Annals of Mathematics **174** (2011), 757–857
http://dx.doi.org/10.4007/annals.2011.174.2.2

Description of two soliton collision for the quartic gKdV equation

By Yvan Martel and Frank Merle

Abstract

In this paper, we give the first description of the collision of two solitons for a nonintegrable equation in a special regime. We consider solutions of the quartic gKdV equation $\partial_t u + \partial_x(\partial_x^2 u + u^4) = 0$, which behave as $t \to -\infty$ like

$$u(t,x) = Q_{c_1}(x - c_1 t) + Q_{c_2}(x - c_2 t) + \eta(t,x),$$

where $Q_c(x - ct)$ is a soliton and $\|\eta(t)\|_{H^1} \ll \|Q_{c_2}\|_{H^1} \ll \|Q_{c_1}\|_{H^1}$.

The global behavior of $u(t)$ is given by the following stability result: for all $t \in \mathbb{R}$, $u(t,x) = Q_{c_1(t)}(x - y_1(t)) + Q_{c_2(t)}(x - y_2(t)) + \eta(t,x)$, where $\|\eta(t)\|_{H^1} \ll \|Q_{c_2}\|_{H^1}$ and $\lim_{t\to+\infty} c_1(t) = c_1^+$, $\lim_{t\to+\infty} c_2(t) = c_2^+$.

In the case where $u(t)$ is a pure 2-soliton solution as $t \to -\infty$ (i.e. $\lim_{t\to-\infty} \|\eta(t)\|_{H^1} = 0$), we obtain $c_1^+ > c_1$, $c_2^+ < c_2$ and for the residual part, $\lim_{t\to+\infty} \|\eta(t)\|_{H^1} > 0$. Therefore, in contrast with the integrable KdV equation (or mKdV equation), no global pure 2-soliton solution exists and the collision is inelastic. A different notion of global 2-soliton is then proposed.

1. Introduction

We consider the generalized Korteweg-de Vries (gKdV) equations:

$$\partial_t u + \partial_x(\partial_x^2 u + u^p) = 0, \quad x, t \in \mathbb{R}, \tag{1.1}$$

in the subcritical case, i.e. for $p = 2$, 3 or 4. Our main results concern the nonintegrable case $p = 4$. An extension to the case of a general nonlinearity $f(u)$ for which the traveling waves are stable is considered in [26].

It is well-known that the Cauchy problem for equation (1.1) is globally well-posed in the energy space $H^1(\mathbb{R})$ (see Kenig, Ponce and Vega [15]): for any $u_0 \in H^1(\mathbb{R})$, there exists a unique solution $u(t) \in C(\mathbb{R}, H^1(\mathbb{R}))$ of (1.1) with $u(0) = u_0$, uniformly bounded in $H^1(\mathbb{R})$. Moreover, the following quantities

This research was supported in part by the Agence Nationale de la Recherche (ANR ONDENONLIN).

are conserved (whenever they are well-defined):

$$\int u(t) = \int u(0), \quad \int u^2(t) = \int u^2(0), \tag{1.2}$$

$$E(u(t)) = \frac{1}{2}\int u_x^2(t) - \frac{1}{p+1}\int u^{p+1}(t) = \frac{1}{2}\int u_x^2(0) - \frac{1}{p+1}\int u^{p+1}(0). \tag{1.3}$$

Recall that for $p=2,3,4$, global well-posedness follows from local well-posedness, (1.2)–(1.3) and the Gagliardo-Nirenberg inequality: for all $v \in H^1$, $\int |v|^{p+1} \leq C\left(\int v^2\right)^{\frac{p+3}{4}}\left(\int v_x^2\right)^{\frac{p-1}{4}}$.

There exist explicit traveling wave solutions of (1.1). Denote by Q the unique even solution of

(1.4)

$$Q > 0, \quad Q''+Q^p = Q, \quad Q \in H^1(\mathbb{R}) \text{ given by } Q(x) = \left(\frac{p+1}{2\cosh^2\left(\frac{p-1}{2}\,x\right)}\right)^{\frac{1}{p-1}},$$

and, for any $c > 0$, let

$$Q_c(x) = c^{\frac{1}{p-1}}Q(\sqrt{c}\,x) \quad \text{be solution of } Q_c'' + Q_c^p = c\,Q_c. \tag{1.5}$$

Then, for any $\delta_0 \in \mathbb{R}$, $c > 0$, the functions $R_{c,\delta_0}(t,x) = Q_c(x - \delta_0 - ct)$ are single soliton solutions of (1.1). These solutions have been intensively studied, especially in the integrable cases $p = 2$ and $p = 3$ in equation (1.1).

1.1. *Known results on soliton and multi-soliton solutions.*

a. *Integrable case $p = 2,3$: N-solitons for the* KdV *and* mKdV *equations.* Pioneering works of Fermi, Pasta and Ulam [10] and Zabusky and Kruskal [39] exhibited several remarkable phenomena related to soliton collision from the numerical point of view. Subsequently, Lax ([17]) developed a mathematical framework to study these problems, known now as complete integrability and the theory of Lax pairs. Many other developements appeared, such as the inverse scattering transform (for a review on this theory, we refer for example to Miura [28]).

This nonlinear transformation exhibits one of the most striking properties of the KdV and mKdV equations, which is the existence of pure N-soliton solutions (Hirota [13]). Namely, let $p = 2$ or $p = 3$, and let $c_1 > \cdots > c_N > 0$, $\delta_1, \ldots, \delta_N \in \mathbb{R}$. There exists an explicit multi-soliton solution $U(t,x)$ of (1.1) that satisfies

$$\left\|U(t,x) - \sum_{j=1}^N Q_{c_j}(. - c_jt - \delta_j)\right\|_{H^1} \underset{t\to-\infty}{\longrightarrow} 0,$$
$$\left\|U(t,x) - \sum_{j=1}^N Q_{c_j}(. - c_jt - \delta_j')\right\|_{H^1} \underset{t\to+\infty}{\longrightarrow} 0,$$

for some δ'_j such that the shifts $\Delta_j = \delta'_j - \delta_j$ depend on the (c_k). Explicit formulas for such solutions were derived using the inverse scattering transform. For example, the following function $U_{1,c}$, which is a solution of (1.1) with $p = 2$, is a 2-soliton solution $(0 < c < 1)$:

(1.6)
$$U_{1,c}(t,x) = 6\frac{\partial^2}{\partial x^2}\log\Big(1+e^{x-t}+e^{\sqrt{c}(x-ct)}+\alpha e^{x-t}e^{\sqrt{c}(x-ct)}\Big) \text{ with } \alpha = \left(\frac{1-\sqrt{c}}{1+\sqrt{c}}\right)^2.$$

The N-solitons are fundamental in studying the properties of general solutions of the KdV equation because of the following (Kruskal [16], Eckhaus and Schuur [9], [33], Cohen [5]):

Decomposition property ([9], [33], [5]). *Let $u(t)$ be a solution of* (1.1) *with $p = 2$. Suppose that $u(0) \in C^4(\mathbb{R})$ satisfies for $k \in \{0, \dots, 4\}$, for all $x \in \mathbb{R}$, $\left|(\partial^k u/\partial x^k)(0,x)\right| \le C/(1+|x|^{10})$. Then, there exist $N \in \mathbb{N}$, $\delta_1, \dots, \delta_N$ and $c_1 > \dots > c_N > 0$ such that for all $x > 0$,*

$$u(t,x) - \sum_{j=1}^{N} Q_{c_j}(x - \delta_j - c_j t) \to 0 \quad \text{as } t \to +\infty.$$

This result means that the asymptotic behavior for large time of any sufficiently regular and decaying solution is governed by a finite number of solitons.

b. PDE *results for the subcritical generalized* KdV *equations* $(p = 2, 3, 4)$. First, we recall the following well-known orbital stability result.

Stability of soliton for the gKdV equation ([1], [2], [4], [37]). *Let $1 < p < 5$. Let $u(t)$ be an H^1 solution of the* gKdV *equation* (1.1). *For all $\epsilon > 0$, there exists $\delta > 0$, such that if $\|u(0) - Q\|_{H^1} \le \delta$, then there exists a continuous function $t \in \mathbb{R} \mapsto \rho(t)$, such that $\|u(t) - Q(. - \rho(t))\|_{H^1(\mathbb{R})} \le \epsilon$.*

By invariance by scaling and translation of the gKdV equation, the result is the same for $Q_{c_0}(x-\delta_0)$, for any $c_0 > 0$, $\delta_0 \in \mathbb{R}$. The proof of this result relies only on the conservation laws (1.2)–(1.3) and the variational characterization of $Q(x)$ (see [4], [37]).

The family of solitons $(R_{c,\delta_0}(t,x))$ is actually asymptotically stable, for equation (1.1) in the subcritical case: $p = 2$, 3 or 4.

Asymptotic stability for the gKdV equation ([21], [22]). *Let $u(t)$ be an H^1 solution of* (1.1). *There exists $\alpha > 0$ such that if $\|u(0) - Q\|_{H^1} \le \alpha$, then there exist c^+ with $|c^+ - 1| = O(\alpha)$ and a C^1 function $\rho : [0, +\infty) \to \mathbb{R}$ such that*

(1.7) $$w(t,x) = u(t,x) - Q_{c^+}(x - \rho(t)) \quad \textit{satisfies} \quad \lim_{t\to+\infty} \|w(t)\|_{H^1(x > \frac{1}{10}t)} = 0.$$

Moreover, $\lim_{t\to+\infty} \frac{d\rho}{dt}(t) = c^+$.

We comment on the above notion of asymptotic stability. This result means that by taking α small enough, we know the behavior of $u(t)$ as $t \to +\infty$ in the energy topology H^1 but restricted to the space time region $x > \frac{1}{10}t$. Actually, the result can be stated for any region $x > \beta t$, for $\beta > 0$, provided α is small enough depending on β. Such regions of convergence are in some sense sharp since there exist solutions which behave asymptotically, as $t \to +\infty$, as the sum $Q(x-t) + Q_c(x-ct)$, where $c > 0$ is arbitrarily small. To get around this difficulty, one may also consider weighted spaces, as Pego and Weinstein [30], who proved the first result of asymptotic stability of solitons of (gKdV).

Note that the above result, proved only for $p = 2$, 3 and 4 in [21], [22] also holds for (1.1) with a general nonlinearity $f(u)$, see the more recent papers [20] and [23]. Stability and asymptotic stability results above can be extended to the sum of N solitons (and then to multi-solitons), when the various solitons are decoupled; see [27]. Moreover, assuming $\int_{x>0} x^2u^2 < +\infty$ implies that $\lim_{t\to+\infty}(\rho(t) - c^+t)$ exists (see [25] and §4.2 of the present paper).

Let us introduce the notion of asymptotic N-soliton solutions and pure N-soliton solution.

Definition 1. (1) A solution $u(t)$ of (1.1) is an *asymptotic N-soliton solution at $-\infty$* if there exist $c_1^- > \cdots > c_N^- > 0$ and $\rho_1^-(t), \ldots, \rho_N^-(t)$ such that

$$\lim_{t\to-\infty} \left\| u(t) - \sum_{j=1}^{N} Q_{c_j^-}(. - \rho_j^-(t)) \right\|_{H^1(\mathbb{R})} = 0. \tag{1.8}$$

(2) A solution $u(t)$ of (1.1) is an *asymptotic N-soliton solution at $+\infty$* if there exist $c_1^+ > \cdots > c_N^+ > 0$ and $\rho_1^+(t), \ldots, \rho_N^+(t)$ such that

$$\lim_{t\to+\infty} \left\| u(t) - \sum_{j=1}^{N} Q_{c_j^+}(. - \rho_j^+(t)) \right\|_{H^1(\mathbb{R})} = 0.$$

(3) An H^1 solution $u(t)$ of (1.1) is a *pure N-soliton solution* if $u(t)$ is an asymptotic N-soliton solution at both $+\infty$ and $-\infty$.

We recall the following existence result.

Asymptotic N-soliton solutions for the gKdV equation ([19]). *Let $p = 2$, 3 or 4. Let $N \geq 1$, $c_1 > \cdots > c_N > 0$, and $\delta_1, \ldots, \delta_N \in \mathbb{R}$. There exists a unique H^1 solution U of* (1.1) *such that*

$$\lim_{t\to-\infty} \left\| U(t) - \sum_{j=1}^{N} Q_{c_j}(. - \delta_j - c_jt) \right\|_{H^1(\mathbb{R})} = 0. \tag{1.9}$$

See Proposition 5.1 for more properties on $U(t)$. A similar statement holds true as $t \to +\infty$, since equation (1.1) is invariant under the transformation $x \to -x$, $t \to -t$.

This result means that there exist asymptotic N-soliton solutions at $-\infty$ for $p = 4$, similarly as in the integrable cases $p = 2, 3$. However, for $p = 4$, no information is known concerning the collision phenomenon or the behavior as $t \to +\infty$ for such solutions.

Recent works have completed the above asymptotic results. Côte ([6], [7]) has proved, for $p = 4$, 5, the existence of solutions satisfying

$$\lim_{t \to +\infty} \left\| u(t,x) - \sum_{j=1}^{N} Q_{c_j^+}(. - \delta_j - c_j t) - \mathcal{W}(t) v_0 \right\|_{H^1} = 0,$$

where $\mathcal{W}(t)$ is the linear Airy group and v_0 is a given function with suitable properties.

Tao [35] has established a well-posedness and scattering result (small data) for (1.1) with $p = 4$ in the critical space $\dot{H}^{-1/6}(\mathbb{R})$. As a corollary of the estimates in [35] and of the asymptotic stability result above, it follows that if u_0 is close to Q in $\dot{H}^{-1/6} \cap H^1$, then there exists $v_0 \in \dot{H}^{-1/6} \cap H^1$ such that

$$\lim_{t \to +\infty} \left\| u(t,x) - Q_{c^+}(. - \rho(t)) - \mathcal{W}(t) v_0 \right\|_{H^1} = 0.$$

1.2. *Motivation of the problem.* We consider in this paper the problem of collision of two solitary waves.

In the integrable case $p = 2, 3$, the explicit 2-soliton solutions provide a precise description of the collision phenomenon, including, for example, computations of the resulting shifts on the trajectories of the solitons (see e.g. [28]). In the nonintegrable situation, the collision problem has been an open question since the 70's. The first works concerning solitons in the field of nonlinear partial differential equations proved general existence and stability properties of single solitary waves. More recently, the PDE community has focused on interactions between single solitary waves and dispersion. So far, except in some integrable situations where explicit formulas for N-solitons are known, the question of the collision of two traveling waves for nonlinear PDE is completely open.

For generalized KdV equations, similarly as in the integrable case, one may conjecture that any general solution (under suitable assumptions) decomposes when time goes to $+\infty$ as a sum of decoupled solitons plus a dispersive residue. A natural question is to try to relate the decomposition as $t \to +\infty$ to the one as $t \to -\infty$ by understanding the interactions of the various parts of the solution. In this framework, the collision of two solitons seems to be the simplest case of interaction between nonlinear objects and thus a relevant question to understand the global behavior of solutions. This question is also related to a well-known open problem of Moser, which asks for the construction of N-solitary wave solutions of the problem of free surface water waves.

In addition to the integrability theory, the problem of interaction of nonlinear waves has been studied since the 60's from both experimental and numerical points of view.

First, Fermi, Pasta and Ulam [10], Zabusky and Kruskal [39] and Zabusky [38] have introduced nonlinear systems and computed interaction of nonlinear objects by numerics. Later, the theory of integrability justified these numerics as explained above. Since then, many other systems have been studied numerically in this context; we just quote a few works below. Bona et al. [3], and Kalisch and Bona [14] perform numerical studies of the problem of collision of two solitary waves for the Benjamin and the BBM equations. Shih [34] studied the case of the gKdV equation (1.1) with some half-integer values of p. Li and Sattinger [18] investigated the collision problem for the case of the Ion Acoustic Plasma equation, and Craig et al. [8] reported on numerics for the Euler equation with free surface. In all these works, the numerics show that, unlike for the pure solitons of the integrable case, the collision of two solitary waves fails to be elastic but only by a very small dispersion which is difficult to see numerically.

There is also an extensive literature devoted to experiments on water tanks. A key question is whether or not the collision between two solitary waves is elastic (equivalently, whether the collision is pure or generates dispersion). From experiments related to wave propagation in shallow water (see Weidman and Maxworthy [36], Hammack et al. [11], Craig et al. [8]), it seems that collisions are inelastic but very close to be elastic, for solitary waves of different amplitude.

Let us now review some more recent mathematical results related to these problems. First, Haragus and Sattinger [12] have studied perturbation of the KdV equation around the explicit N-soliton solutions, in particular the invertibility of the linearized operator around these solutions. Second, Mizumachi [29] for equation (1.1) with $p = 4$ has treated the case of two solitons with close sizes, in a situation of repulsive interaction without collision (using scattering techniques). Finally, the multi-soliton solutions of the NLS (nonlinear Schrödinger) model, with special nonlinearity and under spectral assumptions (ruling out the existence of small solitary waves), have been studied by Perelman [31] and Rodnianski, Schlag and Soffer [32]. Using Galilean invariance, speeds and sizes are independent (in particular, high speed is possible for size one solitary waves). Thus, one can consider the case where the collision has a negligeable effect on the solitary waves due to a very small time of interaction. In all these works, the interaction of two nonlinear objects in a nonperturbative case is not considered. In addition, up to now, no example of inelastic collision is known rigorously to exist.

1.3. *Main results.* Our main results in this paper concern the problem of collision of two solitons for (1.1) in the (nonintegrable) case $p = 4$. We consider the situation where one soliton, Q_{c_1}, is supposed to be large with respect to the other one, Q_{c_2}; thus we assume $c = c_2/c_1 \ll 1$. This is not a perturbative setting, related to the integrable case or to a linearized equation. In addition, the techniques of this paper can be applied in a general context for (1.1), $p = 2, 3, 4$ or with a general nonlinearity $f(u)$ (see [26]). In this situation, we are able to compute the interaction term during the collision up to any order of c, which allows us to describe very precisely the collision phenomenon.

First, this approach allows us to prove that for $p = 4$, there do not exist pure 2-soliton solutions in the regime $c_2 \ll c_1$: an asymptotic 2-soliton solution at $-\infty$ cannot be an asymptotic 2-soliton solution at $+\infty$.

Theorem 1.1 (Nonexistence of a pure 2-soliton solution for $p = 4$). *Let $c_1 > c_2 > 0$. There exists $\epsilon_0 > 0$ such that if $c = \frac{c_2}{c_1} < \epsilon_0$, then there exists no pure 2-soliton solution of (1.1) with speeds c_1, c_2 at $-\infty$.*

More precisely, let δ_1^-, $\delta_2^- \in \mathbb{R}$, and let $u(t)$ be the unique H^1 solution of (1.1) such that

$$\lim_{t\to-\infty} \|u(t) - Q_{c_1}(. - \delta_1^- - c_1 t) - Q_{c_2}(. - \delta_2^- - c_2 t)\|_{H^1} = 0.$$

Then, there exist δ_1^+, δ_2^+, $c_1^+ > c_2^+ > 0$ and $T_0, K > 0$ such that

$$w^+(t,x) = u(t,x) - Q_{c_1^+}(x - \delta_1^+ - c_1^+ t) - Q_{c_2^+}(x - \delta_2^+ - c_2^+ t)$$

satisfies

$$\lim_{t\to+\infty} \|w^+(t)\|_{H^1(x > \frac{1}{10}c_2 t)} = 0, \tag{1.10}$$

$$\tfrac{1}{K} c^{\frac{17}{6}} \le \frac{c_1^+}{c_1} - 1 \le K c^{\frac{11}{6}}, \quad \tfrac{1}{K} c^{\frac{8}{3}} \le 1 - \frac{c_2^+}{c_2} \le K c^{\frac{2}{3}}, \tag{1.11}$$

$$\tfrac{1}{K} c_1^{\frac{7}{12}} c^{\frac{17}{12}} \le \|\partial_x w^+(t)\|_{L^2} + \sqrt{c_1 c}\, \|w^+(t)\|_{L^2} \le K c_1^{\frac{7}{12}} c^{\frac{11}{12}}, \quad \textit{for } t \ge T_0. \tag{1.12}$$

Theorem 1.1 confirms the common belief that the existence of pure 2-soliton solutions, in particular, the elastic collision between two solitons, is a property which is specific to integrable models. However, we observe that the 2-soliton structure persists, in the sense that the slow soliton is not destroyed by the collision and remains approximately of the same size as $t \to +\infty$ (see also Remark 2 after the statement of Theorem 1.2).

The norm of $w^+(t)$ measures the distance of the solution to a pure 2-soliton solution for large time. The bound from below in (1.12) is thus a qualitative version of nonexistence of a pure 2-soliton solution. As a corollary of the proof, asymptotically in time, the minimal distance of any solution to a pure 2 soliton solution at $+\infty$ or at $-\infty$ is $Kc^{\frac{17}{12}}$ at $\pm\infty$, in the same sense as in (1.12). We also see from (1.11) how the speeds and the sizes of Q_{c_1} and Q_{c_2} are altered through the collision; the fast soliton accerelates while the slow soliton slows down.

This result is the first rigorous evidence of the nonintegrability of the equation from the dynamics of the solitary waves.

Remark 1. Using the invariant $\int u(t)$ of equation (1.1) in the framework of Theorem 1.1, one proves that $w^+(t)$ has to contain some dispersive part as $t \to +\infty$, in the sense that it does not converge to a pure sum of small solitons; i.e., $u(t)$ is not an asymptotic N-soliton solution at $+\infty$, for any $N \geq 1$ (see end of §5.1). See also Remark 2(4).

In spite of the nonexistence result above, we prove for $p = 4$ the existence of special solutions (but not unique; see comment 3 in the remarks below) related to the 2-soliton structure. These solutions are another illustration of the persistence of the 2-soliton structure through the collision and provide a different point of view on the collision.

THEOREM 1.2 (Existence of 2-soliton like solutions for $p = 4$). *Let $c_1 > c_2 > 0$. There exists $\epsilon_0 > 0$ such that if $c = \frac{c_2}{c_1} < \epsilon_0$, then there exist an H^1 solution $U(t)$ of* (1.1) *and Δ_1, $\Delta_2 \in \mathbb{R}$, satisfying, for all $t, x \in \mathbb{R}$,*

$$U(-t, -x) = U(t, x), \tag{1.13}$$

and such that the following holds for $w^\pm(t)$:

$$w^-(t, x) = U(t, x) - Q_{c_1}(x - c_1 t + \tfrac{1}{2}\Delta_1) - Q_{c_2}(x - c_2 t + \tfrac{1}{2}\Delta_2),$$
$$w^+(t, x) = U(t, x) - Q_{c_1}(x - c_1 t - \tfrac{1}{2}\Delta_1) - Q_{c_2}(x - c_2 t - \tfrac{1}{2}\Delta_2),$$

1. *Asymptotic behavior at $\pm\infty$:*

$$\lim_{t\to-\infty} \|w^-(t)\|_{H^1(x<\frac{c_2 t}{10})} = 0, \quad \lim_{t\to+\infty} \|w^+(t)\|_{H^1(x>\frac{c_2 t}{10})} = 0, \tag{1.14}$$

where the shifts Δ_1, Δ_2 satisfy $\Delta_1 < 0$, $\Delta_2 < 0$ and

$$\left|c_1^{\frac{1}{2}}\Delta_1 - c^{-\frac{1}{6}}\left(-2\frac{(\int Q)^2}{\int Q^2}\right)\right| + \left|c_1^{\frac{1}{2}}\Delta_2 - \left(\frac{1}{3}\frac{(\int Q)^2}{\int Q^2} - \int Q^3\right)\right| \leq Kc^{\frac{1}{12}}. \tag{1.15}$$

2. *Distance to the sum of two solitons: There exists $T_0 > 0$ such that,*

$$\frac{1}{K}c_1^{\frac{7}{12}}c^{\frac{17}{12}} \leq \|\partial_x w^+(t)\|_{L^2} + \sqrt{c_1 c}\,\|w^+(t)\|_{L^2} \leq Kc_1^{\frac{7}{12}}c^{\frac{17}{12}}, \quad \textit{for all } t \geq T_0. \tag{1.16}$$

Remark 2. (1) From the stability result of one soliton (variational argument), it follows immediately that the soliton Q_{c_1} is preserved up to a certain order through the collision by a slow soliton Q_{c_2}. What is quite surprising, and very similar to the integrable situation, is the fact that the second soliton, which is small, is also preserved by the collision (dynamical arguments). One could have expected the small soliton to be destroyed by such a collision.

Moreover, the solutions constructed in Theorem 1.2 describe precisely the effect of the collision on the two solitons: the speeds at $\pm\infty$ are the same and explicit formulas for the main order of the shifts on the trajectories of Q_{c_1} and Q_{c_2} are available. From the proof of Theorem 1.2, the shifts are a consequence of the collision and are observed in the relatively short period of time around the collision region.

Concerning the shifts, we point out two main differences with the integrable cases:

- The shift Δ_1 on Q_{c_1} and the shift Δ_2 on Q_{c_2} are both negative.
- The shift $\Delta_1 \to -\infty$ as $c = c_2/c_1 \to 0$, which means that the effect of the soliton Q_{c_2} on the trajectory of Q_{c_1} becomes larger when c_2/c_1 is smaller (note also that in this case the period of interaction is larger since the support of Q_{c_2} becomes larger).

(2) By the symmetry property of $U(t)$ (see (1.13)), a statement similar to (1.16) for w^- holds as $t \to -\infty$. Now, let $w_{\rho_1,\rho_2}(t) = U(t) - Q_{c_1}(. - \rho_1) - Q_{c_2}(. - \rho_2)$. Then, from the proof of Theorem 1.2, we also have for $|t|$ large,

$$\frac{1}{K} c_1^{\frac{7}{12}} c^{\frac{17}{12}} \leq \inf_{\rho_1,\rho_2 \in \mathbb{R}} \{\|\partial_x w_{\rho_1,\rho_2}(t)\|_{L^2} + \sqrt{c_1 c}\|w_{\rho_1,\rho_2}(t)\|_{L^2}\} \leq K c_1^{\frac{7}{12}} c^{\frac{17}{12}} \tag{1.17}$$

and

$$\inf_{\rho_1,\rho_2 \in \mathbb{R}} \{\|\partial_x w_{\rho_1,\rho_2}(t)\|_{L^2} + \sqrt{c_1}\|w_{\rho_1,\rho_2}(t)\|_{L^2}\} \leq K c_1^{\frac{7}{12}} c^{\frac{1}{3}}, \text{ for all } t \in \mathbb{R}. \tag{1.18}$$

Estimate (1.18) is sharp. Indeed, at $t = 0$, we have

$$\inf_{\rho_1,\rho_2 \in \mathbb{R}} \{\|\partial_x w_{\rho_1,\rho_2}(0)\|_{L^2} + \sqrt{c_1}\|w_{\rho_1,\rho_2}(0)\|_{L^2}\} \geq K_1 c_1^{\frac{7}{12}} c^{\frac{1}{3}}.$$

For $p = 4$, $\|Q_c\|_{L^2} = c^{\frac{1}{12}}\|Q\|_{L^2}$, this is to be compared with (1.17)–(1.16) giving sharp estimates of the distance of $U(t)$ to the sum of two solitons.

(3) By time and translation invariances, for all $\delta_1,\ \delta_2 \in \mathbb{R}$, one derives from Theorem 1.2 the existence of a solution $\varphi_{\delta_1,\delta_2}$ such that

$$\lim_{t\to-\infty} \|U_{\delta_1,\delta_2}(t) - Q_{c_1}(. - c_1 t - \delta_1 + \tfrac{1}{2}\Delta_1) \\ - Q_{c_2}(. - c_2 t - \delta_2 + \tfrac{1}{2}\Delta_2)\|_{H^1(x<\frac{c_2 t}{10})} = 0,$$

$$\lim_{t\to+\infty} \|U_{\delta_1,\delta_2}(t) - Q_{c_1}(. - c_1 t - \delta_1 - \tfrac{1}{2}\Delta_1) \\ - Q_{c_2}(. - c_2 t - \delta_2 - \tfrac{1}{2}\Delta_2)\|_{H^1(x>\frac{c_2 t}{10})} = 0.$$

From the proof of Theorem 1.2, there exist infinitely many solutions $U(t)$ satisfying the conclusions of Theorem 1.2 for given $c_1 > c_2 > 0$, δ_1, δ_2 (even with the symmetry assumption (1.13)). Indeed, it is enough to perturb the initial data $U(0)$ in a suitable way to obtain a solution with similar properties (see proof of Theorem 1.2).

Finally, we point out that the solution $U(t)$ which we have constructed belongs to H^s for all $s \geq 0$.

(4) Remark 1 also applies to the solution $U(t)$ constructed in Theorem 1.2; i.e., $U(t)$ has some dispersive part as $t \to \pm\infty$. Using Tao [35] (specific for $p = 4$), we should obtain some more information on the solution since $U(0) \in L^{\frac{3}{2}}(\mathbb{R})$. Indeed, $U(t)$ is conjectured to satisfy, for some $v_0 \in H^1$,

(1.19)

$$\lim_{t\to-\infty} \|U(t) - Q_{c_1}(. - c_1 t + \tfrac{1}{2}\Delta_1) - Q_{c_2}(. - c_2 t + \tfrac{1}{2}\Delta_2) - \mathcal{W}(t)v_0\|_{H^1} = 0,$$
$$\lim_{t\to+\infty} \|U(t) - Q_{c_1}(. - c_1 t - \tfrac{1}{2}\Delta_1) - Q_{c_2}(. - c_2 t - \tfrac{1}{2}\Delta_2) - \mathcal{W}(t)v_0\|_{H^1} = 0,$$

where $K_1 c_1^{\frac{7}{12}} c^{\frac{17}{12}} \leq \|\partial_x v_0\|_{L^2} + \sqrt{c_1 c}\, \|v_0\|_{L^2} \leq K_2 c_1^{\frac{7}{12}} c^{\frac{17}{12}}$.

We further conjecture that there exists a universal v_0, minimizer of a certain functional related to energy quantities (for example $\int (\partial_x v_0)^2 + c_1\, c \int v_0^2$). This function v_0 should have additional special properties, such as smoothness and exponential decay in space.

(5) Precise information concerning the solution $U(t)$ at $t = 0$ can be obtained from the proof of Theorem 1.2. See, in particular, Theorem 2.1.

Finally, the behavior of such solutions is proved to be stable in H^1, which means that if a solution $u(t)$ of (1.1) is close to the solution U constructed above at $t = 0$, then $u(t)$ has a 2-soliton structure for all time.

THEOREM 1.3 (Stability of the 2-soliton structure for $p = 4$). *Let $c_1 > c_2 > 0$. Assume that $c = \frac{c_2}{c_1} < \epsilon_0$ is small enough and let $U(t)$ be the function constructed in Theorem* 1.2. *Let $u(t)$ be an H^1 solution of* (1.1) *such that for some $\delta > 0$,*

$$\|\partial_x u(0) - \partial_x \varphi(0)\|_{L^2} + \sqrt{c_1}\|u(0) - \varphi(0)\|_{L^2} \leq c_1^{\frac{7}{12}} c^{\delta + \frac{7}{12}}.$$

Then, there exist $\rho_1(t)$, $\rho_2(t) \in \mathbb{R}$ and $c_1^{\pm}$, $c_2^{\pm} > 0$ such that

1. *Global in time stability*: $w(t,x) = u(t,x) - Q_{c_1}(x - \rho_1(t)) - Q_{c_2}(x - \rho_2(t))$ *satisfies*

$$\|\partial_x w(t)\|_{L^2} + \sqrt{c_1}\|w(t)\|_{L^2} \leq K c_1^{\frac{7}{12}} (c^{\delta + \frac{1}{12}} + c^{\frac{1}{3}}), \quad \textit{for all } t \in \mathbb{R}.$$

2. *Asymptotic stability*:

$$\lim_{t\to-\infty} \|u(t) - Q_{c_1^-}(. - \rho_1(t)) - Q_{c_2^-}(. - \rho_2(t))\|_{H^1(x < \frac{c_1 t}{10})} = 0,$$
$$\lim_{t\to+\infty} \|u(t) - Q_{c_1^+}(. - \rho_1(t)) - Q_{c_2+}(. - \rho_2(t))\|_{H^1(x > \frac{c_1 t}{10})} = 0,$$

$$\left|\frac{c_1^\pm}{c_1}-1\right|\leq Kc^{\frac{7}{12}}(c^\delta+c^{\frac{1}{3}}),\quad \left|\frac{c_2^\pm}{c_2}-1\right|\leq K(c^\delta+c^{\frac{1}{3}}).$$

Remark 3. Theorem 1.3 shows that the various properties exhibited in Theorem 1.2 are stable under perturbation of the initial data (during and after the collision). This constructs, in particular, a large set of initial data having globally in time a 2-soliton structure (as for the integrable case). The stability property can also be proved assuming $u(T_0)$ close to $\varphi(T_0)$ for some T_0 (see the proof of Theorem 1.3).

1.4. *Strategy of the proofs.* Let us sketch the main steps of the proof of Theorem 1.1. The proofs of Theorems 1.2 and 1.3 follow a similar scheme.

(1) By scaling invariance, assume $c_1 = 1$ and $c_2 = c \ll 1$. The first and main step is to construct an approximate solution to the problem in the collision region $[-T_c, T_c]$, where $T_c \gg c^{-\frac{1}{2}}$. The approximate solution has the following specific form:

$$v(t,x)=Q(y)+Q_c(y_c)+\sum_{k,\ell}c^\ell\left(Q_c^k(y_c)A_{k,\ell}(y)+(Q_c^k)'(y_c)B_{k,\ell}(y)\right),$$

where y and y_c are two independent variables and the functions $A_{k,\ell}$, $B_{k,\ell}$ are to be determined. One important point in the decomposition is to choose suitable variables y and y_c for the solitons Q and Q_c. The choice of y_c in (2.2) is straightforward and corresponds to the trajectory of the small soliton without perturbation; indeed, the shift on the trajectory of Q_c will be deduced from a recomposition of the above series and a simple Taylor expansion (see (3) below). The choice of the variable y in (2.4)–(2.5) is more subtle and avoids usual problems due to secular terms. Here secular terms have the same form as nonlinear perturbation terms and the degrees of freedom due to the choice of the parameters $a_{k,\ell}$ in (2.5) are the key to prove the existence of suitable functions $A_{k,\ell}$ and $B_{k,\ell}$.

(2) Second, we consider the unique solution $u(t)$ of (1.1) which satisfies

$$\lim_{t\to-\infty}\|u(t)-Q_{c_1}(.-c_1t)-Q_{c_2}(.-c_2t)\|_{H^1}=0.$$

It is straightforward to compare $u(-T_c)$ and $v(-T_c)$, that is, before the collision. Using sharp asymptotic and perturbation arguments and taking c small enough, we prove that the solution $u(t)$ is close to the approximate solution $v(t)$ on $[-T_c, T_c]$, so that the description of the collision given by $v(t)$ is relevant on $[-T_c, T_c]$.

In addition, from the information on $u(T_c)$ and by asymptotic arguments ([25], [21], [27] and [22]) related to sharp monotonicity properties, we fully describe the solution $u(t)$ in large time, that is, for $t > T$.

(3) Finally, we prove the inelastic character of the collision for $p = 4$ by a further analysis of the approximate solution. The defect is due to a nonzero extra term in the approximate solution after recomposition of the series. Indeed, we have the following expansion of $v(T_c)$:

$$v(T_c) \sim Q(y) + Q_c(y_c) - b_{1,0}Q_c'(y_c) - \frac{1}{2}b_{2,0}(Q_c^2)'(y_c) + O_{H^1}(c),$$

for some explicit constants $b_{1,0}$ and $b_{2,0}$. Note that $\|(Q_c^2)'(y_c)\|_{L^2} \sim Kc^{\frac{11}{12}}$, so that this term is relevant in the above expansion. Whereas the term $-b_{1,0}Q_c'(y_c)$ can be combined with $Q_c(y_c)$ to obtain a shift on the trajectory of Q_c (by the Taylor expansion $Q_c(y_c - b_{1,0}) = Q_c(y_c) - b_{1,0}Q_c'(y_c) + O_{H^1}(c)$), the term $b_{2,0}(Q_c^2)'(y_c)$ is the principal nonmatching term. A decisive point in the computations is to obtain $b_{2,0} \neq 0$; this computation is performed in [24]. Thus, the defect is a direct consequence of the algebra underlying the construction of the approximate solution. Again, stability and monotonicity arguments allow us to propagate the defect on the solution $u(t)$ for all time $t > T_c$ and to prove that $u(t)$ is not a global pure 2-soliton solution, with lower bounds on the defect.

The paper is organized as follows. Section 2 is devoted to the construction of the approximate solution $v(t)$. Section 3 is concerned with the recomposition of $v(t)$ after the collision. We mainly focus on the case $p = 4$. For $p = 2$, we only compare at the main orders the function $v(t)$ to the explicit 2-soliton solutions. In Section 4 we recall and adapt some asymptotic results from [25]. Section 5 is devoted to the proofs of the main results, i.e., Theorems 1.1, 1.2 and 1.3. Some technical proofs are presented in several appendices.

Acknowledgments. The authors would like to thank the anonymous referee for useful suggestions concerning this paper.

2. **Construction of an approximate 2-soliton**

In the proof of the main results (Theorems 1, 2 and 3), we restrict ourselves to the case $c_1 = 1$, $c_2 = c$ small by a scaling argument. Therefore, in this section, we concentrate on this case.

Let $p = 2, 3, 4$ and define

$$T_c = c^{-\frac{1}{2}-\frac{1}{100}} \quad \text{and} \quad q = \frac{1}{p-1} - \frac{1}{4}. \tag{2.1}$$

In this section, for any $n_0 \in \mathbb{N}$, for $0 < c < c_0$ small enough, we construct a function $v_n(t,x) = v(t,x)$ which satisfies the following two properties:

- $v(t, x-t)$ is a solution of the gKdV equation (1.1) on $[-T_c, T_c]$ up to an error term of polynomial order c^{n_0},

$$\forall t \in [-T_c, T_c], \quad \|\partial_t v + \partial_x(\partial_x^2 v - v + v^p)\|_{H^1(\mathbb{R})} \leq K(n_0)c^{n_0}.$$

- The principal contributions to $v(-T_c)$ and $v(+T_c)$ are the sum of two solitons Q and Q_c respectively before and after their collision.

The function $v(t)$ is the new fundamental object of this paper. Its existence and properties will lead to the main results stated in the introduction.

Our approach is to consider c as a small parameter and look for such a function v in terms of expansions in powers of c, both in the functions and the space variables. More precisely, the construction of the function $v(t,x)$ is related to the method of separation of variables: the variable y of the large soliton $Q(y)$ is separated from the variable y_c of the small soliton $Q_c(y_c)$.

First, we set

$$y_c = x + (1-c)t \quad \text{and} \quad R_c(t,x) = Q_c(y_c); \tag{2.2}$$

note that $R_c(t)$ is then a solution of $\partial_t R_c + \partial_x(\partial_x^2 R_c - R_c + R_c^p) = 0$.

We look for a function $v(t,x)$ having the structure

$$v(t,x) = Q(y) + Q_c(y_c) + W(t,x). \tag{2.3}$$

We choose the function W and the variable y under the form of series. Let $k_0 \geq 1$, $\ell_0 \geq 0$ and define

$$\Sigma_0 = \{(k,\ell),\ 1 \leq k \leq k_0,\ 0 \leq \ell \leq \ell_0\}.$$

For real unknown parameters $(a_{k,\ell})_{(k,\ell)\in\Sigma_0}$, we consider the variable y of the form

$$y = x - \alpha(y_c) = x - \alpha\big(x + (1-c)t\big) \quad \text{and} \quad R(t,x) = Q(y), \tag{2.4}$$

where

$$\alpha(s) = \int_0^s \beta(s')ds' \quad \text{and} \quad \beta(s) = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^\ell Q_c^k(s). \tag{2.5}$$

The form of W is

$$W(t,x) = \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(y_c)A_{k,\ell}(y) + (Q_c^k)'(y_c)B_{k,\ell}(y)\right), \tag{2.6}$$

where the functions $A_{k,\ell}$, $B_{k,\ell}$, as the parameters $(a_{k,\ell})$, are to be determined. Note that the functions $c^\ell Q_c^k$ and $c^\ell (Q_c^k)'$ used to define the series play the role of a set of nonlinear eigenfunctions for the interaction problem. Thus, the structure of W will allow us to compute the interaction terms at any order of power of c. Moreover, choosing the variable y as above will allow us to understand the effect of the soliton Q_c on the position of Q, that is, the shift phenomenon which appears through the interaction of two solitons.

THEOREM 2.1 (Construction of an approximate solution of the gKdV equation). *Let $p = 2$, 3 or 4. For all $k \geq 1$, $\ell \geq 0$, there exist $a_{k,\ell} \in \mathbb{R}$ and C^∞ functions $A_{k,\ell}$, $B_{k,\ell} : \mathbb{R} \to \mathbb{R}$ such that, for any $0 < c < 1$, for any $k_0 \geq 1$ and for any $\ell_0 \geq 0$, the function $v(t)$ defined by*

$$v(t,x) = Q(y) + Q_c(y_c) + \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(y_c) A_{k,\ell}(y) + (Q_c^k)'(y_c) B_{k,\ell}(y) \right), \tag{2.7}$$

where $y_c = x + (1-c)t$, $y = x - \alpha(y_c)$ and $\alpha(s) = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^\ell \int_0^s Q_c^k(s')ds'$, satisfies

1. *The function $v(t, x - t)$ is an approximate solution: $S(t)$ defined by*

$$S(t,x) = \partial_t v + \partial_x(\partial_x^2 v - v + v^p) \tag{2.8}$$

   satisfies, for all $j \geq 0$,

$$\forall t \in [-T_c, T_c], \quad \|\partial_x^{(j)} S(t)\|_{L^2(\mathbb{R})} \leq K c^{n_0}, \tag{2.9}$$

   where $n_0 = (\frac{1}{2} - \frac{1}{100}) \min\left(\frac{k_0}{p-1}, 1 + \ell_0\right)$ and $K = K(j, k_0, \ell_0) > 0$.
2. *The function $v(t)$ belongs to $H^1(\mathbb{R})$ for all $t \in \mathbb{R}$ and satisfies, for $K = K(k_0, \ell_0) > 0$,*

$$\forall t \in [-T_c, T_c], \quad \|v(t) - R(t) - R_c(t)\|_{H^1(\mathbb{R})} \leq K c^{\frac{1}{p-1}}. \tag{2.10}$$

Remarks. (a) *Size comparison in* (2.10). First, note that

$$\|Q_c\|_{L^2} = c^q \|Q\|_{L^2}, \quad \|Q_c'\|_{L^2} = c^{q+\frac{1}{2}} \|Q'\|_{L^2} \quad \text{and} \quad \|Q_c\|_{L^\infty} = c^{q+\frac{1}{4}} \|Q\|_{L^\infty}. \tag{2.11}$$

Since $\frac{1}{p-1} = q + \frac{1}{4}$, (2.10) says that $v(t) - R(t) - R_c(t)$ is smaller in H^1 norm than $R_c(t)$ by a factor $c^{1/4}$. Thus, in H^1, $v(t) = R(t) + R_c(t) +$ smaller order terms in c.

Remark that the L^∞ norm is not adequate in this framework; indeed, we also have $\|v(t) - R(t) - R_c(t)\|_{L^\infty} \leq K\|v(t) - R(t) - R_c(t)\|_{H^1} \leq K c^{\frac{1}{p-1}} \leq K\|Q_c\|_{L^\infty}$. Moreover, from (2.7) and from the fact $A_{1,0} \neq 0$ (see proofs), we have for $t \sim 0$, $\|v(t) - R(t) - R_c(t)\|_{L^\infty} \sim \|Q_c\|_{L^\infty}$. Observe also that $\|Q_c'\|_{L^2}$ is smaller than $\|Q_c\|_{L^2}$ for c small. Throughout this paper, the norm that really matters in the various estimates is the L^2 norm.

Note that (2.10) is only a first estimate concerning the relation between v and the sum of two solitons. This estimate does not take into account the shift of the soliton Q_c, and thus cannot be sharp. In Sections 3 and 4, by recompositing v at $t = \pm T_c$, we will prove a better estimate for $v(t) - Q(y) - Q_c(y_c \pm \Delta_c)$, for some Δ_c and for $t = \pm T_c$ (see Proposition 3.1). Estimate (2.9) is also not optimal, especially for small k_0 and ℓ_0 (but $n_0 \to +\infty$ as $k_0, \ell_0 \to +\infty$).

Note also that $k_0 \geq 5$ and $\ell_0 \geq 1$ in Theorem 2.1 would be enough to prove the main results of this paper. Nevertheless, the result as stated for all

k_0, ℓ_0 clearly indicates that there is no algebraic obstruction to the complete understanding of the interaction process, and we expect it to be useful in future works.

(b) *The time interval* $[-T_c, T_c]$ *contains the interaction region.* Since for $t = -T_c$, $y \ll y_c$ and for $t = T_c$, $y \gg y_c$, the interaction of the two solitons Q and Q_c takes place in the time interval $[-T_c, T_c]$. Moreover, since $y_c = y + \alpha(y_c) + (1-c)t$, we have $|y_c| \geq (1-c)|t| - |\alpha(y_c)| - |y|$. Thus, if $\sqrt{c} < 2$, we obtain $\sqrt{c}|y_c| \geq (1-c)\sqrt{c}|t| - \sqrt{c}|\alpha(y_c)| - \frac{1}{2}|y|$, and by neglecting $\sqrt{c}|\alpha(y_c)|$, we obtain for $|t| \geq T_c$,

$$0 \leq R(t)R_c(t) \leq Kc^{\frac{1}{p-1}} e^{-\frac{|y|}{2}} \frac{1}{e^{\frac{1}{2}c^{-\frac{1}{100}}}},$$

which is an exponentially small term when c is small, which says that the interaction between Q and Q_c is very weak for such t.

(c) *Decomposition of* W. The function constructed in Theorem 2.1 is not unique. For given k_0 and ℓ_0 there exist, in fact, several such functions v corresponding to the fact that the decomposition at $t = 0$, for example, is not unique.

We refer to Proposition 2.3 for more properties of the functions $A_{k,\ell}$ and $B_{k,\ell}$ introduced in Theorem 2.1.

Note that choosing $k_0 = \ell_0 = +\infty$ in this expression of v would formally give an exact solution of the gKdV equation at least for $t \in [-T_c, T_c]$. However, one has to verify that the resulting series in (2.6) converges in some appropriate sense, which is an open problem.

We give a first interpretation of the function v constructed in Theorem 2.1.

Integrable case ($p = 2$ and 3)**:** In this case, one of the functions v constructed in Theorem 2.1 coincides in its principal terms of perturbation theory to the explicit 2-soliton solution.

Nonintegrable case ($p = 4$)**:** In this case, explicit 2-soliton solutions are not *a priori* known and indeed will be proved to not exist later in this paper. The function v is a *completely new object.* Note that this object, up to the order c^{n_0}, plays the same role as a 2-soliton solution in the collision region. This will allow us to prove the main results of this paper.

The proof of Theorem 2.1 is organized as follows. In Section 2.1, we claim that the decomposition of $v(t)$ is preserved by gKdV equation; see Proposition 2.1). The main part of the proof of Proposition 2.1 is given in Appendix A.

In Section 2.2, we derive the systems $(\Omega_{k,\ell})$ to be solved at each rank (k, ℓ). Next, we solve a model system (Ω) related to $(\Omega_{k,\ell})$. In particular, we choose a special structure for the functions $A_{k,\ell}$ and $B_{k,\ell}$ which follows from the resolution of the model system.

Then we solve by induction on (k,ℓ) all the systems $(\Omega_{k,\ell})$, for $1 \le k \le k_0$, $0 \le \ell \le \ell_0$. This determines $(a_{k,\ell})$, $(A_{k,\ell})$ and $(B_{k,\ell})$ for all $1 \le k \le k_0$, $0 \le \ell \le \ell_0$ in the expression of v. Thus, at this point the function $v(t)$ is fixed.

Finally, in Section 2.3 we prove some properties of $v(t)$ and estimate the size of $S(t)$ in terms of powers of c.

For k, k', ℓ, $\ell' \in \mathbb{N}$, we denote

$$(k',\ell') \prec (k,\ell) \quad \text{if } k' < k \text{ and } \ell' \le \ell \text{ or if } k' \le k \text{ and } \ell' < \ell.$$

We denote by $\mathcal{Y}$ the set of functions $f \in C^\infty(\mathbb{R})$ such that

$$\forall j \in \mathbb{N},\ \exists K_j,\, r_j > 0,\ \forall x \in \mathbb{R}, \quad |f^{(j)}(x)| \le K_j(1+|x|)^{r_j} e^{-|x|}.$$

Note that the set $\mathcal{Y}$ is stable by sum, multiplication and differentiation.

