HAL open science

THE SHORT PULSE EQUATION BY A RIEMANN-HILBERT APPROACH

Anne Boutet de Monvel, Dmitry Shepelsky, Lech Zielinski

To cite this version:

Anne Boutet de Monvel, Dmitry Shepelsky, Lech Zielinski. THE SHORT PULSE EQUATION BY A RIEMANN-HILBERT APPROACH. 2016. hal-03662513

HAL Id: hal-03662513

https://hal.science/hal-03662513v1

Preprint submitted on 9 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THE SHORT PULSE EQUATION BY A RIEMANN–HILBERT APPROACH

ANNE BOUTET DE MONVEL*, DMITRY SHEPELSKY†, AND LECH ZIELINSKI‡

Abstract. We develop a Riemann–Hilbert approach to the inverse scattering transform method for the short pulse (SP) equation

$$u_{xt} = u + \tfrac{1}{6}(u^3)_{xx}$$

with zero boundary conditions (as $|x| \to \infty$). This approach is directly applied to a Lax pair for the SP equation. It allows us to give a parametric representation of the solution to the Cauchy problem. This representation is then used for studying the long-time behavior of the solution as well as for retrieving the soliton solutions. Finally, the analysis of the long-time behavior allows us to formulate, in spectral terms, a sufficient condition for the wave breaking.

arXiv:1608.02249v2 [nlin.SI] 6 Sep 2016

1. Introduction

It is well-known that for describing the slow modulation of the amplitude of a weakly nonlinear wave packet in a moving medium, the nonlinear Schrödinger (NLS) equation is one of the universal nonlinear integrable models. It has been used with great success in nonlinear optics to describe the propagation of sufficiently broad pulses, or slowly varying wave trains whose spectra are narrowly localized around the carrier frequency. However, high-speed fiber-optic communication demands ultra-short pulses. With this respect, certain technological progress for creating them has been achieved; but it is important that in these conditions, the description of the evolution of these pulses lies beyond the usual approximations leading to the NLS equation.

The short pulse (SP) equation

$$u_{xt} = u + \frac{1}{6}(u^3)_{xx} \tag{1.1}$$

was proposed by Schäfer and Wayne [32, 12] as an alternative (to the NLS equation) model for approximating the evolution of ultra-short intense infrared pulses in silica optics. It was shown in [12] by numerical simulations that the SP equation can be successfully used for describing pulses with broad spectrum.

The short pulse equation is formally integrable: it is the compatibility condition for a pair of linear equations (the Lax pair), see [30]:

$$\Phi_x = U\Phi \tag{1.2a}$$

$$\Phi_t = V\Phi, \tag{1.2b}$$

where U and V are 2×2 matrices dependent on the spectral parameter λ:

$$U = \begin{pmatrix} \lambda & \lambda u_x \\ \lambda u_x & -\lambda \end{pmatrix}, \tag{1.3a}$$

$$V = \begin{pmatrix} \frac{\lambda}{2}u^2 + \frac{1}{4\lambda} & \frac{\lambda}{2}u^2 u_x - \frac{1}{2}u \\ \frac{\lambda}{2}u^2 u_x + \frac{1}{2}u & -\frac{\lambda}{2}u^2 - \frac{1}{4\lambda} \end{pmatrix}. \tag{1.3b}$$

The variants of application of the inverse scattering transform (IST) method to the SP equation, known in the literature, rely on establishing end exploiting the relationship between the

Date: September 7, 2016.

2010 *Mathematics Subject Classification.* Primary: 35Q53; Secondary: 37K15, 35Q15, 35B40, 35Q51, 37K40.

Key words and phrases. Short pulse equation, short wave equation, Camassa–Holm type equation, inverse scattering transform, Riemann–Hilbert problem.

SP equation and other integrable equations (like the sine-Gordon equation, see [31]). But such relations turn out to be rather complicated and implicit, which, in particular, makes it difficult to apply them for studying initial value problems with general initial data.

In the present paper we propose a direct approach to the problem of integration of the SP equation, which is based on applying the inverse scattering transform method, in the form of an associated Riemann–Hilbert (RH) problem. This means that the construction of this RH problem is made in terms of dedicated solutions of the Lax pair equations associated directly to the SP equation.

It is interesting to notice that the short pulse equation can be viewed as the short wave approximation to another integrable equation

$$m_t + \big((u^2 - u_x^2)m\big)_x + u_x = 0, \qquad m := u - u_{xx}, \tag{1.4}$$

usually referred to as the "modified Camassa–Holm equation", and also known as the "Fuchssteiner–Olver–Rosenau–Qiao" equation [17, 18, 27, 29]. Indeed, introducing the new variables

$$x' = \frac{x}{\varepsilon}, \quad t' = t\varepsilon, \quad u' = \frac{u}{\varepsilon^2},$$

passing to the limit $\varepsilon \to 0$ and retaining the main terms reduce (1.4) to (1.1). With this respect we notice that the same short-wave limit applied to the Camassa–Holm (CH) equation

$$u_t - u_{txx} + 2u_x + 3uu_x = 2u_x u_{xx} + uu_{xxx} \tag{1.5}$$

leads to the so-called short wave (SW) equation

$$u_{txx} - 2u_x + 2u_x u_{xx} + uu_{xxx} = 0. \tag{1.6}$$

The RH approach to the inverse scattering method for the CH equation and the SW equation were presented in [3, 5] and in [8], respectively. In what follows we will see that the development of this approach for the SP equation, on one hand, shares many common features with that for the SW equation, but on the other hand, has important differences.

Various aspects of the SP equation have been addressed in the literature, including the construction of solitary wave solutions [19, 24, 31] and periodic solutions [25]. Well-posedness of the Cauchy problem has been studied in [13, 28, 32]. Certain sufficient conditions for wave breaking have been found in [23]. An integrable hierarchy of equations associated with the SP equation is discussed in [9]. The bi-Hamiltonian structure of the SP equation is presented in [10].

In this paper we present a RH problem formalism for the inverse scattering approach to the initial value problem for the SP equation:

$$u_{xt} = u + \frac{1}{6}(u^3)_{xx}, \quad t > 0, \quad -\infty < x < +\infty, \tag{1.7a}$$

$$u(x,0) = u_0(x), \qquad -\infty < x < +\infty. \tag{1.7b}$$

We assume that $u_0(x)$ decays to 0 sufficiently fast:

$$u_0(x) \to 0, \qquad x \to \pm\infty,$$

and we seek a solution $u(x,t)$ decaying to 0 for all $t > 0$:

$$u(x,t) \to 0, \qquad x \to \pm\infty.$$

In Section 2 we present appropriate Lax pairs associated with the SP equation, whose dedicated solutions are used in Section 3 for formulating a matrix Riemann–Hilbert problem suitable for solving the Cauchy problem (1.7). Then we give (Theorem 3.4) a representation of the solution $u(x,t)$ of the problem (1.7) in terms of the solution of this RH problem evaluated at a distinguished point of the complex plane of the spectral parameter. In Section 4 we discuss the construction of soliton solutions using the formalism of the RH problem. In Section 5 we study the long time asymptotics of the solution of the Cauchy problem (1.7). This study is then used in Section 5 to provide a sufficient condition for wave breaking of the solution of the Cauchy problem at a finite time.

2. Lax pairs and eigenfunctions

The RH formalism for integrable nonlinear equations makes use of the fact that it is possible to construct dedicated solutions of the linear equations from the associated Lax pair, which, being considered all together, are well-controlled, as functions of the spectral parameter, in the whole extended complex plane. These solutions are differently defined for different domains in the complex plane, and are related to each other at the boundaries between these domains. The latter fact is then interpreted as the "analyticity defect" for a (matrix-valued) function of the spectral parameter viewed as a function in the whole complex plane, and the inverse problem of the IST method for solving the Cauchy problem for the nonlinear equation in question is realized as a Riemann–Hilbert-type problem of reconstructing a piece-wise analytic function from the known "analyticity defects" of this function, in the form of jump conditions across certain contours supplemented by residue conditions (if any) at certain points in the complex plane of the spectral parameter.

An efficient approach to constructing such solution of the differential equations from the Lax pair is to pass to integral equations, whose solutions are particular solutions to the Lax pair equations. For this purpose, it is convenient to transform the Lax pair equations to a certain form, which is standard for establishing analytic properties of solutions near the singular points (w.r.t. the spectral parameter) of the Lax pair equations.

Notice that the coefficient matrices U and V are traceless, which provides that the determinant of a matrix solution to (1.2) composed from two vector solutions is independent of x and t.

