. Anal.* 162.3 (2002), pp. 247–285.
- [81] T. Taniuti. “Reductive Perturbation Method and Far Fields of Wave Equations”. In: *Progress of Theoretical Physics Supplement* 55 (Jan. 1974), pp. 1–35.
- [82] B. Thaller. *The Dirac equation*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1992, pp. xviii+357.
- [83] R. Wegner. “Global-in-time well-posedness of the one-dimensional hydrodynamic Gross-Pitaevskii equations without vacuum”. In: *Z. Angew. Math. Phys.* 74.5 (2023), p. 194.
- [84] G. B. White. “Splitting of the Dirac operator in the nonrelativistic limit”. In: *Ann. Inst. H. Poincaré Phys. Théor.* 53.1 (1990), pp. 109–121.
- [85] G. B. Whitham. *Linear and nonlinear waves*. Pure and Applied Mathematics (New York). Reprint of the 1974 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. xviii+636.
- [86] G. B. Whitham. “Non-linear dispersive waves”. In: *Proc. Roy. Soc. Lond. Ser. A* 283 (1965), pp. 238–261.
- [87] V. E. Zakharov and A. B. Shabat. “Interactions between solitons in a stable medium”. In: *Sov. Phys. JETP* 37 (1973), pp. 823–828.

Robert Wegner

Institute for Analysis
 Karlsruhe Institute of Technology
 Englerstraße 2
 76131 Karlsruhe, Germany
robert.wegner@kit.edu