2.1. *Preservation of the decomposition* (2.7) *by the equation.* The motivation for choosing W of the form (2.6) is the stability of the family of functions

$$\left\{c^\ell Q_c^k,\ c^\ell (Q_c^k)',\ k \ge 1,\ \ell \ge 0\right\} \tag{2.12}$$

by multiplication and differentiation (see Lemma 2.1). A consequence is that the term $S(t,x) = \partial_t v + \partial_x(\partial_x^2 v - v + v^p)$ has the same decomposition as the function v in terms of functions (2.12). In particular, due to the special choices of the variables y and y_c, secular terms are of same nature as the nonlinear perturbation terms, and do not create specific difficulties in the decomposition of S. Actually, such terms will be the key of the resolution of the systems of $A_{k,\ell}$ and $B_{k,\ell}$.

Let

$$\mathcal{L}w = -\partial_x^2 w + w - pQ^{p-1}w. \tag{2.13}$$

Proposition 2.1 (Decomposition of $S(t,x)$). *Let v be as in* (2.7). *Let $K_0 = (p+1)k_0 + 12$ and $L_0 = (p+1)\ell_0 + 4$. Then*

$$\begin{aligned} S(t,x) = &\sum_{(k,\ell)\in\Sigma_0} c^\ell Q_c^k(y_c)\Big[a_{k,\ell}(-3Q+2Q^p)'(y) - (\mathcal{L}A_{k,\ell})'(y)\Big] \\ &+ \sum_{(k,\ell)\in\Sigma_0} c^\ell (Q_c^k)'(y_c)\Big[a_{k,\ell}(-3Q'')(y) \\ &\qquad\qquad + \left(3A''_{k,\ell} + pQ^{p-1}A_{k,\ell}\right)(y) - (\mathcal{L}B_{k,\ell})'(y)\Big] \\ &+ \sum_{\substack{1\le k\le K_0 \\ 0\le \ell\le L_0}} c^\ell \left(Q_c^k(y_c)F_{k,\ell}(y) + (Q_c^k)'(y_c)G_{k,\ell}(y)\right), \end{aligned}$$

where $F_{k,\ell}$ and $G_{k,\ell}$ are functions defined on $\mathbb{R}$ satisfying

(i) *Dependence property of F and G: For any k,ℓ, the expressions of $F_{k,\ell}$ and $G_{k,\ell}$ depend only on $(a_{k',\ell'})$ and $(A_{k',\ell'})$, $(B_{k',\ell'})$ for k', ℓ' such that $(k',\ell') \prec (k,\ell)$.*

(ii) *Parity property of F and G: Let $k \in \{1,\dots,K_0\}$, $\ell \in \{0,\dots,L_0\}$. Assume that for any (k',ℓ') such that $(k',\ell') \prec (k,\ell)$ $A_{k',\ell'}$ is even and $B_{k',\ell'}$ is odd, then $F_{k,\ell}$ is odd and $G_{k,\ell}$ is even.*

Moreover,

- *If $p=2$, then*

$$\begin{aligned}
F_{1,0} &= 2Q', \quad G_{1,0} = 2Q,\\
F_{2,0} &= (-A_{1,0}+A_{1,0}^2)' - (3B_{1,0}''+2QB_{1,0})\\
&\quad - a_{1,0}(Q+3A_{1,0}''+2QA_{1,0})' + 3a_{1,0}^2 Q^{(3)},\\
G_{2,0} &= A_{1,0}+A_{1,0}^2 + (-2B_{1,0}+A_{1,0}B_{1,0})'\\
&\quad - \frac{a_{1,0}}{2}(9A_{1,0}'+3B_{1,0}''+2QB_{1,0})' + \frac{3}{2}a_{1,0}^2 Q''.
\end{aligned}$$

- *If $p=4$, then*

$$\begin{aligned}
F_{1,0} &= (4Q^3)', \quad G_{1,0} = 4Q^3,\\
F_{2,0} &= (6Q^2(1+A_{1,0})^2)' - a_{1,0}(4Q^3+3A_{1,0}''+4Q^3A_{1,0})' + 3a_{1,0}^2 Q^{(3)},\\
G_{2,0} &= 6Q^2(1+A_{1,0})^2 + (6Q^2B_{1,0}(1+A_{1,0}))'\\
&\quad - \frac{a_{1,0}}{2}(9A_{1,0}'+3B_{1,0}''+4Q^3B_{1,0})' + \frac{3}{2}a_{1,0}^2 Q''.
\end{aligned}$$

See Proposition 2.3, Lemma B.1 and Claim 2.4 for additional properties of $F_{k,\ell}$ and $G_{k,\ell}$.

Proof of Proposition 2.1. A large part of the proof of Proposition 2.1 is given in Appendix A. We present here some preliminary results.

We begin by proving that the family of functions (2.12) is stable by multiplication and differentiation.

LEMMA 2.1 (Properties of Q and Q_c). 1. *The function Q is even and belongs to $\mathcal{Y}$.*

2. *For any $k \in \mathbb{N}^*$,*

$$\begin{aligned}
Q_c'' &= cQ_c - Q_c^p, \quad (Q_c')^2 = cQ_c^2 - \frac{2}{p+1}{Q_c}^{p+1},\\
(Q_c^k)'' &= ck^2 Q_c^k - \frac{k(2k+p-1)}{p+1} Q_c^{k+p-1},\\
(Q_c^k)^{(3)} &= ck^2 (Q_c^k)' - \frac{k(2k+p-1)}{p+1} (Q_c^{k+p-1})',
\end{aligned}$$

$$(Q_c^k)^{(4)} = c^2k^4Q_c^k - c\frac{k(2k+p-1)}{p+1}(k^2+(k+p-1)^2)Q_c^{k+p-1} + k(k+p-1)\frac{(2k+p-1)(2k+3p-3)}{(p+1)^2}Q_c^{k+2p-2}.$$

3. *For any* k_1, $k_2 \in \mathbb{N}^*$,

$$(Q_c^{k_1})'Q_c^{k_2} = \frac{k_1}{k_1+k_2}(Q_c^{k_1+k_2})',$$
$$(Q_c^{k_1})'(Q_c^{k_2})' = ck_1k_2Q_c^{k_1+k_2} - \frac{2k_1k_2}{p+1}Q_c^{k_1+k_2+p-1}.$$

Proof of Lemma 2.1. It is clear from (1.4) that Q is even and belongs to $\mathcal{Y}$. From the equation of Q_c, i.e. $Q_c'' = cQ_c - Q_c^p$, we easily get the second equation by multiplying by Q_c' and integrating over $(-\infty, x)$.

Next, we have

$$(Q_c^k)'' = k(Q_c^{k-1}Q_c')' = k\left((k-1)Q_c^{k-2}(Q_c')^2 + Q_c^{k-1}Q_c''\right)$$
$$= ck^2Q_c^k - k\left(\frac{2(k-1)}{p+1}+1\right)Q_c^{k+p-1} = ck^2Q_c^k - \frac{k(2k+p-1)}{p+1}Q_c^{k+p-1}.$$

From this we immediately obtain the expression of $(Q_c^k)^{(3)}$. Next, we have

$$(Q_c^k)^{(4)} = ck^2(Q_c^k)'' - \frac{k(2k+p-1)}{p+1}(Q_c^{k+p-1})''$$
$$= ck^2\left(ck^2Q_c^k - \frac{k(2k{+}p{-}1)}{p+1}Q_c^{k+p-1}\right)$$
$$- \frac{k(2k{+}p{-}1)}{p+1}\left(c(k{+}p{-}1)^2Q_c^{k+p-1} - (k{+}p{-}1)\frac{2k{+}3p{-}3}{p+1}Q_c^{k+2p-2}\right).$$

The rest of the proof follows.

Now, we give a preliminary decomposition of $S(t)$. We insert $v = R + R_c + W$ into $S(t,x)$, and rearrange terms:

$$\begin{aligned} S(t,x) &= \partial_t v + \partial_x(\partial_x^2 v - v + v^p) \\ &= \partial_t(R + R_c + W) \\ &\quad + \partial_x\left(\partial_x^2(R+R_c+W) - (R+R_c+W) + (R+R_c+W)^p\right) \\ &= \partial_t R + \partial_x(\partial_x^2 R - R + R^p) + \partial_t R_c + \partial_x(\partial_x^2 R_c - R_c + R_c^p) \\ &\quad + \partial_x((R+R_c)^p - R^p - R_c^p) \\ &\quad + \partial_t W + \partial_x(\partial_x^2 W - W + (R+R_c+W)^p - (R+R_c)^p). \end{aligned}$$

By the equation of Q_c ($Q_c'' = cQ_c - Q_c^p$) and $y_c = x + (1-c)t$, it is straightforward that

$$\partial_t R_c + \partial_x(\partial_x^2 R_c - R_c + R_c^p) = ((1-c)Q_c + Q_c'' - Q_c + Q_c^p)'(y_c) = 0. \tag{2.14}$$

Set

$$\mathcal{L}w = -\partial_x^2 w + w - pQ^{p-1}w, \quad \overline{\mathcal{L}}w = -\partial_x^2 w + w - pR^{p-1}w. \tag{2.15}$$

We decompose $S(t,x)$ as follows:

$$S(t,x) = \mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV}, \tag{2.16}$$

where **I**, **II**, **III** and **IV** are respectively:

- Contribution of terms containing only R: $\mathbf{I} = \partial_t R + \partial_x(\partial_x^2 R - R + R^p)$.
- Nonlinear interaction terms between R and R_c: $\mathbf{II} = \partial_x((R+R_c)^p - R^p - R_c^p)$.
- Linear terms in W: $\mathbf{III} = \partial_t W - \partial_x(\overline{\mathcal{L}}W)$.
- Higher order terms in W: $\mathbf{IV} = \partial_x((R+R_c+W)^p - (R+R_c)^p - pR^{p-1}W)$.

The expansion of **I**, **II**, **III** and **IV** is given in Appendix A, and allows us to finish the proof of Proposition 2.1.

2.2. *Resolution of the systems* $(\Omega_{k,\ell})$. From Proposition 2.1, we observe that if for any $0 \le k \le k_0$, $0 \le \ell \le \ell_0$, $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ satisfies the following system:

$$(\Omega_{k,\ell}) \quad \begin{cases} (\mathcal{L}A_{k,\ell})' + a_{k,\ell}(3Q - 2Q^p)' = F_{k,\ell} \\ (\mathcal{L}B_{k,\ell})' + a_{k,\ell}(3Q'') - 3A''_{k,\ell} - pQ^{p-1}A_{k,\ell} = G_{k,\ell}. \end{cases}$$

Then $S(t,x)$ contains only terms of the form $c^\ell Q_c^k$ or $c^\ell (Q_c^k)'$ with $k \ge k_0 + 1$ or $\ell \ge \ell_0 + 1$.

This observation leads us to consider the model system

$$(\Omega) \quad \begin{cases} (\mathcal{L}A)' + a(3Q - 2Q^p)' = F \\ (\mathcal{L}B)' + a(3Q'') - 3A'' - pQ^{p-1}A = G, \end{cases}$$

where $F(x)$ and $G(x)$ are given functions (with a specific structure; see Proposition 2.2) and $(a, A(x), B(x))$ is to be determined. We study existence of solutions of the system (Ω). Before stating and proving the existence result for the model system (Ω), we introduce some notation and we recall well-known results concerning the operator $\mathcal{L}$.

First, let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined by

$$\forall x \in \mathbb{R}, \quad \varphi(x) = -\frac{Q'(x)}{Q(x)}.$$

Claim 2.1. *The function φ is odd and satisfies the following properties.*

(a) $\lim_{x\to-\infty} \varphi(x) = -1$; $\lim_{x\to+\infty} \varphi(x) = 1$.
(b) *For all* $x \in \mathbb{R}$, $|\varphi'(x)| + |\varphi''(x)| + |\varphi^{(3)}(x)| \le Ce^{-|x|}$.
(c) $\varphi' \in \mathcal{Y}$, $(1 - \varphi^2) \in \mathcal{Y}$.
(d) *For* $p = 2$, $(\mathcal{L}\varphi)' = 2Q - \frac{5}{3}Q^2$. *For* $p = 4$, $(\mathcal{L}\varphi)' = \frac{36}{5}Q^3 - \frac{99}{25}Q^6$.

Proof of Claim 2.1. From the explicit formula $Q(x) = \left(\frac{p+1}{2\cosh^2\left(\frac{p-1}{2}x\right)}\right)^{\frac{1}{p-1}}$, we have

$$Q'(x) = -\tanh\left(\tfrac{p-1}{2}x\right)Q(x),$$

and so $\varphi(x) = \tanh\left(\frac{p-1}{2}x\right)$. From $\tanh' = 1 - \tanh^2 = \frac{1}{\cosh^2}$, we obtain (a), (b) and (c).

By $Q'' = Q - Q^p$ and $(Q')^2 = Q^2 - \frac{2}{p+1}Q^{p+1}$, we have

$$\varphi' = -\frac{1}{Q^2}(Q''Q - (Q')^2) = \frac{p-1}{p+1}Q^{p-1}, \quad \text{and} \quad \varphi'' = \frac{(p-1)^2}{p+1}Q'Q^{p-2}.$$

Thus, $-\varphi'' - pQ^{p-1}\varphi = \left(-\frac{(p-1)^2}{p+1} + p\right)Q'Q^{p-2} = \frac{3p-1}{p+1}Q'Q^{p-2}$, and

$$\begin{aligned}(\mathcal{L}\varphi)' &= \frac{3p-1}{p+1}Q''Q^{p-2} + \frac{(3p-1)(p-2)}{p+1}(Q')^2Q^{p-3} + \frac{p-1}{p+1}Q^{p-1}\\ &= \frac{3p(p-1)}{p+1}Q^{p-1} - \frac{3(3p-1)(p-1)}{(p+1)^2}Q^{2(p-1)}.\end{aligned}$$

Lemma 2.2 (Properties of $\mathcal{L}$). *Let $p \geq 2$. The operator $\mathcal{L}$ defined in $L^2(\mathbb{R})$ by* (2.13) *is self-adjoint and satisfies the following properties.*

(i) *First eigenfunction*: $\mathcal{L}Q^{\frac{p+1}{2}} = -\frac{1}{4}(p-1)(p+3)Q^{\frac{p+1}{2}}$.

(ii) *Second eigenfunction*: $\mathcal{L}Q' = 0$; *the kernel of $\mathcal{L}$ is $\{\lambda Q', \lambda \in \mathbb{R}\}$.*

(iii) *For any function $h \in L^2(\mathbb{R})$ orthogonal to Q' for the L^2 scalar product, there exists a unique function $f \in H^2(\mathbb{R})$ orthogonal to Q' such that $\mathcal{L}f = h$. Moreover, if h is even* (*respectively, odd*), *then f is even* (*respectively, odd*).

(iv) *Suppose that $f \in H^2(\mathbb{R})$ is such that $\mathcal{L}f \in \mathcal{Y}$. Then $f \in \mathcal{Y}$.*

Proof of Lemma 2.2. From $Q'' = Q - Q^p$ and $(Q')^2 = Q^2 - \frac{2}{p+1}Q^{p+1}$,

$$\frac{d^2}{dx^2}Q^{\frac{p+1}{2}} = \tfrac{p+1}{2}\left[\tfrac{p-1}{2}Q'^2Q^{\frac{p-3}{2}} + Q''Q^{\frac{p-1}{2}}\right] = \left(\tfrac{p+1}{2}\right)^2 Q^{\frac{p+1}{2}} - pQ^{p-1}Q^{\frac{p+1}{2}},$$

and so $\mathcal{L}Q^{\frac{p+1}{2}} = -\left[\left(\frac{p+1}{2}\right)^2 - 1\right]Q^{\frac{p+1}{2}} = -\frac{1}{4}(p-1)(p+3)Q^{\frac{p+1}{2}}$.

The property $\mathcal{L}Q' = 0$ is easily checked. Moreover, the fact that the spectrum of $\mathcal{L}$ is restricted to $\{\lambda Q', \lambda \in \mathbb{R}\}$ was proved by ordinary differential equations techniques (see Weinstein [37, Prop. 2.8(b)]). The third property is a direct consequence of the structure of $\mathcal{L}$ and the Lax-Milgram theorem.

Property (iv) is also a consequence of standard arguments of ordinary differential equations theory. First, we claim the following.

Claim 2.2. *Suppose that $f \in H^2(\mathbb{R})$ satisfies, for $K > 0$ and $r > 0$,*

$$\forall x \in \mathbb{R}, \quad |(f'' - f)(x)| \leq K(1 + |x|^r)e^{-|x|}. \tag{2.17}$$

Then, there exists $K' > 0$ *such that*

$$\forall x \in \mathbb{R}, \quad |f(x)| \leq K'(1+|x|^{r+1})e^{-|x|}. \tag{2.18}$$

Proof of Claim 2.2. We set $g(x) = e^{-x}(f'+f)$. Then $g' = e^{-x}(f''-f)$,

$$\forall x > 0, \quad |g'(x)| \leq K(1+|x|^r)e^{-2x} \quad \text{and}$$

$$|g(x)| \leq K \int_x^{+\infty} (1+s^r)e^{-2s} ds \leq K'(1+x^r)e^{-2x}.$$

Set $h = e^x f$. Then $|h'| = |e^{2x}g| \leq K(1+|x|^r)$. By integration between 0 and x, we obtain for all $x > 0$, $e^x|f(x)| = |h(x)| \leq K''(1+|x|^{r+1})$. The same property is true for $x < 0$, by changing x in $-x$.

We now finish the proof of (iv). Let $f \in H^2(\mathbb{R})$ be such that $\mathcal{L}f \in \mathcal{Y}$. Since $f'' = (-\mathcal{L}f + f - pQ^{p-1}f)$, by induction on j and $Q \in \mathcal{Y}$, it is clear that $f \in C^j(\mathbb{R})$, for all $j \in \mathbb{N}$. Since $(f^{(j)})'' - f^{(j)} = -(\mathcal{L}f + pQ^{p-1}f)^{(j)}$ and $\mathcal{L}f$, $Q \in \mathcal{Y}$, using Claim 2.2 we prove by an induction argument on j that for all j and all x, $|f^{(j)}(x)| \leq K_j(1+|x|^{r_j})e^{-|x|}$. Thus, $f \in \mathcal{Y}$.

The next result of this section concerns the existence of solutions of system (Ω).

Proposition 2.2 (Existence for the model problem (Ω)). *Let* $F(x)$ *and* $G(x)$ *be such that*

$$F = \overline{F} + \widetilde{F} + \varphi\widehat{F}, \quad G = \overline{G} + \widetilde{G} + \varphi\widehat{G},$$

where

- $\overline{F}$, $\overline{G} \in \mathcal{Y}$: $\overline{F}$ *is odd and* $\overline{G}$ *is even*;
- $\widetilde{F}$ *and* $\widehat{G}$ *are odd polynomial functions*: $\widehat{F}$ *and* $\widetilde{G}$ *are even polynomial functions.*

Then, there exist $a \in \mathbb{R}$ *and two functions* $A(x)$, $B(x)$

$$A = \overline{A} + \widetilde{A} + \varphi\widehat{A}, \quad B = \overline{B} + \widetilde{B} + \varphi\widehat{B},$$

where

- $\overline{A}$, $\overline{B} \in \mathcal{Y}$: $\overline{A}$ *is even and* $\overline{B}$ *is odd*;
- $\widetilde{A}$ *and* $\widehat{B}$ *are even polynomial functions*: $\widehat{A}$ *and* $\widetilde{B}$ *are odd polynomial functions*;

satisfying

$$(\Omega) \quad \begin{cases} (\mathcal{L}A)' + a(3Q - 2Q^p)' = F & (\Omega_A) \\ (\mathcal{L}B)' + a(3Q'') - 3A'' - pQ^{p-1}A = G. & (\Omega_B) \end{cases}$$

The degrees of the polynomial functions $\widetilde{A}$, $\widehat{A}$, $\widetilde{B}$ *and* $\widehat{B}$ *are related to the degrees of* $\widetilde{F}$, $\widehat{F}$, $\widetilde{G}$ *and* $\widehat{G}$ *as follows*:

$$\deg \widetilde{A} \leq 1 + \deg \widetilde{F}, \quad \deg \widetilde{B} \leq \max(1 + \deg \widetilde{G}, \deg \widetilde{F}), \tag{2.19}$$

$$\deg \widehat{A} \leq 1 + \deg \widehat{F}, \quad \deg \widehat{B} \leq \max(1 + \deg \widehat{G}, \deg \widehat{F}). \tag{2.20}$$

Moreover,

$$\text{if } \widetilde{F} = 0 \ (\text{respectively, } \widehat{F} = 0), \text{ then } \widetilde{A} = 0 \ (\text{respectively, } \widehat{A} = 0); \tag{2.21}$$

$$\text{if } \widetilde{A}'' = 0 \text{ and } \widetilde{G} = 0, \text{ then } \widetilde{B} = 0; \tag{2.22}$$

$$\text{if } \widehat{A}'' = 0 \text{ and } \widehat{G} = 0, \text{ then } \deg \widehat{B} = 0. \tag{2.23}$$

Remark. Observe that the conclusions of (2.23) and (2.21)–(2.22) are different. In (2.23), only $\deg \widehat{B} = 0$ which allows the possibility that $\widehat{B} = b$, a nonzero constant, even if no polynomial is present in F and G. Without this freedom, the system cannot be solved in general. This remark is essential for two reasons:

1. The fact that possibly $\widehat{B} \neq 0$ whereas $\widetilde{F}$, $\widehat{F}$, $\widetilde{G}$ and $\widehat{G}$ are zero, is responsible for the apparition of polynomial growths in $A_{k,\ell}$ and $B_{k,\ell}$ when solving the systems $(\Omega_{k,\ell})$. Indeed, from the structure of the systems $(\Omega_{k,\ell})$, one cannot find solutions $A_{k,\ell}$, $B_{k,\ell}$ all in $\mathcal{Y}$. It is the reason why we need to allow polynomial growth in the functions A, B, F and G as in Proposition 2.2.
2. In the next section, we will see that the shift on the soliton Q_c resulting from the interaction with the soliton Q is obtained from $\widehat{B}_{1,0} \neq 0$.

Remark. In Proposition 2.2, we find one solution of the system (Ω). We refer to Corollary 3.1 for the uniqueness question.

Proof of Proposition 2.2. We first reduce the proof to the case where there is no polynomial functions in F and G. Then we solve the problem using Lemma 2.2 and choosing the free parameter a.

Step 1. *Reduction to the case without polynomial functions*. Let F and G be two functions satisfying the assumptions of Proposition 2.2. First, we consider $\widetilde{A}$ and $\widehat{A}$ the two (unique) polynomial functions satisfying

$$-\widetilde{A}''(x) + \widetilde{A}(x) = \int_0^x \widetilde{F}(z)dz \quad \text{and} \quad -\widehat{A}''(x) + \widehat{A}(x) = \int_0^x \widehat{F}(z)dz$$

(obtained by resolution of a system in the basis $\{x^r\}_{r\geq 0}$). Observe that $\widetilde{A}$ is even and $\widehat{A}$ is odd. Moreover,

- if $\widetilde{F} = 0$ (respectively, $\widehat{F} = 0$), then $\widetilde{A} = 0$ (respectively, $\widehat{A} = 0$);
- if $\widetilde{F} \neq 0$ (respectively, $\widehat{F} \neq 0$), then $\deg \widetilde{A} = 1 + \deg \widetilde{F}$ (respectively, $\deg \widehat{A} = 1 + \deg \widehat{F}$).

We have

$$(\mathcal{L}\widetilde{A})' = (-\widetilde{A}'' + \widetilde{A} - pQ^{p-1}\widetilde{A})' = \widetilde{F} - p(Q^{p-1}\widetilde{A})',$$

$$\begin{aligned}(\mathcal{L}(\varphi\widehat{A}))' &= \left(-\varphi\widehat{A}'' - 2\varphi'\widehat{A}' - \varphi''\widehat{A} + \varphi\widehat{A} - pQ^{p-1}\varphi\widehat{A}\right)' \\ &= \varphi\widehat{F} + \varphi'\int_0^x \widehat{F} + \left(-2\varphi'\widehat{A}' - \varphi''\widehat{A} - pQ^{p-1}\varphi\widehat{A}\right)'.\end{aligned}$$

For $\overline{A}$ to be chosen later, let $A = \overline{A} + \widetilde{A} + \varphi\widehat{A}$. Then, A solves (Ω_A) if and only if

$$(\mathcal{L}\overline{A})' + (\mathcal{L}\widetilde{A})' + (\mathcal{L}(\varphi\widehat{A}))' + a(3Q - 2Q^p) = \overline{F} + \widetilde{F} + \varphi\widehat{F},$$

or equivalently by the previous calculations $(\mathcal{L}\overline{A})' + a(3Q - 2Q^p) = \mathcal{F}$, where

$$\begin{aligned}\mathcal{F} &= \overline{F} + \widetilde{F} - (\mathcal{L}\widetilde{A})' + \varphi\widehat{F} - (\mathcal{L}(\varphi\widehat{A}))' \\ &= \overline{F} - \varphi'\int_0^x \widehat{F} + \left(2\varphi'\widehat{A}' + \varphi''\widehat{A} + pQ^{p-1}(\widetilde{A} + \varphi\widehat{A})\right)'.\end{aligned} \tag{2.24}$$

Since $\overline{F}$, φ', $Q \in \mathcal{Y}$, and $\widetilde{A}$, $\widehat{A}$ and $\widehat{F}$ are polynomial functions, we get $\mathcal{F} \in \mathcal{Y}$. Moreover, we observe that $\mathcal{F}$ is odd.

We proceed in a similar way for $B(x)$ except for the need of an additional parameter $b \in \mathbb{R}$ and the term $(-3A'')$ in equation (Ω_B). Let $\widetilde{B}$ and $\widehat{B}^*$ be the two (unique) polynomial functions satisfying

$$-\widetilde{B}''(x) + \widetilde{B}(x) = \int_0^x \left(\widetilde{G}(z) + 3\widetilde{A}''(z)\right) dz,$$

$$-(\widehat{B}^*)''(x) + \widehat{B}^*(x) = \int_0^x \left(\widehat{G}(z) + 3\widehat{A}''(z)\right) dz.$$

Observe that $\widetilde{B}$ is odd and $\widehat{B}^*$ is even. Moreover,

- if $\widetilde{A}'' = 0$ and $\widetilde{G} = 0$, then $\widetilde{B} = 0$;
- if $\widetilde{A}'' \neq 0$ or $\widetilde{G} \neq 0$, then $\deg \widetilde{B} = 1 + \max(\deg \widetilde{G}, \deg \widetilde{A}'')$;
- if $\widehat{A}'' = 0$ and $\widehat{G} = 0$, then $\widehat{B}^* = 0$;
- if $\widehat{A}'' \neq 0$ or $\widehat{G} \neq 0$, then $\deg \widehat{B}^* = 1 + \max(\deg \widehat{G}, \deg \widehat{A}'')$.

In all cases, we have

$$\deg \widetilde{B} \leq \max(1 + \deg \widetilde{G}, \deg \widetilde{F}), \quad \deg \widehat{B}^* \leq \max(1 + \deg \widehat{G}, \deg \widehat{F}). \tag{2.25}$$

We have

$$(\mathcal{L}\widetilde{B})' = (-\widetilde{B}'' + \widetilde{B} - pQ^{p-1}\widetilde{B})' = \widetilde{G} + 3\widetilde{A}'' - p(Q^{p-1}\widetilde{B})',$$

$$\begin{aligned}(\mathcal{L}(\varphi\widehat{B}^*))' &= \left(-\varphi(\widehat{B}^*)'' - 2\varphi'(\widehat{B}^*)' - \varphi''\widehat{B}^* + \varphi\widehat{B}^* - pQ^{p-1}\varphi\widehat{B}^*\right)' \\ &= \varphi(\widehat{G} + 3\widehat{A}'') + \varphi'\int_0^x (\widehat{G}(z) + 3\widehat{A}''(z))dz \\ &\quad - \left(2\varphi'(\widehat{B}^*)' + \varphi''\widehat{B}^* + pQ^{p-1}\varphi\widehat{B}^*\right)'.\end{aligned}$$

For $\overline{B}$ and b to be chosen later, let

$$B = \overline{B} + \widetilde{B} + \varphi\widehat{B}, \quad \text{with} \quad \widehat{B} = \widehat{B}^* + b.$$

Then, B solves (Ω_B) if and only if

$$(\mathcal{L}\overline{B})' + (\mathcal{L}\widetilde{B})' + (\mathcal{L}(\varphi\widehat{B}))' + 3aQ'' - 3\overline{A}'' - pQ^{p-1}\overline{A} \\ - 3\widetilde{A}'' - pQ^{p-1}\widetilde{A} - 3(\varphi\widehat{A})'' - pQ^{p-1}(\varphi\widehat{A}) = \overline{G} + \widetilde{G} + \varphi\widehat{G},$$

or equivalently by the previous calculations

$$(\mathcal{L}\overline{B})' + 3aQ'' - 3\overline{A}'' - pQ^{p-1}\overline{A} = \mathcal{G} - b(\mathcal{L}\varphi)',$$

where the function $\mathcal{G}$ is defined by

(2.26)

$$\begin{aligned}\mathcal{G} &= \overline{G} + \widetilde{G} + 3\widetilde{A}'' - (\mathcal{L}\widetilde{B})' + \varphi\widehat{G} + 3(\varphi\widehat{A})'' - (\mathcal{L}(\varphi\widehat{B}^*))' + pQ^{p-1}(\widetilde{A} + \varphi\widehat{A}) \\ &= \overline{G} + 6\varphi'\widehat{A}' + 3\varphi''\widehat{A} - \varphi'\int_0^x (\widehat{G}(z) + 3\widehat{A}''(z))dz \\ &\quad + \left(2\varphi'(\widehat{B}^*)' + \varphi''\widehat{B}^* + pQ^{p-1}(\widetilde{B} + \varphi\widehat{B}^*)\right)'.\end{aligned}$$

Since $\overline{G}$, φ', $Q \in \mathcal{Y}$, and $\widetilde{A}$, $\widetilde{B}$, $\widehat{B}^*$ and $\widehat{G}$ are polynomial functions, $\mathcal{G} \in \mathcal{Y}$ is even.

Thus, in conclusion, the system (Ω) is equivalent to the following system in $(a, b, \overline{A}, \overline{B})$:

$$\begin{cases} (\mathcal{L}\overline{A})' + a(3Q - 2Q^p)' = \mathcal{F} \\ (\mathcal{L}\overline{B})' + a(3Q'') - 3\overline{A}'' - pQ^{p-1}\overline{A} = \mathcal{G} - b(\mathcal{L}\varphi)', \end{cases}$$

where $\mathcal{F} \in \mathcal{Y}$ is odd, given by (2.24), $\mathcal{G} \in \mathcal{Y}$ is even, given by (2.26). Note that $\mathcal{F}$ and $\mathcal{G}$ do not depend on the parameters a and b.

Step 2. *Existence of a solution of system* (Ω). We set $\mathcal{H}(x) = \int_{-\infty}^x \mathcal{F}(z)dz$. Since $\mathcal{F}$ is odd, $\int_{\mathbb{R}} \mathcal{F} = 0$ and so $\mathcal{H} \in \mathcal{Y}$ is even.

To find a solution $(a, b, \overline{A}, \overline{B})$ of $(\overline{\Omega})$, it is sufficient to solve

$$(\overline{\Omega}) \qquad \begin{cases} \mathcal{L}\overline{A} + a(3Q - 2Q^p) = \mathcal{H} \\ (\mathcal{L}\overline{B})' + a(3Q'') - 3\overline{A}'' - pQ^{p-1}\overline{A} = \mathcal{G} - b(\mathcal{L}\varphi)'. \end{cases}$$

Since $\int \mathcal{H}Q' = 0$ (by parity) and $\mathcal{H} \in \mathcal{Y}$, it follows from Lemma 2.2(iii)–(iv) that there exists $\overline{H} \in \mathcal{Y}$, even, such that

$$\mathcal{L}\overline{H} = \mathcal{H}.$$

By Lemma 2.2, there also exists $V_0 \in \mathcal{Y}$, even, such that $\mathcal{L}V_0 = 3Q - 2Q^p$. It follows that, for all a,

$$\overline{A} = \overline{H} - aV_0 \qquad (2.27)$$

is solution of $\mathcal{L}\overline{A} + a(3Q - 2Q^p) = \mathcal{H}$, moreover, $\overline{A}$ is even and $\overline{A} \in \mathcal{Y}$. Note that at this point (a, b) are still free. They will be chosen when solving the second equation.

Now, replacing $\overline{A}$ by $\overline{H} - aV_0$ in this equation, we only need to find $\overline{B}$ such that

$$(\mathcal{L}\overline{B})' = -aZ_0 + D - b(\mathcal{L}\varphi)', \tag{2.28}$$

where

$$D = 3\overline{H}'' + pQ^{p-1}\overline{H} + \mathcal{G}, \quad Z_0 = 3Q'' + 3V_0'' + pQ^{p-1}V_0.$$

It follows from the properties of Q, V_0, $\mathcal{G}$ and $\overline{H}$ that D and Z_0 are even and satisfy Z_0, $D \in \mathcal{Y}$. To solve (2.28), it suffices to find $\overline{B} \in \mathcal{Y}$ such that

$$\mathcal{L}\overline{B} = E, \quad \text{where} \quad E = \int_0^x (D - aZ_0)(z)dz - b\mathcal{L}\varphi. \tag{2.29}$$

We can choose (a, b) such that the function E is orthogonal to Q' and has decay at $+\infty$.

Claim 2.3. (i) *Nondegeneracy*:

$$\int Z_0 Q = \frac{p-5}{4(p-1)} \int Q^2. \tag{2.30}$$

(ii) *Let* $a = \frac{\int DQ}{\int Z_0 Q}$ *and* $b = \int_0^{+\infty}(D - aZ_0)(z)dz$. *Then, E defined by* (2.29) *satisfies*

$$E \in \mathcal{Y}, \quad E \textit{ is odd}, \quad \int EQ' = 0. \tag{2.31}$$

Assuming Claim 2.3, we finish the proof of Proposition 2.2. We fix (a, b) as in Claim 2.3. Then, from (2.31) and Lemma 2.2, it follows that there exists $\overline{B} \in \mathcal{Y}$ such that $\mathcal{L}\overline{B} = E$. Setting

$$A = \overline{A} + \widetilde{A} + \widehat{A}, \quad B = \overline{B} + \widetilde{B} + \widehat{B},$$

we have constructed a solution of system (Ω) with the structure described in Proposition 2.2.

Now, we only have to prove Claim 2.3.

Proof of Claim 2.3. *Proof of* (i). First, we check that

$$V_0 = -\tfrac{1}{p-1}Q - \tfrac{3}{2}xQ'. \tag{2.32}$$

Indeed, $\mathcal{L}Q = -Q'' + Q - pQ^{p-1}Q = -(p-1)Q^p$ and $\mathcal{L}(xQ') = -2Q'' + x\mathcal{L}Q' = -2Q + 2Q^p$. Thus,

$$\mathcal{L}(-\tfrac{1}{p-1}Q - \tfrac{3}{2}xQ') = -\frac{1}{p-1}\mathcal{L}Q - \tfrac{3}{2}\mathcal{L}(xQ') = -Q^p + 3Q - 3Q^p = 3Q - 2Q^p.$$

Second, we compute $\int Z_0 Q$, where $Z_0 = 3Q'' + 3V_0'' + pQ^{p-1}V_0$. By $Q'' = Q - Q^p$. We get

$$\begin{aligned}\int Z_0 Q &= \int \left(3Q'' + 3V_0'' + pQ^{p-1}V_0\right) Q\\ &= 3\int Q^2 - 3\int Q^{p+1} + \int V_0(3Q'' + pQ^p)\\ &= 3\int Q^2 - 3\int Q^{p+1} + \int V_0(3Q + (p-3)Q^p).\end{aligned}$$

We compute the last term, integrating by parts:

$$\begin{aligned}\int V_0(3Q + (p-3)Q^p) &= -\int \left(\tfrac{1}{p-1}Q + \tfrac{3}{2}xQ'\right)(3Q + (p-3)Q^p)\\ &= -3\left(\tfrac{1}{p-1} - \tfrac{3}{4}\right)\int Q^2 + (p-3)\left(\tfrac{1}{p-1} - \tfrac{3}{2(p+1)}\right)\int Q^{p+1}\\ &= \frac{3(3p-7)}{4(p-1)}\int Q^2 + \frac{(p-5)(p-3)}{2(p-1)(p+1)}\int Q^{p+1}.\end{aligned}$$

Finally, using Claim C.1 in Appendix C,

$$\int Z_0 Q = \frac{3(7p-11)}{4(p-1)}\int Q^2 - \frac{(5p-7)(p+3)}{2(p-1)(p+1)}\int Q^{p+1} = \frac{p-5}{4(p-1)}\int Q^2.$$

Proof of (ii). Let a and b be defined as in Claim 2.3. The function E is odd by its definition in (2.29). By integration by parts and decay properties of Q, we have

$$\int EQ' = -\int (D - aZ_0)Q - b\int (\mathcal{L}\varphi)Q' = -\int DQ + a\int Z_0 Q - b\int \varphi(\mathcal{L}Q') = 0,$$

by the definition of a, b and $\mathcal{L}Q' = 0$. By Claim 2.1 and the definition of a, b, we have

$$\lim_{+\infty} E = \int_0^{+\infty} (D - aZ_0)\,dz - b\lim_{+\infty}(\mathcal{L}\varphi) = 0 \quad \text{and so } E \in \mathcal{Y}. \qquad \square$$

Resolution of the systems $(\Omega_{k,\ell})$. Using Propositions 2.1 and 2.2, we solve the systems $(\Omega_{k,\ell})$ by induction on (k,ℓ). We check that at given (k,ℓ), the systems $(\Omega_{k',\ell'})$ being solved for all $(k',\ell') \prec (k,\ell)$, we can apply Proposition 2.2 to $(\Omega_{k,\ell})$. The induction argument can be, for example,

1) initialization: $k = 1$, $\ell = 0$;
2) for $\ell = 0$, all $k \geq 1$, by induction on k;
3) by induction on $\ell \geq 0$, all $k \geq 1$ similarly as in 2).

For future use in the proof of Theorem 2.1, we estimate in the next section the degrees of the polynomials $\widetilde{A}_{k,\ell}$, $\widehat{A}_{k,\ell}$, $\widetilde{B}_{k,\ell}$ and $\widehat{B}_{k,\ell}$ with respect to k and ℓ (see Lemma 2.3).

PROPOSITION 2.3 (Resolution of $(\Omega_{k,\ell})$ by induction on (k,ℓ)). *For all $k \in \{1,\dots,k_0\}$, $\ell \in \{0,\dots,\ell_0\}$, there exists $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ of the form*

(2.33)
$$\begin{aligned}
&A_{k,\ell}(x) = \overline{A}_{k,\ell}(x) + \widetilde{A}_{k,\ell}(x) + \varphi(x)\widehat{A}_{k,\ell}(x),\\
&B_{k,\ell}(x) = \overline{B}_{k,\ell}(x) + \widetilde{B}_{k,\ell}(x) + \varphi(x)\widehat{B}_{k,\ell}(x), \quad \textit{where}\\
&\overline{A}_{k,\ell},\ \overline{B}_{k,\ell} \in \mathcal{Y};\ \overline{A}_{k,\ell} \textit{ is even and } \overline{B}_{k,\ell} \textit{ is odd};\\
&\widetilde{A}_{k,\ell} \textit{ and } \widehat{B}_{k,\ell} \textit{ are even polynomials};\ \widehat{A}_{k,\ell} \textit{ and } \widetilde{B}_{k,\ell} \textit{ are odd polynomials};
\end{aligned}$$

satisfying

$$(\Omega_{k,\ell}) \quad \begin{cases} (\mathcal{L}A_{k,\ell})' + a_{k,\ell}(3Q - 2Q^p)' = F_{k,\ell} \\ (\mathcal{L}B_{k,\ell})' + a_{k,\ell}(3Q'') - 3A''_{k,\ell} - pQ^{p-1}A_{k,\ell} = G_{k,\ell}, \end{cases}$$

where $F_{k,\ell}$, $G_{k,\ell}$ are defined in Proposition 2.1. *Moreover, for all $1 \le k \le p-1$, $\ell = 0$,*

$$\widetilde{A}_{k,0} = \widehat{A}_{k,0} = \widetilde{B}_{k,0} = 0, \quad \widehat{B}_{k,0} = b_{k,0} \in \mathbb{R}. \tag{2.34}$$

Remark. (i) The parity condition on $A_{k,\ell}$, $B_{k,\ell}$ is related to the resolution of the systems $(\Omega_{k',\ell'})$ for $(k,\ell) \prec (k',\ell')$. The use of the function φ is related to the asymmetry of the gKdV equation.

(ii) The resolution of $(\Omega_{k,\ell})$ at each step (k,ℓ) does not give a unique solution. Indeed, from Corollary 3.1, if $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ is solution, then for any $(\gamma_{k,\ell}, \delta_{k,\ell}) \in \mathbb{R}^2$,

$$(a_{k,\ell} + \gamma_{k,\ell}a_{1,0}, A_{k,\ell} + \gamma_{k,\ell}(1 + A_{1,0}), B_{k,\ell} + \gamma_{k,\ell}B_{1,0} + \delta_{k,\ell}Q') \tag{2.35}$$

is also solution, which gives two degrees of freedom at each step. From Corollary 3.1, (2.35) is exactly the set of solutions in the class (2.33). Note that for $p = 4$, it seems that one cannot use the parameters to avoid polynomial growth. For $p = 2$, there is a choice of parameters giving no polynomial growth corresponding to the explicit 2-soliton solutions. See Section 5.2.

Proof of Proposition 2.3. The proof of Proposition 2.3 is based on Proposition 2.2 and on the structure of $F_{k,\ell}$ and $G_{k,\ell}$ (see Lemma B.1). By the induction argument described above, it is enough to check that if $(a_{k',\ell'}, A_{k',\ell'}, B_{k',\ell'})$ satisfies (2.33) for all $(k',\ell') \prec (k,\ell)$, then we can find $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ as in (2.33) and solving $(\Omega_{k,\ell})$. This will follow from Proposition 2.2 and the following claim.

CLAIM 2.4. *Let (k,ℓ) be such that $(k,\ell) \in \Sigma_0$, with $(k,\ell) \neq (1,0)$. Assume that for all $(k',\ell') \prec (k,\ell)$, the functions $A_{k',\ell'}$ and $B_{k',\ell'}$ verify* (2.33). *Then*

$$(2.36)\qquad \begin{aligned} &F_{k,\ell}(x) = \overline{F}_{k,\ell}(x) + \widetilde{F}_{k,\ell}(x) + \varphi(x)\widehat{F}_{k,\ell}(x),\\ &G_{k,\ell}(x) = \overline{G}_{k,\ell}(x) + \widetilde{G}_{k,\ell}(x) + \varphi(x)\widehat{G}_{k,\ell}(x), \quad \textit{where}\\ &\overline{F}_{k,\ell},\ \overline{G}_{k,\ell} \in \mathcal{Y};\ \overline{F}_{k,\ell} \textit{ is odd and } \overline{G}_{k,\ell} \textit{ is even};\end{aligned}$$

$\widetilde{F}_{k,\ell}$ and $\widehat{G}_{k,\ell}$ are odd polynomials; $\widehat{F}_{k,\ell}$ and $\widetilde{G}_{k,\ell}$ are even polynomials.

Moreover, let $2 \leq k \leq p-1$, if for any $1 \leq k' < k$,

$$\deg \widetilde{A}_{k',0} = \deg \widehat{A}_{k',0} = \deg \widetilde{B}_{k',0} = \deg \widehat{B}_{k',0} = 0 \quad \textit{then} \quad F_{k,0},\ G_{k,0} \in \mathcal{Y}.$$

Claim 2.4 is a consequence of the more detailed Lemma B.1 proved in Appendix B.

Case $k = 1$, $\ell = 0$. The system $(\Omega_{1,0})$ is explicit from Proposition 2.1; indeed, $F_{1,0} = p(Q^{p-1})'$ and $G_{1,0} = pQ^{p-1}$. Thus

$$F_{1,0} = \overline{F}_{1,0} \in \mathcal{Y},\ G_{1,0} = \overline{G}_{1,0} \in \mathcal{Y} \text{ and } \widetilde{F}_{1,0} = \widehat{F}_{1,0} = \widetilde{G}_{1,0} = \widehat{G}_{1,0} = 0.$$

It follows from Proposition 2.2 that $(\Omega_{1,0})$ has a solution $a_{1,0}$, $A_{1,0}$, $B_{1,0}$ with the desired properties. Moreover, from (2.21)–(2.23), we obtain $\widetilde{A}_{1,0} = \widehat{A}_{1,0} = \widetilde{B}_{1,0} = 0$ and $\widehat{B}_{1,0} = b_{1,0}$, where $b_{1,0} \in \mathbb{R}$ is a constant. Whether or not $b_{1,0}$ is zero will be determined in Section 3 for each case $p = 2$, 3 and 4.

Case $2 \leq k \leq p-1$, $\ell = 0$. By induction on $1 \leq k \leq p-1$, we solve $(\Omega_{k,0})$ and we prove

$$(2.37)\qquad \widetilde{A}_{k,0} = \widehat{A}_{k,0} = \widetilde{B}_{k,0} = 0, \quad \widehat{B}_{k,0} = b_{k,0} \in \mathbb{R}.$$

Indeed, let $2 \leq k \leq p-1$ and assume that (2.37) is true for all $1 \leq k' < k$. Then, it follows from Claim 2.4 that $F_{k,0}, G_{k,0} \in \mathcal{Y}$, which means that $\widetilde{F}_{k,0} = \widehat{F}_{k,0} = \widetilde{G}_{k,0} = \widehat{G}_{k,0} = 0$. Therefore, from Proposition 2.2, we solve $(\Omega_{k,0})$ with property (2.37) at rank k, which completes the induction argument. Thus (2.34) is proved.

Case $k \geq p$, $\ell \geq 0$ *or* $k \geq 1$, $\ell \geq 1$. By induction on (k, ℓ), we prove that $(\Omega_{k,\ell})$ has a solution $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ satisfying (2.33). First, note that (2.33) holds for $1 \leq k \leq p-1$, $\ell = 0$ by (2.37). From Claim 2.4, we know that $F_{k,\ell}$ and $G_{k,\ell}$ have the required structure to apply Proposition 2.2; thus we obtain a solution $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ with the structure (2.33). Thus, the induction argument is complete and the system $(\Omega_{k,\ell})$ is solved up to (k_0, ℓ_0).

2.3. *End of the proof of Theorem* 2.1. We consider the function $v(t)$ constructed in (2.3)–(2.6) where $(a_{k,\ell})$, $(A_{k,\ell})$ and $(B_{k,\ell})$ are defined in Proposition 2.3. For this choice, we have

$$(2.38)\qquad S(t,x) = \sum_{\substack{1 \leq k \leq 5k_0+12\\ 0 \leq \ell \leq 5\ell_0+4\\ k>k_0 \text{ or } \ell>\ell_0}} c^\ell \left(Q_c^k(y_c) F_{k,\ell}(y) + (Q_c^k)'(y_c) G_{k,\ell}(y) \right).$$

Recall that $q = \frac{1}{p-1} - \frac{1}{4}$ and $T_c = c^{-\frac{1}{2}-\frac{1}{100}}$.

Proposition 2.4 (Estimates on W and S). *Let $k_0 \geq 1$, $\ell_0 \geq 0$. There exists K such that, for any $0 < c < 1$, for any $t \in [-T_c, T_c]$, $W(t)$, $S(t)$ belong to $H^s(\mathbb{R})$ for all $s \geq 1$ and satisfy*

$$\|W(t)\|_{H^1} = \|v(t) - R(t) - R_c(t)\|_{H^1} \leq K c^{\frac{1}{p-1}}, \tag{2.39}$$

$$j = 0, 1, 2, \quad \|\partial_x^{(j)} S(t)\|_{L^2} \leq K c^{n_0}, \tag{2.40}$$

where $n_0 = (\frac{1}{2} - \frac{1}{100}) \min(\frac{k_0}{p-1}, \ell_0 + 1)$.

Before proving Proposition 2.4, we claim several preliminary results. The first result concerns the degrees of the polynomials in the decomposition of $W(t)$.