In order to make the presentation close to that in the cases of other CH-type equations (see [3, 5, 8, 6]), it is convenient to introduce the spectral parameter $k := \mathrm{i}\lambda$.

Now notice that U and V have singularities (in the extended complex k-plane) at $k = 0$ and at $k = \infty$. In order to control the behavior of solutions to (1.2) as functions of the spectral parameter k, we follow a strategy similar to that adopted for the CH equation [3, 5].

Namely, in order to control the large k behavior of solutions of (1.2), we will transform this Lax pair to the following form (see [1, 3, 5]):

$$\hat{\Phi}_x + Q_x\hat{\Phi} = \hat{U}\hat{\Phi}, \tag{2.1a}$$

$$\hat{\Phi}_t + Q_t\hat{\Phi} = \hat{V}\hat{\Phi}, \tag{2.1b}$$

whose coefficients $Q(x,t,k)$, $\hat{U}(x,t,k)$, and $\hat{V}(x,t,k)$ have the following properties:

(i) Q is diagonal and is unbounded as $k \to \infty$.
(ii) $\hat{U} = \mathrm{O}(1)$ and $\hat{V} = \mathrm{O}(1)$ as $k \to \infty$.
(iii) The diagonal parts of $\hat{U}$ and $\hat{V}$ decay as $k \to \infty$.
(iv) $\hat{U} \to 0$ and $\hat{V} \to 0$ as $x \to \pm\infty$.

Since the coefficient matrix U in (1.3a) is the product of the spectral parameter and a matrix independent of the spectral parameter, in order to obtain Q, one has to diagonalize the latter matrix, i.e., to determine $P(x,t)$ such that

$$PUP^{-1} = -Q.$$

The freedom in determining such P can be used in order to satisfy item (iii) above, or, more precisely, to make the diagonal part of $\hat{U}$ identically 0.

Indeed, introducing

$$w := u_x; \quad q := \sqrt{1+w^2}, \tag{2.2}$$

setting

$$P := \sqrt{\frac{1+q}{2q}}\begin{pmatrix} 1 & \frac{w}{1+q} \\ -\frac{w}{1+q} & 1 \end{pmatrix} \tag{2.3}$$

so that $P^{-1} = \sqrt{\frac{1+q}{2q}}\begin{pmatrix} 1 & -\frac{w}{1+q} \\ \frac{w}{1+q} & 1 \end{pmatrix}$, and introducing

$$\hat{\Phi} := P\Phi$$

reduces (1.2a) to (2.1a) with

$$Q_x(x,t,k) = \mathrm{i}kq(x,t)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \equiv \mathrm{i}kq(x,t)\sigma_3 \tag{2.4}$$

and

$$\hat{U} = \hat{U}(x,t) = \frac{u_{xx}}{2q^2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}. \tag{2.5}$$

Accordingly, the t-equation (1.2b) of the Lax pair is transformed into

$$\hat{\Phi}_t = \check{V}\Phi, \tag{2.6}$$

where

$$\check{V} = -\frac{\mathrm{i}ku^2q}{2}\sigma_3 - \frac{1}{4\mathrm{i}kq}\begin{pmatrix}1 & -w\\ -w & -1\end{pmatrix} + \frac{u^2u_{xx}}{4q^2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}. \tag{2.7}$$

Noticing that $\check{V} \to -\frac{1}{4\mathrm{i}k}\sigma_3$ as $x \to \pm\infty$, we write $\check{V}$ as

$$\begin{aligned}\check{V} &= \left(-\frac{\mathrm{i}ku^2q}{2} - \frac{1}{4\mathrm{i}k}\right)\sigma_3 - \frac{1}{4\mathrm{i}kq}\begin{pmatrix}1-q & -w\\ -w & -1+q\end{pmatrix} + \frac{u^2u_{xx}}{4q^2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\\ &= \left(-\frac{\mathrm{i}ku^2q}{2} - \frac{1}{4\mathrm{i}k}\right)\sigma_3 + \hat{V},\end{aligned}$$

where

$$\hat{V} := -\frac{1}{4\mathrm{i}kq}\begin{pmatrix}1-q & -w\\ -w & -1+q\end{pmatrix} + \frac{u^2u_{xx}}{4q^2}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},$$

and thus (2.6) takes the form (2.1b) provided Q is defined in such a way that

$$Q_t = \left(\frac{\mathrm{i}ku^2q}{2} + \frac{1}{4\mathrm{i}k}\right)\sigma_3. \tag{2.8}$$

Now notice that (2.4) and (2.8) are compatible since the compatibility condition $Q_{xt} = Q_{tx}$ reads

$$q_t = \frac{1}{2}(u^2q)_x, \tag{2.9}$$

which is the "conservation law" form of the short pulse equation (1.1) (notice that the "conservation law" form of the SW equation (1.6) has a very similar form: $q_t = \frac{1}{2}(uq)_x$, but in that case, $q = \sqrt{1-u_{xx}}$). Thus Q can be correctly defined by

$$Q(x,t,k) = \left(\mathrm{i}k\hat{x}(x,t) + \frac{t}{4\mathrm{i}k}\right)\sigma_3, \tag{2.10}$$

where

$$\hat{x}(x,t) := x - \int_x^\infty (q(y,t)-1)\mathrm{d}y. \tag{2.11}$$

Introduce

$$\widetilde{\Phi} = \hat{\Phi}\mathrm{e}^Q \tag{2.12}$$

and think about $\widetilde{\Phi}$ as a 2×2 matrix. Then (2.1) can be rewritten as

$$\widetilde{\Phi}_x + [Q_x, \widetilde{\Phi}] = \hat{U}\widetilde{\Phi}, \tag{2.13a}$$

$$\widetilde{\Phi}_t + [Q_t, \widetilde{\Phi}] = \hat{V}\widetilde{\Phi}, \tag{2.13b}$$

where $[\,\cdot\,,\,\cdot\,]$ denotes the matrix commutator.

Now determine the particular (Jost) solutions $\widetilde{\Phi}_\pm(x,t)$ of (2.13) as the 2×2 matrix-valued solutions of the associated Volterra integral equations:

$$\widetilde{\Phi}_\pm(x,t,k) = I + \int_{\pm\infty}^x \mathrm{e}^{Q(y,t,k)-Q(x,t,k)}\hat{U}(y,t,k)\widetilde{\Phi}_\pm(y,t,k)\mathrm{e}^{Q(x,t,k)-Q(y,t,k)}\mathrm{d}y, \tag{2.14}$$

or, taking into account the definition of Q,

$$\begin{aligned}\widetilde{\Phi}_+(x,t,k) &= I - \int_x^\infty \mathrm{e}^{\mathrm{i}k\int_x^y q(\xi,t)\mathrm{d}\xi\,\sigma_3}\hat{U}(y,t,k)\widetilde{\Phi}_+(y,t,k)\mathrm{e}^{-\mathrm{i}k\int_x^y q(\xi,t)\mathrm{d}\xi\,\sigma_3}\mathrm{d}y,\\ \widetilde{\Phi}_-(x,t,k) &= I + \int_{-\infty}^x \mathrm{e}^{-\mathrm{i}k\int_y^x q(\xi,t)\mathrm{d}\xi\,\sigma_3}\hat{U}(y,t,k)\widetilde{\Phi}_-(y,t,k)\mathrm{e}^{\mathrm{i}k\int_y^x q(\xi,t)\mathrm{d}\xi\,\sigma_3}\mathrm{d}y,\end{aligned} \tag{2.15}$$

where I is the identity matrix.