Lemma 2.3. a) *For all $1 \leq k \leq k_0$ and $0 \leq \ell \leq \ell_0$ such that $\frac{k}{p-1} + \ell \leq 2$,*

$$\deg \widetilde{A}_{k,\ell} = \deg \widehat{A}_{k,\ell} = 0. \tag{2.41}$$

b) *For all $1 \leq k \leq k_0$, $0 \leq \ell \leq \ell_0$,*

$$d_{AB}(k,\ell) = \max\left(\deg \widetilde{A}_{k,\ell}, \deg \widehat{A}_{k,\ell}, \deg \widehat{B}_{k,\ell}, \deg \widetilde{B}_{k,\ell}\right) \leq \frac{k-1}{p-1} + \ell. \tag{2.42}$$

Proof of Lemma 2.3. The proof proceeds by induction on (k, ℓ).

Case $k \geq p$, $\ell \geq 0$ or $k \geq 1$, $\ell \geq 1$. By induction on (k, ℓ), we prove that (2.42) holds. First, note that (2.42) holds for $1 \leq k \leq p-1$, $\ell = 0$ by (2.37). Let

$$\xi(k,\ell) = \frac{k-1}{p-1} + \ell. \tag{2.43}$$

Assume

$$\text{for all } (k',\ell') \prec (k,\ell), \quad d_{AB}(k',\ell') \leq \xi(k',\ell') \text{ holds true.} \tag{2.44}$$

From Lemma B.1, we know that $F_{k,\ell}$ and $G_{k,\ell}$ satisfy

$$d_{FG}(k,\ell) \leq \max\left(d_{AB}(k-1,\ell)-1, d_{AB}(k-p+1,\ell), d_{AB}(k,\ell-1), d_{\mathbf{N}}(k,\ell)\right). \tag{2.45}$$

We claim

$$\begin{aligned} &\text{if } k \geq p, \quad d_{AB}(k-p+1,\ell) \leq \xi(k,\ell) - 1,\\ &\text{if } \ell \geq 1, \quad d_{AB}(k,\ell-1) \leq \xi(k,\ell) - 1,\\ &\text{if } k \geq p, \quad d_{\mathbf{N}}(k,\ell) \leq \xi(k,\ell) - 1 \quad (d_{\mathbf{N}} \text{ is defined in Lemma B.1}). \end{aligned}$$

Indeed, assume $k \geq p$. Then by (2.44),

$$d_{AB}(k-p+1,\ell) \leq \xi(k-p+1,\ell) = \xi(k,\ell) - 1. \tag{2.46}$$

Similarly, if $\ell \geq 1$, by (2.44)

$$d_{AB}(k,\ell-1) \leq \xi(k,\ell-1) = \xi(k,\ell) - 1. \tag{2.47}$$

Finally, if $k \geq p$ and if k_j, ℓ_j satisfy $\sum_{j=1}^p k_j \leq k$ and $\sum_{j=1}^p \ell_j \leq \ell$, then (2.44) implies

$$\sum_{j=1}^p d_{AB}(k_j,\ell_j) \leq \sum_{j=1}^p \xi(k_j,\ell_j) = \frac{k-p}{p-1} + \ell = \frac{k-1}{p-1} + \ell - 1 = \xi(k,\ell) - 1.$$

Thus, $d_{\mathbf{N}}(k,\ell) \leq \xi(k,\ell) - 1$.

By (2.45), we obtain $d_{FG}(k,\ell) \leq \xi(k,\ell) - 1$; thus using Proposition 2.2, $(a_{k,\ell}, A_{k,\ell}, B_{k,\ell})$ satisfies (2.19)–(2.20). Therefore

$$\deg_{AB}(k,\ell) \leq \deg_{FG}(k,\ell) + 1 \leq \xi(k,\ell).$$

Thus, the induction argument is complete and the system $(\Omega_{k,\ell})$ is solved up to (k_0,ℓ_0).

We now prove (2.41) to finish the proof of Proposition 2.3.

Case $p \leq k \leq 2(p-1)$, $\ell = 0$. We prove (2.41) for the case $\ell = 0$ by induction on k starting at $k = p$. For $k = p$ and $\ell = 0$ we know that for all $k' < p$, $\widetilde{A}_{k',0} = \widehat{A}_{k',0} = \widetilde{B}_{k',0} = 0$ and $\deg \widehat{B}_{k',0} = 0$. Thus, by Lemma B.1(b), we have $F_{p,0} \in \mathcal{Y}$, and thus by Proposition 2.2, $A_{p,0} \in \mathcal{Y}$, which means $\widetilde{A}_{p,0} = \widehat{A}_{p,0} = 0$. In the statement of Proposition 2.3, we give a weaker statement $\deg \widetilde{A}_{p,0} = \deg \widehat{A}_{p,0} = 0$, since we want that the rest of the estimate to be compatible with nonzero (constant) $\widetilde{A}_{p,0}$ (see §3). The induction argument from p to $2(p-1)$ is done in the same way and we omit it.

Case $1 \leq k \leq p-1$, $\ell = 1$. We also omit this case, since it is similar.

Claim 2.5 (Estimate on $\alpha(s)$). *Let $\alpha(s)$ be the function defined in* (2.5). *Then*

$$\forall s \in \mathbb{R}, \quad |\alpha(s)| \leq K c^{\frac{1}{p-1} - \frac{1}{2}}, \quad |\alpha'(s)| \leq K c^{\frac{1}{p-1}}.$$

In particular, there exists $c_0 > 0$ so that for all $0 < c < c_0$, for all $s \in \mathbb{R}$, $|\alpha'(s)| \leq \frac{1}{2}$.

Remark. From now on, we choose $c > 0$ small enough so that $1 + \alpha' > 1/2$ for all $s \in \mathbb{R}$.

Proof of Claim 2.5. We have

$$|\alpha(s)| \leq \sum_{(k,\ell)\in\Sigma_0} \left| a_{k,\ell} \int_0^s Q_c^k(s')ds' \right| \leq \max_{(k,\ell)\in\Sigma_0} |a_{k,\ell}| \times \sum_{(k,\ell)\in\Sigma_0} c^\ell \int Q_c^k \leq K \int Q_c.$$

Since $Q_c(s') = c^{\frac{1}{p-1}} Q(\sqrt{c}\, s')$, $|\alpha(s)| \leq K \int Q_c = K c^{\frac{1}{p-1} - \frac{1}{2}} \int Q$ and similarly $|\alpha'(s)| \leq K c^{\frac{1}{p-1}}$.

CLAIM 2.6 (H^1-estimates). *Let $0 < c < 1/2$. Let $f \in \mathcal{Y}$ and let P be a polynomial function of degree d. Then, for all $k \geq 1$, $\ell \geq 0$, for all $t \in [-T_c, T_c]$,*

$$\|c^\ell Q_c^k(y_c)f(y)\|_{H^1} + c^{-\frac{1}{2}}\|c^\ell (Q_c^k)'(y_c)f(y)\|_{H^1} \leq K c^{\frac{k}{p-1}+\ell} e^{-(1-c)\sqrt{c}\,|t|},$$

$$\begin{aligned}&\|c^\ell Q_c^k(y_c)P(y)\|_{H^1}\\ &\quad \leq K c^{\xi(k,\ell)+q-\frac{d}{2}(1+\frac{1}{50})} + c^{-\frac{1}{2}}\|c^\ell (Q_c^k)'(y_c)P(y)\|_{H^1} \leq K c^{\xi(k,\ell)+q-\frac{d}{2}(1+\frac{1}{50})}.\end{aligned}$$

Proof of Claim 2.6. Let $f \in \mathcal{Y}$, so that $|f(y)| \leq K|y|^r e^{-|y|}$ on $\mathbb{R}$. By $Q_c(x) = c^{\frac{1}{p-1}} Q(\sqrt{c}x) \leq K c^{\frac{1}{p-1}} e^{-\sqrt{c}|x|}$, we have

$$|c^\ell Q_c^k(y_c)f(y)|^2 \leq K c^{\frac{2k}{p-1}+2\ell} e^{-2k\sqrt{c}|y_c|}|y|^{2r}e^{-2|y|} \leq K c^{\frac{2k}{p-1}+2\ell} e^{-2\sqrt{c}|y_c|}|y|^{2r}e^{-2|y|}.$$

Since $y_c = x + (1-c)t$ and $y = x + \alpha(y_c)$, we have $y_c = y + (1-c)t - \alpha(y_c)$, and so by Claim 2.5,

$$\sqrt{c}|y_c| \geq \sqrt{c}((1-c)|t| - |y| - |\alpha(y_c)|) \geq (1-c)\sqrt{c}|t| - \sqrt{c}|y| - K.$$

Thus,

$$|c^\ell Q_c^k(y_c)f(y)|^2 \leq K c^{\frac{2k}{p-1}+2\ell} e^{-2(1-c)\sqrt{c}|t|}|y|^{2r}e^{-2(1-\sqrt{c})|y|} \leq K c^{\frac{2k}{p-1}+2\ell} e^{-|y|}.$$

By changing the variable, $y = x + \alpha(x + (1-c)t)$, and using Claim 2.5, we have

$$\int e^{-|y|}dx = \int e^{-|y|}\frac{dy}{1+\alpha'(y_c)} \leq 2\int e^{-|y|}dy \leq K.$$

Thus, $\|c^\ell Q_c^k(y_c)f(y)\|_{L^2} \leq K c^{\frac{k}{p-1}+\ell} e^{-(1-c)\sqrt{c}\,|t|}$.

Since $|Q_c'| \leq \sqrt{c}\,Q_c$ (recall $(Q_c')^2 = cQ_c^2 - \frac{2}{p+1}Q_c^{p+1}$), we also get

$$\|c^\ell (Q_c^k)'(y_c)f(y)\|_{L^2} \leq K c^{\frac{k}{p-1}+\ell+\frac{1}{2}} e^{-(1-c)\sqrt{c}\,|t|}.$$

Since $\partial_x(c^\ell Q_c^k(y_c)f(y)) = c^\ell (Q_c^k)'(y_c)f(y) + (1+\alpha'(y_c))c^\ell Q_c^k(y_c)f'(y)$, and $f' \in \mathcal{Y}$, the above estimates and Claim 2.5 give the H^1 estimate on $c^\ell Q_c^k(y_c)f(y)$. The proof of the estimates for $c^\ell (Q_c^k)'(y_c)f(y)$ is similar.

Now, we consider a monomial function $P(y) = y^d$. For all $t \in [-T_c, T_c]$, and by Claim 2.5,

$$y = y_c - (1-c)t + \alpha(y_c) \quad \text{and so} \quad |y| \leq |y_c| + T_c + K c^{\frac{1}{p-1}-\frac{1}{2}} \leq |y_c| + K c^{-\frac{1}{2}(1+\frac{1}{50})}.$$

Therefore,

$$|c^\ell Q_c^k(y_c)P(y)| = c^\ell Q_c^k(y_c)|y|^d \leq K\left(c^\ell |y_c|^d Q_c^k(y_c) + c^{\ell-\frac{d}{2}(1+\frac{1}{50})}Q_c^k(y_c)\right).$$

By $Q_c(x) = c^{\frac{1}{p-1}}Q(\sqrt{c}x)$,

$$\|c^\ell |y_c|^d Q_c^k(y_c)\|_{L^2} = c^{\ell+\frac{k}{p-1}-\frac{d}{2}-\frac{1}{4}}\||x'|^d Q^k(x)\|_{L^2},$$

$$\|c^{\ell-\frac{d}{2}(1+\frac{1}{50})}Q_c^k(y_c)\|_{L^2} = c^{\ell+\frac{k}{p-1}-\frac{d}{2}(1+\frac{1}{50})-\frac{1}{4}}\|Q^k\|_{L^2}.$$

Thus, $\|c^\ell Q_c^k(y_c)P(y)\|_{L^2} \leq Kc^{\ell+\frac{k}{p-1}-\frac{d}{2}(1+\frac{1}{50})-\frac{1}{4}}\|Q^k\|_{L^2}$. The other estimates are obtained in a similar way.

Proof of Proposition 2.4. From Claim 2.6, we claim sharp estimates on the terms in $W(t)$ and $S(t)$. These estimates are applied to prove Proposition 2.4 and will be used again in the rest of this paper.

CLAIM 2.7 (Estimates for terms in $W(t)$). *For all* $t \in [-T_c, T_c]$,

(a) *for all* $1 \leq k \leq p-1$, $\ell = 0$,

$$\|Q_c^k(y_c)A_{k,0}(y)\|_{H^1} \leq Kc^{\frac{k}{p-1}}e^{-(1-c)\sqrt{c}|t|}, \tag{2.48}$$

$$\|c^\ell (Q_c^k)'(y_c)B_{k,0}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|c^\ell \partial_x((Q_c^k)'(y_c)B_{k,0}(y))\|_{L^2} \leq Kc^{\frac{k}{p-1}+\frac{1}{4}}; \tag{2.49}$$

(b) *for all* $1 \leq k \leq k_0$ *and* $0 \leq \ell \leq \ell_0$ *such that* $\frac{k}{p-1} + \ell \leq 2$,

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|c^\ell \partial_x(Q_c^k(y_c)A_{k,\ell}(y))\|_{L^2} \leq Kc^{\xi(k,\ell)+q}, \tag{2.50}$$

$$\|c^\ell (Q_c^k)'(y_c)B_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|c^\ell \partial_x((Q_c^k)'(y_c)B_{k,\ell}(y))\|_{L^2} \leq Kc^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q+\frac{1}{2}}; \tag{2.51}$$

(c) *for all* $1 \leq k \leq k_0$, $0 \leq \ell \leq \ell_0$,

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|c^\ell \partial_x(c^\ell Q_c^k(y_c)A_{k,\ell}(y))\|_{L^2} \leq Kc^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q}, \tag{2.52}$$

$$\begin{aligned}\|c^\ell (Q_c^k)'(y_c)B_{k,\ell}(y)\|_{L^2} &+ \tfrac{1}{\sqrt{c}}\|c^\ell \partial_x(c^\ell (Q_c^k)'(y_c)B_{k,\ell}(y))\|_{L^2} \\ &\leq Kc^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q+\frac{1}{2}}.\end{aligned} \tag{2.53}$$

Proof of Claim 2.7. By Proposition 2.3, we have

$$A_{k,\ell} = \overline{A}_{k,\ell} + \widetilde{A}_{k,\ell} + \varphi\widehat{A}_{k,\ell}, \quad B_{k,\ell} = \overline{B}_{k,\ell} + \widetilde{B}_{k,\ell} + \varphi\widehat{B}_{k,\ell},$$

where $\overline{A}_{k,\ell}$, $\overline{B}_{k,\ell} \in \mathcal{Y}$ and $\widetilde{A}_{k,\ell}$, $\widehat{A}_{k,\ell}$, $\widehat{B}_{k,\ell}$ $\widetilde{B}_{k,\ell}$ are polynomial functions satisfying (Proposition 2.3(c)):

$$\max\left(\deg\widetilde{A}_{k,\ell}, \deg\widehat{A}_{k,\ell}, \deg\widehat{B}_{k,\ell}, \deg\widetilde{B}_{k,\ell}\right) \leq \frac{k-1}{p-1} + \ell = \xi(k,\ell). \tag{2.54}$$

By Claim 2.6, $c^\ell Q_c^k(y_c)A_{k,\ell}(y)$ and $c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y)$ belong to H^1.

Proof of (c). From the estimates of Claim 2.6 applied to $A_{k,\ell}$ and $B_{k,\ell}$, we obtain from (2.54), for all $t \in [-T_c, T_c]$,

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} \leq Kc^{\xi(k,\ell)+q-\frac{1}{2}(1+\frac{1}{50})\xi(k,\ell)} \leq Kc^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q}$$

and

$$\|c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y)\|_{L^2} \leq Kc^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+\frac{1}{2}+q}.$$

The proof for $\|\partial_x(c^\ell Q_c^k(y_c)A_{k,\ell}(y))\|_{L^2}$ is the same, except that since

$$\partial_x(c^\ell Q_c^k(y_c)A_{k,\ell}(y)) = c^\ell (Q_c^k)'(y_c)A_{k,\ell}(y) + c^\ell Q_c^k(y_c)A'_{k,\ell}(y),$$

there is a gain of $\sqrt{c}$ due to derivation of Q_c , and of $c^{-\frac{1}{2}(1+\frac{1}{50})}$ due to derivation of polynomial terms in $A_{k,\ell}$ (see Claim 2.6).

Proof of (a). Note that by Proposition 2.3(a) for all $1 \le k \le p-1$, $\widetilde{A}_{k,0} = \widehat{A}_{k,0} = 0$, which means $A_{k,0} \in \mathcal{Y}$ and thus for such k, by Claim 2.6, for all $t \in \mathbb{R}$,

$$\|Q_c^k(y_c)A_{k,0}(y)\|_{H^1} \le K c^{\frac{k}{p-1}} e^{-(1-c)\sqrt{c}|t|}.$$

For such k and ℓ, $\widetilde{B}_{k,0} = 0$ and $\deg \widehat{B}_{k,0} = 0$, and thus, for all $t \in [-T_c, T_c]$,

$$\|c^\ell (Q_c^k)'(y_c)B_{k,0}(y)\|_{L^2} \le K c^{\frac{k}{p-1}+\frac{1}{4}}.$$

Proof of (b). From Proposition 2.3(b) and Claim 2.6, in the case $\frac{k}{p-1}+\ell \le 2$ we obtain

$$\forall t \in [-T_c, T_c], \quad \|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} \le K c^{\xi(k,\ell)+q},$$
$$\|c^\ell (Q_c^k)'(y_c)B_{k,0}(y)\|_{L^2} \le K c^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q+\frac{1}{2}}.$$

Claim 2.8 (Estimates for terms in $S(t)$). *For all (k,ℓ) satisfying $k_0+1 \le k \le K_0$ or $\ell_0+1 \le \ell \le L_0$, for all $t \in [-T_c, T_c]$,*

$$\|c^\ell Q_c^k(y_c)F_{k,\ell}(y)\|_{H^1} + \|c^\ell (Q_c^k)'(y_c)G_{k,\ell}(y)\|_{H^1} \le K c^{n_0}, \tag{2.55}$$

for $n_0 = \frac{1}{2}(1-\frac{1}{50}) \min\left(\frac{k_0}{p-1}, 1+\ell_0\right)$.

Proof of Claim 2.8. Assume, for example, that $k \ge k_0+1$. By Claim 2.6, for all $t \in [-T_c, T_c]$,

$$\|c^\ell Q_c^k(y_c)F_{k,\ell}(y)\|_{H^1} + \|c^\ell (Q_c^k)'(y_c)G_{k,\ell}(y)\|_{H^1}$$
$$\le K c^{\xi(k,\ell)+q-\frac{1}{2}(1+\frac{1}{50})\xi(k_0+1,\ell)+\frac{1}{2}(1+\frac{1}{50})} \le K c^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)} \le c^{\frac{1}{2}(1-\frac{1}{50})\frac{k_0}{p-1}}. \qquad \square$$

Recall $W(t,x) = \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(y_c)A_{k,\ell}(y) + (Q_c^k)'(y_c)B_{k,\ell}(y)\right)$. We apply the estimates of Claim 2.7 to each term of $W(t)$, for all $t \in [-T_c, T_c]$:

(1) for $1 \le k \le p-1$ and $\ell = 0$, we have

$$\|Q_c^k(y_c)A_{k,0}(y)\|_{H^1} \le K c^{\frac{k}{p-1}} e^{-(1-c)\sqrt{c}|t|} \le K c^{\frac{1}{p-1}};$$

(2) for $k \ge p$ and $\ell \ge 0$, or $k \ge 1$ and $\ell \ge 1$, for $\xi(k,\ell) \ge 1$ we have

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{H^1} \le K c^{\frac{1}{2}(1-\frac{1}{50})\xi(k,\ell)+q} \le K c^{\frac{1}{4}+q} = K c^{\frac{1}{p-1}}$$

and similarly for $\|c^\ell (Q_c^k)'(y_c)B_{k,\ell}(y)\|$.

Thus, for all $t \in [-T_c, T_c]$, $\|W(t)\|_{H^1} \le K c^{\frac{1}{p-1}}$.

By (2.38), for a given $k_0 \ge 1$ and $\ell_0 \ge 0$, the rest $S(t,x)$ contains only terms for $k \ge k_0+1$ or terms for $\ell \ge \ell_0+1$. Thus, from Claim 2.7, for all

$t \in [-T_c, T_c]$, $\|S(t)\|_{H^1} \leq Kc^{n_0}$, where $n_0 = \frac{1}{2}(1-\frac{1}{50})\min(\frac{k_0}{p-1}, \ell_0 + 1)$. The proof for $\|\partial_x^j S(t)\|_{H^1}$, for $j = 1, 2$, is the same.

3. Recomposition of the approximate solution at $\pm T_c$

In this section, we consider the function v defined in Theorem 2.1. We prove further properties of v by solving explicitly the first two systems $(\Omega_{1,0})$ and $(\Omega_{2,0})$. Detailed properties depend on the specific value of $p = 2$ or 4.

3.1. *Explicit resolution of the first systems.*

1. *Resolution of the systems* $(\Omega_{1,0})$, $(\Omega_{2,0})$ *for* $p = 2, 4$. We begin with two technical results.

Claim 3.1 (Expression of V_1). *Let $V_1 \in \mathcal{Y}$, even, be a solution of $\mathcal{L}V_1 = pQ^{p-1}$. Then $V_1 = -2Q - xQ'$ for $p = 2$ and $V_1 = \frac{1}{3}\Big(Q'(\int_0^x Q^2) - 2Q^3\Big)$ for $p = 4$.*

Proof. For $p = 2$, set $V_1 = -2Q - xQ'$. Then, using the equation of Q,

$$\mathcal{L}V_1 = -V_1'' + V_1 - 2QV_1 = (2Q'' + 2Q'' + xQ^{(3)}) - 2Q - xQ' + 4Q^2 + 2xQQ' = 2Q.$$

Now, let $p = 4$. By $\mathcal{L}(fg) = g(\mathcal{L}f) - 2f'g' - fg''$, we have

$$\mathcal{L}(Q'(\textstyle\int_0^x Q^2)) = (\int_0^x Q^2)\mathcal{L}Q' - 2Q''Q^2 - 2(Q')^2Q,$$

but from Lemma 2.2, $\mathcal{L}Q' = 0$, so that by $Q'' = Q - Q^4$ and $(Q')^2 = Q^2 - \frac{2}{5}Q^5$,

$$\mathcal{L}(Q'(\textstyle\int_0^x Q^2)) = -2Q^3 + 2Q^6 - 2Q^3 + \dfrac{4}{5}Q^6 = -4Q^3 + \dfrac{14}{5}Q^6.$$

We also have $\mathcal{L}Q^3 = -3Q''Q^2 - 6(Q')^2Q + Q^3 - 4Q^6 = -8Q^3 + \frac{7}{5}Q^6$. Thus, by combining these two calculations, we get $\frac{1}{3}\mathcal{L}(Q'(\int_0^x Q^2) - 2Q^3) = 4Q^3$.

Claim 3.2 (Computation of $\int Z_1 Q$). *Let $Z_1 = 3V_1'' + pQ^{p-1}V_1 + pQ^{p-1}$. Then*

$$\int Z_1 Q = \frac{p-3}{2(p-1)} \int Q.$$

Proof of Claim 3.2.

$$\int Z_1 Q = \int (3V_1'' + pQ^{p-1}V_1 + pQ^{p-1})Q = \int V_1(3Q'' + pQ^p) + p\int Q^p.$$

Since $\mathcal{L}Q = -(p-1)Q^p$ and $\mathcal{L}(\frac{2}{p-1}Q + xQ') = -2Q$, we have

$$\begin{aligned} 3Q'' + pQ^p &= 3Q + (p-3)Q^p \\ &= \mathcal{L}(-\tfrac{3}{2}(\tfrac{2}{p-1}Q + xQ') - \tfrac{p-3}{p-1}Q) = -\mathcal{L}(\tfrac{p}{p-1}Q + \tfrac{3}{2}xQ'). \end{aligned}$$

Thus,

$$\begin{aligned}\int Z_1 Q &= -\int V_1 \mathcal{L}(\tfrac{p}{p-1}Q + \tfrac{3}{2}xQ') + p\int Q^p \\ &= -\int (\mathcal{L}V_1)(\tfrac{p}{p-1}Q + \tfrac{3}{2}xQ') + p\int Q^p \\ &= -p\int Q^{p-1}(\tfrac{p}{p-1}Q + \tfrac{3}{2}xQ') + p\int Q^p = \frac{p-3}{2(p-1)}\int Q^p,\end{aligned}$$

by integration by parts. Since $Q = Q^p + Q''$, we have $\int Q^p = \int Q$, and the claim follows.

LEMMA 3.1 (Resolution of the first systems for $p = 2, 4$).

- *For* $p = 2$,

$$a_{1,0} = \frac{2}{3}, \quad b_{1,0} = -2, \quad A_{1,0} = -\frac{4}{3}Q, \quad B_{1,0} = -2\varphi.$$

$$a_{2,0} = -\frac{4}{9}, \quad a_{1,1} = \frac{2}{3}, \quad A_{2,0} = -2 + \frac{4}{3}Q, \quad A_{1,1} = 2 - \frac{2}{3}Q - \frac{1}{3}xQ', \quad b_{2,0} = \frac{4}{3}.$$

- *For* $p = 4$,

$$\begin{aligned} a_{1,0} &= -2\frac{\int Q}{\int Q^2}, \quad b_{1,0} = -\frac{1}{2}\int Q^3 + \frac{1}{6}\frac{\left(\int Q\right)^2}{\int Q^2} < 0, \\ A_{1,0} &= \frac{1}{3}(Q'(\textstyle\int_0^x Q^2) - 2Q^3) + 2\frac{\int Q}{\int Q^2}(-\tfrac{1}{3}Q - \tfrac{3}{2}xQ'), \\ b_{2,0} &= -\frac{1}{18}\int Q^2 - \frac{3}{4}\frac{\left(\int Q\right)\left(\int Q^3\right)}{\int Q^2} - \frac{1}{18}\frac{\left(\int Q\right)^3}{\left(\int Q^2\right)^2} < 0. \end{aligned}$$

From Corollary 3.1, there are several solutions. The choice of the solution for $p = 2$ above is related to the exact 2-soliton solutions.

We only solve $(\Omega_{1,0})$ in this paper. The resolution of the next systems is done in [24].

Proof of Lemma 3.1. From Propositions 2.1 and 2.3, the system $(\Omega_{1,0})$ writes, for $p = 2$, 3 and 4,

$$(\Omega_{1,0}) \quad \begin{cases} \mathcal{L}A_{1,0} + a_{1,0}(3Q - 2Q^p) = pQ^{p-1} \\ (\mathcal{L}B_{1,0})' + a_{1,0}(3Q'') - 3A''_{1,0} - pQ^{p-1}A_{1,0} = pQ^{p-1}. \end{cases}$$

Computation of $A_{1,0}$. Recall from Claim 2.3 that $V_0 = -\frac{1}{p-1}Q - \frac{3}{2}xQ'$ and $\mathcal{L}V_0 = 3Q - 2Q^p$. Thus, the function $A_{1,0} = V_1 - a_{1,0}V_0$ solves the first line of $(\Omega_{1,0})$, independently of the value of $a_{1,0}$. By replacing $A_{1,0}$ in the second line of the system $(\Omega_{1,0})$,

$$(\mathcal{L}B_{1,0})' + a_{1,0}Z_0 = Z_1,$$

where

$$Z_0 = 3Q'' + 3V_0'' + pQ^{p-1}V_0, \quad Z_1 = 3V_1'' + pQ^{p-1}V_1 + pQ^{p-1}. \tag{3.1}$$

Computation of $a_{1,0}$. Since $\mathcal{L}Q' = 0$, we have $\int(\mathcal{L}B_{1,0})'Q = 0$ and so $a_{1,0}\int Z_0 Q = \int Z_1 Q$. Recall that by Claim 2.3, $\int Z_0 Q = \frac{p-5}{4(p-1)}\int Q^2$. Assuming this, we obtain

$$a_{1,0} = 2\,\frac{p-3}{p-5}\,\frac{\int Q}{\int Q^2}. \tag{3.2}$$

• $p = 2$. Since $Q = Q^2 + Q''$, we have $\int Q = \int Q^2$ and so (3.2) gives $a_{1,0} = \frac{2}{3}$. Next, $A_{1,0} = V_1 - \frac{2}{3}V_0 = -2Q - xQ' - \frac{2}{3}(-Q - \frac{3}{2}xQ') = -\frac{4}{3}Q$.

By the second line of the system $(\Omega_{1,0})$, we get

$$\begin{aligned}(\mathcal{L}B_{1,0})' &= 2Q - (3a_{1,0}Q'' - 3A_{1,0}'' - 2QA_{1,0})\\ &= 2Q - (2Q'' + 4Q'' + \frac{8}{3}Q^2) = -4Q + \frac{10}{3}Q^2.\end{aligned}$$

From Claim 2.1, we have $(\mathcal{L}\varphi)' = 2Q - \frac{5}{3}Q^2$; thus, $B_{1,0} = -2\varphi$ is solution.

From Proposition 2.1, we write the following two systems for $p = 2$:

$$(\Omega_{2,0})\quad \begin{cases} (\mathcal{L}A_{2,0})' + a_{2,0}(3Q - 2Q^2)' \\ \qquad = (-A_{1,0} + A_{1,0}^2)' - (3B_{1,0}'' + 2QB_{1,0}) \\ \qquad\quad - a_{1,0}(Q + 3A_{1,0}'' + 2QA_{1,0})' + 3a_{1,0}^2 Q^{(3)}, \\ (\mathcal{L}B_{2,0})' + 3a_{2,0}Q'' - 3A_{2,0}'' - 2QA_{2,0} \\ \qquad = A_{1,0} + A_{1,0}^2 + (-2B_{1,0} + A_{1,0}B_{1,0})' \\ \qquad\quad - \frac{1}{2}a_{1,0}(9A_{1,0}' + 3B_{1,0}'' + 2QB_{1,0})' + \frac{3}{2}a_{1,0}^2 Q'', \end{cases}$$

$$(\Omega_{1,1})\quad \begin{cases} (\mathcal{L}A_{1,1})' + a_{1,1}(3Q - 2Q^2)' = 3A_{1,0}' + 3B_{1,0}'' + 2QB_{1,0} \\ (\mathcal{L}B_{1,1})' + 3a_{1,1}Q'' - 3A_{1,1}'' - 2QA_{1,1} = 3B_{1,0}'. \end{cases}$$

The resolution of these two systems is done in [24].

• $p = 4$. From (3.2), we obtain the expression of $a_{1,0}$, and from $A_{1,0} = V_1 - a_{1,0}V_0$ and the expressions of V_1 and V_0, we obtain $A_{1,0}$. Here $a_{1,0} < 0$, which will have a surprising consequence on the shift of Q after collision (see Proposition 3.1).

Next, $B_{1,0}$ is of the form $B_{1,0} = \overline{B}_{1,0} + \varphi b_{1,0}$, where $\overline{B}_{1,0} \in \mathcal{Y}$ and $b_{1,0} \in \mathbb{R}$ from Proposition 2.2. We do not compute $B_{1,0}$ in this case. Thus we only need to compute $b_{1,0}$. By Claim 2.1,

$$2b_{1,0} = \lim_{+\infty} B_{1,0} - \lim_{-\infty} B_{1,0} = \lim_{+\infty}\mathcal{L}B_{1,0} - \lim_{-\infty}\mathcal{L}B_{1,0}.$$

Recall the equation of $B_{1,0}$: $(\mathcal{L}B_{1,0})' = Z_1 - a_{1,0}Z_0$, where $Z_0 = 3Q'' + 3V_0'' + 4Q^3V_0$ and $Z_1 = 4Q^3 + 3V_1'' + 4Q^3V_1$. It follows that $2b_{1,0} = \int Z_1 - a_{1,0}\int Z_0$.

By integration by parts,

$$\int Z_0 = 4\int Q^3 V_0 = -4\int Q^3(\tfrac{1}{3}Q + \tfrac{3}{2}xQ') = \tfrac{1}{6}\int Q^4 = \tfrac{1}{6}\int Q,$$
$$\int Z_1 = 4\int Q^3 + \tfrac{4}{3}\int (Q^3 Q'(\textstyle\int_0^x Q^2) - 2Q^6) = 4\int Q^3 - 3\int Q^6 = -\int Q^3$$

since $3\int Q^6 = 5\int Q^3$ (from the equation of Q). Thus, $2b_{1,0} = -\int Q^3 - \frac{a_{1,0}}{6}\int Q$, which gives the desired formula.

We justify that $b_{1,0} < 0$. By Cauchy-Schwarz' inequality and Claim C.1, we have

$$\int Q = \int Q^4 \leq \Big(\int Q^2\Big)^{1/2}\Big(\int Q^6\Big)^{1/2} = \sqrt{\frac{5}{3}}\Big(\int Q^2\Big)^{1/2}\Big(\int Q^3\Big)^{1/2}.$$

Thus, $\frac{1}{6}\frac{\left(\int Q\right)^2}{\int Q^2} \leq \frac{5}{18}\int Q^3$ and so $b_{1,0} \leq -\frac{2}{9}\int Q^3$. Numerically, $b_{1,0} \sim -0.9$.

System $(\Omega_{2,0})$ for $p = 4$ writes

$$\begin{cases} (\mathcal{L}A_{2,0})' + a_{2,0}(3Q - 2Q^4)' \\ \qquad\qquad = (6Q^2(1 + A_{1,0})^2)' - a_{1,0}(4Q^3 + 3A''_{1,0} + 4Q^3 A_{1,0})' + 3a_{1,0}^2 Q^{(3)}, \\ (\mathcal{L}B_{2,0})' + 3a_{2,0}Q'' - 3A''_{2,0} - 4Q^3 A_{2,0} \\ \qquad\qquad = 6Q^2(1 + A_{1,0})^2 + (6Q^2 B_{1,0}(1 + A_{1,0}))' \\ \qquad\qquad\quad - \frac{1}{2}a_{1,0}(9A'_{1,0} + 3B''_{1,0} + 4Q^3 B_{1,0})' + \frac{3}{2}a_{1,0}^2 Q''. \end{cases}$$

The fact that $b_{2,0} \neq 0$ can easily be checked by solving $(\Omega_{2,0})$ numerically. However, we were able to give an explicit expression of $b_{2,0}$, by long but elementary calculations; see [24].

2. *Determination of all solutions of* (Ω). Now, let us justify the remark following Proposition 2.3 concerning the existence of several solutions of system $(\Omega_{k,\ell})$. At each step of resolution, the number of solutions of $(\Omega_{k,\ell})$ is related to the existence of nontrivial solutions of the system (Ω_0)

$$(\Omega_0) \quad \begin{cases} (\mathcal{L}A_0)' + a_0(3Q - 2Q^p)' = 0 \\ (\mathcal{L}B_0)' + a_0(3Q'') - 3A''_0 - pQ^{p-1}A_0 = 0. \end{cases}$$

Corollary 3.1. *Assume that* (a_0, A_0, B_0) *solves the system* (Ω_0), *where* A_0 *is a* C^∞ *even function, with at most polynomial growth at* ∞, *and* B_0 *is a* C^∞ *odd function, with at most polynomial growth at* ∞. *Then, there exists* $\gamma \in \mathbb{R}$ *and* $\delta \in \mathbb{R}$ *such that*

$$(a_0, A_0, B_0) = (\gamma a_{1,0}, \gamma(1 + A_{1,0}), \gamma B_{1,0} + \delta Q'). \tag{3.3}$$

Conversely, for any $\gamma, \delta \in \mathbb{R}$, (3.3) *defines a solution of* (Ω_0).

Proof of Corollary 3.1. The first line of (Ω_0) is equivalent to

$$\mathcal{L}A_0 + a_0(3Q - 2Q^p) = \gamma,$$

where γ is a constant. Since $\mathcal{L}1 = 1 - pQ^{p-1}$, we have $\mathcal{L}(A_0 - \gamma + a_0V_0 - \gamma V_1) = 0$. Claim 2.2 implies that if $\mathcal{L}f = 0$ where f is a function with at most polynomial growth, then $f \in L^2(\mathbb{R})$, and so $f = \overline{\delta}Q'$. Since A_0 is even and has at most polynomial growth, we obtain

$$A_0 = \gamma V_1 - a_0V_0 + \gamma.$$

The resolution of the second line of the system is similar to the previous calculations on $(\Omega_{1,0})$:

$$(\mathcal{L}B_0)' = \gamma Z_1 - a_0Z_0,$$

which gives a relation between a_0 and γ: $a_0\int Z_0Q = \gamma\int Z_1Q$, which means that $a_0 = \gamma a_{1,0}$, and so $A_0 = \gamma(1 + A_{1,0})$. Thus, $(\mathcal{L}B_0)' = \gamma(Z_0 - a_{1,0}Z_1) = \gamma(\mathcal{L}B_1)'$, and so $\mathcal{L}(B_0 - \gamma B_{1,0}) = 0$ by parity. Therefore, $B_0 = \gamma B_{1,0} + \delta Q'$.

3.2. *Asymptotics of the approximate solution at* $\pm T_c$. So far, we have searched an approximate solution v on $[-T_c, T_c]$ with a structure adapted to the interaction problem. For $t \in [-T_c, T_c]$, $v(t) = Q(y) + Q(y_c) + W(t)$, with $\|W(\pm T_c)\|_{H^1} \leq Kc^{\frac{1}{p-1}} \sim Kc^{\frac{1}{4}}\|Q_c\|_{H^1}$. Nevertheless, since the functions $A_{k,\ell}$, $B_{k,\ell}$ may contain polynomial functions of degree larger than 1, the previous decomposition is not adapted for $t > T_c$.

At $t = T_c$, we note that $y_c \sim x + T_c$ and $y \sim x - \frac{\Delta}{2}$, where $|\Delta|/T_c \ll 1$. Thus $v(T_c)$ is close to the sum of two exponentially decoupled solitons, and for $t > T_c$, one can use asymptotic techniques (see §4) close to 2-soliton solutions, or equivalently close to the sum of two solitons for the proofs. This set of 2-soliton solutions have several parameters, as the size and the position of each soliton. In this section, we understand what is the optimal choice for these parameters. In fact, at the formal level, from the decomposition, the size parameters will not be changed, we will concentrate on the position parameters.

First, we point out that the function $v(t,x)$ is, as the (gKdV) equation, invariant by the transformation $x \to -x$, $t \to -t$. Indeed,

$$y_c(-t,-x) = -y_c(t,x), \quad \alpha(-s) = -\alpha(s), \quad y(-t,-x) = -y(t,x)$$

and

$$\begin{aligned} v(-t,-x) &= Q(-y) + Q_c(-y_c) \\ &\quad + \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(-y_c)A_{k,\ell}(-y) + (Q_c^k)'(-y_c)B_{k,\ell}(-y)\right) = v(t,x), \end{aligned}$$

by the parity properties of the functions Q, Q_c, $A_{k,\ell}$ and $B_{k,\ell}$. Thus it suffices to study the properties of the function v for $t \geq 0$.

Let us present formal computations to recompose $v(T_c)$ in terms of the asymptotic 2-soliton family at $t \to +\infty$. We first observe that Q and Q_c are well-ordered and located far away in the original space variable x at $t = T_c$.

Indeed, if $x > -T_c/2$, then $y_c = x + (1-c)t > T_c/4$ (say, $0 < c < 1/4$); thus the soliton Q_c is at the left of $x = -T_c/2$. Conversely, if $x < -T_c/2$, then $y = x - \alpha(y_c) < -T_c/4$ for c small and thus the soliton Q is at the right of $x = -T_c/2$.

1. *Position of* Q *at* $t = T_c$ (*for* $p = 2, 4$). We determine the position of $Q(y)$, and thus we consider $x > -T_c/2$. For such x, $\sqrt{c}\, y_c \geq \sqrt{c} T_c/4 \gg 1$, and so $\alpha(y_c) = \int_0^{y_c} \beta(s)ds \sim \int_0^{+\infty} \beta(s)ds$. Since $\int_0^\infty Q_c^k(s)ds = \frac{1}{2} c^{\frac{k}{p-1}-\frac{1}{2}} \int Q^k$, we obtain

$$\alpha(y_c) \sim \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^\ell \int_0^\infty Q_c^k(s)ds = \frac{1}{2} \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^{\frac{k}{p-1}+\ell-\frac{1}{2}} \int Q^k.$$

This means that at $t = T_c$, the soliton Q is located at $x = \frac{\Delta}{2}$, where

$$\Delta = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^{\frac{k}{p-1}+\ell-\frac{1}{2}} \int Q^k. \tag{3.4}$$

By symmetry, at $t = -T_c$, the soliton Q is located at $x = -\frac{\Delta}{2}$. Thus, as a consequence of the interaction with the small soliton Q_c, the large soliton Q is shifted by Δ defined by (3.4).

2. *Position of* Q_c *at* $t = T_c$ (*for* $p = 2, 4$). For the soliton Q, we have introduced the variable y depending on x and t which follows the trajectory of Q and in particular the shift phenomenon. On Q_c, the variable $y_c = x+(1-c)t$ does not catch any shift of the trajectory of Q_c. However, in the integrable cases, it is known that the small soliton is also shifted after passing through the interaction. In fact, the shift on Q_c is to be determined by examining the rest of the expansion of v. Since we want to locate the soliton Q_c at $t = T_c$, we consider $x < -T_c/2$. In particular, $y = x - \alpha(y_c) < -T_c/4$, for c small. Recall from Proposition 2.3 that $A_{1,0}, A_{2,0} \in \mathcal{Y}$, and at $t = T_c$, $B_{1,0} \sim \overline{B}_{1,0} - b_{1,0}$, where $\overline{B}_{1,0} \in \mathcal{Y}$. Thus

$$\begin{aligned} Q_c(y_c) + W(T_c) &\sim (1 + A_{1,0}(y))Q_c(y_c) + A_{2,0}(y)Q_c^2(y_c) + B_{1,0}(y)Q_c'(y_c) \\ &\sim Q_c(y_c) - b_{1,0}Q_c'(y_c) \sim Q_c(y_c - b_{1,0}). \end{aligned}$$

Thus,

$$v(T_c, x) \sim Q(x - \tfrac{\Delta}{2}) + Q_c(y_c - b_{1,0}). \tag{3.5}$$

By the symmetry $x \to -x$, $t \to -t$, the value $-2b_{1,0}$ can be interpreted as the first order of the shift Δ_c on the soliton Q_c. Thus, we can set

$$\Delta_c = 2b_{1,0}.$$

3. *The integrable cases* $p = 2, 3$.

• $p = 2$. In this case, we consider the explicit 2-soliton solution with speeds 1 and $0 < c < 1$ defined in (1.6). It is classical to observe that for

t large, for $x \in \mathbb{R}$,

$$U_2(t,x) \sim Q(x-t-\Delta') + Q_c(x-ct), \quad U_2(-t,x) \sim Q(x+t) + Q_c(x+ct+\Delta'_c),$$

where $\Delta' = -\log \alpha(c) > 0$ and $\Delta'_c = -\frac{1}{\sqrt{c}}\Delta'$.

Let us check that the function v can be chosen to match the explicit 2-soliton at the main orders at T_c. We are not able to check all the relations up to any k_0, ℓ_0 by an algebraic argument. However, one can expect that there exists a function v matching precisely at any order the explicit 2-soliton solution.

First, let us check that the shifts are matching $\Delta' \sim 4\sqrt{c} + \frac{4}{3}c\sqrt{c}$, $\Delta'_c = -\frac{1}{\sqrt{c}}\Delta'$. From (3.4),

$$\Delta \sim (a_{1,0}\int Q)\sqrt{c} + \Big(a_{2,0}\int Q^2 + a_{1,1}\int Q\Big)c^{3/2}, \quad \Delta_c \sim 2b_{1,0}.$$

From $\int Q = \int Q^2 = 6$, $a_{1,0} = \frac{2}{3}$, $a_{2,0} + a_{1,1} = \frac{2}{9}$ (Lemma 3.1) and $b_{1,0} = -2$ (Lemma 3.1), Δ' and Δ (and Δ'_c and Δ_c) math at the first order.

Now, we check that $v(T_c)$ matches the 2-soliton solution or equivalently the sum of two solitons at the principal orders. From the decomposition of v at $t = T_c$, we have

$$\begin{aligned} v(T_c) - Q(y) \sim\ & Q_c(y_c) - b_{1,0}Q'_c(y_c) \\ & + \tilde{A}_{2,0}Q_c^2(y_c) - b_{2,0}(Q_c^2)'(y_c) + \tilde{A}_{1,1}cQ_c(y_c) - b_{1,1}c(Q_c)'(y_c). \end{aligned}$$

Since $\widetilde{A}_{2,0} = -\frac{1}{2}b_{1,0}^2 = -2$, $\widetilde{A}_{1,1} = \frac{1}{2}b_{1,0}^2 = 2$, $b_{2,0} = -b_{1,1} = -\frac{1}{6}b_{1,0}^3 = \frac{4}{3}$, from Lemma 3.1, we obtain

$$\begin{aligned} v(T_c) - Q(y) \sim\ & Q_c(y_c) - b_{1,0}Q'_c(y_c) - \tfrac{1}{2}b_{1,0}^2Q_c^2(y_c) \\ & + \tfrac{1}{6}b_{1,0}^3(Q_c^2)'(y_c) + \tfrac{1}{2}b_{1,0}^2cQ_c(y_c) - \tfrac{1}{6}b_{1,0}^3c(Q_c)'(y_c). \end{aligned}$$

But by Taylor expansion, we have

$$\begin{aligned} Q_c(y_c - b_{1,0}) &\sim Q_c(y_c) - b_{1,0}Q'_c(y_c) + \tfrac{1}{2}b_{1,0}^2Q''_c(y_c) - \tfrac{1}{6}b_{1,0}^3Q_c^{(3)}(y_c) \\ &\sim Q_c(y_c) - b_{1,0}Q'_c(y_c) + \tfrac{1}{2}b_{1,0}^2(cQ_c - Q_c^2)(y_c) - \tfrac{1}{6}b_{1,0}^3(cQ_c - Q_c^2)'(y_c), \end{aligned}$$

since $Q''_c(y_c) = cQ_c - Q_c^2$, and $Q^{(3)} = (cQ_c - Q_c^2)'$. Therefore, $v(T_c)$ matches the sum of two translated solitons at this order.

4. *Case* $p = 4$. In this case, we recall that no explicit 2-soliton solution is known, nor was any approximate solution. In the next sections, by analytical methods, we will use the function v to describe any solution close in large time to the sum of two solitons Q, Q_c for c small. Therefore, the function v really describes the interaction between a soliton Q and a soliton Q_c for $p = 4$. In particular, from equation (3.4) and Lemma 3.1,

$$(3.6) \qquad (p=4) \qquad \Delta \sim -2\,\frac{1}{c^{1/6}}\,\frac{\left(\int Q\right)^2}{\int Q^2}.$$

Equation (3.6) is surprising for two reasons. First, the value of the shift is negative. This means that for $p = 4$, the large soliton Q is shifted to the left by interaction with the small soliton Q_c. This is in contrast with the two previously known situations $p = 2$ and $p = 3$, where the shift is positive.

The second surprise is that the shift becomes infinite as $c \to 0$. Therefore, the smaller c is, the larger is the influence of Q_c on the trajectory of Q. To obtain the next order of the shift Δ for $p = 4$, it is sufficient to compute $a_{2,0}$ from Lemma 3.1. However, note that the next order is $c^{1/6}$ ($k = 2$ and $\ell = 0$) and thus it corresponds to a small perturbation of Δ as c is small.

The function v also allows us to determine the shift Δ_c on the small soliton. From Lemma 3.1, it is at the first order $\Delta_c = 2b_{1,0} < 0$. Thus, the small soliton is also shifted to the left through the interaction, for c sufficiently small, as for $p = 2, 3$.

From the decomposition of v and then Taylor expansion, we obtain at $t = T_c$,

$$\begin{aligned} v(T_c) &= Q(x - \tfrac{\Delta}{2}) + Q_c(y_c) - b_{1,0}Q_c'(y_c) - b_{2,0}(Q_c^2)'(y_c) + O(c) \\ &= Q(x - \tfrac{\Delta}{2}) + Q_c(y_c - b_{1,0}) - b_{2,0}(Q_c^2)'(y_c) + O(c). \end{aligned}$$

From Lemma 3.1, we have the fundamental information that $b_{2,0} < 0$ and the term $(Q_c^2)'(y_c)$ above cannot be interpreted as a translation or scaling perturbation term. Thus the approximate solution v at T_c does not match a sum of two translated solitons by a term of order $\|(Q_c^2)'\|_{H^1} \sim Kc^{11/12}$. Note that compared to $(Q_c^2)'$, terms of order $O(c)$ can indeed be neglected. This fact and perturbative analytic arguments around 2-soliton solutions, allow us to prove in Section 5 that there is no pure 2-soliton solution for the nonintegrable case $p = 4$ and to estimate from above and below the size of the nonzero error term created by the interaction.

Now, we give a precise statement concerning v at $\pm T_c$ for $p = 4$ and then for $p = 2$. We prove it only for $p = 4$, the proof for $p = 2$ being similar.