We denote by $\mu^{(1)}$ and $\mu^{(2)}$ the columns of a 2×2 matrix $\mu = \begin{pmatrix}\mu^{(1)} & \mu^{(2)}\end{pmatrix}$. Then it follows from (2.14) that for all (x,t):

(i) $\det\widetilde{\Phi}_\pm \equiv 1$.
(ii) $\widetilde{\Phi}_-^{(1)}$ and $\widetilde{\Phi}_+^{(2)}$ are analytic in $\{k \mid \operatorname{Im} k > 0$ and continuous in $\{k \mid \operatorname{Im} k \geq 0,\ k \neq 0\}$.
(iii) $\widetilde{\Phi}_+^{(1)}$ and $\widetilde{\Phi}_-^{(2)}$ are analytic in $\{k \mid \operatorname{Im} k < 0$ and continuous in $\{k \mid \operatorname{Im} k \leq 0,\ k \neq 0\}$.
(iv) $\begin{pmatrix}\widetilde{\Phi}_-^{(1)} & \widetilde{\Phi}_+^{(2)}\end{pmatrix} \to I$ as $k\to\infty$ in $\{k \mid \operatorname{Im} k \geq 0\}$.
(v) $\begin{pmatrix}\widetilde{\Phi}_+^{(1)} & \widetilde{\Phi}_-^{(2)}\end{pmatrix} \to I$ as $k\to\infty$ in $\{k \mid \operatorname{Im} k \leq 0\}$.
(vi) Symmetries:

$$\overline{\widetilde{\Phi}_\pm(\,\cdot\,,\,\cdot\,,\bar{k})} = \widetilde{\Phi}_\pm(\,\cdot\,,\,\cdot\,,-k) = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\widetilde{\Phi}_\pm(\,\cdot\,,\,\cdot\,,k)\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}. \tag{2.16}$$

Overline means complex conjugation for all k for which the functions above are defined. The latter property is due to the symmetries of the coefficient matrix $\check{U} := \hat{U} - \mathrm{i}q\sigma_3$

$$\overline{\check{U}(\,\cdot\,,\,\cdot\,,\bar{k})} = \check{U}(\,\cdot\,,\,\cdot\,,-k) = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\check{U}(\,\cdot\,,\,\cdot\,,k)\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}. \tag{2.17}$$

Remark 2.1. Introducing the new variable $\hat{x}$ as in (2.11) and taking into account the bijectivity of the map $x \mapsto \hat{x}$ for any $t \geq 0$ (which is due to the fact that $q > 0$), equation (2.1a) reduces to the (non-self-adjoint) Dirac equation for $\hat{\hat{\Phi}}(\hat{x},t,k) := \hat{\Phi}(x(\hat{x},t),t,k)$:

$$\hat{\hat{\Phi}}_{\hat{x}} + \mathrm{i}k\sigma_3\hat{\hat{\Phi}} = \hat{\hat{U}}\hat{\hat{\Phi}}, \tag{2.18}$$

where

$$\hat{\hat{U}} = \frac{u_{xx}}{2q^3}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \tag{2.19}$$

which is the spatial equation from the Lax pair associated with the *focusing* nonlinear Schrödinger (f NLS) equation; see, e.g., [16]. Therefore, the analytical properties of $\tilde{\Phi}_\pm$ stated above are the same as in the case of the f NLS equation.

Remark 2.2. In the case of the SW equation (1.6), the spatial equation from the Lax pair is also the Dirac equation (like (2.18)), but with a self-adjoint potential $\hat{\hat{U}}$, see [8].

The scattering matrix $s(k)$ (independent of (x,t)) is introduced by

$$\widetilde{\Phi}_+(x,t,k) = \widetilde{\Phi}_-(x,t,k)\mathrm{e}^{-Q(x,t,k)\sigma_3}s(k)\mathrm{e}^{Q(x,t,k)\sigma_3}, \qquad k\in\mathbb{R} \tag{2.20}$$

with Q defined by (2.10), which, due to the symmetries (2.17), can be written in terms of two scalar spectral functions, $a(k)$ and $b(k)$:

$$s(k) = \begin{pmatrix}\overline{a(k)} & b(k)\\ -\overline{b(k)} & a(k)\end{pmatrix}, \qquad k\in\mathbb{R}, \tag{2.21}$$

such that $\overline{a(k)} = a(-k)$ and $\overline{b(k)} = b(-k)$. In view of Remark 2.1, the spectral functions have properties similar to those in the case of the focusing NLS equation [16]:

a) $a(k)$ and $b(k)$ are determined by $u(x,0)$ through the solutions $\widetilde{\Phi}_\pm(x,0)$ of equations (2.15), where $\hat{U} = \hat{U}(x,0)$ is defined by (2.5) with u replaced by $u_0(x)$ (and similarly for q).

b) $a(k)$ is analytic in $\{k \mid \operatorname{Im} k > 0\}$ and continuous in $\{k \mid \operatorname{Im} k \geq 0\}$; moreover, $a(k) \to 1$ as $k \to \infty$.
c) $b(k)$ is continuous for $k \in \mathbb{R}$, and $b(k) \to 0$ as $|k| \to \infty$.
d) $|a(k)|^2 + |b(k)|^2 = 1$ for $k \in \mathbb{R}$.
e) Let $\{k_j\}_1^N$ be the set of zeros of $a(k)$. We make the *genericity assumption* that these zeros are finite in number, simple, and no zero is real. Then $\widetilde{\Phi}_-^{(1)}(x,t,k_j)$ and $\widetilde{\Phi}_+^{(2)}(x,t,k_j)$ are linearly dependent; moreover,

$$\widetilde{\Phi}_-^{(1)}(x,t,k_j) = \mathrm{e}^{2\mathrm{i}(k_j\hat{x} - \frac{t}{4k_j})}\widetilde{\Phi}_+^{(2)}(x,t,k_j)\chi_j \tag{2.22}$$

with some constants χ_j.

3. The Riemann–Hilbert problem

3.1. **A RH problem constructed from dedicated eigenfunctions.** The analytic properties of $\widetilde{\Phi}_\pm$ stated above allow rewriting the scattering relation (2.20) as a jump relation for a piecewise meromorphic (w.r.t. k), 2×2-valued function (depending on x and t as parameters). Indeed, define $M(x,t,k)$ by

$$M(x,t,k) = \begin{cases} \begin{pmatrix} \frac{\widetilde{\Phi}_-^{(1)}(x,t,k)}{a(k)} & \widetilde{\Phi}_+^{(2)}(x,t,k) \end{pmatrix}, & \operatorname{Im} k > 0, \\ \begin{pmatrix} \widetilde{\Phi}_+^{(1)}(x,t,k) & \frac{\widetilde{\Phi}_-^{(2)}(x,t,k)}{\overline{a(k)}} \end{pmatrix}, & \operatorname{Im} k < 0. \end{cases} \tag{3.1}$$

Define

$$r(k) := -\frac{\overline{b(k)}}{a(k)} \quad \text{for } k \in \mathbb{R}. \tag{3.2}$$

Then the limiting values $M_\pm(x,t,k)$, $k \in \mathbb{R}$ of M as k is approached from the domains $\pm \operatorname{Im} k > 0$ are related as follows:

$$M_+(x,t,k) = M_-(x,t,k)\mathrm{e}^{-Q(x,t,k)\sigma_3} J_0(k) \mathrm{e}^{Q(x,t,k)\sigma_3}, \quad k \in \mathbb{R}, \tag{3.3}$$

where

$$J_0(k) = \begin{pmatrix} 1 + |r(k)|^2 & \overline{r(k)} \\ r(k) & 1 \end{pmatrix}. \tag{3.4}$$

Taking into account the properties of $\widetilde{\Phi}_\pm$ and $s(k)$, $M(x,t,k)$ satisfies the following properties:

(i) $\det M \equiv 1$.
(ii) Normalization: $M(\,\cdot\,,\,\cdot\,,k) \to I$ as $k \to \infty$.
(iii) Symmetries:

$$\overline{M(\,\cdot\,,\,\cdot\,,\bar{k})} = M(\,\cdot\,,\,\cdot\,,-k) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} M(\,\cdot\,,\,\cdot\,,k) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.5}$$

(iv) $M^{(1)}$ has poles at the zeros k_j of $a(k)$ (in $\{k \mid \operatorname{Im} k > 0\}$), whereas $M^{(2)}$ has poles at the conjugates $\overline{k_j}$ (in $\{k \mid \operatorname{Im} k < 0\}$), $j = 1, 2, \ldots, N$, and the following residue conditions are satisfied:

$$\begin{aligned} \operatorname{Res}_{k=k_j} M^{(1)}(x,t,k) &= \mathrm{i}\gamma_j \mathrm{e}^{2\mathrm{i}\left(k_j x(\hat{x},t) - \frac{t}{4k_j}\right)} M^{(2)}(x,t,k_j), \\ \operatorname{Res}_{k=\bar{k}_j} M^{(2)}(x,t,k) &= \mathrm{i}\bar{\gamma}_j \mathrm{e}^{2\mathrm{i}\left(\bar{k}_j \hat{x}(x,t) - \frac{t}{4k_j}\right)} M^{(1)}(x,t,\bar{k}_j) \end{aligned} \tag{3.6}$$

with some constants γ_j.