Proposition 3.1. *Let $p = 4$. Let $k_0 \geq 5$ and $\ell_0 \geq 1$. There exists a function v as in Theorem* 2.1 *and Proposition* 2.3 *satisfying, for c sufficiently small,*

1. *Approximate solution on $[-T_c, T_c]$: for all $j \geq 1$, there exists $K = K(j) > 0$ such that*

$$\forall t \in [-T_c, T_c], \quad \|\partial_x^j\big(\partial_t v + \partial_x(\partial_x^2 v - v + v^p)\big)\|_{L^2(\mathbb{R})} \leq Kc^{n_0}, \tag{3.7}$$

where $n_0 = \frac{11}{24}\min(\frac{k_0}{3}, \ell_0 + 1)$.

2. *Closeness to the sum of two solitons for* $t = \pm T_c$:

(3.8)
$$\|v(T_c) - \left\{Q(. - \tfrac{\Delta}{2}) + Q_c(. + (1-c)T_c - \Delta_c/2) - b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} \leq Kc,$$
$$\|v(-T_c) - \left\{Q(. + \tfrac{\Delta}{2}) + Q_c(. - (1-c)T_c + \Delta_c/2) + b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} \leq Kc,$$

(3.9)
$$\|\partial_x(v(T_c) - \left\{Q(. - \tfrac{\Delta}{2}) + Q_c(. + (1-c)T_c - \Delta_c/2))\right\}\|_{L^2} \leq Kc^{\frac{17}{12}},$$

where

(3.10)
$$\Delta = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^{\frac{k}{p-1}+\ell-\frac{1}{2}} \int Q^k, \quad \Delta_c = 2b_{1,0} = -\int Q^3 + \frac{1}{3}\frac{\left(\int Q\right)^2}{\int Q^2} < 0.$$

3. *Decay on the right*:

(3.11)
$$\|v(T_c) - Q(. - \tfrac{\Delta}{2})\|_{H^1(x>-T_c/2)} \leq K\exp(-\tfrac{1}{4}c^{-\frac{1}{100}}),$$

(3.12)
$$\forall x \geq 0, \quad |v(0,x)| \leq C\exp(-\tfrac{1}{2}\sqrt{c}x).$$

Remark. Recall that for $p = 4$, $\|Q_c\|_{L^2} = c^{1/12}\|Q\|_{L^2}$. By (3.10) and Lemma 3.1, we have

(3.13)
$$\left|\Delta - \left(-2\frac{\left(\int Q\right)^2}{\int Q^2}\frac{1}{c^{1/6}} + d_1\, c^{1/6} + d_2\, c^{1/2} + d_3\, c^{5/6}\right)\right| \leq Kc^{7/6},$$

where d_1, d_2 and d_3 are universal constants. The other terms in the sum (3.10) are of higher order than $c^{7/6}$; in particular, these terms are not relevant in our estimate. We will not compute d_1, d_2 and d_3, and will just keep the first order term to state the main results (see Theorem 1.2).

Since $\frac{1}{K}c^{11/12} \leq \|(Q_c^2)'(y_c)\|_{H^1} \leq Kc^{11/12}$ and $b_{2,0} \neq 0$, estimates (3.8) imply that

(3.14)
$$\frac{1}{K}c^{\frac{11}{12}} \leq \|v(T_c) - Q(. - \tfrac{\Delta}{2}) - Q_c(. + (1-c)T_c - \Delta_c/2)\|_{H^1} \leq Kc^{\frac{11}{12}},$$

(3.15)
$$\frac{1}{K}c^{\frac{11}{12}} \leq \|v(-T_c) - Q(. + \tfrac{\Delta}{2}) - Q_c(. - (1-c)T_c + \Delta_c/2)\|_{H^1} \leq Kc^{\frac{11}{12}}.$$

Proof of Proposition 3.1. We consider the function v constructed in Theorem 2.1, for $k_0 \geq 5$ and $\ell_0 \geq 1$. Since $p = 4$, we have $q = 1/12$. Thus estimate (3.7) is a consequence of Theorem 2.1 (2.9). Estimate (2.10) still holds for v on $[-T_c, T_c]$, but our objective is to prove (3.14)–(3.15), which is a much sharper estimate for $t = \pm T_c$. We consider only $t = T_c$ by symmetry. We justify the formal approach above.

1. *Estimates on the remaining terms in* $W(t,x)$ *using* Claim 2.7(a)–(b)–(c). We claim, at $t = T_c$,

(3.16)
$$\|v(T_c) - \left\{Q(y) + Q_c(y_c) - b_{1,0}Q_c'(y_c) - b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} \leq Kc.$$

From (2.48)–(2.49), for $k = 1, 2, 3$, $\ell = 0$, at $t = T_c$, we have

$$\|Q_c(y_c)A_{1,0}(y)\|_{H^1} + \|Q_c^2(y_c)A_{2,0}(y)\|_{H^1} + \|Q_c^3(y_c)A_{3,0}(y)\|_{H^1} \le Ke^{-(1-c)\sqrt{c}T_c} \le Kce^{-c^{-q/2}},$$

$$\|(Q_c^3)'(y_c)B_{3,0}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x((Q_c^3)'(y_c)B_{3,0}(y)\|_{L^2} \le Kc^{5/4}.$$

By similar estimates, since $\overline{B}_{1,0}, \overline{B}_{2,0} \in \mathcal{Y}$, we have at $t = T_c$,

$$\|(Q_c)'(y_c)\overline{B}_{1,0}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x((Q_c)'(y_c)\overline{B}_{1,0}(y))\|_{L^2} + \|(Q_c^2)'(y_c)\overline{B}_{2,0}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x((Q_c^2)'(y_c)\overline{B}_{2,0}(y))\|_{L^2} \le Ke^{-c^{-q/2}}.$$

We also check using Claim 2.7 (2.50)–(2.51), that for $4 \le k \le 6 = 2(p-1)$, $\ell = 0$ and for $1 \le k \le 3 = (p-1)$, $\ell = 1$, at $t = T_c$,

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x(c^\ell Q_c^k(y_c)A_{k,\ell}(y))\|_{L^2} + \|c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x(c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y))\|_{L^2} \le Kc^{25/24}.$$

Finally, by Claim 2.7 (2.52)–(2.53), we check that for (k, ℓ) such that $\xi(k, \ell) \ge 2$,

$$\|c^\ell Q_c^k(y_c)A_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x(c^\ell Q_c^k(y_c)A_{k,\ell}(y))\|_{L^2} + \|c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y)\|_{L^2} + \tfrac{1}{\sqrt{c}}\|\partial_x(c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y))\|_{L^2} \le Kc.$$

Thus (3.16) is proved.

2. *Position of the soliton Q at $t = T_c$.* We claim

(a) For $x \ge -T_c/2$ and $t = T_c$,

$$|\alpha(y_c) - \tfrac{\Delta}{2}| \le Ke^{-\frac{1}{4}c^{-q/2}}. \tag{3.17}$$

(b) For $t = T_c$,

$$\|Q(y) - Q(. - \tfrac{\Delta}{2})\|_{H^1} \le Ke^{-\frac{1}{2}c^{-q/2}}. \tag{3.18}$$

We have $|\alpha(y_c) - \frac{\Delta}{2}| \le K \int_{y_c}^{+\infty} Q_c(s)ds$, and, for any $k \ge 1$, for any $y_c > 0$,

$$0 \le \int_{y_c}^{\infty} Q_c(s)ds \le Kc^{1/3} \int_{y_c}^{\infty} e^{-\sqrt{c}\,s}ds = Kc^{-1/6}e^{-\sqrt{c}\,y_c},$$

we obtain

$$\left|\alpha(y_c) - \tfrac{\Delta}{2}\right| \le Kc^{-1/6}e^{-\sqrt{c}\,y_c}.$$

For $x \ge -T_c/2$ and $t = T_c$, we have $y_c = x + (1-c)T_c \ge (\frac{1}{2} - c)T_c$. Thus $\sqrt{c}\,y_c \ge \frac{1}{2}c^{-q/2} - 1$, and so we obtain (a).

Proof of (b). For $x \geq -T_c/2$, using (a), we have

$$\|Q(y) - Q(. - \tfrac{\Delta}{2})\|_{H^1(x>-T_c/2)} \leq Ke^{-\frac{1}{4}c^{-\frac{1}{100}}}.$$

For $x < -T_c/2$, since $y = x - \alpha(y_c)$ and $|\alpha(y_c)| \leq Kc^{-1/6}$, we have $y < -T_c/4$. Thus,

$$\begin{aligned}&\|Q(y) - Q(. - \tfrac{\Delta}{2})\|_{H^1(x<-T_c/2)}\\ &\leq \|Q(y)\|_{H^1(x<-T_c/2)} + \|Q(. - \tfrac{\Delta}{2})\|_{H^1(x<-T_c/2)} \leq Ke^{-\frac{1}{2}c^{-\frac{1}{100}}}.\end{aligned}$$

3. *Position of the soliton Q_c at $t = T_c$.* We claim that

$$\|Q_c(y_c) - b_{1,0}Q_c'(y_c) - Q_c(. - b_{1,0})\|_{H^1} \leq Kc^{13/12}. \tag{3.19}$$

For example, for the L^2-norm, we have

$$\begin{aligned}\|Q_c - b_{1,0}Q_c' - Q_c(. - b_{1,0})\|_{L^2} &= c^q\|Q - \sqrt{c}\,b_{1,0}Q' - Q(. - \sqrt{c}\,b_{1,0})\|_{L^2}\\ &\leq Kc^q(\sqrt{c}\,b_{1,0})^2 = Kc^{1+q}.\end{aligned}$$

Therefore, we obtain (3.8).

4. *Estimate on the right.* Finally, we prove (3.11). It is sufficient to prove that at $t = T_c$,

$$\|Q_c(y_c) + W(t,x)\|_{H^1(x>-T_c/2)} \leq Ke^{-\frac{1}{2}c^{-\frac{1}{100}}}. \tag{3.20}$$

For $x > -T_c/2$ and $t = T_c$, we have $y_c = x + (1-c)T_c > (1/2 - c)T_c$ and so $\sqrt{c}\,y_c \geq \frac{1}{2}c^{-\frac{1}{100}} - 1$. Thus, it is clear that $\|Q_c(y_c)\|_{H^1(x>-T_c/2)} \leq Ke^{-\frac{1}{2}c^{-\frac{1}{100}}}$. All the other terms in $W(t,x)$ are checked to satisfy the same estimate, using the control on the degrees of the polynomial functions $\widetilde{A}_{k,\ell}$, $\widehat{A}_{k,\ell}$ and $\widetilde{B}_{k,\ell}$, $\widehat{B}_{k,\ell}$ as in the proof of Claim 2.7.

The pointwise estimate (3.12) for $x > 0$ is clear from the decay properties of Q and Q_c. Thus Proposition 3.1 is proved.

Finally, we present without proof a similar result for $p = 2$.

PROPOSITION 3.2. *Let $p = 2$. Let $k_0 \geq 2$ and $\ell_0 \geq 1$. There exist $K > 0$ and a function v as in Theorem* 2.1 *and Proposition* 2.3 *satisfying, for c sufficiently small,*

1. *Approximate solution on $[-T_c, T_c]$: for all $j = 0, 1, 2$ such that*

$$\forall t \in [-T_c, T_c], \quad \|\partial_x^j\big(\partial_t v + \partial_x(\partial_x^2 v - v + v^p)\big)\|_{L^2(\mathbb{R})} \leq Kc^2. \tag{3.21}$$

2. *Closeness to the sum of two solitons for $t = \pm T_c$:*

$$\|v(T_c) - Q(. - 2\sqrt{c}) - Q_c(. + (1-c)T_c + 2)\|_{H^1(\mathbb{R})} \leq Kc^{3/2}, \tag{3.22}$$

$$\|v(-T_c) - Q(. + 2\sqrt{c}) - Q_c(. - (1-c)T_c - 2)\|_{H^1(\mathbb{R})} \leq Kc^{3/2}. \tag{3.23}$$

4. **Preliminary results for stability of the 2-soliton structure**

In this section, we gather several stability results (essentially refinements of tools developed in [21], [27] and [22]). Section 4.1 concerns the stability of $v(t)$ by the gKdV equation during the interaction. Sections 4.2 and 4.3 concern the large time behavior (after interaction).

4.1. *Dynamic stability in the interaction region.* For any c small enough, we consider a function $v(t)$ of the form

$$v(t,x) = Q(y) + Q_c(y_c) + \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(y_c)A_{k,\ell}(y) + (Q_c^k)'(y_c)B_{k,\ell}(y)\right), \tag{4.1}$$

where $y_c = x+(1-c)t$, $\quad y = x-\alpha(y_c)$ and $\alpha(s) = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^\ell \int_0^s Q_c^k(s')ds'$, and $(a_{k,\ell})$, $(A_{k,\ell})$, $(B_{k,\ell})$ satisfy the properties of Proposition 2.3. Set $S(t) = \partial_t v + \partial_x(\partial_x^2 v - v + v^p)$.

Proposition 4.1 (Exact solution close to the approximate solution v). *Let $p = 2$, 3 or 4. Let $\theta > \frac{1}{p-1}$. There exists $c_5 > 0$ such that the following holds for any $0 < c < c_5$. Suppose that*

$$\text{for } j = 1,2,3,\ \forall t \in [-T_c, T_c], \quad \left\|\partial_x^j S(t)\right\|_{L^2(\mathbb{R})} \le K\frac{c^\theta}{T_c} = Kc^{\theta+\frac{1}{2}+\frac{1}{100}} \tag{4.2}$$

and that for some $T_0 \in [-T_c, T_c]$,

$$\|u(T_0) - v(T_0)\|_{H^1(\mathbb{R})} \le Kc^\theta, \tag{4.3}$$

where $u(t)$ is an H^1 solution of the (gKdV) *equation* (1.1). *Then, there exist $K_0 = K_0(\theta, K)$ and a function $\rho : [-T_c, T_c] \to \mathbb{R}$ such that, for all $t \in [-T_c, T_c]$,*

$$\|u(t) - v(t, . - \rho(t))\|_{H^1} \le K_0 c^\theta, \quad |\rho'(t) - 1| \le K_0 c^\theta. \tag{4.4}$$

Remark. By usual techniques related to the resolution of the Cauchy problem, one obtains for approximate solutions a divergence of order e^{T_c} for a time interval $[0, T_c]$. Here, such an estimate would not be sufficient since $T_c = c^{-\frac{1}{2}+\frac{1}{100}} \gg c^{-1/2}$. In this proof, we use the Hamiltonian properties of the gKdV equation. More precisely, the proof is based on the fact that v is close to Q (c is small), and on refined stability analysis around Q (on the one hand standard arguments of long time stability (see Weinstein [37]) and on the other hand some algebraic cancellations in the energy functional). This leads us to a simple ODE estimate in time on the error term.

Note that $\theta > \frac{1}{p-1}$ is arbitrary in Proposition 4.1. Moreover, from the algebraic argument (Theorem 2.1), there exists v such that (4.2) holds for any θ large. This implies that if (for example) $u(0) = v(0)$, then $\|u(T_c) - v(T_c)\|_{H^1} \le K(\theta)c^\theta$, for any θ large. Therefore, the approximate function v and its properties (for example the shift properties) are sharp up to any order c^θ, and provide a sharper description of the collision problem as $\theta \to +\infty$.

Proof of Proposition 4.1. We prove the result on $[T_0, T_c]$. By using the transformation $x \to -x$, $t \to -t$, the proof is the same on $[-T_c, T_0]$. Let $K^* > 1$ be a constant to be fixed later. Since $\|u(T_0) - v(T_0)\|_{H^1} \le c^\theta$, by continuity in time in $H^1(\mathbb{R})$, there exists $T^* > T_0$ such that

$$T^* = \sup\Big\{T \in [T_0, T_c] \text{ s.t. } \forall t \in [T_0, T], \exists r(t) \in \mathbb{R} \\ \text{with } \|u(t) - v(t, . - r(t))\|_{H^1} \le K^* c^\theta\Big\}.$$

Note that the translation direction is degenerate and without the freedom in the translation parameter, the result would not be correct. The objective is to prove that $T^* = T_c$ for K^* large. For this, we argue by contradiction, assuming that $T^* < T_c$ and reaching a contradiction with the definition of T^* by proving independent estimates on $\|u(t) - v(t, . - r)\|_{H^1}$ on $[T_0, T^*]$.

First, we claim some estimates related to v.

Claim 4.1 (Preliminary estimates). *The following hold*:

$$\|\partial_t v(t)\|_{L^\infty} \le K c^{\frac{1}{p-1}}, \tag{4.5}$$

$$\|\partial_t v(t) + \alpha'(y_c) Q'(y)\|_{L^2} \le K c^{q+\frac{1}{2}}, \quad \|\partial_t v(t) + \alpha'(y_c) Q'(y)\|_{L^\infty} \le K c^{m_0}, \tag{4.6}$$

$$\|v^{p-2} - Q^{p-2}(y)\|_{L^\infty} \le K c^{\frac{1}{p-1}}, \tag{4.7}$$

$$\|\partial_x v - Q'(y)\|_{L^2} \le K c^{\frac{1}{p-1}}, \tag{4.8}$$

$$\|\alpha''(y_c)\|_{L^\infty} + \frac{1}{c}\|\alpha^{(4)}(y_c)\|_{L^\infty} \le K c^{\frac{1}{2} + \frac{1}{p-1}}, \tag{4.9}$$

where $m_0 = \min\left(\frac{2}{p-1}, \frac{1}{p-1} + \frac{1}{2}\right)$.

Proof of Claim 4.1. (4.5)—(4.6): We differentiate formula (4.1) with respect to t:

$$\begin{aligned}\partial_t v(t) = &-(1-c)\alpha'(y_c)Q'(y) + (1-c)Q_c'(y_c) \\ &+ \sum_{(k,\ell)\in\Sigma_0} c^\ell \left((1-c)(Q_c^k)'(y_c)A_{k,\ell}(y) - (1-c)\alpha'(y_c)Q_c^k(y_c)A_{k,\ell}'(y)\right) \\ &+ \sum_{(k,\ell)\in\Sigma_0} c^\ell \left((1-c)(Q_c^k)''(y_c)B_{k,\ell}(y) - (1-c)\alpha'(y_c)(Q_c^k)'(y_c)B_{k,\ell}'(y)\right).\end{aligned}$$

By the same estimates as in the proofs of Proposition 2.4 and Claim 2.7, and by $|\alpha'(y_c)| \le K c^{\frac{1}{p-1}}$ (see Claim 2.5), we have $\|\partial_t v(t)\|_{L^\infty} \le K\|Q_c\|_{L^\infty} \le K c^{\frac{1}{p-1}}$, and (4.6).

From the expression of v and estimates as in the proof of Proposition 2.4, we obtain (4.7).

(4.8): Differentiating (4.1) with respect to x:

$$\partial_x v(t) = Q'(y) - \alpha'(y_c)Q'(y) + Q_c'(y) + o(c^{\frac{1}{p-1}}).$$

(4.9): $|\alpha''(s)| \leq K \sum_{1\leq k\leq k_0} |(Q_c^k)'(s)| \leq K\|Q_c'\|_{L^\infty} \leq Kc^{\frac{1}{2}+\frac{1}{p-1}}$.

Step 1. *Choice of the translation parameter and control of the Q' direction.*

LEMMA 4.1 (Modulation). *There exists a C^1 function $\rho : [T_0, T^*] \to \mathbb{R}$ such that, for all $t \in [T_0, T^*]$, the function $z(t)$ defined by $z(t) = u(t, x+\rho(t)) - v(t,x)$ satisfies, for all $t \in [T_0, T^*]$, $\int z(t)Q'(y)dx = 0$, and for K independent of K^*,*

$$\|z(t)\|_{H^1} \leq 2K^* c^\theta, \quad |\rho(T_0)| + \|z(T_0)\|_{H^1} \leq Kc^\theta,$$
$$|\rho'(t) - 1| \leq K\|z(t)\|_{H^1} + K\|S(t)\|_{H^1}. \tag{4.10}$$

Proof of Lemma 4.1. The existence of $\rho(t)$ is obtained at fixed time $t \in [T_0, T^*]$. Let (recall $y = x - \alpha(y_c)$)

$$\zeta(U, r) = \int (U(x+r) - v(t,x))Q'(y)dx.$$

Then $\frac{\partial\zeta}{\partial r}(U,r) = \int U'(x+r)Q'(y)dx$, so that from Claim 4.1, for c small enough,

$$\frac{\partial\zeta}{\partial r}(v,0) = \int(\partial_x v)(t,x)Q'(y)dx > \int(Q'(y))^2dx - Kc^{\frac{1}{p-1}} > \frac{1}{2}\int(Q')^2$$

(note that $\int(Q'(y))^2dx = \int(Q'(y))^2\frac{dy}{1-\alpha_c'(y_c)} > \frac{3}{4}\int(Q'(y))^2dy$). Since $\zeta(v,0) = 0$, for U close to $v(t)$ in L^2 norm, the existence of a unique $\rho(U)$ satisfying $\zeta(U(x-\rho(U)),\rho(U)) = 0$ is a consequence of the Implicit Function Theorem.

From the definition of T^*, it follows that there exists $\rho(t) = \rho(u(t))$, such that $\zeta(u(x-\rho(t)),\rho(t)) = 0$. We set

$$z(t,x) = u(t, x+\rho(t)) - v(t,x); \tag{4.11}$$

then $\int z(t)Q' = 0$ follows from the definition of $\rho(t)$ and (4.10) from the Implicit Function Theorem and the definition of K^*. Moreover, since $\|u(T_0) - v(T_0)\| \leq c^\theta$, we have $|\rho(T_0)| + \|z(T_0)\|_{H^1} \leq Kc^\theta$, where K is independent of K^*.

Let us prove that

$$|\rho'(t) - 1| \leq K\|z(t)\|_{H^1} + K\|S(t)\|_{H^1}. \tag{4.12}$$

From the definition of $z(t)$, $u(t)$ being a solution of the (gKdV) equation, we obtain

$$\begin{aligned}
&\partial_t z + \partial_x(\partial_x^2 z - z + (z+v)^p - v^p)\\
&\quad = -\left\{\partial_t v + \partial_x(\partial_x^2 v - v + v^p)\right\} + (\rho'(t) - 1)\partial_x u\\
&\quad = -S(t) + (\rho'(t)-1)\partial_x(v+z).
\end{aligned} \tag{4.13}$$

Since $\int z(t,x)Q'(y)dx = 0$, by $y = x - \alpha(y_c)$ and $y_c = x + (1-c)t$, we have

$$0 = \frac{d}{dt}\int zQ'(y)dx = \int \partial_t z Q'(y) - (1-c)\int \alpha'(y_c)zQ''(y).$$

Thus, integrating by parts,

$$(4.14)\quad (\rho'(t)-1)\int (v+z)\partial_x(Q'(y)) \\ = \int z(\partial_x^3-\partial_x)(Q'(y)) + \int ((z+v)^p - v^p)\,\partial_x(Q'(y)) \\ - \int S(t)Q'(y) - (1-c)\int \alpha'(y_c) z Q''(y).$$

Therefore $|(\rho'(t)-1)\int (v+z)\partial_x(Q'(y))| \le K\|z(t)\|_{H^1} + K\|S(t)\|_{H^1}$. The term $-\int (v+z)\partial_x Q'(y)$ has a positive lower bound:

$$\int (v+z)\partial_x(Q'(y)) = \int (1-\alpha'(y_c))(v+z)Q''(y) \\ = \int Q(y)Q''(y) + \int (v - Q(y) + z)Q''(y) \\ - \int \alpha'(y_c)(v+z)Q''(y).$$

Since $-\int Q(y)Q''(y)dx > \frac{3}{4}\int (Q'(y))^2 dy > 0$ and since the other terms are small for c small, we have $-\int (v+z)\partial_x Q'(y) > \frac{1}{2}\int (Q')^2$. Note that $\rho(t)$ is C^1 since $Q(y)$ and v are C^∞ and $z(t)$ is continuous in $H^1(\mathbb{R})$. (4.12) is proved.

Step 2. *L^2 norm conservation and control of the direction $\int zQ(y)$.* The use of the L^2 norm conservation replaces a modulation argument in the scaling parameter.

LEMMA 4.2 (Control of the Q direction). *For all $t \in [T_0, T^*]$,*

$$(4.15)\qquad \left|\int z(t)Q(y)\right| \le Kc^\theta + Kc^q\|z(t)\|_{L^2} + \|z(t)\|_{L^2}^2.$$

Proof of Lemma 4.2. Remark that since $v(t)$ is an approximate solution of (1.1), its L^2 norm has a small variation. Indeed, by multiplying the equation $S(t) = \partial_t v + \partial_x(\partial_x^2 v - v + v^p)$ by v and integrating, we obtain $\left|\frac{1}{2}\frac{d}{dt}\int v^2\right| = |\int S(t,x)v(t,x)dx| \le K\|S(t)\|_{L^2}$. Thus,

$$(4.16)\quad \forall t \in [T_0, T^*],\quad \left|\int v^2(t) - \int v^2(T_0)\right| \le KT_c \sup_{t\in[-T_c,T_c]} \|S(t)\|_{H^1} \le Kc^\theta.$$

Since $u(t)$ is a solution of the (gKdV) equation, we have

$$(4.17)\qquad \int u^2(t) = \int (v(t)+z(t))^2 = \int u^2(T_0) = \int (v(T_0)+z(T_0))^2.$$

By expanding (4.17) and using (4.16) and (4.10), we obtain

$$2\left|\int v(t)z(t)\right| \le Kc^\theta + 2\left|\int v(T_0)z(T_0)\right| + \|z(T_0)\|_{L^2}^2 + \|z(t)\|_{L^2}^2 \le Kc^\theta + \|z(t)\|_{L^2}^2.$$

Using this and $\|v(t)-Q(y)\|_{L^2} \leq Kc^q$, we obtain

$$\left|\int z(t)Q(y)\right| \leq \left|\int z(t)v\right| + \left|\int z(t)(v-Q(y))\right| \leq Kc^\theta + Kc^q\|z(t)\|_{L^2} + \|z(t)\|_{L^2}^2.$$

Step 3. *Introduction of a energy functional for* $z(t)$. We set

$$\mathcal{F}(t) = \frac{1}{2}\int \left((\partial_x z)^2 + (1+\alpha'(y_c))z^2)\right) - \frac{1}{p+1}\int \left((v+z)^{p+1} - v^{p+1} - (p+1)v^p z\right).$$

The above definition is similar to a linearized energy

$$\frac{1}{2}\left(\int((\partial_x z)^2 + z^2) - p\int Q^{p-1}z^2\right).$$

However, the terms $\int \alpha'(y_c)z^2$ and the nonlinear terms were added to remove some diverging terms in $\mathcal{F}'$. This is the new ingredient of the proof of Proposition 4.1.

We first claim that the functional $\mathcal{F}(t)$ indeed controls the size of $z(t)$ in H^1 up to the direction $Q(y)$, extending the similar classical result for the linearized energy.

Claim 4.2 (Coercivity of $\mathcal{F}$). *There exists* $\kappa_0 > 0$ *such that*

$$\|z(t)\|_{H^1}^2 \leq \kappa_0\mathcal{F}(t) + \kappa_0\left|\int z(t)Q(y)\right|^2. \tag{4.18}$$

The proof of Claim 4.2 is given in Appendix D.1

Next, we claim the following control of the variation of $\mathcal{F}(t)$ through time.

Lemma 4.3 (Control of the variation of the energy functional).

$$\mathcal{F}(T^*) - \mathcal{F}(T_0) \leq Kc^{2\theta}\left((K^*)^2(1+K^*)c^{q/2} + K^*\right), \tag{4.19}$$

where K *is independent of* c *and* K^*.

Proof of Lemma 4.3. We have

$$\mathcal{F}'(t) = \int \partial_t z\left(-\partial_x^2 z + z - ((v+z)^p - v^p)\right) + \int \partial_t z\alpha'(y_c)z + \int\left\{\frac{1}{2}(1-c)\alpha''(y_c)z^2 - \partial_t v\left((v+z)^p - v^p - pv^{p-1}z\right)\right\} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3.$$

Now, we claim

$$\left|\mathbf{F}_1 + (\rho'(t)-1)\int \alpha'(y_c)Q'(y)z\right| \leq Kc^{q+\frac{1}{2}}\|z(t)\|_{H^1}^2 + K\|z(t)\|_{L^2}\left(\|\partial_x^2 S(t)\|_{L^2} + \|S(t)\|_{L^2}\right), \tag{4.20}$$

(4.21)
$$\left|\mathbf{F}_2 - (\rho'(t)-1)\int \alpha'(y_c)Q'(y)z + \frac{p(p-1)}{2}\int \alpha'(y_c)z^2Q'(y)Q^{p-2}(y)\right|$$
$$\leq K\|z(t)\|_{H^1}^2\left(c^{m_0} + c^{\frac{1}{p-1}}\|z(t)\|_{H^1}\right) + K\|z(t)\|_{H^1}\left(\|\partial_x^2 S(t)\|_{L^2} + \|S(t)\|_{L^2}\right),$$

(4.22)
$$\left|\mathbf{F}_3 - \frac{p(p-1)}{2}\int \alpha'(y_c)Q'(y)Q^{p-2}(y)\,z^2\right| \leq Kc^{m_0}\|z\|_{H^1}^2 + Kc^{\frac{1}{p-1}}\|z(t)\|_{H^1}^3.$$

Assuming (4.20)–(4.22), we conclude the proof of the lemma.

Note that $q+\frac{1}{2} = \frac{1}{p-1}+\frac{1}{4} \leq m_0$. From the cancellations of the main terms of $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, and then from (4.10), (4.2), we get

$$\begin{aligned}|\mathcal{F}'(t)| &\leq K\|z(t)\|_{H^1}^2\left(c^{q+\frac{1}{2}} + c^{\frac{1}{p-1}}\|z(t)\|_{H^1}\right)\\ &\quad + K\|z(t)\|_{H^1}\left(\|\partial_x^2 S(t)\|_{L^2} + \|S(t)\|_{L^2}\right)\\ &\leq K\left[(K^*)^2c^{2\theta}(c^{q+\frac{1}{2}} + K^*c^{\frac{1}{p-1}+\theta}) + K^*c^{\frac{1}{2}(1+\frac{1}{50})+2\theta}\right].\end{aligned}$$

Now, $q \geq \frac{1}{50}$ and $\theta + \frac{1}{p-1} \geq \frac{2}{p-1} \geq q+\frac{1}{2} \geq \frac{1}{2}(1+\frac{1}{50}) + \frac{q}{2}$ imply

$$|\mathcal{F}'(t)| \leq Kc^{\frac{1}{2}(1+\frac{1}{50})+2\theta}\left((K^*)^2c^{q/2}(1+K^*) + K^*\right).$$

Integrating on the time interval $[T_0, T^*]$ where $T^* - T_0 \leq 2T_c = 2c^{\frac{1}{2}(1+\frac{1}{50})}$, we obtain

$$|\mathcal{F}(T^*) - \mathcal{F}(T_0)| \leq Kc^{2\theta}\left((K^*)^2(1+K^*)c^{q/2} + K^*\right).$$

Proof of (4.20). We replace $\partial_t z$ by its expression

$$\begin{aligned}\mathbf{F}_1 &= -\int S(t)\left(-\partial_x^2 z + z - ((v+z)^p - v^p)\right)\\ &\quad + (\rho'(t)-1)\int \partial_x(v+z)\left(-\partial_x^2 z + z - ((v+z)^p - v^p)\right) = \mathbf{g}_1 + \mathbf{g}_2.\end{aligned}$$

By integration by parts and the Cauchy-Schwarz' inequality, we have

$$|\mathbf{g}_1| \leq K\|z(t)\|_{L^2}\left(\|\partial_x^2 S(t)\|_{L^2} + \|S(t)\|_{L^2}\right).$$

Since $\int \partial_x(v+z)(v+z)^p = 0$, and by the definition of $S(t)$,

$$\begin{aligned}\mathbf{g}_2 &= (\rho'(t)-1)\int \partial_x(v+z)(-\partial_x^2 z + z + v^p)\\ &= (\rho'(t)-1)\int(\partial_x v(-\partial_x^2 z + z) + \partial_x z v^p)\\ &= (\rho'(t)-1)\int z\partial_x(-\partial_x^2 v + v - v^p) = (\rho'(t)-1)\int z(\partial_t v - S(t)).\end{aligned}$$

By (4.6) and (4.10), we obtain

$$\begin{aligned}\Big|\mathbf{g}_2 + (\rho'(t)-1)\int \alpha'(y_c)Q'(y)z\Big| &\le K|\rho'(t)-1|\|z(t)\|_{L^2}\left(\|\partial_t v-\alpha'(y_c)Q'(y)\|_{L^2}+\|S(t)\|_{L^2}\right)\\ &\le K\|z(t)\|_{L^2}(\|z(t)\|_{H^1}+\|S(t)\|_{L^2})(c^{q+\frac12}+\|S(t)\|_{L^2}).\end{aligned}$$

Proof of (4.21). Note that the term $\mathbf{F}_2$ was introduced on purpose in the expression of $\mathcal{F}$ to cancel the main terms in $\mathbf{F}_1$ and $\mathbf{F}_3$:

$$\begin{aligned}\mathbf{F}_2 = &\int \alpha'(y_c)z\partial_x(-\partial_x^2 z + z - ((z+v)^p - v^p))\\ &-\int \alpha'(y_c)zS(t) + (\rho'(t)-1)\int \alpha'(y_c)\partial_x(v+z)z = \mathbf{g}_3+\mathbf{g}_4.\end{aligned}$$

First,

$$\mathbf{g}_4 = -\int \alpha'(y_c)zS(t) + (\rho'(t)-1)\int \alpha'(y_c)\partial_x v\, z - \frac12(\rho'(t)-1)\int z^2\alpha''(y_c).$$

By (4.8)–(4.9) and (4.10), we have

$$\Big|\mathbf{g}_4 - (\rho'(t)-1)\int \alpha'(y_c)Q'(y)z\Big| \le Kc^{m_0}\|z(t)\|_{H^1}(\|z(t)\|_{H^1}+\|S(t)\|_{H^1}).$$

Second, for the term $\mathbf{g}_3$, we integrate by parts, to obtain

$$\mathbf{g}_3 = -\int \alpha''(y_c)(\tfrac32(\partial_x z)^2+\tfrac12 z^2) + \int \alpha^{(4)}(\tfrac12 z^2) - \int \alpha'(y_c)z\partial_x((z+v)^p-v^p). \tag{4.23}$$

Using the estimate on $\alpha''(y_c)$ and $\alpha^{(4)}(y_c)$ in Claim 4.1, we obtain

$$\Big|-\int \alpha''(y_c)(\tfrac32(\partial_x z)^2+\tfrac12 z^2) + \int \alpha^{(4)}(\tfrac12 z^2)\Big| \le Kc^{\frac{1}{p-1}+\frac12}\|z(t)\|_{H^1}^2.$$

In the last term of (4.23), cubic and higher order terms are controlled by $Kc^{\frac{1}{p-1}}\|z(t)\|_{H^1}^3$. The quadratic term is

$$\int \alpha'(y_c)z\partial_x(-pv^{p-1}z) = \frac{p}{2}\int \alpha''(y_c)z^2v^{p-1} - \frac{p}{2}\int \alpha'(y_c)z^2\partial_x(v^{p-1}) = \mathbf{g}_5+\mathbf{g}_6.$$

As before, $|\mathbf{g}_5| \le Kc^{\frac12+\frac{1}{p-1}}\|z(t)\|_{H^1}^2$. Finally, by (4.7)–(4.8),

$$\Big|\mathbf{g}_6 + \frac{p(p-1)}{2}\int \alpha'(y_c)z^2Q'(y)Q^{p-2}(y)\Big| \le Kc^{\frac{2}{p-1}}\|z(t)\|_{H^1}^2.$$

Proof of (4.22). First note that $\left|\frac12(1-c)\int \alpha''(y_c)z^2\right| \le Kc^{\frac12+\frac{1}{p-1}}\|z(t)\|_{L^2}^2$. We now estimate

$$\begin{aligned}-\int \partial_t v\left((v+z)^p - v^p - pv^{p-1}z - \tfrac{p(p-1)}{2}v^{p-2}z^2\right)&\\ -\frac{p(p-1)}{2}\int \partial_t v\, v^{p-2}z^2 &= \mathbf{g}_7+\mathbf{g}_8.\end{aligned}$$

By (4.5), we have $|\mathbf{g}_7| \leq Kc^{\frac{1}{p-1}} \|z(t)\|_{H^1}^3$. By (4.6), (4.7), and $|\alpha'(y_c)| \leq Kc^{\frac{1}{p-1}}$, we have

$$\left| \mathbf{g}_8 - \frac{p(p-1)}{2} \int \alpha'(y_c) Q'(y) Q^{p-2} z^2 \right| \leq Kc^{m_0} \|z\|_{H^1}^2.$$

Step 4. *Conclusion of the proof.* By Claim 4.2 and Lemmas 4.1–4.2, we have

$$\left| \int z(T^*) Q(y) \right| \leq Kc^\theta + Kc^q \|z(T^*)\|_{L^2} + \|z(T^*)\|_{L^2}^2.$$

Thus by Claim 4.2, $\|z(T^*)\|_{H^1}^2 \leq K\mathcal{F}(T^*) + K(c^\theta + c^q \|z(T^*)\|_{L^2} + \|z(T^*)\|_{L^2}^2)^2$. It follows that for c small enough, $\|z(T^*)\|_{H^1}^2 \leq (K+1)\mathcal{F}(T^*) + Kc^{2\theta}$.

Next, by Lemma 4.3 and $|\mathcal{F}(T_0)| \leq Kc^{2\theta}$, we obtain

$$\begin{aligned} \|z(T^*)\|_{H^1}^2 &\leq (K+1)(\mathcal{F}(T^*) - \mathcal{F}(T_0)) + Kc^{2\theta} \\ &\leq K_1 c^{2\theta} \left((K^*)^2 (1+K^*) c^{q/2} + K^* + 1 \right), \end{aligned}$$

where K_1 is independent of c and K^*. Choose $c^* = c^*(K^*)$ such that

$$(K^*)^2 (1+K^*) (c^*)^{q/2} < 1.$$

Then, for $0 < c < c^*$,

$$\|z(T^*)\|_{H^1}^2 \leq K_1 c^{2\theta} (2 + K^*).$$

Next, fix K^* such that $K_1(2+K^*) < \frac{1}{2}(K^*)^2$. Then $\|z(T^*)\|_{H^1}^2 \leq \frac{1}{2}(K^*)^2 c^{2\theta}$. This contradict the definition of T^*, thus proving that $T^* = T_c$. Thus estimate (4.4) is proved on $[T_0, T_c]$.

4.2. *Stability and asymptotic stability for large time.* In this section, we consider the stability of the 2-soliton structure after the collision. These questions have been considered in [25]. See also [21], [27] and [19]. Denote for $v \in H^1(\mathbb{R})$, $\|v\|_{H^1_c} = \left(\int_{\mathbb{R}} ((v'(x))^2 + cv^2(x))\, dx \right)^{\frac{1}{2}}$, which corresponds to the natural norm to study the stability of Q_c.

Proposition 4.2 (Stability of two decoupled solitons, [25]). *There exists $K > 0$, $\alpha_0 > 0$, $c_0 > 0$ such that, for any $0 < c < c_0$, $0 < \alpha < \alpha_0$, the following holds. Let $u(t)$ be an H^1 solution of* (1.1) *such that, for some $t_1 \in \mathbb{R}$ and $X_0 \geq \frac{1}{2} T_c$,*

$$\|u(t_1) - Q - Q_c(.+X_0)\|_{H^1} \leq \alpha c^{q+\frac{1}{2}}. \tag{4.24}$$

Then there exist C^1 functions $\rho_1(t)$, $\rho_2(t)$ defined on $[t_1, +\infty)$ such that

1. *Stability*:

$$\sup_{t \geq t_1} \|u(t) - (Q(.-\rho_1(t)) + Q_c(.-\rho_2(t)))\|_{H^1_c} \leq K\alpha c^{q+\frac{1}{2}} + K\exp(-c^{-\frac{1}{400}}), \tag{4.25}$$

$$\forall t \geq t_1,\ \tfrac{1}{2} \leq \rho_1'(t) - \rho_2'(t) \leq \tfrac{3}{2}, \tag{4.26}$$
$$|\rho_1(t_1)| \leq K\alpha c^{q+\frac{1}{2}}, \quad |\rho_2(t_1) + X_0| \leq K\alpha.$$

2. *Convergence of $u(t)$: There exist $c_1^+, c_2^+ > 0$ such that*

$$\lim_{t\to+\infty} \|u(t) - Q_{c_1^+}(x-\rho_1(t)) - Q_{c_2^+}(x-\rho_2(t))\|_{H^1(x>ct/10)} = 0, \tag{4.27}$$

(4.28)
$$|c_1^+ - 1| \leq K\alpha c^{q+\frac{1}{2}} + K\exp(-c^{-\frac{1}{400}}), \quad \left|\frac{c_2^+}{c} - 1\right| \leq K\alpha + K\exp(-c^{-\frac{1}{400}}).$$

3. *Assume further that $\int_{x>0} x^2\, u^2(t_1,x)dx < K_0$. Then, there exist δ_1^+ and δ_2^+ such that*

$$\lim_{t\to+\infty} \rho_1(t) - c_1^+ t = \delta_1^+, \quad \lim_{t\to+\infty} \rho_2(t) - c_2^+ t = \delta_2^+. \tag{4.29}$$

For $p = 4$, if for $\kappa > 0$,

$$\alpha < \kappa c^{\frac{1}{3}} \quad \text{and} \quad \int_{x>\frac{11}{12}|\ln c|} x^2\, u^2(t_1,x)dx < \kappa c^{\frac{5}{4}}, \tag{4.30}$$

then

$$|\delta_1^+ - \rho_1(t_1)| \leq Kc^{\frac{5}{8}}, \quad |\delta_2^+ - \rho_2(t_1)| \leq Kc^{\frac{1}{12}}. \tag{4.31}$$

The proof of Proposition 4.2 is based on energy arguments, monotonicity results on local quantities, and a Virial argument; see [23] and [25].

Remark. To obtain the convergence of the translation parameters, one has to add an extra assumption on the initial data such as (4.30). Indeed, in the energy space, one can construct an explicit example where convergence does not hold (see [22]).

4.3. *Decomposition and monotonicity result.* We recall a more precise stability result related to the usual decomposition of the solution $u(t)$. See the proof of Proposition 4 in [25]. Define

$$\psi(x) = \tfrac{2}{\pi}\arctan(\exp(-\tfrac{x}{4})), \text{ so that } \lim_{+\infty}\psi = 0,\ \lim_{-\infty}\psi = 1. \tag{4.32}$$

Claim 4.3 ([25]). *Under the assumptions of Proposition 4.2, there exist C^1 functions $\rho_1(t)$, $\rho_2(t)$, $c_1(t)$ and $c_2(t)$, defined on $[t_1,+\infty)$, such that $\eta(t,x)$ and $g(t)$ defined by*

$$\eta(t,x) = u(t,x) - R_1(t,x) - R_2(t,x), \tag{4.33}$$
$$\textit{where for } j=1,2,\ R_j(t,x) = Q_{c_j(t)}(x-\rho_j(t)),$$

$$g(t) = \int \left(\eta_x^2(t,x) + (c+\psi(x-m(t)))\,\eta^2(t,x)\right)dx, \tag{4.34}$$

satisfy for all $t \in [t_1, +\infty)$, $\int R_j(t)\eta(t) = \int (x-\rho_j(t))R_j(t)\eta(t) = 0$, $j=1,2$,

(4.35)
$$\|\eta(t)\|_{H_c^1}^2 \le g(t) \le Kg(t_1) + K\exp(-c^{-\frac{1}{400}}) \le K\alpha^2 c^{2q+1} + K\exp(-c^{-\frac{1}{400}}),$$

$$\left|\frac{c_1(t)}{c_1(t_1)} - 1\right| + c^{2q+1}\left|\frac{c_2(t)}{c_2(t_1)} - 1\right| \le Kg(t_1) + K\exp(-c^{-\frac{1}{400}}), \tag{4.36}$$

$$|c_1(t)-1| + c^{q+\frac{1}{2}}\left|\frac{c_2(t)}{c} - 1\right| \le K\alpha c^{q+\frac{1}{2}} + K\exp(-c^{-\frac{1}{400}}). \tag{4.37}$$

Now, we recall monotonicity results for quantities defined in $\eta(t)$, to be used in the proof of Theorem 1.2. For $0 \le t_0 \le t$, $x_0 \ge 0$, $j=1,2$, let

$$\begin{aligned}
\mathcal{M}_j(t) &= \int \eta^2 \psi_j,\\
\mathcal{E}_j(t) &= \int \Big[\frac{1}{2}\eta_x^2 - \frac{1}{p+1}\left((R_1{+}R_2{+}\eta)^{p+1} {-} (p{+}1)R_1^p\eta\right.\\
&\qquad\qquad \left. -(p{+}1)R_2^p\eta {-} (R_1{+}R_2)^{p+1}\right)\Big]\psi_j,
\end{aligned}$$

where $\psi_1(t,x) = \psi(\tilde{x})$, $\tilde{x} = x - \rho_1(t) + x_0 + \frac{1}{2}(t-t_0)$, and $\psi_2(t,x) = \psi(\sqrt{c}\tilde{x}_c)$, $\tilde{x}_c = x - \rho_2(t) + x_0 + \frac{c}{2}(t-t_0)$.

Claim 4.4 ([25]). *Let* $x_0 > 0$, $t_0 > 0$. *For all* $t \ge t_0$,

$$\begin{aligned}
&\frac{d}{dt}\left(c_1^{2q}(t)\int Q^2 + \mathcal{M}_1(t)\right) \le Ke^{-\frac{1}{16}(t-t_0+x_0)}g_1(t) + Ke^{-\frac{1}{32}\sqrt{c}(t+T_c)},\\
&\frac{d}{dt}\left(-\frac{2q}{2q+1}c_1^{2q+1}(t)\int Q^2 + 2\mathcal{E}_1(t) + \frac{1}{100}\left(c_1^{2q}(t)\int Q^2 + \mathcal{M}_1(t)\right)\right)\\
&\qquad \le Ke^{-\frac{1}{16}(t-t_0+x_0)}g_1(t) + Ke^{-\frac{1}{32}\sqrt{c}(t+T_c)},\\
&\frac{d}{dt}\left(\left(c_1^{2q}(t) + c_2^{2q}(t)\right)\int Q^2 + \mathcal{M}_2(t)\right)\\
&\qquad \le Ke^{-\frac{c\sqrt{c}}{16}(t-t_0)}e^{-\frac{\sqrt{c}}{16}x_0}\sqrt{c}\,g_2(t) + Ke^{-\frac{1}{32}\sqrt{c}(t+T_c)},\\
&\frac{d}{dt}\left(-\frac{2q}{2q{+}1}\left(c_1^{2q+1}(t){+}c_2^{2q+1}(t)\right)\int Q^2 + 2\mathcal{E}_2(t)\right.\\
&\qquad\left. + \frac{c}{100}\left(\left(c_1^{2q}(t){+}c_2^{2q}(t)\right)\int Q^2 + \mathcal{M}_2(t)\right)\right)\\
&\qquad \le Ke^{-\frac{c\sqrt{c}}{16}(t-t_0)}e^{-\frac{\sqrt{c}}{16}x_0}c^{\frac{3}{2}}g_2(t) + Ke^{-\frac{1}{32}\sqrt{c}(t+T_c)}.
\end{aligned}$$

5. Proofs of the main results ($p = 4$)

First, let us remark that for Theorems 1.1, 1.2 and 1.3 (concerning the case $p=4$), by considering $\tilde{u}(t,x) = \lambda^{\frac{1}{3}}u(\lambda^{\frac{3}{2}}t, \lambda^{\frac{1}{2}}x)$ with $\lambda = \frac{1}{c_1}$ instead of

$u(t)$, we can restrict ourselves to the case $c_1 = 1$ and $c_2 = c \leq \epsilon_0 \ll 1$ without loss of generality. We consider $0 < c < c_0$, where c_0 is small enough so that Sections 2, 3 and 4 apply for any $0 < c' < 2c_0$.