The idea of the Riemann–Hilbert problem approach in the inverse scattering method consists in considering the jump relation (3.3) complemented by the normalization condition $M \to I$ as $k \to \infty$, and by the residue conditions (3.6), as the factorization problem of finding $M(x,t,k)$ (and, consequently, $u(x,t)$) from the jump matrix in (3.3) and the residue conditions (3.6) at the singularities of M. As in the case of any Camassa–Holm-type equation, when realizing this idea, one faces the problem that the determination of the jump matrix, which is $\mathrm{e}^{-Q} J_0(k) \mathrm{e}^{Q}$, involves

not only objects uniquely determined by the initial data $u(x,0)$ (the functions $a(k)$ and $b(k)$ involved in $J_0(k)$ and the constants involved in the residue conditions), but also $Q = Q(x,t,k)$, which is obviously not determined by $u(x,0)$ (it involves $u(x,t)$ for $t \geq 0$). This problem can be resolved by considering a RH problem depending, instead of (x,t), on the parameters $\hat{x}$ and t. Then the jump matrix and the residue conditions become explicit in the new variables $\hat{x}$ and t. Indeed, introducing

$$\hat{M}(\hat{x},t,k) := M(x(\hat{x},t),t,k),$$

the jump condition takes the form

$$\hat{M}_+(\hat{x},t,k) = \hat{M}_-(\hat{x},t,k)J(\hat{x},t,k), \qquad k \in \mathbb{R}, \tag{3.7}$$

where

$$J(\hat{x},t,k) := \mathrm{e}^{-\hat{Q}(\hat{x},t,k)}J_0(k)\mathrm{e}^{\hat{Q}(\hat{x},t,k)} \tag{3.8}$$

with $J_0(k)$ as in (3.4) and

$$\hat{Q}(\hat{x},t,k) := \left(\mathrm{i}k\hat{x} + \frac{t}{4\mathrm{i}k}\right)\sigma_3, \tag{3.9}$$

Accordingly, the residue conditions (3.6) take the form

$$\begin{aligned} \mathrm{Res}_{k=k_j}\, \hat{M}^{(1)}(\hat{x},t,k) &= \mathrm{i}\gamma_j \mathrm{e}^{2\mathrm{i}\left(k_j\hat{x} - \frac{t}{4k_j}\right)}\hat{M}^{(2)}(\hat{x},t,k_j), \\ \mathrm{Res}_{k=\bar{k}_j}\, \hat{M}^{(2)}(\hat{x},t,k) &= \mathrm{i}\bar{\gamma}_j \mathrm{e}^{-2\mathrm{i}\left(\bar{k}_j\hat{x} - \frac{t}{4k_j}\right)}\hat{M}^{(1)}(\hat{x},t,\bar{k}_j). \end{aligned} \tag{3.10}$$

Recall that the jump and residue conditions for $\hat{M}(\hat{x},t,k)$ were obtained above assuming that there exists a solution $u(x,t)$ of the SP equation decaying to 0 as $x \to \pm\infty$ for any fixed $t > 0$. On the other hand, the conditions (3.7)–(3.10) can be viewed as a factorization problem of Riemann–Hilbert (RH) type:

RH problem. Given $\{r(k),\, k \in \mathbb{R}; \{k_j, \gamma_j\}_1^N\}$, find a piece-wise (w.r.t. $\mathbb{R}$) meromorphic function $\hat{M}(\hat{x},t,k)$ satisfying the conditions (3.7)–(3.10) complemented by the normalization condition

$$\hat{M}(\hat{x},t,k) \to I \quad \text{ as } k \to \infty. \tag{3.11}$$

Remark 3.1 (symmetries). Since the jump matrix J satisfies the symmetry conditions described in (3.5), it follows from the uniqueness of the solution $\hat{M}$ of the RH problem that this solution satisfies (3.5) as well.

Remark 3.2 (unique solvability). The structure of the jump matrix and the residue conditions are the same as in the case of the focusing NLS equation (only the dependence on $\hat{x}$ and t, which are just the *parameters* for the RH problem, is different), which implies that for all $\hat{x}$ and t, there exists a unique solution of the RH problem (3.7)–(3.11) provided that $r(k)$ can be represented as $r(k) = \int_{-\infty}^{\infty} \hat{r}(s)\mathrm{e}^{\mathrm{i}ks}\mathrm{d}s$ with some $\hat{r}(s) \in L^1(-\infty,\infty)$, see [16].

3.2. **Recovering the solution of the Cauchy problem from the associated RH problem.** Now our goal is to show that $u(x,t)$ can be recovered in terms of $\hat{M}(\hat{x},t,k)$, which is considered as the solution of the Riemann–Hilbert problem (3.7)–(3.11) (notice that the data for this problem are uniquely determined by $u(x,0)$), evaluated at $k = 0$. Indeed this value of k is specific for (1.2a), since the coefficient matrix U in (1.2a) vanish identically at $k = 0$.

In order to have a good control of the behavior of $\hat{M}(\hat{x},t,k)$ as $k \to 0$, it is convenient to rewrite the Lax pair (1.2) in the form

$$\Phi_x + \mathrm{i}k\sigma_3\Phi = U_0\Phi \tag{3.12a}$$

$$\Phi_t + \frac{1}{4\mathrm{i}k}\sigma_3\Phi = V_0\Phi, \tag{3.12b}$$

where

$$U_0 = -\mathrm{i}kw \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{3.13a}$$

$$V_0 = -\frac{\mathrm{i}ku^2}{2} \begin{pmatrix} 1 & w \\ w & -1 \end{pmatrix} + \frac{u}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.13b}$$

Notice that $U_0 \to 0$ and $V_0 \to 0$ as $|x| \to \infty$. Besides, it is important that $U_0(x,t,0) \equiv 0$.

Introduce

$$Q_0(x,t,k) := \left(\mathrm{i}kx + \frac{t}{4\mathrm{i}k}\right)\sigma_3 \tag{3.14}$$

and

$$\widetilde{\Phi}_0 = \Phi \mathrm{e}^{Q_0}. \tag{3.15}$$

Then (3.12) can be rewritten as

$$\widetilde{\Phi}_{0x} + [Q_{0x}, \widetilde{\Phi}_0] = U_0 \widetilde{\Phi}_0, \tag{3.16a}$$

$$\widetilde{\Phi}_{0t} + [Q_{0t}, \widetilde{\Phi}_0] = V_0 \widetilde{\Phi}_0. \tag{3.16b}$$

The Jost solutions $\widetilde{\Phi}_{0\pm}(x,t,k)$ of (3.16) are determined, similarly to above, as the solutions of associated Volterra integral equations:

$$\widetilde{\Phi}_{0\pm}(x,t,k) = I + \int_{\pm\infty}^{x} \mathrm{e}^{\mathrm{i}k(y-x)} U_0(y,t,k) \widetilde{\Phi}_{0\pm}(y,t,k) \mathrm{e}^{\mathrm{i}k(x-y)} \mathrm{d}y. \tag{3.17}$$

Now, since $U_0(x,t,0) \equiv 0$, we have the following important property:

$$\widetilde{\Phi}_{0\pm}(x,t,0) \equiv I \tag{3.18}$$

for all x and t. Moreover, directly using (3.17) one obtains

Proposition 3.3. *As $k \to 0$,*

$$\widetilde{\Phi}_{0\pm}(x,t,k) = I - \mathrm{i}ku(x,t) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \mathrm{O}(k^2).$$

Further, we notice that $\widetilde{\Phi}_\pm$ and $\widetilde{\Phi}_{0\pm}$, being related to the same system of equations (1.2), are related as

$$\widetilde{\Phi}_\pm(x,t,k) = P(x,t) \widetilde{\Phi}_{0\pm}(x,t,k) \mathrm{e}^{-Q_0(x,t,k)} C_\pm(k) \mathrm{e}^{Q(x,t,k)}, \tag{3.19}$$

where $C_\pm(k)$ are some matrices independent of x and t. Passing to the limits $x \to \pm\infty$ determines $C_\pm(k)$:

$$C_+(k) = I, \qquad C_-(k) = \mathrm{e}^{\mathrm{i}k\alpha}\sigma_3,$$

where $\alpha = \int_{-\infty}^{\infty} (q(y,t) - 1)\mathrm{d}y$; notice that in view of (2.9), α is constant (does not depend on t).