5.1. *Nonexistence of pure* 2*-soliton solutions. Proof of Theorem* 1.1. This section is devoted to the proof of Theorem 1.1. First, we recall the following:

PROPOSITION 5.1. *Let* $p = 2, 3$ *or* 4. *Let* $0 < c < c_0$, *for* c_0 *small enough.*

1. *Existence and exponential decay*: *Let* $\delta_1, \delta_2 \in \mathbb{R}$. *There exists a unique solution* $U_{c,\delta_1,\delta_2} = U \in C(\mathbb{R}, H^1(\mathbb{R}))$ *of* (1.1) *such that*

$$\lim_{t\to-\infty} \|U(t) - Q(. - t - \delta_1) - Q_c(. - ct - \delta_2)\|_{H^1(\mathbb{R})} = 0. \tag{5.1}$$

Moreover, $U(t)$ *satisfies, for all* $t \leq \frac{\delta_2-\delta_1}{1-c} - \frac{T_c}{32}$,

$$\|U(t) - Q(. - t - \delta_1) - Q_c(. - ct - \delta_2)\|_{H^1(\mathbb{R})} \leq K e^{\frac{1}{4}\sqrt{c}((1-c)t-(\delta_2-\delta_1))}. \tag{5.2}$$

2. *Uniqueness of the asymptotic* 2*-soliton solution at* $-\infty$: *If* $u(t)$ *is an* H^1 *solution of* (1.1) *satisfying*

$$\lim_{t\to-\infty} \|u(t) - Q(. - \rho_1(t)) - Q_c(. - \rho_2(t))\|_{H^1(\mathbb{R})} = 0, \tag{5.3}$$

for $\rho_1,\ \rho_2 : \mathbb{R} \to \mathbb{R}$, *then* $u(t)$ *satisfies* (5.1) *for some* $\delta_1,\ \delta_2$, *and so* $u(t) \equiv U_{c,\delta_1,\delta_2}(t)$.

This result was essentially proved in [19], using tools from [21] and [27]. However, some statements in Proposition 5.1 are slightly more precise than the main result in [19], so we justify them in Appendix D.2.

The main ingredient of the proof of Theorem 1.1 is the following proposition related to the approximate solution constructed in Section 2. We keep the notation of Section 3, in particular, $v(t,x)$, $b_{2,0}$ and V_1.

PROPOSITION 5.2. *Let* $p = 4$. *Let* Δ *and* Δ_c *be defined by* (3.10). *Let*

$$v_\#(t,x) = v(t,x) + w_\#(t,x), \quad \textit{where} \quad w_\#(t,x) = -(Q_c^2)'(y_c) b_{2,0}(1 + V_1(y)), \tag{5.4}$$

and

$$S_\#(t,x) = \partial_t v_\# + \partial_x(\partial_x^2 v_\# - v_\# + v_\#^4), \tag{5.5}$$

where v *is the function constructed in Proposition* 3.1. *Then, for all* $0 < c < c_0$, *for* c_0 *sufficiently small,*

1. *Approximate solution*: *for* $j = 0, 1, 2$,

$$\forall t \in [-T_c, T_c], \quad \|\partial_x^j S_\#(t)\|_{L^2} \leq K c^{3/2}. \tag{5.6}$$

2. *Closeness to a pure two soliton at* $t = -T_c$:

$$\|v_\#(-T_c) - Q(.+\tfrac{\Delta}{2}) - Q_c(.-(1-c)T_c+\tfrac{\Delta_c}{2})\|_{H^1} \leq Kc. \tag{5.7}$$

3. *Nonmatching with a pure two-soliton solution at $t = T_c$:*

$$\|v_\#(T_c) - \big\{Q(.-\tfrac{\Delta}{2}) + Q_c(.+(1-c)T_c-\tfrac{\Delta_c}{2}) \\ -2b_{2,0}(Q_c^2)'(.+(1-c)T_c-\tfrac{\Delta_c}{2})\big\}\|_{H^1} \leq Kc. \tag{5.8}$$

Remark. Recall that $\|(Q_c^2)'(y_c)\|_{H^1} > Kc^{11/12}$ and $b_{2,0} < 0$. Thus at T_c, the function $v_\#$ differs from a two-soliton solution of a factor $c^{11/12}$. At $-T_c$ it is close to a two-soliton solution up to a factor c and it is an approximate solution of the gKdV equation in the sense (5.6). This will be sufficient to prove Theorem 1.1 applying Proposition 4.1.

The function $v_\#(t,x)$ is not exactly of the form imposed by Proposition 2.3. Indeed, the function $1 + V_1$ is even, and thus the function $w_\#(t,x)$ does not have the required structure. This will have no consequence in applying Proposition 4.1, which does not rely on the parity structure. In contrast, the presence of $w_\#$ in $v_\#$ is definitely a problem in following the procedure of Proposition 2.3. Indeed, this term creates a new term $F_{5,0}$ which has a nonzero even part, not orthogonal to Q, which is a problem in determining a suitable $A_{5,0}$. Thus, we cannot improve (5.6) up to any power. However, (5.6) is sufficient for our purposes, and the function $v_\#$ is closer to a 2-soliton solution at $t = -T_c$ than the function v itself.

It would be interesting to investigate further improvements of the function $v_\#$ since it would help understanding the behavior for $t > 0$ of solutions which are pure two-soliton solutions at $t \to -\infty$.

Proof of Proposition 5.2. We have

$$\begin{aligned} S_\#(t,x) &= \partial_t v_\# + \partial_x(\partial_x^2 v_\# - v_\# + v_\#^4) \\ &= S(t,x) + \partial_x((v + w_\#)^4) - v^4 - 4Q^3 w_\#) + \partial_t w_\# - \partial_x(\overline{\mathcal{L}} w_\#), \end{aligned}$$

where $\overline{\mathcal{L}}$ is defined in (2.15).

1a. *Estimate of the linear part in $S_\#$.* We estimate $\partial_t w_\# - \partial_x(\overline{\mathcal{L}} w_\#)$, where $w_\#(t,x) = (Q_c^2)'(y_c) b_{2,0}(1 + V_1(y))$. Recall that from Claim 3.1,

$$\mathcal{L}(1 + V_1) = 1 - 4Q^3 + \mathcal{L}V_1 = 1,$$

and thus $(\mathcal{L}(1 + V_1))' = 0$. Claim A.4 gives an explicit expression for $\partial_t w_\# - \partial_x(\overline{\mathcal{L}} w_\#)$, where the first term in the second-hand member is zero. For the other terms, we use

$$(1 + V_1)' = V_1' \in \mathcal{Y},\ Q^3(1 + V_1) \in \mathcal{Y}, \quad \|(Q_c^2)'\|_{L^\infty} = Kc^{7/6},\ \|\beta\|_{L^\infty} \leq Kc^{1/3}, \\ \|(Q_c^2)''\|_{L^\infty} = Kc^{5/3}, \quad \|(Q_c^2)^{(4)}\|_{L^2} = Kc^{29/12},$$

so that $\|\partial_t w_\# - \partial_x(\overline{\mathcal{L}} w_\#)\|_{L^2} \leq Kc^{3/2}$.

We obtain, for all $j = 0, 1, 2$, $\|\partial_x^j \left(\partial_t w_\# - \partial_x(\overline{\mathcal{L}} w_\#)\right)\|_{L^2} \leq K_j c^{3/2}$.

1b. *Estimate of the nonlinear part in* $S_\#$. Note that

$$(v+w_\#)^4 - v^4 - 4Q^3 w_\# = 4(v^3-Q^3)w_\# + 6v^2 w_\#^2 + 4v w_\#^3 + w_\#^4,$$

so that

$$\begin{aligned}\partial_x\left[(v+w_\#)^4 - v^4 - 4Q^3 w_\#\right] &= 4\partial_x(v^3-Q^3)w_\# + 4(v^3-Q^3)\partial_x w_\# \\ &\quad + 6\partial_x(v^2 w_\#^2) + 4\partial_x(v w_\#^3) + \partial_x(w_\#^4).\end{aligned}$$

Moreover,

$$\begin{aligned}\partial_x(v^3-Q^3) &= \partial_x(v-Q)(v^2+vQ+Q^2) + (v-Q)\partial_x(v^2+vQ+Q^2),\\ \partial_x(v^2+vQ+Q^2) &= \partial_x(3Q^2 + (v^2-Q^2) + (v-Q)Q).\end{aligned}$$

Thus,

$$\|\partial_x(v^3-Q^3)\|_{L^2} \le K(\|\partial_x(v-Q)\|_{L^2} + \|v-Q\|_{L^\infty}) \le Kc^{1/3}.$$

We also have

$$\begin{aligned}\|w_\#\|_{L^\infty} &\le Kc^{7/6}, & \|v^3-Q^3\|_{L^2} &\le Kc^{1/2},\\ \|\partial_x w_\#\|_{L^\infty} &\le Kc^{5/3}, & \|w_\#^2\|_{L^\infty} &\le Kc^{5/3}.\end{aligned}$$

Thus,

$$\|\partial_x\left[(v+w_\#)^4 - v^4 - 4Q^3 w_\#\right]\|_{L^2} \le Kc^{3/2}.$$

Similarly, for $j=1,2$, $\|\partial_x^{(j+1)}\left[(v+w_\#)^4 - v^4 - 4Q^3 w_\#\right]\|_{L^2} \le K_j c^{3/2}$.

Taking k_0, ℓ_0 large enough, so that $\|\partial_x^{(j)} S\|_{L^2} \le Kc^{3/2}$, by Proposition 3.1, we have proved $\|\partial_x^{(j)} S_\#\|_{L^2} \le Kc^{3/2}$.

2. *Analysis at* $t=\pm T_c$. By the proof of Proposition 3.1 (see (3.8)), we have

$$\begin{aligned}\|v(-T_c) - \left\{Q(.+\tfrac{\Delta}{2}) + Q_c(.-(1-c)T_c + \Delta_c/2) + b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} &\le Kc,\\ \|v(T_c) - \left\{Q(.-\tfrac{\Delta}{2}) + Q_c(.+(1-c)T_c - \Delta_c/2) - b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} &\le Kc.\end{aligned}$$

Note that by the definition of $v_\#$ and Claim 2.6, we have

$$\|v(\pm T_c) - \left(v_\#(\pm T_c) - b_{2,0}(Q_c^2)'(y_c)\right)\|_{H^1} = \|b_{2,0}(Q_c')^2(y_c)V_1(y)\|_{H^1} \le Kc^{7/6}.$$

Thus,

$$\begin{aligned}\|v_\#(-T_c) - \left\{Q(.+\tfrac{\Delta}{2}) + Q_c(.-(1-c)T_c + \Delta_c/2)\right\}\|_{H^1} &\le Kc\\ \|v_\#(T_c) - \left\{Q(.-\tfrac{\Delta}{2}) + Q_c(.+(1-c)T_c - \Delta_c/2) - 2b_{2,0}(Q_c^2)'(y_c)\right\}\|_{H^1} &\le Kc.\end{aligned}$$

By

$$\begin{aligned}&\|(Q_c^2)'(y_c) - (Q_c^2)'(.+(1-c)T_c - \Delta_c/2)\|\|_{H^1}\\ &\qquad = \|(Q_c^2)' - (Q_c^2)'(.-\Delta_c/2)\|\|_{H^1} \le Kc^{\frac{7}{6}},\end{aligned}$$

since Δ_c is a constant independent of c, we obtain the result.

Proof of Theorem 1.1.

Step 1. *Proof of nonexistence of a pure* 2*-soliton solution.* First, we claim that if there exists a global 2-soliton solution, then the speeds parameters at $+\infty$, $c_1^+ < c_2^+$ and at $-\infty$, $c_1^- < c_2^-$ satisfy $c_1^+ = c_1^-$ and $c_2^+ = c_2^-$. Indeed, by the conservation of mass and energy, and a strong limit in $H^1(\mathbb{R})$, the following holds ($q = \frac{1}{p-1} - \frac{1}{4}$):

$$(c_1^+)^{2q} + (c_2^+)^{2q} = (c_1^-)^{2q} + (c_2^-)^{2q}, \quad (c_1^+)^{2q+1} + (c_2^+)^{2q+1} = (c_1^-)^{2q+1} + (c_2^-)^{2q+1}.$$

Set $\gamma = \frac{2q+1}{2q}$, $b = \left(\frac{c_1^+}{c_1^-}\right)^{2q}$, $a^+ = \frac{c_1^+}{c_2^+} < 1$, $a^- = \frac{c_1^-}{c_2^-} < 1$. The first identity yields $b(1+a^+) = 1 + a^-$, and the second identity yields $b^\gamma(1+(a^+)^\gamma) = 1 + (a^-)^\gamma$. Thus,

$$\left(\frac{1+a^+}{1+a^-}\right)^\gamma = \frac{1+(a^+)^\gamma}{1+(a^-)^\gamma} \quad \text{and} \quad \frac{1+(a^+)^\gamma}{(1+a^+)^\gamma} = \frac{1+(a^-)^\gamma}{(1+a^-)^\gamma}.$$

The function $x \mapsto \frac{1+x^\gamma}{(1+x)^\gamma}$ is strictly decreasing on $[0,1]$; thus $a^+ = a^-$ and $b = 1$.

(i) *Behavior at* $-\infty$. Let $u(t)$ be an asymptotic 2-soliton solution at $-\infty$ with speed parameters 1 and c (c small enough), in the sense of Definition 1. Then, by the uniqueness part of Proposition 5.1, there exists δ_1^-, $\delta_2^- \in \mathbb{R}$ such that, for all $t \leq \frac{\delta_2^- - \delta_1^-}{(1-c)} - \frac{1}{32}T_c$,

$$\|u(t) - Q(x-t-\delta_2^-) - Q_c(x-ct-\delta_1^-)\|_{H^1} \leq Ke^{\frac{1}{4}\sqrt{c}((1-c)t-(\delta_2^- - \delta_1^-))}. \tag{5.9}$$

Let

$$T_c^- = T_c + \frac{\delta_1^- - \delta_2^-}{1-c} + \frac{1}{2}\frac{\Delta - \Delta_c}{1-c} \geq -\frac{\delta_2^- - \delta_1^-}{1-c} + \frac{1}{4}T_c \quad \text{and} \quad a = \tfrac{\Delta}{2} - (T_c^- - \delta_2^-).$$

Recall from (3.10) that $|\Delta| \leq Kc^{\frac{1}{6}}$ and Δ_c is a constant independent of c. Then, applying (5.9) to $t = -T_c^-$, we obtain

$$\begin{aligned}&\|u(-T_c^-, .+a) - Q(.+\tfrac{\Delta}{2}) - Q_c(.-(1-c)T_c + \tfrac{\Delta_c}{2})\|_{H^1}\\ &\qquad \leq Ke^{\frac{1}{4}\sqrt{c}(-(1-c)T_c^- - (\delta_2^- - \delta_1^-))} \leq Ke^{-\frac{1}{4}\sqrt{c}((1-c)T_c + \frac{1}{2}(\Delta_c - \Delta))} \leq Kc,\end{aligned}$$

for c small enough. By translation in time and space, we can assume $T_c^- = T_c$ and $a = 0$, so that

$$\|u(-T_c) - Q(.+\tfrac{\Delta}{2}) - Q_c(.-(1-c)T_c + \tfrac{\Delta_c}{2})\|_{H^1} \leq Kc,$$

(i.e., we consider $\tilde{u}(t,x) = u(t - T_c^- + T_c, x+a)$ instead of $u(t,x)$, and we still call it $u(t)$).

(ii) *Behavior at* $t = T_c$. By (5.7) and the above estimate, we deduce

$$\|u(-T_c) - v_\#(-T_c)\|_{H^1} \leq Kc.$$

Now, we apply Proposition 4.1 for $v_\#$ concerning the interaction region, with $\theta = 1 - \frac{1}{100}$ and $T_0 = -T_c$. Thus,

$$\forall t \in [-T_c, T_c], \quad \|u(t) - v_\#(t, . - \rho(t))\|_{H^1} \leq Kc^{1-\frac{1}{100}},$$

for some $\rho(t)$ satisfying $|\rho'(t)| \leq Kc^{1-\frac{1}{100}}$. In particular,

$$\|u(T_c) - v_\#(T_c, . - \rho(T_c))\|_{H^1)} \leq Kc^{1-\frac{1}{100}},$$

and so by Proposition 5.2, we obtain for a_-, $b_- \in \mathbb{R}$ such that $a_- - b_- > \frac{1}{2}T_c$,

$$\|u(T_c) - Q(. - a_-) - Q_c(. - b_-) - 2b_{2,0}(Q_c^2)'(. - b_-)\|_{H^1} \leq Kc^{1-\frac{1}{100}}. \tag{5.10}$$

(iii) *Behavior as* $t \to +\infty$. First, since $\|(Q_c^2)'\|_{H^1} \leq Kc^{\frac{11}{12}}$, estimate (5.10) implies that for $t = T_c$,

$$\|u(T_c) - Q(. - a_-) - Q_c(. - b_-)\|_{H^1} \leq Kc^{\frac{11}{12}}. \tag{5.11}$$

We apply Proposition 4.2 to $u(t)$ (stability of the 2-soliton structure after interaction) with $\alpha = Kc^{\frac{1}{3}}$, so that, for $w(t) = u(t) - Q(. - \rho_1(t)) - Q_c(. - \rho_2(t))$,

$$\forall t \geq T_c, \quad \sqrt{c}\|w(t)\|_{H^1} \leq \|w_x(t)\| + \sqrt{c}\|w(t))\|_{L^2} \leq Kc^{\frac{11}{12}}, \tag{5.12}$$

with $\rho_1(t)$, $\rho_2(t)$ satisfying

$$|\rho_1(T_c) - a_-| \leq Kc^{\frac{11}{12}}, \; |\rho_2(T_c) - b_-| \leq Kc^{\frac{1}{3}}, \;\; \forall t \geq T_c, \; \rho_1(t) - \rho_2(t) \geq \frac{T_c}{2} + \frac{1}{2}(t - T_c). \tag{5.13}$$

Assume now that $u(t)$ is also an asymptotic 2-soliton solution at $+\infty$. By Proposition 5.1 (applied to $+\infty$), there exist δ_1^+, δ_2^+ such that, for all $t \geq \frac{\delta_2^+ - \delta_1^+}{1-c} + \frac{T_c}{32}$,

$$\|u(t) - Q(. - t - \delta_1^+) - Q_c(. - ct - \delta_2^+)\|_{H^1} \leq Ke^{-\frac{\sqrt{c}}{4}((1-c)t - (\delta_2^+ - \delta_1^+))}. \tag{5.14}$$

We define

$$T_c^+ = \frac{\delta_2^+ - \delta_1^+}{1-c} + \frac{T_c}{32}.$$

By (5.12) and (5.14), we have for all $t \geq \max(T_c, T_c^+)$,

$$|\rho_1(t) - (t + \delta_1^+)| \leq Kc^{\frac{5}{12}}, \quad |\rho_2(t) - (ct + \delta_2^+)| \leq Kc^{-\frac{1}{6}}. \tag{5.15}$$

This is proved by considering the smallness of the L^2 norm of $Q(. - \rho_1(t)) + Q_c(. - \rho_2(t)) - Q(. - t - \delta_1^+) - Q_c(. - ct - \delta_2^+)$ in the two regions $x > \frac{1}{2}(\rho_1(t) + \rho_2(t))$ and $x < \frac{1}{2}(\rho_1(t) + \rho_2(t)))$ and the fact that for a small

$$|a| \leq K\|Q - Q(. - a)\|_{L^2} \quad |a| \leq Kc^{-\frac{7}{12}}\|Q_c - Q_c(. - a)\|_{L^2}. \tag{5.16}$$

Let us prove that $T_c > T_c^+$. By contradiction, if $T_c^+ \geq T_c$, then by (5.15) we have

$$|\rho_1(T_c^+) - \rho_2(T_c^+)| \leq |(1-c)T_c^+ + \delta_1^+ - \delta_2^+| + Kc^{-\frac{1}{6}} = \frac{T_c}{32}(1-c) + Kc^{-\frac{1}{6}} \leq \frac{T_c}{30}.$$

From (5.13),

$$|\rho_1(T_c^+) - \rho_2(T_c^+)| \geq \frac{1}{4}T_c + \frac{1}{2}(T_c^+ - T_c) \geq \frac{1}{4}T_c.$$

We obtain a contradiction from these two estimates and thus $T_c > T_c^+$.

(iv) *Conclusion of the proof.* Let $a_+ = T_c + \delta_1^+$ and $b_+ = cT_c + \delta_2^+$. By (5.14), we know

$$\|u(T_c) - Q(. - a_+) - Q_c(. - b_+)\|_{H^1} \leq Ke^{-\frac{\sqrt{c}}{4}(1-c)(T_c - T_c^+ + \frac{1}{32}T_c)} \leq Kc.$$

Thus, from (5.10)–(5.11) and Proposition 4.2,

$$\begin{aligned} |a_- - a_+| &\leq |a_- - \rho_1(T_c)| + |\rho_1(T_c) - (T_c + \delta_1^+)| \leq Kc^{\frac{11}{12}}, \\ |b_- - b_+| &\leq |b_- - \rho_2(T_c)| + |\rho_2(T_c) - (cT_c + \delta_2^+)| \leq Kc^{\frac{1}{3}}, \end{aligned} \tag{5.17}$$

and

$$\begin{aligned} \|(Q(. - a_-) - Q(. - a_+)) &+ (Q_c(. - b_-) - Q_c(. - b_+)) \\ &+ 2b_{2,0}(Q_c^2)'(. - b_-)\|_{H^1} \leq Kc^{1-\frac{1}{100}}. \end{aligned}$$

Considering the L^2 norm in the region $x < \frac{1}{2}(a_+ + b_+)$, we obtain

$$\|Q_c + 2b_{2,0}(Q_c^2)' - Q_c(. - (b_+ - b_-))\|_{L^2} \leq Kc^{1-\frac{1}{100}},$$

where $\min(a_-, a_+) > \max(b_-, b_+) + \frac{1}{4}T_c$. By scaling, it gives for $b_c = \sqrt{c}(b_+ - b_-)$,

$$\|Q + 2b_{2,0}c^{5/6}(Q^2)' - Q(. - b_c)\|_{L^2} \leq Kc^{\frac{11}{12} - \frac{1}{100}},$$

where $|b_c| \leq Kc^{\frac{5}{6}}$. Thus, $Q(x) - Q(x - b_c) = \lambda_c c^{5/6} Q'(x) + c^{5/6}o(1)$, where $|\lambda_c| \leq K$, so that

$$\|\lambda_c Q' - 2b_{2,0}(Q^2)'\|_{L^2} = o(1),$$

which is a contradiction with the fact $b_{2,0} \neq 0$ (Lemma 3.1).

Step 2. *Behavior as* $t \to +\infty$ *of* $u(t)$. As in the previous step, we consider the solution $u(t)$ which is an asymptotic 2-soliton solution at $-\infty$ i.e. satisfying (5.9). Recall that we have just proved:

– $u(t)$ is not an asymptotic 2-soliton solution at $+\infty$.
– There exist $\rho_1(t)$, $\rho_2(t)$ such that $w(t,x) = u(t,x) - (Q(x - \rho_1(t)) + Q_c(x - \rho_2(t)))$, satisfies (5.12), (5.13); in particular,

$$\forall t \geq T_c, \quad \|w(t)\|_{H^1} \leq c^{-\frac{1}{2}} \|w(t)\|_{H^1_c} \leq Kc^{\frac{5}{12}}. \tag{5.18}$$

(i) *Stability properties of* $u(t)$ *for* $t \geq T_c$. First, we claim

$$\int_{x>0} x^2 u^2(T_c, x)dx \leq K. \tag{5.19}$$

This follows directly from integration of the following estimate. For all $x_0 > 0$,

$$\int_{x>x_0} u^2(T_c, x+\rho_1(T_c))dx \leq Ke^{-\frac{1}{16}x_0} + K\exp(-c^{-\frac{1}{400}})\, e^{-\frac{1}{16}\sqrt{c}x_0}. \tag{5.20}$$

Let us prove (5.20). On the one hand, by monotonicity arguments on $u(t)$ as in Lemma 1 of [19],

$$\begin{aligned} \int u^2(T_c,x)\psi(x-\rho_1(T_c)-x_0)dx \\ \leq \int u^2(-T_c,x)\psi(x-\rho_1(-T_c)-x_0-\tfrac{T_c}{2})dx + Ke^{-\frac{1}{16}x_0}. \end{aligned} \tag{5.21}$$

On the other hand, using $\mathcal{I}_{\sigma,y_0}$ for $\sigma = c$, $y_0 = \rho_1(-T_c) + x_0 + \frac{T_c}{2}$, we get for any $t < -T_c$,

$$\begin{aligned} &\int u^2(-T_c,x)\psi(\sqrt{c}(x-\rho_1(-T_c)-x_0-\tfrac{T_c}{2}))dx \\ &\quad\leq \int u^2(t,x)\psi(\sqrt{c}(x-\rho_1(-T_c)-x_0-\tfrac{T_c}{2}-\tfrac{c}{4}t))dx + Ke^{-\frac{1}{16}\sqrt{c}(x_0+\frac{1}{2}T_c)}. \end{aligned}$$

By (5.9) and letting $t \to -\infty$, we obtain

$$\int u^2(-T_c,x)\psi(\sqrt{c}(x-\rho_1(-T_c)-x_0-\tfrac{T_c}{2}))dx \leq K\exp(-c^{-\frac{1}{400}})\, e^{-\frac{1}{16}\sqrt{c}x_0}. \tag{5.22}$$

Therefore, from (5.21) and (5.22),

$$\int_{x>x_0} u^2(T_c,x+\rho_1(T_c))dx \leq \frac{1}{2}\int u^2(T_c,x)\psi(x-\rho_1(T_c)-x_0)dx$$

and $\psi(\sqrt{c}y) \geq \frac{1}{2}\psi(y)$, we obtain (5.20).

Now, from (5.19) and (5.11), we can apply Proposition 4.2 to $u(.+T_c)$, for $t \geq 0$ (with $\alpha = Kc^{\frac{1}{3}}$). It follows that there exists $c_1^+,\ c_2^+ > 0$, $\delta_1^+, \delta_2^+ \in \mathbb{R}$ such that $w^+(t) = u(t) - Q_{c_1^+}(.-\delta_1^+ - c_1^+t) - Q_{c_2^+}(.-\delta_2^+ - c_2^+t)$ satisfies

$$\lim_{t\to+\infty} \|w^+(t)\|_{H^1(x>ct/10)} = 0 \quad \text{with} \quad |c^+ - 1| \leq Kc^{\frac{11}{12}}, \quad \left|\frac{c_2^+}{c} - 1\right| \leq Kc^{\frac{1}{3}}. \tag{5.23}$$

Note also that from the stability (5.11) and (5.23), we obtain the following upper bound on $w^+(t)$ for t large enough:

$$\begin{aligned} \|w^+(t)\|_{H^1} &\leq \|w^+(t)\|_{H^1(x<\frac{1}{10}ct)} + \|w^+(t)\|_{H^1(x>\frac{1}{10}ct)} \\ &\leq \|w(t)\|_{H^1(x<\frac{1}{10}ct)} + o(1) \leq Kc^{\frac{5}{12}}. \end{aligned}$$

Therefore, to finish the proof of Theorem 1.1, we only have to prove the lower bounds on $w^+(t)$, $c_1^+ - 1$ and $1 - \frac{c_2^+}{c}$.

(ii) *Lower bounds on the defects.* Let $\eta(t)$, $g(t)$ and $c_j(t)$ $(j=1,2)$ be defined from $u(t)$ for $t \geq T_c$ as in Claim 4.3 and satisfying

$$\|\eta(t)\|_{H^1(x\geq\frac{c}{10}t)} \to 0, \quad c_j(t) \to c_j^+ \quad \text{as } t\to+\infty \ (j=1,2). \tag{5.24}$$

In particular, it is sufficient to prove the lower bounds on $\eta(t)$ to obtain lower bounds on $w^+(t)$ for large time. We claim

$$\forall t \geq T_c, \quad \|\eta(t)\|_{H^1_c} \geq K_1 c^{\frac{17}{12}} \quad (K_1>0). \tag{5.25}$$

Proof of (5.25). To prove this lower bounds using the defect $(Q_c^2)'$ in (5.10), we need to apply an argument of stability backwards in time, locally around the soliton $R_2(t)$. For this, we will use monotonicity type results on $\eta(t)$ as in Claim 4.4.

First, we claim

$$\int_{x\leq\rho_2(T_c)+\frac{1}{4}T_c} \eta^2(T_c,x)dx \geq K_0 c^{\frac{11}{6}} \quad (K_0>0). \tag{5.26}$$

Proof of (5.26). Let $\varepsilon>0$ to be fixed later and assume for the sake of contradiction that $\int_{x\leq\rho_2(T_c)+\frac{1}{4}T_c} \eta^2(T_c,x)dx \leq \varepsilon^2 c^{\frac{11}{6}}$. Recall from (5.10) that

$$\|u(T_c) - Q(.-a_-) - Q_c(.-b_-) + 2b_{2,0}(Q_c^2)'(.-b_-)\|_{L^2} \leq Kc. \tag{5.27}$$

Thus, as in Step 1(iv), we obtain for c small enough,

$$\|Q_c(.-b_-) - 2b_{2,0}(Q_c^2)'(.-b_-) - Q_{c_2(T_c)}(.-b_+)\|_{L^2} \leq K\varepsilon c^{\frac{11}{12}},$$

and after scaling,

$$\|Q - 2b_{2,0}c^{\frac{5}{6}}(Q^2)' - Q_\Lambda(.-b_c)\|_{L^2} \leq K\varepsilon c^{\frac{5}{6}},$$

for $\Lambda = \frac{c_2(T_c)}{c}$, $b_c = \sqrt{c}(b_+ - b_-)$. From orthogonality of even and odd functions in L^2 and parity of $\frac{d^k}{dc^k}Q_c$ for any $k\geq 0$, we obtain

$$\|Q - 2b_{2,0}c^{\frac{5}{6}}(Q^2)' - Q(.-b_c)\|_{L^2} \leq K\varepsilon c^{\frac{5}{6}},$$

which is a contradiction for ε small enough, as in Step 1(iv) $(b_{2,0}\neq 0)$. Thus, (5.26) is proved.

Let $\varepsilon>0$ to be fixed later and assume for the sake of contradiction that for some $t' \geq T_c$,

$$\|\eta(t')\|_{H^1_c} \leq \varepsilon c^{\frac{17}{12}}. \tag{5.28}$$

Let $\tilde\psi_2(t,x) = 1 - \psi(\sqrt{c}(x-\rho_2(t) - \frac{1}{4}T_c - \frac{1}{2}(t'-t))$, where ψ is defined in (4.32) and

$$\tilde{\mathcal{M}}_2(t) = \int \eta^2 t)\tilde\psi_2,$$

$$\tilde{\mathcal{E}}_2(t) = \int \Big[\tfrac{1}{2}\eta_x^2 - \tfrac{1}{5}((R_1+R_2+\eta)^5 - 5R_1^4\eta - 5R_2^4\eta - (R_1+R_2)^5)\Big]\tilde\psi_2.$$

From (5.28) and the properties of R_1, R_2, we have $c\tilde{\mathcal{M}}_2(t')+|\tilde{\mathcal{E}}_2(t')| \le K\varepsilon^2 c^{\frac{17}{6}}$. Thus from Claim 4.4, integrated on $[T_c, t']$, we have

$$(c_2^{2q}(T_c) - c_2^{2q}(t'))\int Q^2 \le -\tilde{\mathcal{M}}_2(T_c) + K\varepsilon^2 c^{\frac{17}{6}},$$

$$\Big(\frac{2q}{2q+1}(c_2^{2q+1}(T_c) - c_2^{2q+1}(t')) - \tfrac{c}{100}(c_2^{2q}(T_c) - c_2^{2q}(t'))\Big)\int Q^2 \\ \ge 2\mathcal{E}_2(T_c) + \tfrac{c}{100}\mathcal{M}_2(T_c) - K\varepsilon^2 c^{\frac{17}{6}}.$$

From this, using the coercive functional of $\eta(t)$: $\tilde{\mathcal{E}}_2(t) + \frac{1}{2}c_2(t')\tilde{\mathcal{M}}_2(t)$, and proceeding as in [25, App. B.3], we obtain successively

$$|c_2(T_c) - c_2(t')| \le K\int(\eta_x^2 + c\eta^2)(T_c)\tilde{\psi}_2 + K\varepsilon^2 c^{\frac{17}{6}},$$

$$\int(\eta_x^2 + c\eta^2)(T_c)\tilde{\psi}_2 \le K\varepsilon^2 c^{\frac{17}{6}} + K|c_2(T_c) - c_2(t')|^2 \le K'\varepsilon^2 c^{\frac{17}{6}},$$

which contradicts (5.26) for ε small enough.

Finally, we prove (1.11), that is, the upper and lower bounds on $c_1^+ - 1$ and $1 - \frac{c_2^+}{c}$, using the two conservation laws, written as $t \to \pm\infty$ and the upper and lower bounds on $w^+(t)$. By (5.9) and (5.23), we have, for t large,

$$\int u^2(0) = \int Q^2 + \int Q_c^2 = \int Q_{c_1^+}^2 + \int Q_{c_2^+}^2 + \int (w^+)^2(t) + o(1),$$
$$E(u(0)) = E(Q) + E(Q_c) = E(Q_{c_1^+}) + E(Q_{c_2^+}) + E(w^+(t)) + o(1).$$

By the Gagliardo-Nirenberg inequality and the estimate $\|w^+(t)\|_{H^1} \le Kc^{\frac{5}{12}}$, we have $\int(w^+)^5 \le K\|w^+\|_{H^1}^3\int(w^+)^2 \le Kc^{\frac{5}{4}}\int(w^+)^2$ and thus, for t large enough,

$$\Big|E(w^+(t)) - \tfrac{1}{2}\int(w_x^+(t))^2\Big| \le Kc^{\frac{5}{4}}\int(w^+(t))^2.$$

Thus, by Claim C.1, for t large, we obtain

$$\Big|(c^{2q} - (c_2^+)^{2q}) + (1 - (c_1^+)^{2q}) - \frac{1}{\int Q^2}\int(w^+(t))^2\Big| \le Kc^4, \tag{5.29}$$

$$\Big|(c^{2q+1} - (c_2^+)^{2q+1}) + (1 - (c_1^+)^{2q+1}) + \frac{1}{2|E(Q)|}\int(w_x^+(t))^2\Big| \\ \le Kc^{\frac{5}{4}}\int(w^+(t))^2 + Kc^4. \tag{5.30}$$

Let $a = (c^{2q+1} - (c_2^+)^{2q+1})/(c^{2q+1} - c\,(c_2^+)^{2q})$; then $\frac{1}{2}\frac{2q+1}{2q} \le a \le \frac{3}{2}\frac{2q+1}{2q}$. Multiplying (5.29) by $c\,a$ and summing (5.30), we obtain, for c small enough,

$$K(c_1^+ - 1) \ge (c_1^+)^{2q+1} - 1 \ge K\int(w_x^+)^2 + c(w^+)^2)(t) - Kc^4 \ge K_0 c^{\frac{17}{6}}.$$

Similarly, set $b = (1-(c_1^+)^{2q})/(1-(c_1^+)^{2q+1})$, then $\frac{1}{2} \le b \le \frac{3}{2}$, and multiplying (5.30) by $-b$ and summing (5.29), we obtain, for c small enough ($q = \frac{1}{12}$),

$$Kc^{\frac{1}{6}}\left(1 - \frac{c_2^+}{c}\right) \ge c^{2q} - (c_2^+)^{2q} \ge K \int ((w_x^+)^2 + (w^+)^2)(t) \ge Kc^{\frac{17}{6}}.$$

Arguing similarly and using the upper bound (5.12) on w^+, we also obtain the upper bounds in (1.11) (in particular, using the conservation laws, we improve the estimates (5.23) which were obtained by a stability argument). This completes the proof of Theorem 1.1.

Proof of Remark 1. The remark is based on the fact that for $p = 4$,

$$\int Q_c = c^{-\frac{1}{6}} \int Q.$$

In the framework of the proof of Theorem 1.1, we consider $u(t)$ the asymptotic 2-soliton solution at $-\infty$ with speed parameters 1 and c (c small enough). Let us prove by contradiction that $u(t)$ is not an asymptotic N-soliton solution at $+\infty$.

Assume that $\|u(t) - \sum_{j=1}^N Q_{c_j^+}(. - \delta_j^+ - c_j^+ t)\|_{H^1} \to 0$ as $t \to +\infty$, where $c_2^+ > c_3^+ > \cdots > c_N^+$. Using the methods of [27], [19] and the fact that $u(t)$ is an asymptotic N-soliton solution both at $\pm\infty$, we have, for some $T_0 > 0$ large enough,

$$\forall t \ge T_0, \forall x \in \mathbb{R}, \quad |u(t,x)| \le K \sum_{j=1}^N Q_{c_j^+}^{\frac{1}{4}}(. - \delta_j^+ - c_j^+ t),$$

which proves that $u(t) \in L^1(\mathbb{R})$, and, in particular, $\int u(t) = I_0$ is well-defined and constant in time. Moreover, $u(t) - \sum_{j=1}^N Q_{c_j^+}(. - \delta_j^+ - c_j^+ t) \to 0$ as $t \to +\infty$ in $L^1(\mathbb{R})$, from the H^1 convergence. A similar convergence in L^1 holds at $-\infty$.

On the one hand, at $-\infty$, $I_0 = \lim_{t\to-\infty} \int u(t) = \int Q_{c_1} + \int Q_{c_2} = (c_1^{-\frac{1}{6}} + c_2^{-\frac{1}{6}}) \int Q$. On the other hand, at $+\infty$, $I_0 = \sum_{j=1}^N (c_j^+)^{-\frac{1}{6}} \int Q$. Since by Theorem 1.1, $\|w^+(t)\|_{L^2} \le Kc_2^{\frac{7}{12}}$, we have $c_3^+ \ll (c_2^+)^4$. Thus, $I_0 \gg (c_2^+)^{-\frac{2}{3}} \int Q$, which is a contradiction, for c_2 small.

5.2. *Existence of a* 2*-soliton-like solution. Proof of Theorem* 1.2. We consider first the case $c_1 = 1$ and $c_2 = c$, the general case following from a scaling argument. For any $c > 0$ small enough, we consider $u_c(t)$ the global solution of

$$\partial_t u_c + \partial_x(\partial_x^2 u_c + u_c^4) = 0, \quad u_c(0,x) = v_c(0,x),$$

where $v_c(t)$ is the approximate solution constructed in Proposition 3.1, for k_0, ℓ_0 large enough but fixed. Recall also that Δ and Δ_c are defined in Proposition 3.1. By the parity property of $x \to v_c(0,x)$ and since equation (1.1) is invariant under the transformation $x \to -x$, $t \to -t$, the solution $u_c(t)$ has the following symmetry:

$$u_c(t,x) = u_c(-t,-x). \tag{5.31}$$

Thus, we shall only study $u_c(t)$ for $t \geq 0$. We claim the following concerning $u_c(t)$.

Proposition 5.3. *There exist $c_0 > 0$ such that for all $0 < c < c_0$, there exist $c_1^+(c)$, $c_2^+(c) > 0$, and $\delta_1^+(c)$, $\delta_2^+(c) \in \mathbb{R}$ such that*

$$w^+(t,x) = u_c(t,x) - Q_{c_1^+(c)}(x - c_1^+(c)t - \delta_1^+(c)) - Q_{c_2^+(c)}(x - c_2^+(c)t - \delta_2^+(c)).$$

1. *Asymptotic behavior*:

$$\lim_{t\to+\infty} \|w^+(t)\|_{H^1(x>ct/10)} = 0, \tag{5.32}$$

$$|\delta_1^+(c) - \tfrac{1}{2}\Delta| \leq Kc^{\frac{3}{8}}, \quad |\delta_2^+(c) - \tfrac{1}{2}\Delta_c| \leq Kc^{\frac{1}{12}}, \tag{5.33}$$
$$|c_1^+(c) - 1| \leq Kc^{\frac{11}{12}}, \qquad \left|\frac{c_2^+(c)}{c} - 1\right| \leq Kc^{\frac{1}{3}},$$

$$\text{for } t \text{ large,} \quad \frac{1}{K}c^{\frac{17}{12}} \leq \|w^+(t)\|_{H^1_c} \leq K \min_{\rho_1,\rho_2\in\mathbb{R}} \|w_{\rho_1,\rho_2}(t)\|_{H^1_c} \leq K^2 c^{\frac{17}{12}}, \tag{5.34}$$

$$\text{where} \quad w_{\rho_1,\rho_2}(t,x) = u(t,x) - Q_{c_1^+(c)}(x-\rho_1) - Q_{c_2^+(c)}(x-\rho_2). \tag{5.35}$$

2. *$c \mapsto c_j^+(c)$ for $j = 1, 2$ are continuous.*

Proof of Theorem 1.2 *assuming Proposition* 5.3. We claim that a rescaled version of $u_{\tilde{c}}(t)$ for some $\tilde{c} \sim c$ satisfies the conclusions of Theorem 1.2.

From Proposition 5.3, the function $h(c) = \frac{c_2^+(c)}{c_1^+(c)}$ is continuous on $(0, c_0]$; moreover $\frac{1}{2}c \leq h(c) \leq \frac{3}{2}c$. It follows that $h((0,c_0])$ is an interval containing $(0, \frac{1}{2}c_0]$. Thus, for any $c \in (0, \frac{1}{2}c_0]$, there exists $\tilde{c}$ such that

$$\frac{1}{2}c \leq \tilde{c} \leq 2c, \quad h(\tilde{c}) = c. \tag{5.36}$$

Let

$$U_{1,c}(t,x) = U(t,x) = c_1^{-\frac{1}{3}}(\tilde{c}) u_{\tilde{c}}\Big(c_1^{-\frac{3}{2}}(\tilde{c})t, c_1^{-\frac{1}{2}}(\tilde{c})x\Big). \tag{5.37}$$

From Proposition 5.3, (5.37), (5.36) and (5.31), it follows that U satisfies (1.14). Moreover, (1.16) follows from (5.34).

Let $c_1 > 0$ and $c_2 > 0$ such that $c = \frac{c_2}{c_1} < \epsilon_0$ small. Let

$$U_{c_1,c_2}(t,x) = c_1^{\frac{1}{3}} U_{1,c}\Big(c_1^{\frac{3}{2}}t, c_1^{\frac{1}{2}}x\Big), \quad \Delta_j = \Delta_j(c_1,c_2) = c_1^{-\frac{1}{2}}\delta_j^+(\tilde{c}),\ j = 1, 2.$$

Then U_{c_1,c_2} verifies the conclusion of Theorem 1.2. Note, in particular, that (1.15) follows from (5.33) and (3.10), (3.13).

Proof of Proposition 5.3. In Steps 1 and 2 of this proof, we omit the c dependency.

Step 1. *Control of the modulation of $u(t)$ for $t \geq T_c$.* Applying Proposition 4.1 for $t \in [0, T_c]$, with $\theta = n_0 - \frac{1}{2} - \frac{1}{100}$, we obtain, for some $\rho(t)$,

$$\forall t \in [0, T_c], \quad \|u(t) - v(t, . - \rho(t))\|_{H^1} \leq K c^\theta, \tag{5.38}$$

where $|\rho'(t) - 1| \leq K c^\theta$, $\rho(0) = 0$, and so $|\rho(T_c) - T_c| \leq K c^{\theta - \frac{1}{2} - \frac{1}{100}}$ by $T_c = c^{-\frac{1}{2} - \frac{1}{100}}$.

By (3.11) and (5.38), and then by $\|(Q_c^2)'\|_{H_c^1} = K c^{\frac{17}{12}}$ and (3.8)–(3.9), we have, for $\theta \geq 2$,

$$\|u(T_c) - Q(. - a) - Q_c(. - b)\|_{H^1(x > T_c/4)} \leq K c^\theta \leq K c^2, \tag{5.39}$$

(5.40)
$$\sqrt{c}\|u(T_c) - Q(. - a) - Q_c(. - b)\|_{H^1} \leq \|u(T_c) - Q(. - a) - Q_c(. - b)\|_{H_c^1} \leq K c^{\frac{17}{12}},$$

for $a = \frac{1}{2}\Delta + \rho(T_c)$, $b = (1 - c)T_c + \frac{1}{2}\Delta_c + \rho(T_c)$, so that $a - b \geq \frac{1}{2} T_c$.

Therefore, from Claim 4.3 and Proposition 4.2, we have the decomposition of $u(t)$ in terms of $\eta(t)$, $c_j(t)$, $\rho_j(t)$ $(j = 1, 2)$ defined for all $t \geq T_c$.

LEMMA 5.1. *For all $t \geq T_c$, $\frac{1}{K} c^{\frac{17}{12}} \leq \|\eta(t)\|_{H_c^1} \leq K c^{\frac{17}{12}}$.*

Proof of Lemma 5.1. (i) *Upper bounds by stability properties.* We use Claim 4.3, which is a refinement of Proposition 4.2 (see proof of Proposition 2 in [25]). Let $g(t)$ be defined from $\eta(t)$ by (4.34). Remark from (5.39) and the proof of Claim 4.3 in [25], that $|c_1(t) - 1| + |a_- - \rho_1(t)| \leq K c^2$ and

$$\|\eta(T_c)\|_{H^1(x > T_c/4)} \leq K c^2. \tag{5.41}$$

Similarly, we obtain $\|\eta(T_c)\|_{H_c^1(x < T_c/4)} \leq K c^{\frac{17}{12}}$ from (5.40). Thus,

$$\sqrt{g(T_c)} \leq K \|\eta(T_c)\|_{H_c^1(x < T_c/4)} + \|\eta(T_c)\|_{H^1(x > T_c/4)} \leq K c^{\frac{17}{12}}.$$

By Claim 4.3, for all $t \geq T_c$, $\|\eta(t)\|_{H_c^1} \leq \sqrt{g(t)} \leq K(\sqrt{g(T_c)} + \exp(-c^{-\frac{1}{400}})) \leq K c^{\frac{17}{12}}$.

(ii) *Lower bounds by backwards stability.* See the proof of (5.25) (Theorem 1.1).

Step 2. *Proof of asymptotic stability.* From properties of v, we claim the following:

$$\int_{x > \frac{11}{12}|\ln c|} x^2 u^2(T_c, x + T_c + \tfrac{1}{2}\Delta) dx \leq K c^{\frac{5}{4}}, \tag{5.42}$$

$$|\rho_1(T_c) - T_c - \tfrac{\Delta}{2}| \leq K c^{\frac{11}{12}}, \quad |\rho_2(T_c) - c T_c - \tfrac{\Delta_c}{2}| \leq K c^{\frac{1}{3}}. \tag{5.43}$$

See Appendix D.3 for the proof of (5.42) and (5.43). Note that the proof of (5.42) is based on monotonicity arguments on $z(t) = u(t) - v(t, . - \rho(t))$ as defined in (4.11) in the proof of Proposition 4.1.

From (5.40)–(5.42), we apply Proposition 4.2 to $u(. + T_c)$ with $\alpha = Kc^{\frac{1}{3}}$. There exist c_1^+, $c_2^+ > 0$, δ_1^+, $\delta_2^+ \in \mathbb{R}$ such that

$$c_j(t) \to c_j^+, \quad \rho_j(t) - c_j^+ t \to \delta_j^+, \quad \text{as } t \to +\infty,\ j = 1, 2, \tag{5.44}$$

and

$$\lim_{t \to +\infty} \|u(t) - w^+(t)\|_{H^1(x > ct/10)} = 0, \tag{5.45}$$

where $w^+(t, x) = Q_{c_1^+}(x - c_1^+ t - \delta_1^+) - Q_{c_2^+}(x - c_2^+ t - \delta_2^+)$,

$$|c_1^+ - 1| \le Kc^{\frac{11}{12}}, \qquad \left| \frac{c_2^+}{c} - 1 \right| \le Kc^{\frac{1}{3}}, \tag{5.46}$$

$$|\delta_1^+ + c_1^+ T_c - \rho_1(T_c)| \le Kc^{\frac{5}{8}}, \quad |\delta_2^+ + c_2^+ T_c - \rho_2(T_c)| \le Kc^{\frac{1}{12}}. \tag{5.47}$$

From (5.43) and (5.47), we finish the computation of δ_j^+. For δ_1^+, inserting (5.43) in (5.47), we obtain: $|\delta_1^+ - (1 - c_1^+)T_c - \frac{1}{2}\Delta| \le Kc^{\frac{5}{8}}$. Since $|1 - c_1^+|T_c \le Kc^{\frac{11}{12}} T_c \le Kc^{\frac{3}{8}}$, we conclude that $|\delta_1^+ - \frac{1}{2}\Delta| \le Kc^{\frac{3}{8}}$. Similarly for δ_2^+, we obtain from (5.43) and (5.47), $|\delta_2^+ - \frac{1}{2}\Delta_c| \le Kc^{\frac{1}{12}}$.

From (5.44), $\|\eta(t) - w^+(t)\|_{H^1_c} \to 0$ as $t \to +\infty$ and thus, from Lemma 5.1, we obtain $\frac{1}{K} c^{\frac{17}{12}} \le \|w^+(t)\|_{H^1_c} \le Kc^{\frac{17}{12}}$ for t large. From (5.45), $\|w^+(t)\|_{H^1_c} \le \min_{\rho_1, \rho_2} \|w_{\rho_1, \rho_2}(t)\|_{H^1_c} + o(1)$ for t large, where $w_{\rho_1,\rho_2}(t)$ is defined in (5.35), and thus (5.34) follows. This concludes the proof of the first part of Proposition 5.3.