Combining Proposition 3.3 with (3.19) one gets

$$\begin{aligned} \widetilde{\Phi}_+(x,t,k) &= P(x,t)\left(I - \mathrm{i}k\left(u(x,t)\sigma_1 + \int_x^\infty (q(y,t)-1)\mathrm{d}y\ \sigma_3\right) + \mathrm{O}(k^2)\right), \\ \widetilde{\Phi}_-(x,t,k) &= P(x,t)\left(I - \mathrm{i}k\left(u(x,t)\sigma_1 - \int_{-\infty}^x (q(y,t)-1)\mathrm{d}y\ \sigma_3\right) + \mathrm{O}(k^2)\right) \end{aligned} \tag{3.20}$$

as $k \to 0$. Using these expansions in (2.20) we expand $s(k)$ at $k = 0$, then by (2.21) we obtain

$$a(k) = 1 + \mathrm{i}k\alpha + \mathrm{O}(k^2), \qquad b(k) = \mathrm{O}(k^2), \quad k \to 0. \tag{3.21}$$

Finally, substituting (3.20) and (3.21) into (3.1) gives

$$M(x,t,k) = P(x,t)\left(I - \mathrm{i}k\left(u(x,t)\sigma_1 + \int_x^\infty (q(y,t)-1)\mathrm{d}y\ \sigma_3\right) + \mathrm{O}(k^2)\right), \quad k \to 0, \tag{3.22}$$

which, in view of (2.11), reads

$$M(x,t,k) = P(x,t)\left(I - \mathrm{i}k\begin{pmatrix} x - \hat{x} & u \\ u & \hat{x} - x \end{pmatrix} + \mathrm{O}(k^2)\right), \quad k \to 0. \tag{3.23}$$

This relation (3.23) leads to the following result.

Theorem 3.4 (representation result)**.** *Assume that the Cauchy problem* (1.7) *for the SP equation has a solution* $u(x,t)$*s. Let* $\{r(k),\, k \in \mathbb{R};\, \{k_j, \gamma_j\}_1^N\}$ *be the spectral data determined by* $u_0(x)$*, and let* $\hat{M}(\hat{x}, t, k)$ *be the solution of the associated RH problem* (3.7)–(3.11)*. Then, evaluating* $\hat{M}$ *as* $k \to 0$*, we get a parametric representation for the solution* $u(x,t)$ *of the Cauchy problem* (1.7)*:*

$$u(x,t) = \hat{u}(\hat{x}(x,t), t) \tag{3.24a}$$

with

$$x(\hat{x}, t) = \hat{x} + f_1(\hat{x}, t), \tag{3.24b}$$

$$\hat{u}(\hat{x}, t) = f_2(\hat{x}, t), \tag{3.24c}$$

where

$$\begin{pmatrix} f_1 & f_2 \\ f_2 & -f_1 \end{pmatrix} (\hat{x}, t) := \lim_{k \to 0} \frac{\mathrm{i}}{k} \left(\hat{M}^{-1}(\hat{x}, t, 0) \hat{M}(\hat{x}, t, k) - I \right). \tag{3.24d}$$

Remark 3.5. The representation result of Theorem 3.4 can be interpreted in two ways:

i) If there is a global (in time) classical solution $u(x,t)$ of the Cauchy problem (1.7), then (3.24) gives a parametric representation of this solution for all t.
ii) If wave breaking occurs at a finite time, then the bijectivity of the map $\hat{x} \mapsto x$ described by (3.24b) is broken for certain values of t. Then (3.24b) and (3.24c) present a continuation of the solution of the Cauchy problem (1.7) after the wave breaking. Particularly, if the bijectivity of the map $\hat{x} \mapsto x$ is restored for all times t greater than a certain T, (3.24b) and (3.24c) present a solution that, after undergoing a sequence of wave breakings, retrieves the form of a classical solution to the SP equation (1.7a).

Remark 3.6. We emphasize that the wave breaking mechanism for the SP equation is exclusively related to the break of bijectivity in (3.24b) while the solution on the variables $(\hat{x}, t)$ always exists globally, see Remark 3.2. This is quite different comparing with other Camassa–Holm-type equations; particularly, this is different from the case of the SW equation (1.6), whose RH formalism is very close to that for the SP equation, including the dependence of the RH problem on the parameters $\hat{x}$ and t, see [8]. The reason is that for all equations mentioned above, the RH formalism relies on the so-called sign condition to be satisfied by the initial data. This condition, on one hand, provides the existence of a global in time solution to the Cauchy problem, and on the other hand, plays a crucial role in introducing the new spatial variable $\hat{x}$, see [3, 5, 6, 8]. For instance, this condition reads $-u_{0xx} + 1 > 0$ in the case of the SW equation (1.6), or $u_0 - u_{0xx} + 1 > 0$ in the case of the CH equation (1.5). In the case of the SP equation, the analogous condition would read $1 + u_{0x}^2 > 0$, see the definition of q in (2.2), which is obviously automatically satisfied.

3.3. **From the RH problem to a solution of the SP equation.** The representation result of Theorem 3.4 has been actually obtained under assumption of existence of a solution $u(x,t)$ to the Cauchy problem (1.7). On the other hand, an important element of the inverse scattering approach to nonlinear equations is the possibility to check directly that a solution of the RH problem with any appropriate $r(k)$ (ensuring the unique solvability of the RH problem) gives rise to a solution of the nonlinear equation in question. The idea consists in direct checking that the solution of the RH problem, properly normalized, satisfies a system of differential equations (w.r.t. the *outer parameters* x (or $\hat{x}$) and t), which can be interpreted as the Lax pair for the nonlinear equation. For example, see [16] for the nonlinear Schrödinger equation. For equations of the Camassa–Holm type (and their short wave limits), the procedure is more involved; see, e.g., [7] for the case of the short wave limit of the Degasperis–Procesi equation. For the SP equation, the following theorem holds.

Theorem 3.7. *Let $\{r(k),\ k\in\mathbb{R};\ \{k_j,\gamma_j\}_1^N\}$ be a data set such that the RH problem* (3.7)–(3.11) *it determines has a unique solution $\hat M(\hat x,t,k)$. Define f_1, f_2 by* (3.24d)*. Introduce $x(\hat x,t)$ and $\hat u(\hat x,t)$ as in* (3.24b)–(3.24c)*, and*

$$\hat q(\hat x,t):=\frac{1}{\alpha^2(\hat x,t)-\beta^2(\hat x,t)},\qquad \hat w(\hat x,t):=\frac{2\alpha(\hat x,t)\beta(\hat x,t)}{\alpha^2(\hat x,t)-\beta^2(\hat x,t)},\tag{3.25}$$

where

$$\begin{pmatrix}\alpha(\hat x,t) & \beta(\hat x,t)\\ -\beta(\hat x,t) & \alpha(\hat x,t)\end{pmatrix}:=\hat M(\hat x,t,0).\tag{3.26}$$

Then the following equations (between functions of $(\hat x,t)$) hold:

(a) $x_{\hat x}=\dfrac{1}{\hat q}$;

(b) $\hat u_{\hat x}=\dfrac{\hat w}{\hat q}$;

(c) $\hat q_t=\hat u\hat q\hat w$.

Remark 3.8. We have already noted (Remark 3.1) that $\hat M$ satisfies the symmetries (3.5). The specific forms of the l.h.s. of (3.24d) and (3.26) follow from these symmetries, and the functions $f_1(\hat x,t)$, $f_2(\hat x,t)$, $x(\hat x,t)$, $\hat u(\hat x,t)$, $\alpha(\hat x,t)$, and $\beta(\hat x,t)$ are all real-valued for the same reasons. Moreover, $\hat q>1$ because $\alpha^2+\beta^2=\det\hat M(0)=1$.