Step 3. *Continuity of* $c_1^+(c)$ *and* $c_2^+(c)$. Now, we prove that the maps $c \mapsto c_1^+(c)$ is continuous. Let us denote by $\eta_c(t)$, $c_{c,j}(t)$, $c_j^+(c)$, the parameters in the decomposition of $u_c(t)$. We claim:

Claim 5.1. *For all* $t \ge T_c$,

$$|c_1^+(c) - c_{c,1}(t)| \le K_0 \int (\eta_{c,x}^2 + \eta_c^2)(t, x) \psi(x - \rho_1(t) + \tfrac{t}{4}) dx + K_0 e^{-\frac{1}{64}\sqrt{c}t}. \tag{5.48}$$

Assuming this claim, let us complete the proof of continuity of $c_1^+(c)$. Let $0 < \bar{c} < c_0$ and let $\varepsilon > 0$. Since $\|\eta_{\bar{c}}(t)\|_{H^1(x > ct/10)} \to 0$ as $t \to +\infty$, there exits $T_\varepsilon > 0$ such that

$$K_0 \int (\eta_{\bar{c},x}^2 + \eta_{\bar{c}}^2)(T_\varepsilon, x) \psi(x - \rho_1(T_\varepsilon) + \tfrac{T_\varepsilon}{4}) dx + K_0 e^{-\frac{1}{64}\sqrt{c}T_\varepsilon} \le \varepsilon.$$

We fix $T_\varepsilon > 0$ to such a value. Then, by continuous dependence in H^1 of the $u_c(t)$ solution of (1.1) upon the initial data (see [15]), and the fact that

$u_c(0) = v_c(0)$ is continuous upon the parameter c, there exists $\delta > 0$ such that if $|c - \bar{c}| \leq \delta$, then

$$K_0 \int (\eta_{c,x}^2 + \eta_c^2)(T_\varepsilon, x)\psi(x - \rho_1(T_\varepsilon) + \tfrac{T_\varepsilon}{4})dx + K_0 e^{-\frac{1}{64}\sqrt{c}T_\varepsilon} \leq 2\varepsilon,$$

$$|c_{\bar{c},1}(T_\varepsilon) - c_{c,1}(T_\varepsilon)| \leq \varepsilon.$$

From Claim 5.1, applied to η_c and $\eta_{\bar{c}}$, we have $|c_1^+(c) - c_{c,1}(T_\varepsilon)| \leq 2\varepsilon$ and $|c_1^+(\bar{c}) - c_{\bar{c},1}(T_\varepsilon)| \leq \varepsilon$. Therefore, $|c_1^+(\bar{c}) - c_1^+(c)| \leq 4\varepsilon$. Thus, $c \mapsto c_1^+(c)$ is continuous. We argue similarly for $c \mapsto c_2^+(c)$ using a claim similar to Claim 5.1 on $|c_2^+(c) - c_{c,2}(t)|$ (related to $\mathcal{M}_2(t)$ and $\mathcal{E}_2(t)$) and the previous result on $c_1^+(c)$. This concludes the proofs of Proposition 5.3 and of Theorem 1.2.

Proof of Claim 5.1. The proof follows closely some arguments in [25]. For $T_c \leq t_0 \leq t$, let $\mathcal{M}_1(t)$ and $\mathcal{E}_1(t)$ be defined in Secion 4.3, with $x_0 = \frac{t_0}{4}$. From the conclusions of Claim 4.4 integrated on $[t_0, t]$, we obtain

$$(c_1^{2q}(t) - c_1^{2q}(t_0)) \int Q^2 \leq (\mathcal{M}_1(t_0) - \mathcal{M}_1(t)) + Ke^{-\frac{1}{64}\sqrt{c}t_0},$$

$$\Big(\frac{2q}{2q+1}(c_1^{2q+1}(t) - c_1^{2q+1}(t_0)) - \frac{1}{100}(c_1^{2q}(t) - c_1^{2q}(t_0))\Big) \int Q^2$$
$$\geq 2\mathcal{E}_1(t) - 2\mathcal{E}_1(t_0) + \frac{1}{100}(\mathcal{M}_1(t) - \mathcal{M}_1(t_0)) - Ke^{-\frac{1}{64}\sqrt{c}t_0}.$$

Note, in particular, that $\int_{t_0}^t e^{-\frac{1}{16}(t-t_0+x_0)} g_1(t)dt \leq Ke^{-\frac{1}{16}x_0} \leq Ke^{-\frac{1}{64}t_0}$. Letting $t \to +\infty$, by the asymptotic stability, this gives

$$((c_1^+)^{2q} - c_1^{2q}(t_0)) \int Q^2 \leq \mathcal{M}_1(t_0) + Ke^{-\frac{1}{64}\sqrt{c}t_0},$$

$$\Big(\frac{2q}{2q+1}((c_1^+)^{2q+1} - c_1^{2q+1}(t_0)) - \frac{1}{100}((c_1^+)^{2q} - c_1^{2q}(t_0))\Big) \int Q^2$$
$$\geq -2\mathcal{E}_1(t_0) - \frac{1}{100}\mathcal{M}_1(t_0) - Ke^{-\frac{1}{64}\sqrt{c}t_0}.$$

Thus, we obtain

$$|c_1^+ - c_1(t_0)| \leq K \int (\eta_x^2 + \eta^2)(t_0, x)\psi(x - \rho_1(t_0) + \tfrac{t_0}{4})dx - Ke^{-\frac{1}{64}\sqrt{c}t_0}.$$

5.3. *Stability of the 2-soliton structure. Proof of Theorem* 1.3. Without loss of generality, we prove Theorem 1.3 in the case $c_1 = 1$ and $c_2 = c$. We assume

$$\|u(0) - U(0)\|_{H^1} \leq Kc^{\delta + \frac{7}{12}},$$

for $\delta > 0$, where U is the solution constructed in Theorem 1.2. Let $\tilde{c} > 0$ small satisfy $\frac{c_2^+(\tilde{c})}{c_1^+(\tilde{c})} = c$ and $\lambda = 1/c_1^+(\tilde{c})$. Then,

$$\|\lambda^{\frac{1}{3}} u(0, \sqrt{\lambda}x) - \lambda^{\frac{1}{3}} U(0, \sqrt{\lambda}x)\|_{H^1} \leq Kc^{\delta + \frac{7}{12}}.$$

By construction of $U(t)$ in Theorem 1.2, $\lambda^{\frac{1}{3}}U(0,\sqrt{\lambda}x)=v(0)$, where v is the approximate solution introduced in Proposition 3.1 corresponding to $\tilde{c}$ for k_0, ℓ_0 large enough. Since the solution of (1.1) corresponding to $\lambda^{\frac{1}{3}}u(0,\sqrt{\lambda}x)$ is $\lambda^{\frac{1}{3}}u(\lambda t,\sqrt{\lambda}x)$, it is enough to prove the theorem in the case

$$\|u(0)-v(0)\|_{H^1}\le Kc^{\delta+\frac{7}{12}}. \tag{5.49}$$

By invariance of (1.1) by the transformation $x\to -x$, $t\to -t$, it is enough to prove the result for $t\ge 0$.

(i) *Estimates on* $[0,T_c]$. By (5.49) and Proposition 4.1, we obtain, for all $t\in[-T_c,T_c]$, for some $\rho(t)$,

$$\|u(t)-v(t,x-\rho(t))\|_{H^1}\le Kc^{\delta+\frac{7}{12}}.$$

From Proposition 3.1, we deduce, for some a, b, with $a-b\ge \frac{1}{2}T_c$,

$$\|u(T_c)-Q(.-a)-Q_c(.-b)\|_{H^1}\le K(c^{\delta+\frac{7}{12}}+c^{\frac{11}{12}}). \tag{5.50}$$

(ii) *Estimates on* $[T_c,+\infty)$. By (5.50) and Propositions 4.2 and 4.2, for all $t\in[T_c,+\infty)$, there exist $\rho_1(t)$, $\rho_2(t)$ and c_1^+, c_2^+, such that (recall that for $p=4$, $q+\frac{1}{2}=\frac{7}{12}$)

$$\|u(t)-Q_{c_1^+}(.-\rho_1(t))-Q_{c_2^+}(.-\rho_2(t))\|_{H^1}\le K(c^{\delta+\frac{1}{12}}+c^{\frac{5}{12}}),$$
$$\left|c_1^+-1\right|\le K(c^{\delta+\frac{7}{12}}+c^{\frac{11}{12}}),\quad \left|\frac{c_2^+}{c}-1\right|\le K(c^{\delta}+c^{\frac{1}{3}}).$$

Appendix A. Proof of Proposition 2.1

To prove Proposition 2.1, we decompose each of the terms **I**, **II**, **III** and **IV** obtained in (2.16) in series of $c^\ell Q_c^k$, $c^\ell (Q_c^k)'$. In this decomposition (for future use in solving the systems $(\Omega_{k,\ell})$), we will separate terms depending on (k,ℓ) and terms depending on $(k',\ell')\prec(k,\ell)$.

Claim A.1. 1. *For* $r>0$, $Q_c^r(y_c)\beta(y_c)=\sum_{\substack{1+r\le k\le k_0+r\\ 0\le\ell\le\ell_0}} c^\ell Q_c^k(y_c)a_{k-r,\ell}$.

2. *Decomposition of* β'', β^2, $\beta'\beta$ *and* β^3. *There exist* $a_{k,\ell}^{1*}$, $a_{k,\ell}^{2*}$, $a_{k,\ell}^{3*}$ *and* $a_{k,\ell}^{4*}$ *depend on* $(a_{k',\ell'})$ *for* $(k',\ell')\prec(k,\ell)$ *such that*

$$\beta''(y_c)=\sum_{\substack{1\le k\le k_0+p-1\\ 0\le\ell\le\ell_0+1}} c^\ell Q_c^k(y_c)a_{k,\ell}^{1*},\quad \beta^2(y_c)=\sum_{\substack{2\le k\le 2k_0\\ 0\le\ell\le 2\ell_0}} c^\ell Q_c^k(y_c)a_{k,\ell}^{2*},$$
$$\beta'(y_c)\beta(y_c)=\sum_{\substack{2\le k\le 2k_0\\ 0\le\ell\le 2\ell_0}} c^\ell (Q_c^k)'(y_c)a_{k,\ell}^{3*},\quad \beta^3(y_c)=\sum_{\substack{3\le k\le 3k_0\\ 0\le\ell\le 3\ell_0}} c^\ell Q_c^k(y_c)a_{k,\ell}^{4*}.$$

Proof of Claim A.1. The first formula follows immediately from the decomposition of $\beta(y_c)$:

$$\beta(y_c) = \sum_{(k,\ell)\in\Sigma_0} a_{k,\ell}\, c^\ell Q_c^k(y_c). \tag{A.1}$$

Decomposition of β''. Using Lemma 2.1,

$$\begin{aligned}
\beta''(y_c) &= \sum_{(k,\ell)\in\Sigma_0} c^\ell (Q_c^k)''(y_c) a_{k,\ell} \\
&= \sum_{\substack{1\le k\le k_0\\ 0\le \ell\le \ell_0}} c^\ell \left(ck^2 Q_c^k(y_c) - \frac{k(2k+p-1)}{p+1} Q_c^{k+p-1}(y_c)\right) a_{k,\ell} \\
&= \sum_{\substack{1\le k\le k_0\\ 1\le \ell\le \ell_0+1}} c^\ell Q_c^k(y_c) k^2 a_{k,\ell-1} \\
&\quad + \sum_{\substack{p\le k\le k_0+p-1\\ 0\le \ell\le \ell_0}} c^\ell Q_c^k(y_c) \left(-\frac{(k-p+1)(2k-p+1)}{p+1} a_{k-p+1,\ell}\right).
\end{aligned}$$

Thus, $\beta''(y_c) = \sum_{\substack{1\le k\le k_0+p-1\\ 0\le \ell\le \ell_0+1}} c^\ell Q_c^k(y_c) a^{1*}_{k,\ell}$, where ($\mathbf{1}$ denoting the characteristic function)

$$a^{1*}_{k,\ell} = k^2 a_{k,\ell-1}\mathbf{1}_{\left\{\substack{1\le k\le k_0\\ 1\le \ell\le \ell_0+1}\right\}} + \frac{(k-p+1)(2k-p+1)}{p+1} a_{k-p+1,\ell}\mathbf{1}_{\left\{\substack{p\le k\le k_0+p-1\\ 0\le \ell\le \ell_0}\right\}}. \tag{A.2}$$

Thus, the coefficient $a^{1*}_{k,\ell}$ depend on some $(a_{k',\ell'})$ only for k', ℓ' such that $(k',\ell') \prec (k,\ell)$ (more precisely, either $k'\le k$ and $\ell'\le \ell-1$ or $k'\le k-p+1$ and $\ell'\le \ell$).

Decomposition of β^2. By (A.1),

$$\beta^2(y_c) = \sum_{\substack{1\le k_1,k_2\le k_0\\ 0\le \ell_1,\ell_2\le \ell_0}} c^{\ell_1+\ell_2} Q_c^{k_1+k_2}(y_c) a_{k_1,\ell_1} a_{k_2,\ell_2} = \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell Q_c^k(y_c) a^{2*}_{k,\ell},$$

where

$$a^{2*}_{k,\ell} = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} a_{k_1,\ell_1} a_{k-k_1,\ell-\ell_1}. \tag{A.3}$$

Note that the expression of $a^{2*}_{k,\ell}$ above involves a_{k_1,ℓ_1} with $k_1\le k-1$ and $a_{k-k_1,\ell-\ell_1}$ with $k-k_1\le k-1$ since $k_1\ge 1$. Thus it is checked that $a_{k,\ell}$ does not appear in the expression of $a^{2*}_{k,\ell}$.

Decomposition of $\beta'(y_c)\beta(y_c)$.

$$\beta'(y_c)\beta(y_c) = \sum_{\substack{1\le k_1,k_2\le k_0\\ 0\le \ell_1,\ell_2\le \ell_0}} c^{\ell_1+\ell_2}(Q_c^{k_1+k_2})'(y_c)\frac{k_1}{k_1+k_2}a_{k_1,\ell_1}a_{k_2,\ell_2}$$
$$= \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^{\ell}(Q_c^{k})'(y_c)a^{3*}_{k,\ell},$$

where

$$a^{3*}_{k,\ell} = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \frac{k_1}{k}a_{k_1,\ell_1}a_{k-k_1,\ell-\ell_1}. \tag{A.4}$$

Decomposition of $\beta^3(y_c)$. By $\beta^3(y_c)=\beta(y_c)\beta^2(y_c)$ and the decomposition of β^2,

$$\beta^3(y_c) = \left(\sum_{\substack{1\le k_1\le k_0\\ 0\le \ell_1\le \ell_0}} c^{\ell_1}Q_c^{k_1}(y_c)a_{k_1,\ell_1}\right)\times\left(\sum_{\substack{2\le k_2\le 2k_0\\ 0\le \ell_2\le 2\ell_0}} c^{\ell_2}Q_c^{k_2}(y_c)a^{*}_{k_2,\ell_2}\right)$$
$$= \sum_{\substack{3\le k\le 3k_0\\ 0\le \ell\le 3\ell_0}} c^{\ell}Q_c^{k}(y_c)a^{4*}_{k,\ell},$$

where

$$a^{4*}_{k,\ell} = \sum_{\substack{\max(k-2k_0,1)\le k_1\le \min(k-2,k_0)\\ \max(\ell-2\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} a_{k_1,\ell_1}a^{2*}_{k-k_1,\ell-\ell_1}. \tag{A.5}$$

A.1. *Decomposition of* $\mathbf{I}=\partial_t R+\partial_x(\partial_x^2 R - R + R^p)$.

Lemma A.1 (Equation of $R(t)$).

$$\mathbf{I} = \sum_{(k,\ell)\in\Sigma_0} c^{\ell}\left(Q_c^k(y_c)a_{k,\ell}(-3Q+2Q^p)'(y) + (Q_c^k)'(y_c)a_{k,\ell}(-3Q'')(y)\right)$$
$$+ \sum_{\substack{1\le k\le \max(3k_0,k_0+p-1)\\ 0\le \ell\le \max(3\ell_0,\ell_0+1)}} c^{\ell}\left(Q_c^k(y_c)F^{\mathbf{I}}_{k,\ell}(y) + (Q_c^k)'(y_c)G^{\mathbf{I}}_{k,\ell}(y)\right), \tag{A.6}$$

where $F^{\mathbf{I}}_{k,\ell}$ *and* $G^{\mathbf{I}}_{k,\ell}$ *are functions defined on* $\mathbb{R}$ *satisfying*

(i) $F^{\mathbf{I}}_{k,\ell}$, $G^{\mathbf{I}}_{k,\ell}\in\mathcal{Y}$;

(ii) $F^{\mathbf{I}}_{k,\ell}$ *and* $G^{\mathbf{I}}_{k,\ell}$ *depend only on* $(a_{k',\ell'})$ *for* k', ℓ' *such that* $(k',\ell')\prec(k,\ell)$;

(iii) $F^{\mathbf{I}}_{k,\ell}$ *is odd and* $G^{\mathbf{I}}_{k,\ell}$ *is even.*

Moreover, $F^{\mathbf{I}}_{1,0}=0$, *and for all* $\ell\ge 0$, $G^{\mathbf{I}}_{1,\ell}=0$, *and*

- *if* $p=2$, *then* $F^{\mathbf{I}}_{2,0}=a_{1,0}Q'+3a_{1,0}^2Q^{(3)}$, $\quad G^{\mathbf{I}}_{2,0}=\frac{3}{2}a_{1,0}^2Q''$;
- *if* $p=4$, *then* $F^{\mathbf{I}}_{2,0}=3a_{1,0}^2Q^{(3)}$, $\quad G^{\mathbf{I}}_{2,0}=\frac{3}{2}a_{1,0}^2Q''$.

Claim A.2. *Let $h(t,x)=g(y)=g(x-\alpha(y_c))$, where g is a C^3 function. Then,*

$$\begin{aligned}
&\partial_t h(t,x) = -(1-c)\beta(y_c)g'(y), \quad \partial_x h(t,x) = (1-\beta(y_c))g'(y),\\
&\partial_x^2 h(t,x) = (1-2\beta(y_c)+\beta^2(y_c))g''(y) - \beta'(y_c)g'(y),\\
&\partial_x^3 h(t,x) = (1-3\beta(y_c)+3\beta^2(y_c)-\beta^3(y_c))g^{(3)}(y)\\
&\qquad\qquad + (-3\beta'(y_c)+3\beta'(y_c)\beta(y_c))g''(y) - \beta''(y_c)g'(y).
\end{aligned}$$

Proof of Claim A.2. Recall that $y_c = x+(1-c)t$ and $\alpha'(s)=\beta(s)$. Thus,

$$\partial_t h(t,x) = -\frac{\partial y_c}{\partial t}\alpha'(y_c)g'(y) = -(1-c)\beta(y_c)g'(y),$$

$$\partial_x h(t,x) = \left(1-\frac{\partial y_c}{\partial x}\alpha'(y_c)\right)g'(y) = (1-\beta(y_c))g'(y).$$

Next, $\partial_x^2 h(t,x) = (1-\beta(y_c))^2 g''(y) - \beta'(y_c)g'(y)$, and so

$$\begin{aligned}
\partial_x^3 h(t,x) &= -2\beta'(y_c)(1-\beta(y_c))g''(y) + (1-\beta(y_c))^3 g^{(3)}(y)\\
&\quad - \beta''(y_c)g'(y) - \beta'(y_c)(1-\beta(y_c))g''(y)\\
&= (1-\beta(y_c))^3 g^{(3)}(y) - 3\beta'(y_c)(1-\beta(y_c))g''(y) - \beta''(y_c)g'(y).
\end{aligned}$$

Proof of Lemma A.1.

Expression of $\mathbf{I}$. We claim

$$\begin{aligned}
\mathbf{I} &= \beta(y_c)(-3Q+2Q^p)'(y) + \beta'(y_c)(-3Q'')(y) + c\beta(y_c)Q'(y) + \beta''(y_c)(-Q')(y)\\
&\quad + \beta^2(y_c)(3Q^{(3)})(y) + \beta'(y_c)\beta(y_c)(3Q'')(y) + \beta^3(y_c)(-Q^{(3)})(y)\\
&= \mathbf{I}_1+\mathbf{I}_2+\mathbf{I}_3+\mathbf{I}_4+\mathbf{I}_5+\mathbf{I}_6+\mathbf{I}_7.
\end{aligned}$$

Indeed, since $R(t,x)=Q(y)$, by Claim A.2, we have

$$\begin{aligned}
&\partial_t R(t,x) = -(1-c)\beta(y_c)Q'(y),\\
&\partial_x^3 R(t,x) = (1-3\beta(y_c)+3\beta^2(y_c)-\beta^3(y_c))Q^{(3)}(y)\\
&\qquad\qquad + (-3\beta'(y_c)+3\beta'(y_c)\beta(y_c))Q''(y) - \beta''(y_c)Q'(y).\\
&-\partial_x R(t,x) = -(1-\beta(y_c))Q'(y), \quad \partial_x(R^p) = (1-\beta(y_c))(Q^p)'(y).
\end{aligned}$$

Thus, by arranging terms by increasing order of derivatives and powers of $\beta(y_c)$, we get

$$\begin{aligned}
\mathbf{I} &= \partial_t R + \partial_x(\partial_x^2 R - R + R^p)\\
&= (Q''-Q+Q^p)'(y) + \beta(y_c)(-3Q''-Q^p+cQ)'(y) + \beta'(y_c)(-3Q'')(y)\\
&\quad + \beta''(y_c)(-Q')(y) + \beta^2(y_c)(3Q^{(3)})(y)\\
&\quad + \beta'(y_c)\beta(y_c)(3Q'')(y) + \beta^3(y_c)(-Q^{(3)})(y).
\end{aligned}$$

By the equation of Q, i.e. $Q''-Q+Q^p=0$, the claim is proved.

Decomposition of $\mathbf{I}_1$ *and* $\mathbf{I}_2$. These two terms give (A.6):

$$\mathbf{I}_1 = \beta(y_c)(-3Q + 2Q^p)'(y) = \sum_{(k,\ell)\in\Sigma_0} c^\ell Q_c^k(y_c) a_{k,\ell}(-3Q + 2Q^p)'(y),$$

$$\mathbf{I}_2 = \beta'(y_c)(-3Q'')(y) = \sum_{(k,\ell)\in\Sigma_0} c^\ell (Q_c^k)'(y_c) a_{k,\ell}(-3Q'')(y).$$

Decomposition of $\mathbf{I}_3 = c\beta(y_c)Q'(y)$.

$$\mathbf{I}_3 = \sum_{(k,\ell)\in\Sigma_0} c^{\ell+1} Q_c^k(y_c) a_{k,\ell} Q'(y)$$
$$= \sum_{\substack{1\le k\le k_0\\ 1\le \ell\le \ell_0+1}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{I}_3}(y), \quad \text{where} \quad F_{k,\ell}^{\mathbf{I}_3} = a_{k,\ell-1}Q'.$$

Decomposition of $\mathbf{I}_4$, $\mathbf{I}_5$, $\mathbf{I}_6$ *and* $\mathbf{I}_7$. For these terms, we use Claim A.1:

$$\mathbf{I}_4 = \sum_{\substack{1\le k\le k_0+p-1\\ 0\le \ell\le \ell_0+1}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{I}_4}(y), \quad \text{where} \quad F_{k,\ell}^{\mathbf{I}_4}(y) = a_{k,\ell}^{1*}(-Q'(y)).$$

$$\mathbf{I}_5 = \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{I}_5}(y), \quad \text{where} \quad F_{k,\ell}^{\mathbf{I}_5}(y) = a_{k,l}^{2*}(3Q^{(3)})(y).$$

$$\mathbf{I}_6 = \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell (Q_c^k)'(y_c) G_{k,\ell}^{\mathbf{I}_6}(y), \quad \text{where} \quad G_{k,\ell}^{\mathbf{I}_6}(y) = a_{k,\ell}^{3*}(3Q'')(y).$$

$$\mathbf{I}_7 = \sum_{\substack{3\le k\le 3k_0\\ 0\le \ell\le 3\ell_0}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{I}_7}(y), \quad \text{where} \quad F_{k,\ell}^{\mathbf{I}_7}(y) = a_{k,\ell}^{4*}(-Q^{(3)})(y).$$

We check that $F_{k,\ell}^{\mathbf{I}_3}$, $F_{k,\ell}^{\mathbf{I}_4}$, $F_{k,\ell}^{\mathbf{I}_5}$, $G_{k,\ell}^{\mathbf{I}_6}$ and $F_{k,\ell}^{\mathbf{I}_7}$ satisfy properties (i), (ii) and (iii). Set $F_{k,\ell}^{\mathbf{I}} = F_{k,\ell}^{\mathbf{I}_3} + F_{k,\ell}^{\mathbf{I}_4} + F_{k,\ell}^{\mathbf{I}_5} + F_{k,\ell}^{\mathbf{I}_7}$ and $G_{k,\ell}^{\mathbf{I}} = G_{k,\ell}^{\mathbf{I}_6}$; they satisfy (i), (ii) and (iii).

To finish the proof of Lemma A.1, we compute $F_{1,0}^{\mathbf{I}}$, $G_{1,\ell}^{\mathbf{I}}$, $F_{2,0}^{\mathbf{I}}$ and $F_{2,0}^{\mathbf{I}}$.

$k = 1$: We check that $F_{1,0}^{\mathbf{I}_3} = 0$, $F_{1,0}^{\mathbf{I}_4} = a_{1,0}^{1*}(-Q') = 0$, $F_{1,0}^{\mathbf{I}_5} = F_{1,0}^{\mathbf{I}_7} = 0$, so that $F_{1,0}^{\mathbf{I}} = 0$. Moreover, for any $\ell \ge 0$, we have $G_{1,0}^{\mathbf{I}} = G_{1,0}^{\mathbf{I}_6} = 0$.

$k = 2$: We check $F_{2,0}^{\mathbf{I}_3} = 0$. The term $F_{2,0}^{\mathbf{I}_4} = a_{2,0}^{1*}(-Q')$ depends on the value of p. From (A.2), if $p = 2$ then $F_{2,0}^{\mathbf{I}_4} = a_{1,0}Q'$, and if $p = 3$ or 4, then $F_{2,0}^{\mathbf{I}_4} = 0$. By (A.3), we have $F_{2,0}^{\mathbf{I}_5} = 3a_{2,0}^{2*}Q^{(3)} = 3a_{1,0}^2 Q^{(3)}$ and by (A.5), $F_{2,0}^{\mathbf{I}_7} = -a_{2,0}^{4*}Q^{(3)} = 0$. Thus, if $p = 2$, we obtain $F_{2,0}^{\mathbf{I}} = a_{1,0}Q' + 3a_{1,0}^2 Q^{(3)}$, and if $p = 3$ or 4, we obtain $F_{2,0}^{\mathbf{I}} = 3a_{1,0}^2 Q^{(3)}$.

Similarly, $G_{2,0}^{\mathbf{I}} = G_{2,0}^{\mathbf{I}_6} = a_{2,0}^{3*}(3Q'') = \frac{3}{2}a_{1,0}^2 Q''$.

A.2. *Decomposition of* $\mathbf{II} = \partial_x((R+R_c)^p - R^p - R_c^p)$.

LEMMA A.2 (Interaction term between R and R_c).

$$\mathbf{II} = \sum_{\substack{1\le k\le k_0+p-1\\ 0\le \ell\le \ell_0}} c^\ell \left(Q_c^k(y_c) F^{\mathbf{II}}_{k,\ell}(y) + (Q_c^k)'(y_c) G^{\mathbf{II}}_{k,\ell}(y)\right), \tag{A.7}$$

where for any $k\ge 1$, $\ell \ge 0$, $F^{\mathbf{II}}_{k,\ell}$, $G^{\mathbf{II}}_{k,\ell}$ *satisfy properties* (i), (ii) *and* (iii) *as in Lemma* A.1. *Moreover,* $F_{1,0} = p(Q^{p-1})'$, $G_{1,0} = pQ^{p-1}$, $F^{\mathbf{II}}_{1,\ell} = G^{\mathbf{II}}_{1,\ell} = 0$, *for any* $\ell \ge 1$.

- *If* $p=2$, *then*

$$F^{\mathbf{II}}_{k,\ell} = -2a_{k-1,\ell}Q', \text{ for any } k\in\{2,k_0+1\},\ \ell\in\{0,\ell_0\}.$$

- *If* $p=4$, *then*

$$F^{\mathbf{II}}_{2,0} = (-4a_{1,0}Q^3 + 6Q^2)',\ G^{\mathbf{II}}_{2,0} = 6Q^2,\ G^{\mathbf{II}}_{2,\ell} = 0, \text{ for any } \ell\ge 1,$$
$$F^{\mathbf{II}}_{3,0} = (-4a_{2,0}Q^3 - 6a_{1,0}Q^2 + 4Q)',\ G^{\mathbf{II}}_{3,0} = 4Q,\ G^{\mathbf{II}}_{3,\ell} = 0, \text{ for any } \ell\ge 1.$$

Proof of Lemma A.2

- $p=2$. Recall that $R(t,x) = Q(y)$ and $R_c(t,x) = Q_c(y_c)$. By Claims A.2 and A.1, we have

$$\begin{aligned}
\mathbf{II} &= 2\partial_x(R\,R_c) = 2(1-\beta(y_c))Q'(y)Q_c(y_c) + 2Q(y)Q_c'(y_c)\\
&= Q_c(y_c)2Q'(y) + Q_c'(y_c)2Q(y) + Q_c(y_c)\beta(y_c)(-2Q'(y))\\
&= Q_c(y_c)2Q'(y) + Q_c'(y_c)2Q(y) + \sum_{\substack{2\le k\le k_0+1\\ 0\le\ell\le\ell_0}} c^\ell Q_c^k(y_c) a_{k-1,\ell}(-2Q')(y).
\end{aligned}$$

- $p=4$. As before,

$$\begin{aligned}
\mathbf{II} &= \partial_x(4R^3R_c + 6R^2R_c^2 + 4RR_c^3)\\
&= Q_c(y_c)(4Q^3)'(y) + Q_c'(y_c)(4Q^3)(y) + Q_c(y_c)\beta(y_c)(-4Q^3)'(y)\\
&\quad + Q_c^2(y_c)(6Q^2)'(y) + (Q_c^2)'(y_c)(6Q^2)(y) + Q_c^2(y_c)\beta(y_c)(-6Q^2)'(y)\\
&\quad + Q_c^3(y_c)(4Q)'(y) + (Q_c^3)'(y_c)(4Q)(y) + Q_c^3(y_c)\beta(y_c)(-4Q)'(y)\\
&= Q_c(y_c)(4Q^3)'(y) + Q_c'(y_c)(4Q^3)(y)\\
&\quad + Q_c^2(y_c)(-4a_{1,0}Q^3 + 6Q^2)'(y) + (Q_c^2)'(y_c)(6Q^2)(y)\\
&\quad + \sum_{1\le\ell\le\ell_0} c^\ell Q_c^2(y_c) a_{1,\ell}(-4Q^3)'(y)\\
&\quad + Q_c^3(y_c)(-4a_{2,0}Q^3 - 6a_{1,0}Q^2 + 4Q)'(y) + (Q_c^3)'(y_c)(4Q)(y)\\
&\quad + \sum_{1\le\ell\le\ell_0} c^\ell Q_c^3(y_c)(-4a_{2,\ell}Q^3 - 6a_{1,\ell}Q^2)'(y)
\end{aligned}$$

$$+\sum_{\substack{4\le k\le k_0+1\\0\le\ell\le\ell_0}} c^\ell Q_c^k(y_c)(-4a_{k-1,\ell}Q^3-6a_{k-2,\ell}Q^2-4a_{k-3,\ell}Q)'(y)$$
$$+\sum_{0\le\ell\le\ell_0} c^\ell\Big(Q_c^{k_0+2}(y_c)(-6a_{k_0,\ell}Q^2-4a_{k_0-1,\ell}Q)'(y) + Q_c^{k_0+3}(y_c)(-4a_{k_0,\ell}Q)'(y)\Big).$$

A.3. *Decomposition of* $\mathbf{III}=\partial_x W-\partial_x(\overline{\mathcal{L}}W)$.

Lemma A.3 (Linear terms in W).

$$\text{(A.8)}\qquad \mathbf{III}=\sum_{(k,\ell)\in\Sigma_0} c^\ell\Big(Q_c^k(y_c)(-\mathcal{L}A_{k,\ell})'(y) + (Q_c^k)'(y_c)\big(3A''_{k,\ell}+pQ^{p-1}A_{k,\ell}-(\mathcal{L}B_{k,\ell})'\big)(y)\Big)$$
$$+\sum_{\substack{1\le k\le 4k_0+2p-2\\0\le\ell\le 4\ell_0+2}} c^\ell\left(Q_c^k(y_c)F^{\mathbf{III}}_{k,\ell}(y)+(Q_c^k)'(y_c)G^{\mathbf{III}}_{k,\ell}(y)\right),$$

where for any $k\ge1$, $\ell\ge0$, $F^{\mathbf{III}}_{k,\ell}$ *and* $G^{\mathbf{III}}_{k,\ell}$ *satisfy*

(i) *Dependence property*: $F^{\mathbf{III}}_{k,\ell}$ *and* $G^{\mathbf{III}}_{k,\ell}$ *depend only on* $(a_{k',\ell'})$ *and* $(A_{k',\ell'})$, $(B_{k',\ell'})$ *for* k', ℓ' *such that* $(k',\ell')\prec(k,\ell)$.

(ii) *Parity property*: *Let* $k\in\{1,\dots,4k_0+2p-2\}$, $\ell\in\{0,\dots,4\ell_0+2\}$. *Assume that for any* $(k',\ell')\prec(k,\ell)$, $A_{k',\ell'}$ *is even and* $B_{k',\ell'}$ *is odd, then* $F^{\mathbf{III}}_{k,\ell}$ *is odd and* $G^{\mathbf{III}}_{k,\ell}$ *is even.*

Moreover, $F^{\mathbf{III}}_{1,0}=G^{\mathbf{III}}_{1,0}=0$.

- *If* $p=2$, *then*

$$F^{\mathbf{III}}_{2,0}=a_{1,0}(-3A''_{1,0}-2QA_{1,0})'-(3A'_{1,0}+3B''_{1,0}+2QB_{1,0}),$$
$$G^{\mathbf{III}}_{2,0}=\frac{a_{1,0}}{2}(-9A'_{1,0}-3B''_{1,0}-2QB_{1,0})'-(A_{1,0}+3B'_{1,0}).$$

- *If* $p=4$, *then*

$$F^{\mathbf{III}}_{2,0}=a_{1,0}(-3A''_{1,0}-pQ^{p-1}A_{1,0})',$$
$$G^{\mathbf{III}}_{2,0}=\frac{a_{1,0}}{2}(-9A'_{1,0}-3B''_{1,0}-pQ^{p-1}B_{1,0})'.$$

First, we claim two preliminary results concerning $\mathbf{III}$.

Claim A.3. *Let* $k\in\mathbb{N}$ *and let* $A(x)$ *be a class* C^3 *function. Let* $w(t,x)=Q_c^k(y_c)A(y)$. *Then*

$$\begin{aligned}
\partial_t w - \partial_x(\overline{\mathcal{L}}w) &= Q_c^k(y_c)(-\mathcal{L}A)'(y)\\
&+ Q_c^k(y_c)\beta(y_c)(-3A'' - pQ^{p-1}A + cA)'(y) + Q_c^k(y_c)\beta'(y_c)(-3A'')(y)\\
&+ Q_c^k(y_c)\beta''(y_c)(-A')(y) + Q_c^k(y_c)\beta^2(y_c)(3A^{(3)})(y)\\
&+ Q_c^k(y_c)\beta'(y_c)\beta(y_c)(3A'')(y) + Q_c^k(y_c)\beta^3(y_c)(-A^{(3)})(y)\\
&+ (Q_c^k)'(y_c)(3A'' + pQ^{p-1}A - cA)(y)\\
&+ (Q_c^k)'(y_c)\beta(y_c)(-6A'')(y)\\
&+ (Q_c^k)'(y_c)\beta'(y_c)(-3A')(y) + (Q_c^k)'(y_c)\beta^2(y_c)(3A'')(y)\\
&+ (Q_c^k)''(y_c)(3A')(y) + (Q_c^k)''(y_c)\beta(y_c)(-3A')(y) + (Q_c^k)^{(3)}(y_c)A(y).
\end{aligned}$$

Claim A.4. *Let $k \in \mathbb{N}$ and let $B(x)$ be a class C^3 function. Let $w(t,x) = (Q_c^k)'(y_c)B(y)$. Then*

$$\begin{aligned}
\partial_t w - \partial_x(\overline{\mathcal{L}}w) &= (Q_c^k)'(y_c)(-\mathcal{L}B)'(y)\\
&+ (Q_c^k)'(y_c)\beta(y_c)(-3B'' - pQ^{p-1}B + cB)'(y) + (Q_c^k)'(y_c)\beta'(y_c)(-3B'')(y)\\
&+ (Q_c^k)'(y_c)\beta''(y_c)(-B')(y) + (Q_c^k)'(y_c)\beta^2(y_c)(3B^{(3)})(y)\\
&+ (Q_c^k)'(y_c)\beta'(y_c)\beta(y_c)(3B'')(y) + (Q_c^k)'(y_c)\beta^3(y_c)(-B^{(3)})(y)\\
&+ (Q_c^k)''(y_c)(3B'' + pQ^{p-1}B - cB)(y)\\
&+ (Q_c^k)''(y_c)\beta(y_c)(-6B'')(y)\\
&+ (Q_c^k)''(y_c)\beta'(y_c)(-3B')(y) + (Q_c^k)''(y_c)\beta^2(y_c)(3B'')(y)\\
&+ (Q_c^k)^{(3)}(y_c)(3B')(y) + (Q_c^k)^{(3)}(y_c)\beta(y_c)(-3B')(y) + (Q_c^k)^{(4)}(y_c)B(y).
\end{aligned}$$

Proof of Claim A.3. Let $\mathcal{A}(t,x) = A(y) = A(x - \alpha(y_c))$, and $w(t,x) = Q_c^k(y_c)\mathcal{A}(t,x)$. We first give the expression of $\partial_t w - \partial_x(\overline{\mathcal{L}}w)$ in terms of the partial derivatives of $\mathcal{A}$. First,

$$\partial_t w = (1-c)(Q_c^k)'(y_c)\mathcal{A} + Q_c^k(y_c)\partial_t\mathcal{A}.$$

Since $\overline{\mathcal{L}}(fg) = g\overline{\mathcal{L}}f - 2f'g' - g''$, we have $\overline{\mathcal{L}}w = Q_c^k(y_c)(\overline{\mathcal{L}}\mathcal{A}) - 2(Q_c^k)'(y_c)\partial_x\mathcal{A} - (Q_c^k)''(y_c)\mathcal{A}$, and so

$$\begin{aligned}
-\partial_x(\overline{\mathcal{L}}w) = &-Q_c^k(y_c)\partial_x(\overline{\mathcal{L}}\mathcal{A}) - (Q_c^k)'(y_c)(\overline{\mathcal{L}}\mathcal{A}) + 2(Q_c^k)''(y_c)\partial_x\mathcal{A}\\
&+ 2(Q_c^k)'(y_c)\partial_x^2\mathcal{A} + (Q_c^k)^{(3)}(y_c)\mathcal{A} + (Q_c^k)''(y_c)\partial_x\mathcal{A}.
\end{aligned}$$

Thus, by arranging terms by increasing order of derivatives of $Q_c^k(y_c)$, we get

(A.9)

$$\begin{aligned}
\partial_t w - \partial_x(\overline{\mathcal{L}}w) = &\, Q_c^k(y_c)(\partial_t\mathcal{A} - \partial_x(\overline{\mathcal{L}}\mathcal{A})) + (Q_c^k)'(y_c)((1-c)\mathcal{A} - (\overline{\mathcal{L}}\mathcal{A}) + 2\partial_x^2\mathcal{A})\\
&+ (Q_c^k)''(y_c)(3\partial_x\mathcal{A}) + (Q_c^k)^{(3)}(y_c)\mathcal{A}.
\end{aligned}$$

Second, we use Claim A.2 to express the partial derivatives of $\mathcal{A}$ in terms of derivatives of A. We have

$$\begin{aligned}
\partial_t\mathcal{A}-\partial_x(\overline{\mathcal{L}}\mathcal{A}) &= -(1-c)\beta(y_c)A'(y)+(1-3\beta(y_c)+3\beta^2(y_c)-\beta^3(y_c))A^{(3)}(y)\\
&\quad+(-3\beta'(y_c)+3\beta'(y_c)\beta(y_c))A''(y)-\beta''(y_c)A'(y)\\
&\quad+(1-\beta(y_c))(-A+pQ^{p-1}A)'(y)\\
&=(1-3\beta(y_c)+3\beta^2(y_c)-\beta^3(y_c))A^{(3)}(y)\\
&\quad+(-3\beta'(y_c)+3\beta'(y_c)\beta(y_c))A''(y)\\
&\quad+(1-c\beta(y_c)+\beta''(y_c))(-A')(y)+(1-\beta(y_c))(pQ^{p-1}A)'(y).
\end{aligned}$$

Thus, by arranging terms by increasing order of derivatives and powers of $\beta(y_c)$, we get

$$\begin{aligned}
&\partial_t\mathcal{A}-\partial_x(\overline{\mathcal{L}}\mathcal{A})=(-\mathcal{L}A)'(y)+\beta(y_c)(-3A''-pQ^{p-1}A+cA)'(y)+\beta'(y_c)(-3A'')(y)\\
&\quad+\beta''(y_c)(-A')(y)+\beta^2(y_c)(3A^{(3)})(y)+\beta'(y_c)\beta(y_c)(3A'')(y)+\beta^3(y_c)(-A^{(3)})(y).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&(1-c)\mathcal{A}-(\overline{\mathcal{L}}\mathcal{A})+2\partial_x^2\mathcal{A}=-c\mathcal{A}+3\partial_x^2\mathcal{A}+pQ^{p-1}(y)\mathcal{A}\\
&=-cA(y)+3(1-2\beta(y_c)+\beta^2(y_c))A''(y)-3\beta'(y_c)A'(y)+pQ^{p-1}(y)A(y)\\
&=(3A''+pQ^{p-1}A-cA)(y)+\beta(y_c)(-6A'')(y)+\beta'(y_c)(-3A')(y)+\beta^2(y_c)(3A'')(y),
\end{aligned}$$

and

$$3\partial_x\mathcal{A}=3A'(y)-3\beta(y_c)A'(y).$$

Inserting all this into (A.9), we obtain Claim A.3. The proof of Claim A.4 is the same.

Proof of Lemma A.3. We recall $W(t,x)=\sum_{(k,\ell)\in\Sigma_0}c^\ell Q_c^k(y_c)A_{k,\ell}(y)+c^\ell(Q_c^k)'(y_c)B_{k,\ell}(y)$. To expand **III**, we use Claims A.3–A.4 on $W(t,x)$. We obtain $\mathbf{III}=\sum_{(k,\ell)\in\Sigma_0}c^\ell\ \mathrm{III}(k,\ell)$, where

$$\begin{aligned}
\mathrm{III}(k,\ell)&=Q_c^k(y_c)(-\mathcal{L}A_{k,\ell})'(y)+(Q_c^k)'(y_c)(3A''_{k,\ell}+pQ^{p-1}A_{k,\ell}-(\mathcal{L}B_{k,\ell})')(y)\\
&\quad+c(Q_c^k)'(y_c)(-A_{k,\ell})(y)\\
&\quad+\beta(y_c)Q_c^k(y_c)(-3A''_{k,\ell}-pQ^{p-1}A_{k,\ell})'(y)\\
&\quad+\beta(y_c)(Q_c^k)'(y_c)(-6A'_{k,\ell}-3B''_{k,\ell}-pQ^{p-1}B_{k,\ell})'(y)\\
&\quad+c\beta(y_c)Q_c^k(y_c)(A'_{k,\ell})(y)+c\beta(y_c)(Q_c^k)'(y_c)(B'_{k,\ell})(y)\\
&\quad+\beta'(y_c)Q_c^k(y_c)(-3A''_{k,\ell})(y)+\beta'(y_c)(Q_c^k)'(y_c)(-3A'_{k,\ell}-3B''_{k,\ell})(y)\\
&\quad+\beta''(y_c)Q_c^k(y_c)(-A'_{k,\ell})(y)+\beta''(y_c)(Q_c^k)'(y_c)(-B'_{k,\ell})(y)\\
&\quad+\beta^2(y_c)Q_c^k(y_c)(3A^{(3)}_{k,\ell})(y)+\beta^2(y_c)(Q_c^k)'(y_c)(3A''_{k,\ell}+3B^{(3)}_{k,\ell})(y)\\
&\quad+\beta'(y_c)\beta(y_c)Q_c^k(y_c)(3A''_{k,\ell})(y)+\beta'(y_c)\beta(y_c)(Q_c^k)'(y_c)(3B''_{k,\ell})(y)
\end{aligned}$$

$$
\begin{aligned}
&+\beta^3(y_c)Q_c^k(y_c)(-A_{k,\ell}^{(3)})(y)+\beta^3(y_c)(Q_c^k)'(y_c)(-B_{k,\ell}^{(3)})(y)\\
&+(Q_c^k)''(y_c)(3A_{k,\ell}'+3B_{k,\ell}''+pQ^{p-1}B_{k,\ell})(y)\\
&+(Q_c^k)^{(3)}(y_c)(A_{k,\ell}+3B_{k,\ell}')(y)+c(Q_c^k)''(y_c)(-B_{k,\ell})(y)\\
&+\beta(y_c)(Q_c^k)''(y_c)(-3A_{k,\ell}'-6B_{k,\ell}'')(y)+\beta(y_c)(Q_c^k)^{(3)}(y_c)(-3B_{k,\ell}')(y)\\
&+\beta'(y_c)(Q_c^k)''(y_c)(-3B_{k,\ell}')(y)\\
&+\beta^2(y_c)(Q_c^k)''(y_c)(3B_{k,\ell}'')(y)+(Q_c^k)^{(4)}(y_c)B_{k,\ell}(y)\\
=\ &\mathrm{III}_1+\mathrm{III}_2+\mathrm{III}_3+\mathrm{III}_4+\mathrm{III}_5+\mathrm{III}_6\\
&+\mathrm{III}_7+\mathrm{III}_8+\mathrm{III}_9+\mathrm{III}_{10}+\mathrm{III}_{11}+\mathrm{III}_{12}+\mathrm{III}_{13}.
\end{aligned}
$$

For $j\in\{1,\dots,13\}$, we denote $\mathbf{III}_j=\sum_{(k,\ell)\in\Sigma_0}\mathrm{III}_j$.

Decomposition of $\mathbf{III}_1$. This term gives (A.8):

$$
\begin{aligned}
\mathbf{III}_1=\sum_{(k,\ell)\in\Sigma_0} c^\ell\Big(&Q_c^k(y_c)(-\mathcal{L}A_{k,\ell})'(y)\\
&+(Q_c^k)'(y_c)(3A_{k,\ell}''+pQ^{p-1}A_{k,\ell}-(\mathcal{L}B_{k,\ell})')(y)\Big).
\end{aligned}
$$

For the other terms, by elementary calculations, we obtain

$$
\mathbf{III}_2=\sum_{\substack{1\le k\le k_0\\1\le\ell\le\ell_0+1}} c^\ell(Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_2}(y),\quad\text{where}\quad G_{k,\ell}^{\mathbf{III}_2}(y)=(-A_{k,\ell-1})(y);
$$

$$
\mathbf{III}_3=\sum_{\substack{2\le k\le 2k_0\\0\le\ell\le 2\ell_0}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_3}(y)+\sum_{\substack{2\le k\le 2k_0\\0\le\ell\le 2\ell_0}} c^\ell(Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_3}(y),
$$

where

(A.10)

$$
\begin{aligned}
F_{k,\ell}^{\mathbf{III}_3}(y)&=\sum_{\substack{\max(k-k_0,1)\le k_1\le\min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le\ell_1\le\min(\ell,\ell_0)}} a_{k_1,\ell_1}(-3A_{k-k_1,\ell-\ell_1}''-pQ^{p-1}A_{k-k_1,\ell-\ell_1})'(y),\\
G_{k,\ell}^{\mathbf{III}_3}(y)&=\sum_{\substack{\max(k-k_0,1)\le k_1\le\min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le\ell_1\le\min(\ell,\ell_0)}} a_{k_1,\ell_1}\frac{k-k_1}{k}\\
&\qquad\times(-6A_{k-k_1,\ell-\ell_1}'-3B_{k-k_1,\ell-\ell_1}''-pQ^{p-1}B_{k-k_1,\ell-\ell_1})'(y).
\end{aligned}
$$

From (A.10), we easily check property (i) since in the sum defining $F_{k,\ell}^{\mathbf{III}_3}$, we have $k_1\le k-1$ and $k-k_1\le k-1$; moreover, $0\le\ell_1\le\ell$ and $\ell-\ell_1\le\ell$. The parity statement (ii) is also easily checked, as in the rest of this proof. Thus $F_{k,\ell}^{\mathbf{III}_3}$ satisfies properties (i) and (ii).

$$
\mathbf{III}_4=\sum_{\substack{2\le k\le 2k_0\\1\le\ell\le 2\ell_0+1}} c^\ell\left(Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_4}(y)+(Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_4}(y)\right),
$$

where

$$F^{\mathbf{III}_4}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,\ell_0)}} a_{k_1,\ell_1} A'_{k-k_1,\ell-\ell_1-1}(y),$$

$$G^{\mathbf{III}_4}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,\ell_0)}} a_{k_1,\ell_1} \frac{k-k_1}{k} B'_{k-k_1,\ell-\ell_1-1}(y);$$

$$\mathbf{III}_5 = \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell (Q_c^k)'(y_c) G^{\mathbf{III}_5}_{k,\ell}(y) + \sum_{\substack{2\le k\le 2k_0+p-1\\ 0\le \ell\le 2\ell_0+1}} c^\ell Q_c^k(y_c) F^{\mathbf{III}_5}_{k,\ell}(y),$$

where

$$G^{\mathbf{III}_5}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \frac{k_1}{k} a_{k_1,\ell_1}(-3A''_{k-k_1,\ell-\ell_1})(y),$$

$$\begin{aligned} F^{\mathbf{III}_5}_{k,\ell}(y) = &\sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,\ell_0)}} k_1(k-k_1) a_{k_1,\ell_1} \\ &\qquad \times (-3A'_{k-k_1,\ell-\ell_1-1} - 3B''_{k-k_1,\ell-\ell_1-1})(y) \\ &+ \sum_{\substack{\max(k-k_0-p+1,1)\le k_1\le \min(k-p,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \frac{2k_1(k-k_1-p+1)}{p+1} a_{k_1,\ell_1} \\ &\qquad \times (3A'_{k-k_1-p+1,\ell-\ell_1} + 3B''_{k-k_1-p+1,\ell-\ell_1})(y) \end{aligned} \tag{A.11}$$

satisfy properties (i) and (ii). In (A.11), the first sum term has no contribution for k such that $\max(k-k_0,1) > \min(k-1,k_0)$ (i.e. $k=1$ or $k>2k_0$) and similarly for the condition on ℓ. We will use this notation in all sums appearing in this proof.

For the next terms, we use Claim A.1:

$$\mathbf{III}_6 = \sum_{\substack{2\le k\le 2k_0+p-1\\ 0\le \ell\le 2\ell_0+1}} c^\ell \left(Q_c^k(y_c) F^{\mathbf{III}_6}_{k,\ell}(y) + (Q_c^k)'(y_c) G^{\mathbf{III}_6}_{k,\ell}(y)\right),$$

where

$$F^{\mathbf{III}_6}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0+p-1)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0+1)}} a^{1*}_{k_1,\ell_1}(-A'_{k-k_1,\ell-\ell_1})(y)$$

$$G^{\mathbf{III}_6}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0+p-1)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0+1)}} \frac{k-k_1}{k} a^{1*}_{k_1,\ell_1}(-B'_{k-k_1,\ell-\ell_1})(y);$$

$$\mathbf{III}_7 = \sum_{\substack{3\le k\le 3k_0\\ 0\le \ell\le 3\ell_0}} c^\ell \left(Q_c^k(y_c) F^{\mathbf{III}_7}_{k,\ell}(y) + (Q_c^k)'(y_c) G^{\mathbf{III}_7}_{k,\ell}(y)\right),$$

where

$$F^{\mathbf{III}_7}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,2)\le k_1\le \min(k-1,2k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,2\ell_0)}} a^{2*}_{k_1,\ell_1}(3A^{(3)}_{k-k_1,\ell-\ell_1})(y),$$

$$G^{\mathbf{III}_7}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,2)\le k_1\le \min(k-1,2k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,2\ell_0)}} \frac{k-k_1}{k} a^{2*}_{k_1,\ell_1}(3A''_{k-k_1,\ell-\ell_1}+3B^{(3)}_{k-k_1,\ell-\ell_1})(y);$$

$$\mathbf{III}_8 = \sum_{\substack{3\le k\le 3k_0+p-1\\ 0\le \ell\le 3\ell_0+1}} c^\ell \left(Q^k_c(y_c) F^{\mathbf{III}_8}_{k,\ell}(y) + (Q^k_c)'(y_c) G^{\mathbf{III}_8}_{k,\ell}(y)\right),$$

where

$$G^{\mathbf{III}_8}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,2)\le k_1\le \min(k-1,2k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,2\ell_0)}} \frac{k_1}{k} a^{3*}_{k_1,\ell_1}(3A''_{k-k_1,\ell-\ell_1}(y)),$$

$$\begin{aligned} F^{\mathbf{III}_8}_{k,\ell}(y) = & \sum_{\substack{\max(k-k_0,2)\le k_1\le \min(k-1,2k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,2\ell_0)}} k_1(k-k_1) a^{3*}_{k_1,\ell_1}(3B''_{k-k_1,\ell-\ell_1-1}(y)) \\ & + \sum_{\substack{\max(k-k_0-p+1,2)\le k_1\le \min(k-p,2k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,2\ell_0)}} \\ & - \frac{2k_1(k-k_1-p+1)}{p+1} a^{3*}_{k_1,\ell_1}(3B''_{k-k_1-p+1,\ell-\ell_1}(y)); \end{aligned}$$

$$\mathbf{III}_9 = \sum_{\substack{4\le k\le 4k_0\\ 0\le \ell\le 4\ell_0}} c^\ell \left(Q^k_c(y_c) F^{\mathbf{III}_9}_{k,\ell}(y) + (Q^k_c)'(y_c) G^{\mathbf{III}_9}_{k,\ell}(y)\right),$$

where

$$F^{\mathbf{III}_9}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,3)\le k_1\le \min(k-1,3k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,3\ell_0)}} a^{4*}_{k_1,\ell_1}(-A^{(3)}_{k-k_1,\ell-\ell_1})(y)$$

$$G^{\mathbf{III}_9}_{k,\ell}(y) = \sum_{\substack{\max(k-k_0,3)\le k_1\le \min(k-1,3k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,3\ell_0)}} \frac{k-k_1}{k} a^{4*}_{k_1,\ell_1}(-B^{(3)}_{k-k_1,\ell-\ell_1})(y).$$

Using the expression of $(Q^k_c)''$ from Lemma 2.1,

$$\begin{aligned} \mathbf{III}_{10} = & \sum_{\substack{1\le k\le k_0\\ 1\le \ell\le \ell_0+1}} \left(c^\ell Q^k_c(y_c) F^{\mathbf{III}^1_{10}}_{k,\ell}(y) + c^\ell (Q^k_c)'(y_c) G^{\mathbf{III}^1_{10}}_{k,\ell}(y)\right) \\ & + \sum_{\substack{p\le k\le k_0+p-1\\ 0\le \ell\le \ell_0}} \left(c^\ell Q^k_c(y_c) F^{\mathbf{III}^2_{10}}_{k,\ell}(y) + c^\ell (Q^k_c)'(y_c) G^{\mathbf{III}^2_{10}}_{k,\ell}(y)\right), \end{aligned}$$

where

$$F_{k,\ell}^{\mathbf{III}_{10}^1}(y) = k^2(3A'_{k,\ell-1} + 3B''_{k,\ell-1} + pQ^{p-1}B_{k,\ell-1})(y),$$

$$F_{k,\ell}^{\mathbf{III}_{10}^2}(y) = -\frac{(k-p+1)(2k-p+1)}{p+1} \times (3A'_{k-p+1,\ell} + 3B''_{k-p+1,\ell} + pQ^{p-1}B_{k-p+1,\ell})(y),$$

$$G_{k,\ell}^{\mathbf{III}_{10}^1}(y) = k^2(A_{k,\ell-1} + 3B'_{k,\ell-1})(y),$$

$$G_{k,\ell}^{\mathbf{III}_{10}^2}(y) = -\frac{(k-p+1)(2k-p+1)}{p+1}(A_{k-p+1,\ell} + 3B'_{k-p+1,\ell})(y).$$

We set $F_{k,\ell}^{\mathbf{III}_{10}} = F_{k,\ell}^{\mathbf{III}_{10}^1} + F_{k,\ell}^{\mathbf{III}_{10}^2}$, $G_{k,\ell}^{\mathbf{III}_{10}} = G_{k,\ell}^{\mathbf{III}_{10}^1} + G_{k,\ell}^{\mathbf{III}_{10}^2}$.

$$\mathbf{III}_{11} = \sum_{\substack{1\le k\le k_0\\ 2\le \ell\le \ell_0+2}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{11}^1}(y) + \sum_{\substack{p\le k\le k_0+p-1\\ 1\le\ell\le\ell_0+1}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{11}^2}(y),$$

where

$$F_{k,\ell}^{\mathbf{III}_{11}^1}(y) = k^2(-B_{k,\ell-2})(y),$$

$$F_{k,\ell}^{\mathbf{III}_{11}^2}(y) = \frac{(k-p+1)(2k-p+1)}{p+1}B_{k-p+1,\ell-1}(y).$$

We set $F_{k,\ell}^{\mathbf{III}_{11}} = F_{k,\ell}^{\mathbf{III}_{11}^1} + F_{k,\ell}^{\mathbf{III}_{11}^2}$.

$$\mathbf{III}_{12} = \sum_{\substack{2\le k\le 2k_0\\ 1\le\ell\le 2\ell_0+1}} \Big(c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{12}^1}(y) + c^\ell (Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_{12}^1}(y)\Big) + \sum_{\substack{p+1\le k\le 2k_0+p-1\\ 0\le\ell\le 2\ell_0}} \Big(c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{12}^2}(y) + c^\ell (Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_{12}^2}(y)\Big),$$

where

$$F_{k,\ell}^{\mathbf{III}_{12}^1}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le\min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le\ell_1\le\min(\ell-1,\ell_0)}} a_{k_1,\ell_1}(k-k_1)^2 \times(-3A'_{k-k_1,\ell-\ell_1-1} - 6B''_{k-k_1,\ell-\ell_1-1})(y),$$

$$F_{k,\ell}^{\mathbf{III}_{12}^2}(y) = \sum_{\substack{\max(k-k_0-p+1,1)\le k_1\le\min(k-p,k_0)\\ \max(\ell-\ell_0,0)\le\ell_1\le\min(\ell,\ell_0)}} a_{k_1,\ell_1}\frac{(k-k_1-p+1)(2k-2k_1-p+1)}{p+1} \times(3A'_{k-k_1-p+1,\ell-\ell_1} + 6B''_{k-k_1-p+1,\ell-\ell_1})(y),$$

$$G_{k,\ell}^{\mathbf{III}_{12}^1}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le\min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le\ell_1\le\min(\ell-1,\ell_0)}} \frac{(k-k_1)^3}{k}a_{k_1,\ell_1}(-3B'_{k-k_1,\ell-\ell_1-1})(y),$$

$$G_{k,\ell}^{\mathbf{III}_{12}^2}(y) = \sum_{\substack{\max(k-k_0-p+1,1)\le k_1\le \min(k-p,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \frac{(k-k_1-p+1)^2(2k-2k_1-p+1)}{k(p+1)} a_{k_1,\ell_1} \times 3B'_{k-k_1-p+1,\ell-\ell_1}(y).$$

We set $F_{k,\ell}^{\mathbf{III}_{12}} = F_{k,\ell}^{\mathbf{III}_{12}^1} + F_{k,\ell}^{\mathbf{III}_{12}^2}$, $G_{k,\ell}^{\mathbf{III}_{12}} = G_{k,\ell}^{\mathbf{III}_{12}^1} + G_{k,\ell}^{\mathbf{III}_{12}^2}$.

The last term $\mathbf{III}_{13}$ is the sum of three different terms.

The contribution of $\beta'(y_c)(Q_c^k)''(y_c)(-3B'_{k,\ell}(y))$ is

$$\sum_{\substack{2\le k\le 2k_0\\ 1\le \ell\le 2\ell_0+1}} c^\ell (Q_c^k)'(y_c) G_{k,\ell}^{\mathbf{III}_{13}^1}(y) + \sum_{\substack{p+1\le k\le 2k_0+p-1\\ 0\le \ell\le 2\ell_0}} c^\ell (Q_c^k)'(y_c) G_{k,\ell}^{\mathbf{III}_{13}^2}(y),$$

where

$$G_{k,\ell}^{\mathbf{III}_{13}^1}(y) = \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,\ell_0)}} \frac{k_1(k-k_1)^2}{k} a_{k_1,\ell_1}(-3B'_{k-k_1,\ell-\ell_1-1})(y)$$

and

$$G_{k,\ell}^{\mathbf{III}_{13}^2}(y) = \sum_{\substack{\max(k-k_0-p+1,1)\le k_1\le \min(k-p,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \frac{k_1(k-k_1-p+1)(2k-2k_1-p+1)}{k(p+1)} a_{k_1,\ell_1} \times (3B'_{k-k_1-p+1,\ell-\ell_1})(y).$$

Using (A.3), the contribution of $\beta^2(y_c)(Q_c^k)''(y_c)(3B''_{k,\ell}(y))$ is

$$\sum_{\substack{3\le k\le 3k_0\\ 1\le \ell\le 3\ell_0+1}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{III}_{13}^1}(y) + \sum_{\substack{p+2\le k\le 3k_0+p-1\\ 0\le \ell\le 3\ell_0}} c^\ell Q_c^k(y_c) F_{k,\ell}^{\mathbf{III}_{13}^2}(y),$$

where

$$F_{k,\ell}^{\mathbf{III}_{13}^1}(y) = \sum_{\substack{\max(k-k_0,2)\le k_1\le \min(k-1,2k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,2\ell_0)}} (k-k_1)^2 a^{2*}_{k_1,\ell_1}(3B''_{k-k_1,\ell-\ell_1-1})(y),$$

$$F_{k,\ell}^{\mathbf{III}_{13}^2}(y) = \sum_{\substack{\max(k-k_0-p+1,2)\le k_1\le \min(k-p,2k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,2\ell_0)}} \frac{(k-k_1-p+1)(2k-2k_1-p+1)}{p+1} \times a^{2*}_{k_1,\ell_1}(-3B''_{k-k_1-p+1,\ell-\ell_1})(y).$$

From Lemma 2.1, the contribution of $(Q_c^k)^{(4)}(y_c)B(y)$ is

$$\sum_{\substack{1\le k\le k_0\\ 2\le \ell\le \ell_0+2}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{13}^3}(y)+\sum_{\substack{p\le k\le k_0+p-1\\ 1\le \ell\le \ell_0+1}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{13}^4}(y)+\sum_{\substack{2p-1\le k\le k_0+2p-2\\ 0\le \ell\le \ell_0}} c^\ell Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{13}^5}(y),$$

where $F_{k,\ell}^{\mathbf{III}_{13}^3}(y)=k^4B_{k,\ell-2}(y)$,

$$F_{k,\ell}^{\mathbf{III}_{13}^4}(y)=-\frac{(k-p+1)(2k-p+1)}{p+1}((k-p+1)^2+k^2)B_{k-p+1,\ell-1}(y),$$

$$F_{k,\ell}^{\mathbf{III}_{13}^5}(y)=(k-2p+2)(k-p+1)\frac{(2k-3p+3)(2k-p+1)}{(p+1)^2}B_{k-2p+2,\ell}(y).$$

We set $F_{k,\ell}^{\mathbf{III}_{13}}=F_{k,\ell}^{\mathbf{III}_{13}^1}+F_{k,\ell}^{\mathbf{III}_{13}^2}+F_{k,\ell}^{\mathbf{III}_{13}^3}+F_{k,\ell}^{\mathbf{III}_{13}^4}+F_{k,\ell}^{\mathbf{III}_{13}^5}$, $G_{k,\ell}^{\mathbf{III}_{13}}=G_{k,\ell}^{\mathbf{III}_{13}^1}+G_{k,\ell}^{\mathbf{III}_{13}^2}$, so that

$$\mathbf{III}_{13}=\sum_{\substack{1\le k\le \max(3k_0+p-1,k_0+2p-2)\\ 0\le \ell\le \max(3\ell_0+1,\ell_0+2)}} c^\ell\left(Q_c^k(y_c)F_{k,\ell}^{\mathbf{III}_{13}}(y)+(Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{III}_{13}}(y)\right).$$

Finally, we set $F_{k,\ell}^{\mathbf{III}}=\sum_{j=3}^{13}F_{k,\ell}^{\mathbf{III}_j}$, $G_{k,\ell}^{\mathbf{III}}=\sum_{j=2}^{13}G_{k,\ell}^{\mathbf{III}_j}$.

We now finish the proof of Lemma A.3 by computing explicitely $F_{1,0}^{\mathbf{III}}$, $G_{1,0}^{\mathbf{III}}$, $F_{2,0}^{\mathbf{III}}$ and $G_{2,0}^{\mathbf{III}}$. We first check $F_{1,0}^{\mathbf{III}}=G_{1,0}^{\mathbf{III}}=0$. For $F_{2,0}^{\mathbf{III}}$, we make the following observations:

- $F_{2,0}^{\mathbf{III}_3}=a_{1,0}(-3A_{1,0}''-pQ^{p-1}A_{1,0})'$; $F_{2,0}^{\mathbf{III}_4}=F_{2,0}^{\mathbf{III}_5}=0$; $F_{2,0}^{\mathbf{III}_6}=0$ since $a_{1,0}^{1*}=0$;
- $F_{2,0}^{\mathbf{III}_7}=F_{2,0}^{\mathbf{III}_8}=F_{2,0}^{\mathbf{III}_9}=F_{2,0}^{\mathbf{III}_{10}^1}=0$;
- For $p=2$, we have $F_{2,0}^{\mathbf{III}_{10}^2}=-(3A_{1,0}'+3B_{1,0}''+2QB_{1,0})$, for $p=4$, we have $F_{2,0}^{\mathbf{III}_{10}^2}=0$.
- All the remaining terms in $F_{2,0}^{\mathbf{III}}$ are checked to be zero.

Similarly, we check that the only nonzero contributions to $G_{2,0}^{\mathbf{III}}$ are

- $G_{2,0}^{\mathbf{III}_3}=\frac{1}{2}a_{1,0}(-6A_{1,0}'-3B_{1,0}''-pQ^{p-1}B_{1,0})'$; $G_{2,0}^{\mathbf{III}_5}=\frac{1}{2}a_{1,0}(-3A_{1,0}')$;
- if $p=2$, $G_{2,0}^{\mathbf{III}_9}=-(A_{1,0}+3B_{1,0}')$, and if $p=4$, $G_{2,0}^{\mathbf{III}_9}=0$.

Thus, summing up, Lemma A.3 is proved.

A.4. *Expansion of* $\mathbf{IV}=\partial_x\left((R+R_c+W)^p-(R+R_c)^p-pR^{p-1}W\right)$.

Lemma A.4 (Nonlinear terms in W).

$$\mathbf{IV}=\sum_{\substack{2\le k\le (p+1)k_0+12\\ 0\le \ell\le (p+1)\ell_0+4}} c^\ell\left(Q_c^k(y_c)F_{k,\ell}^{\mathbf{IV}}(y)+(Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{IV}}(y)\right),$$

where $F^{\mathbf{IV}}_{k,\ell}$ *and* $G^{\mathbf{IV}}_{k,\ell}$ *are functions defined on* $\mathbb{R}$ *satisfying*

(i) *Dependence property*: $F^{\mathbf{IV}}_{k,\ell}$ *and* $G^{\mathbf{IV}}_{k,\ell}$ *depend only on* $(a_{k',\ell'})$ *and* $(A_{k',\ell'})$, $(B_{k',\ell'})$ *for* k', ℓ' *such that* $(k',\ell') \prec (k,\ell)$.

(ii) *Parity property*: *Let* $k \in \{1,\dots,(p+1)k_0+12\}$, $\ell \in \{0,\dots,(p+1)\ell_0+4\}$. *Assume that for any* (k',ℓ') *such that* $(k',\ell') \prec (k,\ell)$, $A_{k',\ell'}$ *is even and* $B_{k',\ell'}$ *is odd, then* $F^{\mathbf{IV}}_{k,\ell}$ *is odd and* $G^{\mathbf{IV}}_{k,\ell}$ *is even.*

Moreover,

- *If* $p=2$, *then*

$$F^{\mathbf{IV}}_{2,0} = \left(2A_{1,0} + A_{1,0}^2\right)', \quad G^{\mathbf{IV}}_{2,0} = 2A_{1,0} + A_{1,0}^2 + \left(B_{1,0} + A_{1,0}B_{1,0}\right)'. \tag{A.12}$$

- *If* $p=4$, *then*

$$F^{\mathbf{IV}}_{2,0} = \left(12A_{1,0}Q^2 + 6A_{1,0}^2Q^2\right)', \tag{A.13}$$

$$G^{\mathbf{IV}}_{2,0} = 12A_{1,0}Q^2 + 6A_{1,0}^2Q^2 + \left(6B_{1,0}Q^2 + 6A_{1,0}B_{1,0}Q^2\right)'. \tag{A.14}$$

Proof of Lemma A.4. Set $\mathbf{N} = (R + R_c + W)^p - (R+R_c)^p - pR^{p-1}W$. First, we determine $F^{\mathbf{N}}_{k,\ell}$ and $G^{\mathbf{N}}_{k,\ell}$ such that

$$\mathbf{N} = \sum_{k\geq p,\ \ell\geq 0} c^\ell \left(Q_c^k(y_c)F^{\mathbf{N}}_{k,\ell}(y) + (Q_c^k)'(y_c)G^{\mathbf{N}}_{k,\ell}(y)\right). \tag{A.15}$$

Second, we differentiate formula (A.15) with respect to x to get the decomposition of **IV**. We treat only the case $p=4$, the case $p=2$ is similar and easier.

- $p=4$:

$$\begin{aligned}\mathbf{N} &= 4\left((R+R_c)^3 - R^3\right)W + 6(R+R_c)^2W^2 + 4(R+R_c)W^3 + W^4\\ &= 12R^2R_cW + 12RR_c^2W + 4R_c^3W + 6R^2W^2 + 12RR_cW^2 + 6R_c^2W^2\\ &\quad + 4RW^3 + 4R_cW^3 + W^4\\ &= \mathbf{N}_1 + \mathbf{N}_2 + \mathbf{N}_3 + \mathbf{N}_4 + \mathbf{N}_5 + \mathbf{N}_6 + \mathbf{N}_7 + \mathbf{N}_8 + \mathbf{N}_9.\end{aligned}$$

Terms $\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{N}_3$:

$$\begin{aligned}\mathbf{N}_1 &= Q_c(y_c)(12WQ^2(y))\\ &= \sum_{\substack{2\leq k\leq k_0+1\\ 0\leq\ell\leq\ell_0}} c^\ell \left(Q_c^k(y_c)(12A_{k-1,\ell}(y)Q^2(y)) + (Q_c^k)'(y_c)\frac{k-1}{k}\left(12B_{k-1,\ell}Q^2\right)(y)\right);\\ \mathbf{N}_2 &= Q_c^2(y_c)(12WQ(y))\\ &= \sum_{\substack{3\leq k\leq k_0+2\\ 0\leq\ell\leq\ell_0}} c^\ell \left(Q_c^k(y_c)(12A_{k-2,\ell}(y)Q(y)) + (Q_c^k)'(y_c)\frac{k-2}{k}\left(12B_{k-2,\ell}Q\right)(y)\right);\end{aligned}$$

$$\begin{aligned}\mathbf{N}_3 &= Q_c^3(y_c)(4W)\\ &= \sum_{\substack{4\le k\le k_0+3\\ 0\le \ell\le \ell_0}} c^\ell \left(Q_c^k(y_c)(4A_{k-3,\ell}(y)) + (Q_c^k)'(y_c)\frac{k-3}{k}\,(4B_{k-3,\ell})\,(y)\right).\end{aligned}$$

For the next three terms, we first need to expand W^2:

$$\begin{aligned}W^2 = &\sum_{\substack{1\le k_1\le k_0\\ 0\le \ell_1\le \ell_0}} c^{\ell_1} \left(Q_c^{k_1}(y_c)A_{k_1,\ell_1}(y) + (Q_c^{k_1})'(y_c)B_{k_1,\ell_1}(y)\right)\\ &\times \sum_{\substack{1\le k_2\le k_0\\ 0\le \ell_2\le \ell_0}} c^{\ell_2} \left(Q_c^{k_2}(y_c)A_{k_2,\ell_2}(y) + (Q_c^{k_2})'(y_c)B_{k_2,\ell_2}(y)\right).\end{aligned}$$

Using Lemma 2.1,

$$\begin{aligned}W^2 = &\sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell Q_c^k(y_c) \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} A_{k_1,\ell_1}(y)A_{k-k_1,\ell-\ell_1}(y)\\ &+ \sum_{\substack{2\le k\le 2k_0\\ 1\le \ell\le 2\ell_0+1}} c^\ell Q_c^k(y_c) \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0-1,0)\le \ell_1\le \min(\ell-1,\ell_0)}} (k_1(k-k_1)B_{k_1,\ell_1}B_{k-k_1,\ell-\ell_1-1})(y)\\ &+ \sum_{\substack{5\le k\le 2k_0+3\\ 0\le \ell\le 2\ell_0}} c^\ell Q_c^k(y_c)\\ &\quad\times \sum_{\substack{\max(k-k_0-3,1)\le k_1\le \min(k-4,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \left(-\tfrac{2}{5}k_1(k-k_1-3)B_{k_1,\ell_1}B_{k-k_1-3,\ell-\ell_1}\right)(y)\\ &+ \sum_{\substack{2\le k\le 2k_0\\ 0\le \ell\le 2\ell_0}} c^\ell (Q_c^k)'(y_c) \sum_{\substack{\max(k-k_0,1)\le k_1\le \min(k-1,k_0)\\ \max(\ell-\ell_0,0)\le \ell_1\le \min(\ell,\ell_0)}} \tfrac{2(k-k_1)}{k}A_{k_1,\ell_1}(y)B_{k-k_1,\ell-\ell_1}(y).\end{aligned}$$

Therefore,

$$W^2 = \sum_{\substack{2\le k\le 2k_0+3\\ 0\le \ell\le 2\ell_0+1}} c^\ell \left(Q_c^k(y_c)A^*_{k,\ell}(y) + (Q_c^k)'(y_c)B^*_{k,\ell}(y)\right), \tag{A.16}$$

where $A^*_{k,\ell}$ and $B^*_{k,\ell}$ can be extracted from the previous formula.

Terms $\mathbf{N}_4$, $\mathbf{N}_5$ *and* $\mathbf{N}_6$:

$$\begin{aligned}\mathbf{N}_4 &= 6W^2Q^2(y)\\ &= \sum_{\substack{2\le k\le 2k_0+3\\ 0\le \ell\le 2\ell_0+1}} c^\ell \left(Q_c^k(y_c)(6A^*_{k,\ell}Q^2)(y) + (Q_c^k)'(y_c)(6B^*_{k,\ell}Q^2)(y)\right);\\ \mathbf{N}_5 &= Q_c(y_c)W^2(12Q(y))\\ &= \sum_{\substack{3\le k\le 2k_0+4\\ 0\le \ell\le 2\ell_0+1}} c^\ell \left(Q_c^k(y_c)(12A^*_{k-1,\ell}Q)(y) + (Q_c^k)'(y_c)\frac{k-1}{k}\left(12B^*_{k-1,\ell}Q\right)(y)\right);\end{aligned}$$

$$\begin{aligned}\mathbf{N}_6 &= Q_c^2(y_c)(6W^2)\\ &= \sum_{\substack{4\le k\le 2k_0+5\\ 0\le \ell\le 2\ell_0+1}} c^\ell \left(Q_c^k(y_c)(4A^*_{k-2,\ell}(y)) + (Q_c^k)'(y_c)\frac{k-2}{k}\left(6B^*_{k-2,\ell}\right)(y)\right).\end{aligned}$$

For the next two terms, we expand $W^3 = WW^2$ using (A.16). We get

$$W^3 = \sum_{\substack{3\le k\le 3k_0+6\\ 0\le \ell\le 3\ell_0+2}} c^\ell \left(Q_c^k(y_c)A^{**}_{k,\ell}(y) + (Q_c^k)'(y_c)B^{**}_{k,\ell}(y)\right), \tag{A.17}$$

where $A^{**}_{k,\ell}$ and $B^{**}_{k,\ell}$ are explicit in terms of $A_{k,\ell}$ and $B_{k,\ell}$, $A^*_{k,\ell}$ and $B^*_{k,\ell}$.

Terms $\mathbf{N}_7$ *and* $\mathbf{N}_8$:

$$\mathbf{N}_7 = 4W^3Q(y) = \sum_{\substack{3\le k\le 3k_0+6\\ 0\le \ell\le 3\ell_0+2}} c^\ell \left(Q_c^k(y_c)(4QA^{**}_{k,\ell})(y) + (Q_c^k)'(y_c)(4QB^{**}_{k,\ell})(y)\right);$$

$$\begin{aligned}\mathbf{N}_8 &= 4Q_c(y_c)W^3(y)\\ &= \sum_{\substack{4\le k\le 3k_0+7\\ 0\le \ell\le 3\ell_0+2}} c^\ell \left(Q_c^k(y_c)(4A^{**}_{k-1,\ell})(y) + (Q_c^k)'(y_c)(4B^{**}_{k-1,\ell})(y)\right).\end{aligned}$$

Term $\mathbf{N}_9 = W^4$: By using $W^4 = W^2W^2$ and (A.16), we get

$$\mathbf{N}_9 = \sum_{\substack{4\le k\le 4k_0+9\\ 0\le \ell\le 4\ell_0+3}} c^\ell \left(Q_c^k(y_c)A^{***}_{k,\ell}(y) + (Q_c^k)'(y_c)B^{***}_{k,\ell}(y)\right),$$

where $A^{***}_{k,\ell}$ and $B^{***}_{k,\ell}$ are explicit in terms of $A^{**}_{k,\ell}$ and $B^{**}_{k,\ell}$.

Next,

$$\begin{aligned}\mathbf{IV} &= \partial_x(\mathbf{N})\\ &= \sum_{\substack{2\le k\le 4k_0+9\\ 0\le \ell\le 4\ell_0+3}} c^\ell \Big[(Q_c^k)'(y_c)F^{\mathbf{N}}_{k,\ell}(y) + Q_c^k(y_c)\left((F^{\mathbf{N}}_{k,\ell})'(y) - \beta(y_c)(F^{\mathbf{N}}_{k,\ell})'(y)\right)\\ &\quad + (Q_c^k)''(y_c)G^{\mathbf{N}}_{k,\ell}(y) + (Q_c^k)'(y_c)\left((G^{\mathbf{N}}_{k,\ell})'(y) - \beta(y_c)(G^{\mathbf{N}}_{k,\ell})'(y)\right)\Big].\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{IV} &= \sum_{\substack{2\le k\le 4k_0+9\\ 0\le \ell\le 4\ell_0+3}} c^\ell Q_c^k(y_c)(F^{\mathbf{N}}_{k,\ell})'(y)\\ &\quad + \sum_{\substack{3\le k\le 5k_0+9\\ 0\le \ell\le 5\ell_0+3}} c^\ell Q_c^k(y_c) \sum_{\substack{\max(k-4k_0-9,1)\le k_1\le \min(k-2,k_0)\\ \max(\ell-4\ell_0-3,0)\le \ell_1\le \min(\ell,\ell_0)}} \left(-a_{k_1,\ell_1}(F^{\mathbf{N}}_{k-k_1,\ell-\ell_1})'(y)\right)\\ &\quad + \sum_{\substack{2\le k\le 4k_0+9\\ 1\le \ell\le 4\ell_0+4}} c^\ell Q_c^k(y_c)k^2G^{\mathbf{N}}_{k,\ell-1}(y)\end{aligned} \tag{A.18}$$

$$+\sum_{\substack{5\leq k\leq 4k_0+12\\ 0\leq \ell\leq 4\ell_0+3}} c^\ell Q_c^k(y_c)\left(-\tfrac{(k-3)(2k-3)}{5}G^{\mathbf{N}}_{k-3,\ell}(y)\right)$$
$$+\sum_{\substack{2\leq k\leq 4k_0+9\\ 0\leq \ell\leq 4\ell_0+3}} c^\ell (Q_c^k)'(y_c)\left(F^{\mathbf{N}}_{k,\ell}(y)+(G^{\mathbf{N}}_{k,\ell})'(y)\right)+\sum_{\substack{3\leq k\leq 5k_0+9\\ 0\leq \ell\leq 5\ell_0+3}} c^\ell (Q_c^k)'(y_c)$$
$$\times \sum_{\substack{\max(k-4k_0-9,1)\leq k_1\leq \min(k-2,k_0)\\ \max(\ell-4\ell_0-3,0)\leq \ell_1\leq \min(\ell,\ell_0)}} \left(-\tfrac{k-k_1}{k}a_{k_1,\ell_1}(G^{\mathbf{N}}_{k-k_1,\ell-\ell_1})'(y)\right).$$

It follows that **IV** can be written as

$$\mathbf{IV}=\sum_{\substack{2\leq k\leq \max(5k_0+9,4k_0+12)\\ 0\leq \ell\leq \max(5\ell_0+3,4\ell_0+4)}} c^\ell\left(Q_c^k(y_c)F^{\mathbf{IV}}_{k,\ell}(y)+(Q_c^k)'(y_c)G^{\mathbf{IV}}_{k,\ell}(y)\right), \tag{A.19}$$

where $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$ can be extracted from the previous calculations. Let us check that $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$ satisfy properties (i) and (ii).

Dependence property (i). In the decomposition of $\mathbf{N}_1$, the function in factor of $c^\ell Q_c^k$ is $12A_{k-1,\ell}Q^2$ and the function in factor of $c^\ell(Q_c^k)'$ is $\frac{k-1}{k}(12B_{k-1,\ell}Q^2)$. In the decomposition of W^2, in factor of $c^\ell Q_c^k$, we have sums where $k_1\leq k-1$ and $k-k_1\leq k-1$ since $k_1\geq 1$; moreover $\ell_1\leq \ell$ and $\ell-\ell_1\leq \ell$. Similar remarks apply to the other terms in $\mathbf{N}$.

Thus, $F^{\mathbf{N}}_{k,\ell}$ and $G^{\mathbf{N}}_{k,\ell}$ contain $(A_{k',\ell'})$ and $(B_{k',\ell'})$ only for $(k',\ell')\prec(k,\ell)$ (in fact $k'\leq k-1$ is always true). From (A.18) it is clear that the same is true for $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$. Note, in a similar way, that when $(a_{k',\ell'})$ is involved in some formula for $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$, it is only for $(k',\ell')\prec(k,\ell)$.

Parity property (ii). Assume that all $(A_{k',\ell'})$ are even and all $(B_{k',\ell'})$ are odd. From the decomposition of the various terms of $\mathbf{N}$, it is easy to observe that all $(F^{\mathbf{N}}_{k,\ell})$ are even and all $(G^{\mathbf{N}}_{k,\ell})$ are odd. Then, formula (A.18) ensures that all $(F^{\mathbf{IV}}_{k,\ell})$ are odd and all $(G^{\mathbf{IV}}_{k,\ell})$ are even.

To complete the proof of Lemma A.4, we only have to compute $F^{\mathbf{IV}}_{2,0}$ and $G^{\mathbf{IV}}_{2,0}$.

By (A.18), we have $F^{\mathbf{IV}}_{2,0}=(F^{\mathbf{N}}_{2,0})'$, and so we are reduced to compute $F^{\mathbf{N}}_{2,0}$. We give below the contribution of each $\mathbf{N}_j$ for $j=1,\dots,9$ to $F^{\mathbf{N}}_{2,0}$:

- For $\mathbf{N}_1$, the contribution is $12A_{1,0}Q^2$;
- The contribution of $\mathbf{N}_4$ is $6A^*_{2,0}Q^2=6A^2_{1,0}Q^2$, by the expression of W^2;
- The contribution of all the other terms $\mathbf{N}_2$, $\mathbf{N}_3$ $\mathbf{N}_5$, $\mathbf{N}_6$, $\mathbf{N}_7$, $\mathbf{N}_8$ and $\mathbf{N}_9$ is zero.

Therefore, $F^{\mathbf{IV}}_{2,0}=(F^{\mathbf{N}}_{2,0})'=\left(12A_{1,0}Q^2+6A^2_{1,0}Q^2\right)'$.

By (A.18), we have $G^{\mathbf{IV}}_{2,0} = F^{\mathbf{N}}_{2,0} + (G^{\mathbf{N}}_{2,0})'$. Since $F^{\mathbf{N}}_{2,0}$ was computed above, we are reduced to compute $G^{\mathbf{N}}_{2,0}$. We give below the contribution of each $\mathbf{N}_j$ for $j = 1, \dots, 9$ to $G^{\mathbf{N}}_{2,0}$:

- For $\mathbf{N}_1$, the contribution is $6B_{1,0}Q^2$;
- The contribution of $\mathbf{N}_4$ is $6B^*_{2,0}Q^2 = 6A_{1,0}B_{1,0}Q^2$, by the expression of W^2;
- The contribution of all the other terms $\mathbf{N}_2$, $\mathbf{N}_3$, $\mathbf{N}_5$, $\mathbf{N}_6$, $\mathbf{N}_7$, $\mathbf{N}_8$ and $\mathbf{N}_9$ is zero.

Therefore, $G^{\mathbf{IV}}_{2,0} = 12A_{1,0}Q^2 + 6A^2_{1,0}Q^2 + \left(6B_{1,0}Q^2 + 6A_{1,0}B_{1,0}Q^2\right)'$.

A.5. *End of the proof of Proposition* 2.1. By Lemmas A.1–A.4, we only have to sum the various contributions of **I**, **II**, **III** and **IV** to prove Proposition 2.1. Setting

$$F_{k,\ell} = F^{\mathbf{I}}_{k,\ell} + F^{\mathbf{II}}_{k,\ell} + F^{\mathbf{III}}_{k,\ell} + F^{\mathbf{IV}}_{k,\ell} \quad \text{and} \quad G_{k,\ell} = G^{\mathbf{I}}_{k,\ell} + G^{\mathbf{II}}_{k,\ell} + G^{\mathbf{III}}_{k,\ell} + G^{\mathbf{IV}}_{k,\ell},$$

we obtain the formula of Proposition 2.1 for $S(t,x)$. Properties (i) and (ii) have been checked on the functions $F^{\mathbf{I}}_{k,\ell}$, $F^{\mathbf{II}}_{k,\ell}$, $F^{\mathbf{III}}_{k,\ell}$, $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{I}}_{k,\ell}$, $G^{\mathbf{II}}_{k,\ell}$, $G^{\mathbf{III}}_{k,\ell}$, $G^{\mathbf{IV}}_{k,\ell}$, and so they are also true on $F_{k,\ell}$ and $G_{k,\ell}$.

The expressions of $F_{1,0}$, $G_{1,0}$, $F_{2,0}$ and $G_{2,0}$ are obtained from Lemmas A.1–A.4. Observe that the only nonzero contribution to $F_{1,0}$ and $G_{1,0}$ comes from $F^{\mathbf{II}}_{1,0}$ and $G^{\mathbf{II}}_{1,0}$; we obtain $F_{1,0} = p(Q^{p-1})'$ and $G_{1,0} = pQ^{p-1}$.

Appendix B. Lemma B.1

Lemma B.1 (Structure of $F_{k,\ell}$ and $G_{k,\ell}$). *Let (k,ℓ) be such that $1 \le k \le K_0$, $0 \le \ell \le L_0$, with $(k,\ell) \ne (1,0)$. Assume that for all $1 \le k' \le k_0$, $0 \le \ell' \le \ell_0$ such that $(k',\ell') \prec (k,\ell)$, the functions $A_{k',\ell'}$ and $B_{k',\ell'}$ verify*

$$A_{k',\ell'} = \overline{A}_{k',\ell'} + \widetilde{A}_{k',\ell'} + \varphi\widehat{A}_{k',\ell'}, \quad B_{k',\ell'} = \overline{B}_{k',\ell'} + \widetilde{B}_{k',\ell'} + \varphi\widehat{B}_{k',\ell'}, \tag{B.1}$$

where

- *$\overline{A}_{k',\ell'}$, $\overline{B}_{k',\ell'} \in \mathcal{Y}$; the function $\overline{A}_{k',\ell'}$ is even and the function $\overline{B}_{k',\ell'}$ is odd.*
- *$\widetilde{A}_{k',\ell'}$ and $\widehat{B}_{k',\ell'}$ are even polynomials; $\widehat{A}_{k',\ell'}$ and $\widetilde{B}_{k',\ell'}$ are odd polynomials.*

Then the functions $F_{k,\ell}$ and $G_{k,\ell}$ obtained in Proposition 2.1 from $(a_{k',\ell'})$, $(A_{k',\ell'})$ and $(B_{k',\ell'})$ are such that

$$F_{k,\ell} = \overline{F}_{k,\ell} + \widetilde{F}_{k,\ell} + \varphi\widehat{F}_{k,\ell}, \quad G_{k,\ell} = \overline{G}_{k,\ell} + \widetilde{G}_{k,\ell} + \varphi\widehat{G}_{k,\ell},$$

where

- *$\overline{F}_{k,\ell}$, $\overline{G}_{k,\ell} \in \mathcal{Y}$; the function $\overline{F}_{k,\ell}$ is odd and the function $\overline{G}_{k,\ell}$ is even.*
- *$\widetilde{F}_{k,\ell}$ and $\widehat{G}_{k,\ell}$ are odd polynomials; $\widehat{F}_{k,\ell}$ and $\widetilde{G}_{k,\ell}$ are even polynomials.*

Moreover, the following hold:

(a) *Let* $2 \le k \le p-1$, $\ell = 0$. *If for any* $1 \le k' < k$,

$$\deg \widetilde{A}_{k',0} = \deg \widehat{A}_{k',0} = \deg \widetilde{B}_{k',0} = \deg \widehat{B}_{k',0} = 0, \quad \textit{then} \quad F_{k,0},\ G_{k,0} \in \mathcal{Y}.$$

(b) *Let* $1 \le k \le k_0$ *and* $0 \le \ell \le \ell_0$ *be such that* $\frac{k}{p-1} + \ell \le 2$. *If for any* $(k',\ell') \prec (k,\ell)$,

$$\deg \widetilde{A}_{k',\ell'} = \deg \widehat{A}_{k',\ell'} = 0 \ \textit{and}\ \deg \widetilde{B}_{k',\ell'} = \deg \widehat{B}_{k',\ell'} \le 1, \ \textit{then} \quad F_{k,\ell} \in \mathcal{Y}.$$

(c) *Let*

$$d_{AB}(k,\ell) = \begin{cases} \max\limits_{\substack{1\le k'\le k\\ 0\le \ell'\le \ell}} \left(\deg \widetilde{A}_{k',\ell'}, \deg \widehat{A}_{k',\ell'}, \deg \widetilde{B}_{k',\ell'}, \deg \widehat{B}_{k',\ell'}\right) & \textit{if } k \ge 1,\ \ell \ge 0,\\ 0 \ \textit{otherwise}, \end{cases}$$

$$d_{\mathbf{N}}(k,\ell) = \begin{cases} \max\limits_{\substack{1\le k_j\le k_0\\ 0\le \ell_j\le \ell_0}} \left(\sum_{j=1}^{p} d_{AB}(k_j,\ell_j) \ \textit{for}\ \sum_{j=1}^{p} k_j \le k, \sum_{j=1}^{p} \ell_j \le \ell\right) & \textit{if } k \ge p,\ \ell \ge 0,\\ 0 \ \textit{otherwise}, \end{cases}$$

$$d_{FG}(k,\ell) = \max(\deg \widetilde{F}_{k,\ell}, \deg \widehat{F}_{k,\ell}, \deg \widetilde{G}_{k,\ell}, \deg \widehat{G}_{k,\ell}) \ \textit{for}\ 1 \le k \le K_0,\ 0 \le \ell \le L_0.$$

Then, for all $1 \le k \le K_0$, $0 \le \ell \le L_0$,

$$d_{FG}(k,\ell) \le \max\left(d_{AB}(k-1,\ell)-1, d_{AB}(k-p+1,\ell), d_{AB}(k,\ell-1), d_{\mathbf{N}}(k,\ell)\right). \tag{B.2}$$

Proof of Lemma B.1. Let (k,ℓ) be such that $1 \le k \le K_0$ and $0 \le \ell \le L_0$, with $(k,\ell) \ne (1,0)$. We suppose that for all $1 \le k' \le k_0$, $0 \le \ell \le \ell_0$ such that $(k',\ell') \prec (k,\ell)$, $(a_{k',\ell'}, A_{k',\ell'}, B_{k',\ell'})$ satisfies the assumptions of Lemma B.1. We consider $F_{k,\ell}$, $G_{k,\ell}$ defined by Proposition 2.1 (recall that for given (k,ℓ), $F_{k,\ell}$ and $G_{k,\ell}$ depend only on $(a_{k',\ell'}, A_{k',\ell'}, B_{k',\ell'})$ for $(k',\ell') \prec (k,\ell)$). From the proof of Proposition 2.1 (Appendix A),

$$F_{k,\ell} = F^{\mathbf{I}}_{k,\ell} + F^{\mathbf{II}}_{k,\ell} + F^{\mathbf{III}}_{k,\ell} + F^{\mathbf{IV}}_{k,\ell}, \quad G_{k,\ell} = G^{\mathbf{I}}_{k,\ell} + G^{\mathbf{II}}_{k,\ell} + G^{\mathbf{III}}_{k,\ell} + G^{\mathbf{IV}}_{k,\ell},$$

where $F^{\mathbf{I}}_{k,\ell}$, $F^{\mathbf{II}}_{k,\ell}$, etc. are the contributions of **I**, **II**, **III** and **IV** in the decomposition of $S(t,x)$; see (2.16).

Contribution of **I** *and* **II**. From Lemmas A.1 and A.2, it follows that $F^{\mathbf{I}}_{k,\ell}$, $F^{\mathbf{II}}_{k,\ell}$, $G^{\mathbf{I}}_{k,\ell}$ and $G^{\mathbf{II}}_{k,\ell}$ belong to $\mathcal{Y}$ and do not depend on $(A_{k',\ell'})$, $(B_{k',\ell'})$ but only on the coefficients $(a_{k',\ell'})$. Moreover, $F^{\mathbf{I}}_{k,\ell}$ and $F^{\mathbf{II}}_{k,\ell}$ are odd, and $G^{\mathbf{I}}_{k,\ell}$ and $G^{\mathbf{II}}_{k,\ell}$ are even. Therefore, they only contribute to $\overline{F}_{k,\ell}$ and $\overline{G}_{k,\ell}$, with the desired parity property.

Contribution of **III**. We use the notation and calculations of the proof of Lemma A.3. Note that $\mathbf{III}_1$ does not contribute to $F^{\mathbf{III}}_{k,\ell}$ and $G^{\mathbf{III}}_{k,\ell}$. Observing the other terms, i.e. $\mathbf{III}_2$, $\mathbf{III}_3^1$, $\mathbf{III}_3^2$, etc. up to $\mathbf{III}_{13}^5$, we note that there are three kinds of terms depending on the structure of the function of the variable y:

T_1: *Terms depending on* $A_{k',\ell'}(y)$ *and* $B_{k',\ell'}(y)$ *without derivative, for* $(k',\ell') \prec (k,\ell)$. A complete list of these terms is given in formula (B.3) below.

T_2: *Terms depending on derivatives of* $A_{k',\ell'}(y)$ *and* $B_{k',\ell'}(y)$ (*up to order* 3) *for* $(k',\ell') \prec (k,\ell)$. Examples of such terms are $F_{k,\ell}^{\mathbf{III}_4}$, $G_{k,\ell}^{\mathbf{III}_4}$, a part of $F_{k,\ell}^{\mathbf{III}_3}$, etc.

T_3: *Terms depending on* $(Q^{p-1}A_{k',\ell'})'(y)$ *and* $(Q^{p-1}B_{k',\ell'})'(y)$ *for* $(k',\ell') \prec (k,\ell)$. Examples of such terms are a part of $F_{k,\ell}^{\mathbf{III}_3}$, $G_{k,\ell}^{\mathbf{III}_3}$, etc.