Proof of Theorem 3.7. The proof of (a)-(c) is based on calculations of $\Psi_{\hat x}\Psi^{-1}$ and $\Psi_t\Psi^{-1}$ where

$$\Psi(\hat x,t,k):=\hat M(\hat x,t;k)\mathrm{e}^{(-\mathrm{i}k\hat x-\frac{t}{4\mathrm{i}k})\sigma_3}.$$

Proof of (a)-(b). We consider $\Psi_{\hat x}\Psi^{-1}$. Starting from the expansion

$$\hat M(\hat x,t,k)=I+M_1/\mathrm{i}k+\mathrm{O}(k^{-2}),\quad k\to\infty$$

and denoting $W:=-[M_1,\sigma_3]$, we get

$$(\Psi_{\hat x}\Psi^{-1})(\hat x,t,k)=-\mathrm{i}k\sigma_3+W(\hat x,t)+\mathrm{O}(k^{-1}),\quad k\to\infty.$$

Moreover, $(\Psi_{\hat x}\Psi^{-1})(\hat x,t,k)+\mathrm{i}k\sigma_3$ has neither jumps no singularities and is bounded in $k\in\mathbb{C}$; hence, by Liouville's theorem,

$$(\Psi_{\hat x}\Psi^{-1})(\hat x,t,k)=-\mathrm{i}k\sigma_3+W(\hat x,t).\tag{3.27}$$

On the other hand, starting from the expansion

$$\Psi(\hat x,t,k)=G_0(\hat x,t)\left(I-\mathrm{i}kG_1(\hat x,t)+\mathrm{O}(k^2)\right)\mathrm{e}^{(-\mathrm{i}k\hat x-\frac{t}{4\mathrm{i}k})\sigma_3},\quad k\to 0,$$

where, by (3.24d) and (3.26))

$$G_0:=\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix},\qquad G_1:=\begin{pmatrix}f_1 & f_2\\ f_2 & -f_1\end{pmatrix},$$

we obtain

$$\Psi_{\hat x}\Psi^{-1}=G_{0\hat x}G_0^{-1}-\mathrm{i}kG_0(G_{1\hat x}+\sigma_3)G_0^{-1}+\mathrm{O}(k^2),\quad k\to 0.$$

Comparing this with (3.27), it follows that

$$G_{1\hat x}=-\sigma_3+G_0^{-1}\sigma_3G_0=\begin{pmatrix}\alpha^2-\beta^2-1 & 2\alpha\beta\\ 2\alpha\beta & \alpha^2-\beta^2-1\end{pmatrix},$$

which, in terms of f_1, f_2, $\hat w$, and $\hat q$, reads

$$f_{1\hat x}=\frac{1}{\hat q}-1,\qquad f_{2\hat x}=\frac{\hat w}{\hat q}\tag{3.28}$$

and thus (a) and (b) hold. By (3.24b) and (3.24c) we indeed have $x_{\hat x}=1+f_{1\hat x}$ and $\hat u_{\hat x}=f_{2\hat x}$.

Proof of (c). Now we consider $\Psi_t\Psi^{-1}$. On one hand,

$$\Psi_t\Psi^{-1} = \mathrm{O}(k^{-1}), \quad k \to \infty.$$

On the other hand,

$$\Psi_t\Psi^{-1} = -\frac{1}{4\mathrm{i}k}G_0\sigma_3G_0^{-1} + \left\{G_{0t} + \frac{1}{4}G_0[G_1,\sigma_3]\right\}G_0^{-1}, \quad k \to 0.$$

Thus, by Liouville's theorem,

$$G_{0t} = -\frac{1}{4}G_0[G_1,\sigma_3] = -\frac{1}{2}\begin{pmatrix}\beta f_2 & -\alpha f_2\\ \alpha f_2 & \beta f_2\end{pmatrix},$$

which, in terms of $\hat{u}$, $\hat{w}$ and $\hat{q}$, reads $\hat{q}_t = \hat{u}\hat{q}\hat{w}$ and thus item (c) of Theorem 3.7 holds. □

Corollary 3.9. *With the same assumptions and notations as in Theorem 3.7 we introduce*

$$u(x,t) := \hat{u}(\hat{x}(x,t),t), \quad q(x,t) := \hat{q}(\hat{x}(x,t),t).$$

Then the three equations (a)–(c) *from Theorem 3.7 reduce to*

$$q_t = \frac{1}{2}(u^2q)_x, \tag{3.29a}$$

$$q = \sqrt{1+u_x^2}, \tag{3.29b}$$

which is the SP equation in the conservation law form.

Proof. First, it follows from (a) that $\hat{x}_x(x,t) = q(x,t)$ and from (b) that $\hat{u}_{\hat{x}}(\hat{x}(x,t),t) = \frac{w}{q}(x,t)$, where $w(x,t) := \hat{w}(\hat{x}(x,t),t)$. Hence, the identity $u_x(x,t) = \hat{u}_{\hat{x}}(\hat{x}(x,t),t)\hat{x}_x(x,t)$ gives

$$w = u_x. \tag{3.30}$$

Thus, (3.29b) reads $q = \sqrt{1+w^2}$, or $\hat{q} = \sqrt{1+\hat{w}^2}$, which follows from the definitions (3.25) of $\hat{q}$ and $\hat{w}$.

In order to derive (3.29a), we first notice that (c) can be written in the conservation law form

$$\left(\frac{1}{\hat{q}}\right)_t = -\frac{1}{2}\left(\hat{u}^2\right)_{\hat{x}}. \tag{3.31}$$

Indeed,

$$\left(\frac{1}{\hat{q}}\right)_t = -\frac{\hat{q}_t}{\hat{q}^2} = -\frac{\hat{u}\hat{w}}{\hat{q}} = -\hat{u}\hat{u}_{\hat{x}} = -\frac{1}{2}\left(\hat{u}^2\right)_{\hat{x}},$$

where (c) and then (b) have been used. Now, we calculate $x_t(\hat{x},t)$ starting from (a), then using (3.31):

$$x_t(\hat{x},t) = -\frac{\partial}{\partial t}\left(\int_{\hat{x}}^{\infty}\frac{\mathrm{d}\xi}{\hat{q}(\xi,t)}\right) = \frac{1}{2}\int_{\hat{x}}^{\infty}\left(\hat{u}^2\right)_{\xi}(\xi,t)\mathrm{d}\xi = -\frac{1}{2}\hat{u}^2(\hat{x},t).$$

Substituting this into the identity $\hat{q}_t = q_xx_t + q_t$ (between functions of $(\hat{x},t)$) and using (c) gives $q_t = \hat{u}\hat{q}\hat{w} + \frac{1}{2}q_x\hat{u}^2$, which reads $q_t = uqw + \frac{1}{2}q_xu^2$ in terms of functions of (x,t). Using (3.30) yields (3.29a):

$$q_t = uqu_x + \frac{1}{2}q_xu^2 = \frac{1}{2}(u^2q)_x.$$ □

4. Solitons

In the general case, solving a Riemann–Hilbert problem reduces to solving a coupled system of integral equations (generated by the jump condition) and algebraic equations (generated by the residue conditions). In this framework, pure soliton solutions arise in the case where the jump condition is trivial ($J \equiv I$) and thus the solution of the RH problem, being a rational function of the spectral parameter, reduces to solving a system of linear algebraic equations only. The dimension of this system is determined by the number of poles.

In the case of the SP equation, it is natural to distinguish between the solutions associated with pure imaginary zeros of $a(k)$ and those associated with zeros k_j with a nonzero real part.

4.1. First, consider the case, where $a(k)$ has a single, pure imaginary zero at $k = \mathrm{i}\nu$, $\nu > 0$. Then $\hat{M}$ has two simple poles: one at $k = \mathrm{i}\nu$ and the second one at $k = -\mathrm{i}\nu$. From the normalization condition it follows that $\hat{M}$ has the form

$$\hat{M} = \begin{pmatrix} \frac{k-B_{11}}{k-\mathrm{i}\nu} & \frac{B_{12}}{k+\mathrm{i}\nu} \\ \frac{B_{21}}{k-\mathrm{i}\nu} & \frac{k-B_{22}}{k+\mathrm{i}\nu} \end{pmatrix},$$

where B_{ij} are functions of $\hat{x}$ and t to be determined from the residue conditions.

Using the symmetry condition $\overline{\hat{M}}(-\overline{k}) = \hat{M}(k)$ we conclude that $B_{ij} = -\overline{B}_{ij}$, $i, j \in \{1, 2\}$ whereas the symmetry $\hat{M}(-k) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \hat{M}(k) \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ implies $B_{11} = -B_{22}$ and $B_{12} = B_{21}$.

Introducing the real-valued functions b_1 and b_2 by $B_{11} = -B_{22} = \mathrm{i}b_1$ and $B_{12} = B_{21} = \mathrm{i}b_2$, M can be written as

$$\hat{M} = \begin{pmatrix} \frac{k-\mathrm{i}b_1}{k-\mathrm{i}\nu} & \frac{\mathrm{i}b_2}{k+\mathrm{i}\nu} \\ \frac{\mathrm{i}b_2}{k-\mathrm{i}\nu} & \frac{k+\mathrm{i}b_1}{k+\mathrm{i}\nu} \end{pmatrix}. \tag{4.1}$$

Denoting $e_1 := \mathrm{e}^{-2\nu\hat{x} - \frac{t}{2\nu} + \log|\gamma|}$ and taking into account that in this case $\gamma \in \mathbb{R}$, the residue conditions (3.10) take the form

$$\mathrm{Res}_{k=\mathrm{i}\nu} \hat{M}^{(1)} = \mathrm{i}\,\mathrm{sign}(\gamma) e_1 \hat{M}^{(2)}(\mathrm{i}\nu). \tag{4.2}$$

Taking into account (4.1), these conditions lead to the system of equations for b_1 and b_2:

$$\begin{cases} \mathrm{i}\nu - \mathrm{i}b_1 = \mathrm{i}\,\mathrm{sign}(\gamma) e_1 \frac{b_2}{2\nu} \\ \mathrm{i}b_2 = \mathrm{i}\,\mathrm{sign}(\gamma) e_1 \frac{\nu + b_1}{2\nu} \end{cases}$$

from which b_1 and b_2 can be determined as follows:

$$b_1 = \frac{\nu(4\nu^2 - e_1^2)}{4\nu^2 + e_1^2}, \qquad b_2 = \frac{4\nu^2\,\mathrm{sign}(\gamma) e_1}{4\nu^2 + e_1^2}. \tag{4.3}$$

Thus we have solved the RH problem. In accordance with Theorem 3.4, the expansion of $\hat{M}(\,\cdot\,,\,\cdot\,,k)$ at $k = 0$ gives

$$\hat{M}(k) = \begin{pmatrix} \frac{b_1}{\nu} + \mathrm{i}k\frac{\nu - b_1}{\nu^2} & \frac{b_2}{\nu} + \mathrm{i}k\frac{b_2}{\nu^2} \\ -\frac{b_2}{\nu} + \mathrm{i}k\frac{b_2}{\nu^2} & \frac{b_1}{\nu} + \mathrm{i}k\frac{b_1 - \nu}{\nu^2} \end{pmatrix} + \mathrm{O}(k^2).$$

Particularly,

$$\hat{M}(0) = \begin{pmatrix} \frac{b_1}{\nu} & \frac{b_2}{\nu} \\ \frac{-b_2}{\nu} & \frac{b_1}{\nu} \end{pmatrix}$$

and thus

$$\hat{M}(0)^{-1}\hat{M}(k) = I - \mathrm{i}k \begin{pmatrix} \frac{1}{\nu} - \frac{b_1}{b_1^2 + b_2^2} & -\frac{b_2}{b_1^2 + b_2^2} \\ -\frac{b_2}{b_1^2 + b_2^2} & \frac{b_1}{b_1^2 + b_2^2} - \frac{1}{\nu} \end{pmatrix} + \mathrm{O}(k^2).$$

Finally, using (3.24), we arrive at

Theorem 4.1 (one-soliton)**.** *One-soliton solutions $u(x,t)$ of the SP equation* (1.1) *can be expressed, in parametric form, as follows:*

$$u(x,t) = \hat{u}(\hat{x}(x,t), t),$$

where

$$\begin{aligned} \hat{u}(\hat{x}, t) &= -\frac{4\,\mathrm{sign}(\gamma) e_1(\hat{x}, t)}{4\nu^2 + e_1^2(\hat{x}, t)}, \\ x(\hat{x}, t) &= \hat{x} + \frac{2}{\nu}\frac{e_1^2(\hat{x}, t)}{4\nu^2 + e_1^2(\hat{x}, t)}, \end{aligned} \tag{4.4}$$

with

$$e_1(\hat{x}, t) := \mathrm{e}^{-2\nu\hat{x} - \frac{t}{2\nu} + \log|\gamma|}.$$

Here, $\nu > 0$ and $\gamma \in \mathbb{R}$ are the soliton parameters.

Introducing

$$\phi(\hat{x},t) = 2\nu\left(\hat{x} + \frac{t}{4\nu^2} - y_0\right), \qquad y_0 = \frac{1}{2\nu}\log\frac{|\gamma|}{2\nu},$$

the soliton formulas (4.4) can be written as

$$\begin{aligned} \hat{u}(\hat{x},t) &= -\frac{\operatorname{sign}\gamma}{\nu}\frac{1}{\cosh\phi(\hat{x},t)}, \\ x(\hat{x},t) &= \hat{x} + \frac{1}{\nu}\left(1 - \tanh\phi(\hat{x},t)\right). \end{aligned} \tag{4.5}$$

We notice that the one-soliton solution described above is always a multivalued function having the form of a loop. Indeed, $\frac{\partial x}{\partial \hat{x}} = 2\tanh^2\phi - 1$, which changes sign twice as ϕ is varying from $-\infty$ to $+\infty$.

For $\nu = \frac{1}{2}$ and $\gamma = -1$, (4.5) reads

$$\begin{aligned} \hat{u}(\hat{x},t) &= \frac{2}{\cosh(\hat{x}+t)}, \\ x(\hat{x},t) &= \hat{x} - 2\tanh(\hat{x}+t) + 2, \end{aligned} \tag{4.6}$$

and thus we retrieve the formulas for the soliton solution presented in [31] (comparing with [31], the additional constant in (4.6) provides that $x - \hat{x} \to 0$ as $\hat{x} \to +\infty$, cf. (2.11)), where they were obtained using the connection between the short pulse equation and the sine-Gordon equation.

4.2. Now consider the case, where $a(k)$ has a pair of zeros: $a(k_0) = 0 = a(-\bar{k}_0)$ with

$$k_0 = \mu + \mathrm{i}\nu, \quad \mu > 0,\ \nu > 0.$$

In this case, the symmetries (3.5) and the normalization condition lead to

$$\hat{M} = \frac{1}{2}\begin{pmatrix} \frac{k-b_1}{k-k_0} + \frac{k+\bar{b}_1}{k+k_0} & \frac{b_2}{k-k_0} - \frac{\bar{b}_2}{k+k_0} \\ -\frac{\bar{b}_2}{k-k_0} + \frac{b_2}{k+k_0} & \frac{k-\bar{b}_1}{k-k_0} + \frac{k+b_1}{k+k_0} \end{pmatrix} \tag{4.7}$$

(cf. (4.1)), where $b_1(\hat{x},t)$ and $b_2(\hat{x},t)$ can be found solving the system of linear equations resulting from the residue conditions (3.10) at $k = k_0$:

$$\operatorname{Res}_{k=k_0}\hat{M}^{(1)}(\hat{x},t,k) = \mathrm{i}\gamma_0 \mathrm{e}^{2\mathrm{i}\left(k_0\hat{x} - \frac{t}{4k_0}\right)}\hat{M}^{(2)}(\hat{x},t,k_0) \tag{4.8}$$

(the other residue conditions at $-\bar{k}_0$, $-k_0$, and $\bar{k}_0$ then follow from the symmetry condition).

Introducing $\gamma = |\gamma|\mathrm{e}^{\mathrm{i}\arg\gamma}$ and writing $\gamma_0\mathrm{e}^{2\mathrm{i}\left(k_0\hat{x} - \frac{t}{4k_0}\right)}$ in (4.8) as

$$\gamma_0\mathrm{e}^{2\mathrm{i}\left(k_0\hat{x} - \frac{t}{4k_0}\right)} = \mathrm{e}^{-\frac{\nu}{|k_0|}\left(2|k_0|\hat{x} + \frac{t}{2|k_0|} - \frac{|k_0|\log|\gamma|}{\nu}\right)}\mathrm{e}^{\frac{\mathrm{i}\mu}{|k_0|}\left(2|k_0|\hat{x} - \frac{t}{2|k_0|} + \frac{|k_0|\arg\gamma}{\mu}\right)}$$

suggest introducing

$$\begin{aligned} \phi &= \frac{\nu}{\sqrt{\nu^2+\mu^2}}\left(2|k_0|\hat{x} + \frac{t}{2|k_0|} - \frac{|k_0|\log|\gamma|}{\nu}\right), \\ \psi &= \frac{\mu}{\sqrt{\nu^2+\mu^2}}\left(2|k_0|\hat{x} - \frac{t}{2|k_0|} + \frac{|k_0|\arg\gamma}{\mu}\right), \end{aligned}$$

in terms of which the solution of the SP equation (after solving (4.7), (4.8) for b_1 and b_2) is given by (see [23, 31])

$$\hat{u}(\hat{x},t) = \frac{2\mu\nu}{\nu^2+\mu^2}\frac{\nu\sin\psi\sinh\phi + \mu\cos\psi\cosh\phi}{\nu^2\sin^2\psi + \mu^2\cosh^2\phi}, \tag{4.9a}$$

$$x(\hat{x},t) = \hat{x} + \frac{\mu\nu}{\nu^2+\mu^2}\left(\frac{\nu\sin(2\psi) - \mu\sinh(2\phi)}{\nu^2\sin^2\psi + \mu^2\cosh^2\phi} + \frac{2}{\mu}\right). \tag{4.9b}$$

Observing that (see [23])

$$\frac{\partial x}{\partial \hat{x}} = 1 - \frac{8\mu^2\nu^2 \sin^2\psi \cosh^2\phi}{(\nu^2+\mu^2)(\nu^2\sin^2\psi + \mu^2\cosh^2\phi)^2} = \cos\left(4\arctan\frac{\nu\sin\psi}{\mu\cosh\phi}\right),$$

we see that if $\left|\frac{\nu}{\mu}\right| < \tan\frac{\pi}{8}$, then $\frac{\partial x}{\partial \hat{x}} > 0$ for all x, and thus (4.9) represents a smooth solution – the breather. On the other hand, if $\left|\frac{\nu}{\mu}\right| > \tan\frac{\pi}{8}$, then $\frac{\partial x}{\partial \hat{x}}$ is not sign-definite, and thus (4.9) represents a multivalued solution.