Terms of type T_3 are easily handled. Indeed, since $A_{k',\ell'}$ and $B_{k',\ell'}$ are of the form (B.1) and since $Q \in \mathcal{Y}$, it follows that $(Q^{p-1}A_{k',\ell'})'$ and $(Q^{p-1}B_{k',\ell'})'$ belong to $\mathcal{Y}$. Therefore, this kind of terms only contribute to $\overline{F}_{k,\ell}$ and $\overline{G}_{k,\ell}$. The parity statement for these terms was already checked in the proof of Lemma A.3.

We now handle terms of type T_2. It suffices to remark that when differentiating terms such as $A_{k',\ell'}$ and $B_{k',\ell'}$ of the form (B.1), we obtain terms of the same form, except that the degrees of the polynomial functions decrease by one or more depending on the order of derivation. Indeed, for example, it follows from (B.1) that

$$A'_{k',\ell'} = (\overline{A}'_{k',\ell'} + \varphi'\widehat{A}_{k',\ell'}) + \widetilde{A}'_{k',\ell'} + \varphi\widehat{A}'_{k',\ell'},$$

and $\overline{A}'_{k',\ell'} + \varphi'\widehat{A}_{k',\ell'} \in \mathcal{Y}$, because of the property $\varphi' \in \mathcal{Y}$. Thus, for example, we get $F_{k,\ell}^{\mathbf{III}_4} = \overline{F}_{k,\ell}^{\mathbf{III}_4} + \widetilde{F}_{k,\ell}^{\mathbf{III}_4} + \varphi\widehat{F}_{k,\ell}^{\mathbf{III}_4}$, where

$$\deg \widetilde{F}_{k,\ell}^{\mathbf{III}_4} \le \max_{(k',\ell')\prec(k,\ell)} (\deg \widehat{A}_{k',\ell'}) - 1 \le \max(d_{AB}(k-1,\ell), d_{AB}(k,\ell-1)) - 1,$$

$$\deg \widehat{F}_{k,\ell}^{\mathbf{III}_4} \le \max_{(k',\ell')\prec(k,\ell)} (\deg \widetilde{A}_{k',\ell'}) - 1 \le \max(d_{AB}(k-1,\ell), d_{AB}(k,\ell-1))) - 1,$$

if $\max(d_{AB}(k-1,\ell), d_{AB}(k,\ell-1))) \ge 1$, and $\widetilde{F}_{k,\ell}^{\mathbf{III}_4} = \widehat{F}_{k,\ell}^{\mathbf{III}_4} = 0$ otherwise.

We obtain similar estimates for all terms of this type. The parity properties are easily checked. For terms of type T_2 with higher order derivatives (in fact, only second and third derivative), the argument is the same.

Finally, we look at terms of type T_1, i.e. depending on $A_{k',\ell'}$ and $B_{k',\ell'}$ without derivative:

(B.3)

$$G_{k,\ell}^{\mathbf{III}_2}, \quad G_{k,\ell}^{\mathbf{III}_{10}^1}, \quad G_{k,\ell}^{\mathbf{III}_{10}^2}, \quad F_{k,\ell}^{\mathbf{III}_{11}^1}, \quad F_{k,\ell}^{\mathbf{III}_{11}^2}, \quad F_{k,\ell}^{\mathbf{III}_{13}^3}, \quad F_{k,\ell}^{\mathbf{III}_{13}^4}, \quad F_{k,\ell}^{\mathbf{III}_{13}^5}.$$

With the assumptions on $A_{k',\ell'}$ and $B_{k',\ell'}$, these terms have the desired structure. We only have to check the estimates on the degrees of the polynomials.

First, we note from the proof of Lemma A.3 that terms $G_{k,\ell}^{\mathbf{III}_2}$, $G_{k,\ell}^{\mathbf{III}_{10}^1}$, $F_{k,\ell}^{\mathbf{III}_{11}^1}$, $F_{k,\ell}^{\mathbf{III}_{11}^2}$, $F_{k,\ell}^{\mathbf{III}_{13}^3}$, $F_{k,\ell}^{\mathbf{III}_{13}^4}$ depend only on $A_{k',\ell'}$ and $B_{k',\ell'}$ with $k' \le k$ and $\ell' \le \ell - 1$. Thus, they appear only for $\ell \ge 1$ and contain polynomials with degrees less than or equal to $d_{AB}(k,\ell-1)$. The other two terms $G_{k,\ell}^{\mathbf{III}_{10}^2}$ and

$F^{\mathbf{III}^5_{13}}_{k,\ell}$ depend on $A_{k',\ell'}$ and $B_{k',\ell'}$ with $k' \leq k-p+1$ and $\ell' \leq \ell$. Thus they appear only for $k \geq p$, and contain polynomials with degrees less than or equal to $d_{AB}(k-p+1,\ell)$.

Thus, in conclusion for the term **III**, we get polynomials of degrees less than

$$\max\left(d_{AB}(k-1,\ell)-1, d_{AB}(k-p+1,\ell), d_{AB}(k,\ell-1)\right).$$

This proves (c) for $d^{\mathbf{III}}_{FG}$.

Let us now prove (a) and (b) for $F^{\mathbf{III}}_{k,\ell}$ and $G^{\mathbf{III}}_{k,\ell}$.

Proof of (a). First, observe that terms of type T_1 (see above) do not appear for $k \leq p-1$ and $\ell = 0$. Thus, for such k, if we assume $\widetilde{A}_{k',0} = \widehat{A}_{k',0} = \widetilde{B}_{k',0} = 0$ and $\widehat{B}_{k',0} = b_{k',0} \in \mathbb{R}$, then $\widetilde{A}'_{k',0} = \widehat{A}'_{k',0} = \widetilde{B}'_{k',0} = \widehat{B}'_{k',0} = 0$ for all $1 \leq k' < k$, and so $\widetilde{F}^{\mathbf{III}}_{k,\ell} = \widehat{F}^{\mathbf{III}}_{k,\ell} = \widetilde{G}^{\mathbf{III}}_{k,\ell} = \widehat{G}^{\mathbf{III}}_{k,\ell} = 0$, which means $F^{\mathbf{III}}_{k,\ell}, G^{\mathbf{III}}_{k,\ell} \in \mathcal{Y}$. This proves (a) for $F^{\mathbf{III}}_{k,\ell}$ and $G^{\mathbf{III}}_{k,\ell}$.

Proof of (b). To justify (b) for $F^{\mathbf{III}}_{k,\ell}$, we first observe that for (k,ℓ) such that $\frac{k}{p-1} + \ell \leq 2$, there is no term of type T_1 contributing to $F^{\mathbf{III}}_{k,\ell}$. Indeed, looking at the expression of all the terms in the list (B.3) in the proof of Proposition 2.1, we see that $F^{\mathbf{III}^1_{11}}_{k,\ell}$, $F^{\mathbf{III}^2_{11}}_{k,\ell}$, $F^{\mathbf{III}^3_{13}}_{k,\ell}$, $F^{\mathbf{III}^4_{13}}_{k,\ell}$, $F^{\mathbf{III}^5_{13}}_{k,\ell}$ involves $B_{k',\ell'}$ for $k' \leq k-2(p-1)$ or $\ell' \leq \ell-2$ or simultaneously $k' \leq k-(p-1)$ and $\ell' \leq \ell-1$. Therefore, these terms do not appear if $\frac{k}{p-1} + \ell \leq 2$. Concerning terms of type T_2, we first note that $B_{k',\ell'}$ appear with at least two derivatives, thus any polynomial function of degree 1 disappears. Second, $A_{k',\ell'}$ are differentiated at least once, and so again any constant term disappears. Thus, there remains no polynomial growth and $F^{\mathbf{III}}_{k,\ell} \in \mathcal{Y}$ for such k, ℓ.

Contribution of **IV**. We focus on the case $p = 4$. The other cases, i.e. $p = 2$ or 3 are similar and easier. We use the notation and calculations of the proof of Lemma A.4, where we have written $\mathbf{IV} = \partial_x(\mathbf{N})$, $\mathbf{N} = (R+R_c+W)^4 - (R+R_c)^4 - 4R^3W$, and where we have decomposed $\mathbf{N}$ into several parts $\mathbf{N}_1, \ldots, \mathbf{N}_9$. Here, we distinguish two kind of terms: first $\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{N}_4$, $\mathbf{N}_5$, $\mathbf{N}_7$, which contain the function $Q(y)$, and second, $\mathbf{N}_3$, $\mathbf{N}_6$, $\mathbf{N}_8$, $\mathbf{N}_9$, which depend only on Q_c and W.

For the first terms, $\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{N}_4$, $\mathbf{N}_5$ and $\mathbf{N}_7$, since $Q \in \mathcal{Y}$, by the structure of W, and the assumptions on $A_{k',\ell'}$ and $B_{k',\ell'}$, the result follows.

For $\mathbf{N}_3$, $\mathbf{N}_6$, $\mathbf{N}_8$ and $\mathbf{N}_9$, we set

$$\mathbf{M} = \mathbf{N}_3 + \mathbf{N}_6 + \mathbf{N}_8 + \mathbf{N}_9 = (Q_c + W)^4 - Q_c^4.$$

In order to have a simple expression when expanding $(Q_c+W)^4$, it is convenient to set

$$\mathbf{A}_{1,0} = 1 + A_{1,0}, \;\; \mathbf{A}_{k,\ell} = A_{k,\ell} \text{ for any } (k,\ell) \neq (1,0), \;\; \mathbf{B}_{k,\ell} = B_{k,\ell} \text{ for any } (k,\ell),$$

$$\deg \mathbf{A}_{k,\ell} = \deg A_{k,\ell}, \quad \deg \mathbf{B}_{k,\ell} = \deg B_{k,\ell}. \tag{B.4}$$

With this notation, we have

$$Q_c + W = \sum_{(k,\ell)\in\Sigma_0} c^\ell \left(Q_c^k(y_c)\mathbf{A}_{k,\ell}(y) + (Q_c^k)'(y_c)\mathbf{B}_{k,\ell}(y)\right).$$

Then,

$$\begin{aligned}(Q_c+W)^4 = \sum_{\substack{(k_j,\ell_j)\in\Sigma_0\\ j=1,2,3,4}} & c^{\ell_1+\ell_2+\ell_3+\ell_4} \\ \times \Big\{ & Q_c^{k_1+k_2+k_3+k_4}(y_c)\ (\mathbf{A}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4})(y) \\ & + 4((Q_c^{k_1})'Q_c^{k_2+k_3+k_4})(y_c)(\mathbf{B}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4})(y) \\ & + 6((Q_c^{k_1})'(Q_c^{k_2})'Q_c^{k_3+k_4})(y_c)(\mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4})(y) \\ & + 4((Q_c^{k_1})'(Q_c^{k_2})'(Q_c^{k_3})'Q_c^{k_4})(y_c)(\mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{B}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4})(y) \\ & + ((Q_c^{k_1})'(Q_c^{k_2})'(Q_c^{k_3})'(Q_c^{k_2})')(y_c)(\mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{B}_{k_3,\ell_3}\mathbf{B}_{k_4,\ell_4})(y)\Big\}.\end{aligned} \tag{B.5}$$

Recall that by Lemma 2.1, we have

$$\begin{aligned}(Q_c^{k_1})'Q_c^{k_2+k_3+k_4} &= \tfrac{k_1}{k_1+k_2+k_3+k_4}(Q_c^{k_1+k_2+k_3+k_4})', \\ (Q_c^{k_1})'(Q_c^{k_2})'Q_c^{k_3+k_4} &= k_1k_2\left(cQ_c^{k_1+k_2+k_3+k_4} - \tfrac{2}{p+1}Q_c^{k_1+k_2+k_3+k_4+3}\right), \\ (Q_c^{k_1})'(Q_c^{k_2})'(Q_c^{k_3})'Q_c^{k_4} &= k_1k_2k_3\left(\frac{c(Q_c^{k_1+k_2+k_3+k_4})'}{k_1+k_2+k_3+k_4}\right. \\ &\qquad\left. - \frac{2(Q_c^{k_1+k_2+k_3+k_4+3})'}{(p+1)(k_1+k_2+k_3+k_4+3)}\right), \\ (Q_c^{k_1})'(Q_c^{k_2})'(Q_c^{k_3})'(Q_c^{k_4})' &= k_1k_2k_3k_4Q_c^{k_1+k_2+k_3+k_4}\left(c^2 - \tfrac{4c}{p+1}Q_c^3 + \tfrac{4}{(p+1)^2}Q_c^6\right).\end{aligned}$$

Therefore, we can write

$$\mathbf{M} = \sum_{\substack{4\le k\le 4k_0+6\\ 0\le\ell\le 4\ell_0+2}} c^\ell\left(Q_c^k(y_c)F_{k,\ell}^{\mathbf{M}}(y) + (Q_c^k)'(y_c)G_{k,\ell}^{\mathbf{M}}(y)\right), \tag{B.6}$$

where at given $k\ge 4$, $\ell\ge 0$, $F_{k,\ell}^{\mathbf{M}}$ contains only terms of the type

$$\begin{gathered}\mathbf{A}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}, \quad \mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}, \\ \mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{B}_{k_3,\ell_3}\mathbf{B}_{k_4,\ell_4},\end{gathered} \tag{B.7}$$

for $\sum_{j=1}^4 k_j\le k$ and $\sum_{j=1}^4 \ell_j\le \ell$, and $G_{k,\ell}^{\mathbf{M}}$ contains only terms of the type

$$\mathbf{B}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}, \quad \mathbf{B}_{k_1,\ell_1}\mathbf{B}_{k_2,\ell_2}\mathbf{B}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}, \tag{B.8}$$

for $\sum_{j=1}^4 k_j\le k$ and $\sum_{j=1}^4 \ell_j\le \ell$.

Therefore, we only have to check the structure of the functions in (B.7) and (B.8). We check the first term $\mathbf{A}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}$. The other terms can be checked similarly.

Recall that $\mathbf{A}_{k_j,\ell_j} = \overline{\mathbf{A}}_{k_j,\ell_j} + \widetilde{\mathbf{A}}_{k_j,\ell_j} + \varphi\widehat{\mathbf{A}}_{k_j,\ell_j}$, where $\overline{\mathbf{A}}_{k_j,\ell_j} \in \mathcal{Y}$, and $\widetilde{\mathbf{A}}_{k_j,\ell_j}$ and $\widehat{\mathbf{A}}_{k_j,\ell_j}$ are polynomials. In the product $\mathbf{A}_{k_1,\ell_1}\mathbf{A}_{k_2,\ell_2}\mathbf{A}_{k_3,\ell_3}\mathbf{A}_{k_4,\ell_4}$, any term in factor to some $\overline{\mathbf{A}}_{k_j,\ell_j}$ is automatically in $\mathcal{Y}$. The other terms are

$$(\widetilde{\mathbf{A}}_{k_1,\ell_1} + \varphi\widehat{\mathbf{A}}_{k_1,\ell_1})(\widetilde{\mathbf{A}}_{k_2,\ell_2} + \varphi\widehat{\mathbf{A}}_{k_2,\ell_2})(\widetilde{\mathbf{A}}_{k_3,\ell_3} + \varphi\widehat{\mathbf{A}}_{k_3,\ell_3})(\widetilde{\mathbf{A}}_{k_4,\ell_4} + \varphi\widehat{\mathbf{A}}_{k_4,\ell_4}).$$

In this product we distinguish two kinds of terms:

- $\Pi_{j=1}^4 \widetilde{\mathbf{A}}_{k_j,\ell_j}$, $\widetilde{\mathbf{A}}_{k_1,\ell_1}\widetilde{\mathbf{A}}_{k_2,\ell_2}(\varphi^2\widehat{\mathbf{A}}_{k_3,\ell_3}\widehat{\mathbf{A}}_{k_4,\ell_4})$ (and similar terms), $\varphi^4\Pi_{j=1}^4\widehat{\mathbf{A}}_{k_j,\ell_j}$. Since $1-\varphi^2, 1-\varphi^4 \in \mathcal{Y}$, these terms are of the form $\overline{F} + \widetilde{F}$, where $\overline{F} \in \mathcal{Y}$ is even and $\widetilde{F}$ is an even polynomial of degree less than or equal to $d_{\mathbf{N}}(k,\ell)$.
- $\Pi_{j=1}^3 \widetilde{\mathbf{A}}_{k_j,\ell_j}(\varphi\widehat{\mathbf{A}}_{k_4,\ell_4})$, $\widetilde{\mathbf{A}}_{k_1,\ell_1}(\varphi^3\Pi_{j=2}^4\widehat{\mathbf{A}}_{k_j,\ell_j})$ (and similar terms). Since $\varphi^3 - \varphi \in \mathcal{Y}$, these terms are of the form $\overline{F} + \varphi\widehat{F}$, where $\overline{F} \in \mathcal{Y}$ and $\widehat{F}$ is a polynomial function of degree are less than $d_{\mathbf{N}}(k,\ell)$.

In conclusion, we obtain

$$F^{\mathbf{M}}_{k,\ell} = \overline{F}^{\mathbf{M}}_{k,\ell} + \widetilde{F}^{\mathbf{M}}_{k,\ell} + \varphi\widehat{F}^{\mathbf{M}}_{k,\ell}, \quad G^{\mathbf{M}}_{k,\ell} = \overline{G}^{\mathbf{M}}_{k,\ell} + \widetilde{G}^{\mathbf{M}}_{k,\ell} + \varphi\widehat{G}^{\mathbf{M}}_{k,\ell},$$

where

- $\overline{F}^{\mathbf{M}}_{k,\ell}, \overline{G}^{\mathbf{M}}_{k,\ell} \in \mathcal{Y}$; $\overline{F}^{\mathbf{M}}_{k,\ell}$ is even and $\overline{G}^{\mathbf{M}}_{k,\ell}$ is odd.
- $\widetilde{F}^{\mathbf{M}}_{k,\ell}$ and $\widehat{G}^{\mathbf{M}}_{k,\ell}$ are even polynomials; $\widehat{F}^{\mathbf{M}}_{k,\ell}$ and $\widetilde{G}^{\mathbf{M}}_{k,\ell}$ are odd polynomials, satisfying

$$d^{\mathbf{M}}_{FG}(k,\ell) = \max(\deg \widetilde{F}^{\mathbf{M}}_{k,\ell}, \deg \widehat{F}^{\mathbf{M}}_{k,\ell}, \deg \widetilde{G}^{\mathbf{M}}_{k,\ell}, \deg \widehat{G}^{\mathbf{M}}_{k,\ell}), \leq d_{\mathbf{N}}(k,\ell). \tag{B.9}$$

The last step for **IV** is to use formulas (A.18) and (A.19) to derive the properties of $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$ from the properties of $F^{\mathbf{N}}_{k,\ell}$ and $G^{\mathbf{N}}_{k,\ell}$. We note that $F^{\mathbf{IV}}_{k,\ell}$ involves some $G^{\mathbf{N}}_{k',\ell'}$ and $(F^{\mathbf{N}}_{k',\ell'})'$ for $k' \leq k$ and $\ell' \leq \ell$ and $G^{\mathbf{IV}}_{k,\ell}$ involves some $(G^{\mathbf{N}}_{k',\ell'})'$ and $F^{\mathbf{N}}_{k',\ell'}$ for $k' \leq k$ and $\ell' \leq \ell$. Thus **IV** contains polynomials with degrees less than $d_{\mathbf{N}}(k,\ell)$, and the parity properties are satisfied, which proves (c) for $d^{\mathbf{IV}}_{FG}$.

Let us now prove (a) and (b) for $F^{\mathbf{IV}}_{k,\ell}$ and $G^{\mathbf{IV}}_{k,\ell}$.

Proof of (a). Note that from (B.5)–(B.6), for any $1 \leq k \leq p-1 = 3$, $F^{\mathbf{M}}_{k,0} = G^{\mathbf{M}}_{k,0} = 0$. Thus, $F^{\mathbf{N}}_{k,0}, G^{\mathbf{N}}_{k,0} \in \mathcal{Y}$ for such k. From (A.18) and (A.19) it follows that $F^{\mathbf{IV}}_{k,0}, G^{\mathbf{IV}}_{k,0} \in \mathcal{Y}$ for all $1 \leq k \leq p-1$. This proves (a) for term **IV**.

Proof of (b). To prove (b) for $F^{\mathbf{IV}}_{k,\ell}$, we need to give a closer look to (A.18) and (A.19). Note that $F^{\mathbf{IV}}_{k,\ell}$ contains only terms of the type $G^{\mathbf{N}}_{k,\ell-1}$, $G^{\mathbf{N}}_{k-3,\ell}$ and $(F^{\mathbf{N}}_{k,\ell})'$. For (k,ℓ) such that $\frac{k}{3} + \ell \leq 2$, this provides terms $G^{\mathbf{N}}_{k',\ell'}$ for $\frac{k'}{3} + \ell \leq 1$. Since $k' \geq 1$, this condition implies $\ell = 0$ and $k' \leq 3 = p-1$. But, we know from (B.5)–(B.6) that $G^{\mathbf{M}}_{k',\ell'} = 0$ for such (k',ℓ'). Next, by (B.5), $F^{M}_{k',1} = 0$ for

$1 \leq k' \leq 3$. Moreover, $F^M_{k',0}$ contains only product of $A_{k_j,0}$ for $k' \leq 6$. Indeed, if we look for example at a term of the form $(Q_c^{k_1})'(Q_c^{k_2})'Q_c^{k_3}Q_c^{k_4}B_{1,0}B_{2,0}A_{3,0}A_{4,0}$, by the formula of $(Q_c^2)'$, it gives a contribution only for $F_{k',\ell'}$, where $k' \geq 7$ or $k' \geq 4$ and $\ell \geq 1$.

Thus, by the assumptions on $A_{k',0}$, $F^M_{k',0}$ contains only constant polynomial functions and so its derivative is in $\mathcal{Y}$.

Appendix C. Identities related to Q

Claim C.1 (Identities for any $p > 1$).

$$\int Q^{p+1} = \frac{2(p+1)}{p+3}\int Q^2, \quad \int (Q')^2 = \frac{p-1}{p+3}\int Q^2,$$

$$\int Q_c^2 = c^{2q}\int Q^2, \quad E(Q_c) = c^{2q+1}, \quad E(Q) = -\frac{5-p}{2(p+3)}c^{2q+1}\int Q^2.$$

Proof of Lemma C.1. These are well-known calculations. We have $Q^p = Q - Q''$ and $\frac{2}{p+1}Q^{p+1} = Q^2 - (Q')^2$. Thus, by integration

$$\int Q^{p+1} = \int Q^2 + \int (Q')^2, \quad \frac{2}{p+1}\int Q^{p+1} = \int Q^2 - \int (Q')^2.$$

Therefore, $\int Q^{p+1} = \frac{2(p+1)}{p+3}\int Q^2$ and $\int (Q')^2 = \int Q^{p+1} - \int Q^2 = \frac{p-1}{p+3}\int Q^2$. Moreover, $E(Q) = \frac{1}{2}\int (Q')^2 - \frac{1}{p+1}\int Q^{p+1} = \frac{p-5}{2(p+3)}\int Q^2$.

Since $Q_c(y) = c^{\frac{1}{p-1}}Q(\sqrt{c}y)$ and $q = \frac{1}{p-1} - \frac{1}{4}$, we have

$$\int Q_c^2(y)dy = c^{\frac{2}{p-1}}\int Q^2(\sqrt{c}y)dy = c^{2q}\int Q^2.$$

Similarly, $\int (Q_c')^2 = c^{2q+1}\int (Q')^2$ and $\int Q_c^{p+1} = c^{2q+1}\int Q^{p+1}$, and so $E(Q_c) = c^{2q+1}E(Q)$.

Appendix D. Proof of some technical results

D.1. *Proof of Claim* 4.2. The proof is based on the following well-known fact: *There exists $\lambda_1 > 0$ such that if $v \in H^1(\mathbb{R})$ satisfies $\int Qv = \int xQv = 0$, then*

$$\int v_x^2 - pQ^{p-1}v^2 + v^2 \geq \lambda_1 \|v\|^2_{H^1}. \tag{D.1}$$

First, we claim from (D.1) that if $\tilde{v} \in H^1(\mathbb{R})$ satisfies $\int xQv = 0$, then

$$\int \tilde{v}_x^2 - pQ^{p-1}\tilde{v}^2 + \tilde{v}^2 \geq \lambda_0\|\tilde{v}\|^2_{H^1} - \frac{1}{\lambda_0}\left(\int \tilde{v}Q\right)^2. \tag{D.2}$$

Set $v = \tilde{v} - \frac{\int \tilde{v}Q}{\int Q^2}Q$. Then $\int Qv = \int xQv = 0$ and from (D.1), $\int v_x^2 - pQ^{p-1}v^2 + v^2 \geq \lambda_1\|v\|^2_{H^1}$. Moreover, $\|v\|^2_{H^1} \geq \|\tilde{v}\|^2_{H^1} - K(\int \tilde{v}Q)^2$ and $\int v_x^2 - pQ^{p-1}v^2 + v^2 \leq \int \tilde{v}_x^2 - pQ^{p-1}\tilde{v}^2 + \tilde{v}^2 + K(\int \tilde{v}Q)^2$. Thus (D.2) follows.

Second, we recall

$$\mathcal{F}(t) = \frac{1}{2}\int \Big((\partial_x z)^2 + (1+\alpha'(y_c))z^2)\Big) \\ - \frac{1}{p+1}\int \Big((v+z)^{p+1} - v^{p+1} - (p+1)v^p z\Big).$$

Since $|\alpha'(s)| \leq Kc^{\frac{1}{p-1}}$, $\|v-Q\|_{L^\infty} \leq Kc^{\frac{1}{p-1}}$ and $\|z\|_{L^\infty} \leq 2K^* c^\theta$, we have from (D.2) and $\int zQ' = 0$, for c small enough,

$$\mathcal{F}(t) \geq \frac{1}{2}\int \Big((\partial_x z)^2 + z^2 - pQ^{p-1}z^2\Big) - K(c^{\frac{1}{p-1}} + K^* c^\theta)\int z^2 \\ \geq \frac{\lambda_0}{4}\int \Big((\partial_x z)^2 + z^2\Big) - \frac{1}{2\lambda_0}\Big(\int zQ\Big)^2.$$

D.2. *Proof of Proposition* 5.1.

D.2.1. For given $0 < c < 1$, δ_1, $\delta_2 \in \mathbb{R}$, the existence of a solution $U(t)$ satisfying (5.1) is a consequence of Theorem 1 in [19]. Therefore, we only have to check (5.2), for c small, which is a more precise estimate than the one in [19], giving explicitly the dependency in c. This is obtained by combining the argument of the proof in [19] and estimates depending on c in the proof of Proposition 4.2 of the present paper.

We work on the time interval $(-\infty, -\tilde{T}_c]$, for $-\tilde{T}_c = \frac{\delta_2-\delta_1}{(1-c)} - \frac{T_c}{32}$. Let $R(t,x) = Q(x-t-\delta_1) + Q_c(x-ct-\delta_2)$. In the spirit of Proposition 3 in [19], we first claim the following:

Proposition D.1. *For $c > 0$ small enough, if there exists $t^* \leq -\tilde{T}_c$ such that for all $t \leq t^*$, $\|u(t) - R(t)\|_{H^1} \leq \exp(-c^{-\frac{q}{4}})$ then for all $t \leq t^*$, $\|u(t) - R(t)\|_{H^1} \leq K_0 e^{\frac{\sqrt{c}}{4}((1-c)t-(\delta_2-\delta_1))}$.*

Assume Proposition D.1. Since $\lim_{t\to-\infty}\|u(t) - R(t)\|_{H^1} = 0$, we can define

$$t^* = \sup\left\{t \leq -\tilde{T}_c \text{ such that } \forall s \leq t^*,\ \|u(s) - R(s)\|_{H^1} \leq \exp(-c^{-r})\right\}.$$

Since $K_0 e^{\frac{\sqrt{c}}{4}(-(1-c)\tilde{T}_c-(\delta_2-\delta_1))} \leq K_0 e^{\frac{\sqrt{c}}{128}T_c} \leq \frac{1}{2}\exp(-c^{-r})$, for c small enough, by a standard continuity argument in H^1, we have $t^* = -\tilde{T}_c$, and thus the result follows from Proposition D.1 applied on $(-\infty, -\tilde{T}_c]$. Therefore, we are reduced to prove Proposition D.1.

Sketch of the proof of Proposition D.1. For more details, we refer to the proof of Proposition 3 in [19]. We decompose the solution $u(t)$ on $(-\infty, t^*]$ by Lemma 4.3, with $\alpha = 0$ and $\rho_1(t) - \rho_2(t) \leq -\frac{t}{2} - \frac{T_c}{64}$. Note that here the two solitons are ordered in a different way, $\rho_2(t) > \rho_1(t)$, where $\rho_1(t)$ is center of Q and $\rho_2(t)$ is center of Q_c.

Then, by [25], we have

$$|c_1(t)-1|+|c_2(t)-c|\le Kg(t)+K\exp\left(\frac{\sqrt{c}}{4}((1-c)t-(\delta_2-\delta_1))\right),$$

where $g(t)$ is defined as in (4.34).

Next, similarly as in [25], we use a monotonicity argument, but since the solitons are ordered in reverse order, we will need the following quantities:

$$\mathcal{M}(t)=\int u^2(t,x)\psi(x-m(t))dx,$$
$$\tilde{\mathcal{E}}(t)=\int\left(\tfrac{1}{2}u_x^2-\tfrac{1}{p+1}u^{p+1}+\tfrac{c}{100}u^2\right)\psi(x-m(t))dx,$$

where $m(t)=\frac{1}{2}(\rho_1(t)+\rho_2(t))$. Similarly as in Lemma 1 of [19], we obtain, for $t'\le t\le t^*$,

$$\mathcal{M}(t)-\mathcal{M}(t')\le K\exp(\tfrac{1}{4}((1-c)t-(\delta_2-\delta_1))),$$
$$\tilde{\mathcal{E}}(t)-\tilde{\mathcal{E}}(t')\le K\exp(\tfrac{1}{4}((1-c)t-(\delta_2-\delta_1))).$$

We set

$$\begin{aligned}\mathcal{F}(t)&=\tfrac{1}{2}\int u^2(t)+E(u(t))+\left(\tfrac{1}{2c}-\tfrac{1}{2}\right)\mathcal{M}(t)\\&\quad+\left(\tfrac{1}{c^2}-1\right)\int\left(\tfrac{1}{2}u_x^2-\tfrac{1}{p+1}u^{p+1}\right)\psi(x-m(t))\\&=\frac{1}{2}\int u^2(t)+E(u(t))+\left(\tfrac{1}{c^2}-1\right)\tilde{\mathcal{E}}(t)+\tfrac{1}{2}\left(\tfrac{1}{c}-1\right)(1-\tfrac{1}{100}(\tfrac{1}{c}+1))\mathcal{M}(t).\end{aligned}$$

By the monotonicity results on $\mathcal{M}$ and $\tilde{\mathcal{E}}$, we have for all $t'\le t\le t^*$,

$$\mathcal{F}(t)-\mathcal{F}(t')\le K\exp(\tfrac{1}{4}((1-c)t-(\delta_2-\delta_1))),$$

and using an expansion of $\mathcal{F}(t)$ from (4.33), and passing to the limit $t'\to-\infty$, we obtain the conclusion of Proposition D.1.

D.2.2. *Sharper uniqueness property.* First, we check that for c small enough, if the solution $u(t)$ satisfies (5.3), then for $-t$ large, $\rho_1(t)-\rho_2(t)\le-\frac{1}{4}|t|$. This is a consequence of the asymptotic stability of one soliton. Indeed, if c is small enough, then for $-t$ large, $u(t)=Q(x-\delta_1)+\varepsilon(t,x)$, and $\varepsilon(t)$ small in H^1. Then, by stability and asymptotic stability of the soliton (see 4.35), there exists λ such that $|\lambda-1|\le\frac{1}{4}$ and $\rho(t)$ with $\frac{3}{2}t<\rho(t)<\frac{t}{2}$ for $-t$ large such that $\|u(t)-Q_\lambda(x-\rho(t))\|_{H^1(x<t/10)}\to 0$ as $t\to-\infty$. Thus, $\|Q(x-\rho_1(t))+Q_c(x-\rho_2(t))-Q_\lambda(x-\rho(t))\|_{H^1(x<t/10)}\to 0$ as $t\to-\infty$. This clearly implies that $\lambda=1$ and $\rho_2(t)>t/10$ for $-t$ large, and thus $\rho_1(t)-\rho_2(t)<t/4$.

Using $\rho_1(t)-\rho_2(t)\le-\frac{1}{4}|t|$, as before by monotonicity arguments, we have $\|u(t)-Q(x-\rho_1(t))-Q_c(x-\rho_2(t)\|_{H^1}\le K\exp(\frac{1}{8}((1-c)t-(\delta_2-\delta_1)))$ for $-t$

large. Therefore, for this solution $u(t)$, we obtain

$$|\rho_1'(t)-1|+|\rho_2'(t)-c|\leq K\exp(\tfrac{1}{8}((1-c)t-(\delta_2-\delta_1))),$$

which proves the convergence of $\rho_1(t)-t$ and $\rho_2(t)-ct$ as $t\to-\infty$. Thus, there exist δ_1, δ_2 such that (5.3) holds. We now apply the uniqueness result of [19] to conclude.

D.3. *Proofs of* (5.42)–(5.43). *Proof of* (5.42). Consider the decomposition of $u(t,x)$ introduced in the proof of Proposition 4.1; i.e., $u(t,x)=v(t,x-\rho(t))+z(t,x-\rho(t))$, $v(t)$ and $S(t)=\partial_t v+\partial_x(\partial_x^2 v-v+v^p)$ satisfy the assumptions of Proposition 4.1. Recall that $z(t)$ satisfies equation (4.13), and $\sup_{[0,T_c]}\|z(t)\|_{H^1}\leq Kc^{\theta}$, where θ is to be fixed large (for $\theta\geq\frac{5}{8}$).

First, we check that

$$\int_{x\geq 0}x^2z^2(T_c,x+\tfrac{1}{2}\Delta)dx\leq Kc^{2\theta} \tag{D.3}$$

implies the result. By the explicit expression of $v(t,x)$ in (4.1), the decay properties of Q and Q_c, and $\|\alpha\|_{L^\infty}\leq Kc^{-\frac{1}{6}}$, ($\alpha$ is defined in §4.1) we have the following pointwise estimates:

(D.4)
$$\forall t\in[0,\tfrac{1}{2}T_c],\ \forall x\geq\tfrac{1}{8}T_c,\ |v(t,x)|+|v_x(t,x)|+|S(t,x)|\leq K\exp(-c^{-r})e^{-\frac{1}{2}\sqrt{c}x},$$

$$\forall t\in[\tfrac{1}{2}T_c,T_c],\ \forall x\geq\tfrac{1}{2}\Delta,\ |v(t,x)|+|v_x(t,x)|+|S(t,x)| \leq K(e^{-\frac{9}{10}(x-\frac{1}{2}\Delta)}+\exp(-c^{-r})e^{-\frac{1}{2}\sqrt{c}x}). \tag{D.5}$$

By $u^2(T_c,x)\leq 2(v^2(T_c,x-\rho(T_c))+z^2(T_c,x-\rho(T_c)))$, $|\rho(T_c)-T_c|\leq 1$, and (D.5) at $t=T_c$, we have

$$\begin{aligned}&\int_{x\geq\frac{11}{12}|\ln c|}x^2u^2(T_c,x+T_c+\tfrac{1}{2}\Delta)dx\\ &\leq K\int_{x\geq\frac{11}{12}|\ln c|}x^2e^{-\frac{9}{5}x}dx+\exp(-\tfrac{1}{2}c^{-r})+\int_{x\geq 0}(x+1)^2z^2(T_c,x+\tfrac{1}{2}\Delta-1)dx\\ &\leq Kc^{\frac{5}{4}}+\int_{x\geq 0}x^2z^2(T_c,x+\tfrac{1}{2}\Delta)dx.\end{aligned}$$

Second, we prove (D.3), which will finish the proof of (5.42). This is proved by monotonicity arguments on $z(t)$. For $x_0>0$, $t\in[0,T_c]$, let (ψ is defined in (4.32))

$$\mathcal{M}_z(t)=\int z^2(t,x)\psi(x_z)dx,\quad\text{where}\quad x_z=x-\tfrac{1}{2}(T_c-t)-\tfrac{1}{2}\Delta-x_0.$$

Using (4.13), we have by direct calculations

$$\begin{aligned}\frac{d}{dt}\mathcal{M}_z(t) = &-3\int z_x^2\psi'(x_z) + \int z^2\psi'''(x_z)\\ &-\frac{1}{2}\int z^2\psi'(x_z) - (\rho'(t)-1)\int z^2\psi'(x_z)\\ &+\int((z+v)^4 - v^4 - z^4)(z\psi'(x_z) + z_x\psi(x_z))\\ &+\frac{4}{5}\int z^5\psi'(x_z) + (\rho'-1)\int v_x z\psi(x_z).\end{aligned}$$

By (4.32), $\|z(t)\|_{H^1} \leq Kc^\theta$ small, and then (4.12), we obtaiin

$$\begin{aligned}\frac{d}{dt}\mathcal{M}_z(t) \leq K(&\sup_{[0,T_c]}(\|z(t)\|_{L^2}^2 + \|S(t)\|_{L^2}^2)\\ &\times(\|v\psi'(x_z)\|_{L^\infty} + \|v_x\psi(x_z)\|_{L^\infty} + \|v_x\psi(x_z)\|_{L^2}).\end{aligned}$$

Therefore, by the properties of ψ and (D.4)–(D.5), we obtain

$$\frac{d}{dt}\mathcal{M}_z(t) \leq K\exp(-c^{-r})e^{-\frac{1}{2}\sqrt{c}x_0} + Kc^{2\theta}e^{-\frac{1}{4}(x_0+\frac{1}{2}(T_c-t))}.$$

Thus, by integating in $t \in [0, T_c]$, we obtain for all $x_0 > 0$,

$$\int_{x>x_0} z^2(T_c, x+\frac{1}{2}\Delta)dx \leq K\exp(-c^{-r})e^{-\frac{1}{2}\sqrt{c}x_0} + Kc^{2\theta}e^{-\frac{1}{4}x_0}.$$

Thus $\int_{x>0} x^2z^2(T_c, x+\frac{1}{2}\Delta)dx \leq Kc^{2\theta}$ and (D.3) follows.

Proof of (5.43). From (5.38), $|\rho(T_c) - T_c| \leq Kc^2$ and $\|v\|_{H^2} \leq K$, we have

$$\|u(T_c) - v(T_c, . - T_c)\|_{H^1} \leq Kc^2.$$

By (3.8)–(3.9), we have

$$\|v(T_c) - Q(. - \tfrac{\Delta}{2}) - Q_c(. + (1-c)T_c - \Delta_c/2))\|_{H_c^1} \leq Kc^{\frac{17}{12}}.$$

Thus by the decomposition of $u(T_c)$, and (5.16), we deduce (5.43).

References

[1] T. B. Benjamin, The stability of solitary waves, *Proc. Roy. Soc.* (*London*) *Ser. A* **328** (1972), 153–183. MR 0338584. http://dx.doi.org/10.1098/rspa.1972.0074.

[2] J. L. Bona, On the stability theory of solitary waves, *Proc. Roy. Soc. London Ser. A* **344** (1975), 363–374. MR 0386438. Zbl 0328.76016. http://dx.doi.org/10.1098/rspa.1975.0106.

[3] J. L. Bona, W. G. Pritchard, and L. R. Scott, Solitary-wave interaction, *Phys. Fluids* **23** (1980), 438–441. http://dx.doi.org/10.1063/1.863011.

[4] T. Cazenave and P.-L. Lions, Orbital stability of standing waves for some nonlinear Schrödinger equations, *Comm. Math. Phys.* **85** (1982), 549–561. MR 0677997. http://dx.doi.org/10.1007/BFO1403504.

[5] A. Cohen, Existence and regularity for solutions of the Korteweg-de Vries equation, *Arch. Rational Mech. Anal.* **71** (1979), 143–175. MR 0525222. http://dx.doi.org/10.1007/BF00248725.

[6] R. Côte, Construction of solutions to the subcritical gKdV equations with a given asymptotical behavior, *J. Funct. Anal.* **241** (2006), 143–211. MR 2264249. Zbl 1157.35091. http://dx.doi.org/10.1016/j.jfa.2006.04.007.

[7] ———, Construction of solutions to the L^2-critical KdV equation with a given asymptotic behaviour, *Duke Math. J.* **138** (2007), 487–531. MR 2322685. Zbl 1130.35112. http://dx.doi.org/10.1215/S0012-7094-07-13835-3.

[8] W. Craig, P. Guyenne, J. Hammack, D. Henderson, and C. Sulem, Solitary water wave interactions, *Phys. Fluids* **18** (2006), 057106, 25. MR 2259317. Zbl 1185.76463. http://dx.doi.org/10.1063/1.2205916.

[9] W. Eckhaus and P. Schuur, The emergence of solitons of the Korteweg-de Vries equation from arbitrary initial conditions, *Math. Methods Appl. Sci.* **5** (1983), 97–116. MR 0690898. Zbl 0518.35074. http://dx.doi.org/10.1002/mma.1670050108.

[10] E. Fermi, J. Pasta, and S. Ulam, Studies of nonlinear problems, I, Los Alamos Report LA1940, reproduced in *Nonlinear Wave Motion*, A.C. Newell, ed., Amer. Math. Soc., Providence, R. I., 1974, pp. 143–156.

[11] J. Hammack, D. Henderson, P. Guyenne, and Y. Ming, Solitary-wave collisions, in *Proceedings of the* 23*rd ASME Offshore Mechanics and Artic Engineering* (A symposium to honor Theodore Yao-Tsu Wu, Vancouver, Canada, June 2004), World Scientific, Singapore, 2004. Zbl 0353.70028.

[12] M. Hărăguş-Courcelle and D. H. Sattinger, Inversion of the linearized Korteweg-de Vries equation at the multi-soliton solutions, *Z. Angew. Math. Phys.* **49** (1998), 436–469. MR 1629549. Zbl 0904.35076. http://dx.doi.org/10.1007/s000000050101.

[13] R. Hirota, Exact solution of the Korteweg-de Vries equation for multiple collisions of solitons, *Phys. Rev. Lett.* **27** (1971), 1192–1194. Zbl 1168.35423.

[14] H. Kalisch and J. L. Bona, Models for internal waves in deep water, *Discrete Contin. Dynam. Systems* **6** (2000), 1–20. MR 1739590. Zbl 1021.76006.

[15] C. E. Kenig, G. Ponce, and L. Vega, Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle, *Comm. Pure Appl. Math.* **46** (1993), 527–620. MR 1211741. Zbl 0808.35128. http://dx.doi.org/10.1002/cpa.3160460405.

[16] M. D. Kruskal, The Korteweg-de Vries equation and related evolution equations, in *Nonlinear Wave Motion* (Proc. AMS-SIAM Summer Sem., Clarkson Coll. Tech., Potsdam, N.Y., 1972)) (A. C. Newell, ed.), *Lectures in Appl. Math.* **15**, Amer. Math. Soc., Providence, R.I., 1974, pp. 61–83. MR 0352741. Zbl 0292.35017.

[17] P. D. Lax, Integrals of nonlinear equations of evolution and solitary waves, *Comm. Pure Appl. Math.* **21** (1968), 467–490. MR 0235310. Zbl 0162.41103. http://dx.doi.org/10.1002/cpa.3160210503.

[18] Y. Li and D. H. Sattinger, Soliton collisions in the ion acoustic plasma equations, *J. Math. Fluid Mech.* **1** (1999), 117–130. MR 1699021. Zbl 0934.35148. http://dx.doi.org/10.1007/s000210050006.

[19] Y. Martel, Asymptotic N-soliton-like solutions of the subcritical and critical generalized Korteweg-de Vries equations, *Amer. J. Math.* **127** (2005), 1103–1140. MR 2170139. Zbl 1090.35158. Available at http://muse.jhu.edu/journals/american_journal_of_mathematics/v127/127.5martel.pdf.

[20] ———, Linear problems related to asymptotic stability of solitons of the generalized KdV equations, *SIAM J. Math. Anal.* **38** (2006), 759–781. MR 2262941. Zbl 1126.35055. http://dx.doi.org/10.1137/050637510.

[21] Y. Martel and F. Merle, Asymptotic stability of solitons for subcritical generalized KdV equations, *Arch. Ration. Mech. Anal.* **157** (2001), 219–254. MR 1826966. Zbl 0981.35073. http://dx.doi.org/10.1007/s002050100138.

[22] ———, Asymptotic stability of solitons of the subcritical gKdV equations revisited, *Nonlinearity* **18** (2005), 55–80. MR 2109467. Zbl 1064.35171. http://dx.doi.org/10.1088/0951-7715/18/1/004.

[23] ———, Asymptotic stability of solitons of the gKdV equations with general nonlinearity, *Math. Ann.* **341** (2008), 391–427. MR 2385662. Zbl 1153.35068. http://dx.doi.org/10.1007/s00208-007-0194-z.

[24] ———, Note on coupled linear systems related to two soliton collision for the quartic gKdV equation, *Rev. Mat. Complut.* **21** (2008), 327–349. MR 2441957. Zbl 1188.35164.

[25] ———, Refined asymptotics around solitons for gKdV equations, *Discrete Contin. Dyn. Syst.* **20** (2008), 177–218. MR 2358258. Zbl 1137.35062.

[26] ———, Stability of two soliton collision for nonintegrable gKdV equations, *Comm. Math. Phys.* **286** (2009), 39–79. MR 2470923. Zbl 1179.35291. http://dx.doi.org/10.1007/s00220-008-0685-0.

[27] Y. Martel, F. Merle, and T.-P. Tsai, Stability and asymptotic stability in the energy space of the sum of N solitons for subcritical gKdV equations, *Comm. Math. Phys.* **231** (2002), 347–373. MR 1946336. http://dx.doi.org/10.1007/s00220-002-0723-2.

[28] R. M. Miura, The Korteweg-de Vries equation: a survey of results, *SIAM Rev.* **18** (1976), 412–459. MR 0404890. Zbl 0333.35021. http://dx.doi.org/10.1137/1018076.

[29] T. Mizumachi, Weak interaction between solitary waves of the generalized KdV equations, *SIAM J. Math. Anal.* **35** (2003), 1042–1080. MR 2049033. Zbl 1054.35083. http://dx.doi.org/10.1137/S003614100240871X.

[30] R. L. Pego and M. I. Weinstein, Asymptotic stability of solitary waves, *Comm. Math. Phys.* **164** (1994), 305–349. MR 1289328. Zbl 0805.35117. http://dx.doi.org/10.1007/BFO2101705.

[31] G. Perelman, Asymptotic stability of multi-soliton solutions for nonlinear Schrödinger equations, *Comm. Partial Differential Equations* **29** (2004), 1051–1095. MR 2097576. http://dx.doi.org/10.1081/PDE-200033754.

[32] I. Rodnianski, W. Schlag, and A. Soffer, On asymptotic stability of n-soliton states of NLS, *Rev. Math. Phys.* **15** (2003), 877–903.

[33] P. C. Schuur, *Asymptotic Analysis of Soliton Problems*, *Lecture Notes in Math.* **1232**, Springer-Verlag, New York, 1986. MR 0874343. Zbl 0643.35003.

[34] L. Y. Shih, Soliton-like interaction governed by the generalized Korteweg-de Vries equation, *Wave Motion* **2** (1980), 197–206. MR 0589165. http://dx.doi.org/10.1016/0165-2125(80)90001-3.

[35] T. Tao, Scattering for the quartic generalised Korteweg-de Vries equation, *J. Differential Equations* **232** (2007), 623–651. MR 2286393. Zbl 1171.35107. http://dx.doi.org/10.1016/j.jde.2006.07.019.

[36] P. D. Weidman and T. Maxworthy, Experiments on strong interactions between solitary waves, *J. Fluids Mech.* **85** (1978), 417–431. http://dx.doi.org/10.1017/S0022112078000713.

[37] M. I. Weinstein, Lyapunov stability of ground states of nonlinear dispersive evolution equations, *Comm. Pure Appl. Math.* **39** (1986), 51–67. MR 0820338. Zbl 0594.35005. http://dx.doi.org/10.1002/cpa.3160390103.

[38] N. J. Zabusky, Solitons and energy transport in nonlinear lattices, *Comput. Phys. Comm.* **5** (1973), 1–10. http://dx.doi.org/10.1016/0010-4655(73)90002-7.

[39] N. J. Zabusky and M. D. Kruskal, Interaction of "solitons" in a collisionless plasma and recurrence of initial states, *Phys. Rev. Lett.* **15** (1965), 240–243.

(Received: June 1, 2007)

Université de Versailles St-Quentin,
Laboratoire de mathématiques de Versailles, Versailles cedex, France
E-mail: yvan.martel@math.uvsq.fr

Université de Cergy-Pontoise, IHES and CNRS,
Laboratoire de mathématiques, Cergy-Pontoise cedex, France
E-mail: frank.merle@math.u-cergy.fr