5. Long-time asymptotics

A major advantage of the representation of the solution u to the Cauchy problem for a nonlinear integrable equation in terms of the solution of an associated Riemann–Hilbert problem is that it can be efficiently used for studying *in details* the long-time behavior of the former problem via the long-time analysis of the latter, applying the nonlinear steepest descent method introduced by Deift and Zhou [15]. For Camassa–Holm-type equations, this approach has been presented in [2, 4, 6, 7]. A key feature of this method is the deformation of the original RH problem according to the "signature table" for the phase function θ in the jump matrix $\hat{J}$ written in the form (cf. (3.8), (3.9))

$$\hat{J}(\hat{x}, t; k) = \mathrm{e}^{-\mathrm{i}t\theta(\hat{\zeta},k)\sigma_3} J_0(k) \mathrm{e}^{\mathrm{i}t\theta(\hat{\zeta},k)\sigma_3}, \tag{5.1}$$

where

$$\theta(\hat{\zeta}, k) = \hat{\zeta}k - \frac{1}{4k}, \tag{5.2}$$

$$\hat{\zeta} := \frac{\hat{x}}{t}. \tag{5.3}$$

The signature table is the distribution of signs of $\operatorname{Im}\theta(\hat{\zeta}, k)$ in the k-plane, depending on the values of $\hat{\zeta}$. In the case of the SP equation,

$$\operatorname{Im}\theta(\hat{\zeta}, k) = \operatorname{Im} k \cdot \Big(\hat{\zeta} + \frac{1}{4|k|^2}\Big).$$

Now we notice that $\hat{J}(\hat{x}, t; k)$ in the present case looks very similar to the case of the SW equation, see [8], including the matrix structure of $\hat{J}_0(k)$ and the form of $\theta(\hat{\zeta}, k)$. Namely, the latter in the case of the SW equation has the form $\theta(\hat{\zeta}, k) = \hat{\zeta}k - \frac{1}{2k}$ and thus the distribution of signs of $\operatorname{Im}\theta(\hat{\zeta}, k)$ is the same modulo the scaling factor $\frac{1}{2}$. As for the structure of $\hat{J}_0(k)$, the difference with the case of the SW equation is that $\bar{r}$ is to be replaced by $-\bar{r}$ while keeping r the same. A direct consequence of this is that the long time analysis in the case of the SP equation repeats the steps made in the case of the SW equation. As for the differences, we notice the following.

1) The basic difference is that in the case of the SP equation, the RH problem involves, in general, residue conditions (absent in the case of the SW equation). Here there is a complete analogy with the NLS equation [16], where the defocusing NLS equation corresponds to the SW equation whereas the focusing NLS equation corresponds to the SP equation. Consequently, if $a(k)$ has zeros, then the solitons associated with the residue conditions dominate the long time behavior of the solution of the Cauchy problem.
2) In the solitonless case ($a(k) \neq 0$ for all k with $\operatorname{Im} k \geq 0$), the main asymptotic term is expressed in terms of the solution of the model RH problem, which is different from that in the SW case. More precisely, the model problem for the SW equation is exactly as in [22, Appendix B] whereas in the case of the SP equation, the jump matrix of the model problem is as in [22, (B.1)], with $\bar{q}$ replaced by $-\bar{q}$ (keeping q the same). Accordingly, the large-z

expansion of the solution of the model problem, which is

$$m^X(q,\hat{k}) = I + \frac{\mathrm{i}}{\hat{k}}\begin{pmatrix} 0 & -\beta^X(q) \\ \overline{\beta^X(q)} & 0 \end{pmatrix} + \mathrm{O}(\hat{k}^{-2}), \qquad \hat{k}\to\infty$$

from [22, (B.2)], is to be replaced by

$$m^X(q,\hat{k}) = I + \frac{\mathrm{i}}{\hat{k}}\begin{pmatrix} 0 & -\beta^X(q) \\ -\overline{\beta^X(q)} & 0 \end{pmatrix} + \mathrm{O}(\hat{k}^{-2}), \qquad \hat{k}\to\infty \tag{5.4}$$

in the case of the SP equation, with

$$\beta^X(q) = \sqrt{-h(q)}\,\mathrm{e}^{\mathrm{i}\left(\frac{\pi}{4}-\arg q+\arg\Gamma(\mathrm{i}\nu(q))\right)} \tag{5.5}$$

(notice that $h(q)<0$ in this case), where Γ is the Euler Gamma function.

Now we are going to give a sketch of the asymptotic analysis in the solitonless case and present an exact asymptotic result. Since the distribution of signs of $\operatorname{Im}\theta(\hat{\zeta},k)$ is as in the case of the SW equation, the long time behavior of u is qualitatively different in the same two ranges of values of

$$\zeta := \frac{x}{t}\,. \tag{5.6}$$

5.1. **Range $\zeta > \varepsilon$.** In this case the set $\{k \mid \operatorname{Im}\theta(\hat{\zeta},k)=0\}$ coincides with the real axis $\operatorname{Im}k=0$ and $\pm\operatorname{Im}\theta>0$ for $\pm\operatorname{Im}k>0$. This suggests the use of the following factorization of the jump matrix for all $k\in\mathbb{R}$:

$$\hat{J} = \begin{pmatrix} 1 & \bar{r}(\bar{k})\mathrm{e}^{-2\mathrm{i}t\theta} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ r(k)\mathrm{e}^{2\mathrm{i}t\theta} & 1 \end{pmatrix}. \tag{5.7}$$

Indeed, the triangular factors in (5.7) can be absorbed into a new RH problem for $\hat{M}^{(1)}(\hat{x},t,k)$ in the same way as in the case of the SW equation [8]:

$$\hat{M}^{(1)} = \begin{cases} \hat{M}\begin{pmatrix} 1 & 0 \\ -r(k)\mathrm{e}^{2\mathrm{i}t\theta} & 1 \end{pmatrix}, & 0<\operatorname{Im}k<\varepsilon, \\ \hat{M}\begin{pmatrix} 1 & \bar{r}(\bar{k})\mathrm{e}^{-2\mathrm{i}t\theta} \\ 0 & 1 \end{pmatrix}, & -\varepsilon<\operatorname{Im}k<0, \\ \hat{M}, & \text{otherwise.} \end{cases}$$

This reduces the RH problem to a RH problem with a jump matrix that decays exponentially (in t) to the identity matrix. Since this RH problem is holomorphic (there is no residue condition), its solution decays fast to I and consequently $\hat{u}(\hat{x},t)$ decays fast to 0 while $\hat{x}$ approaches fast x, and thus the domains $\hat{\zeta}>\varepsilon$ and $\zeta>\varepsilon$ coincide asymptotically.

5.2. **Range $\zeta < -\varepsilon$.** Similarly to [8], in a domain of the form $\hat{\zeta}<-\varepsilon$ for any $\varepsilon>0$, the signature table dictates the use of two factorizations. Let $\pm\hat{\kappa}$ be the points where the distribution of signs is changing:

$$\hat{\kappa} = \frac{1}{2\sqrt{|\hat{\zeta}|}}\,. \tag{5.8}$$

i) For $k\in(-\hat{\kappa},\hat{\kappa})$ we consider again the factorization (5.7)

$$\hat{J} = \begin{pmatrix} 1 & \bar{r}(\bar{k})\mathrm{e}^{-2\mathrm{i}t\theta} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ r(k)\mathrm{e}^{2\mathrm{i}t\theta} & 1 \end{pmatrix}.$$

ii) For $k\in(-\infty,-\hat{\kappa})\cup(\hat{\kappa},\infty)$ we consider a factorization with triangular factors in reverse order:

$$\hat{J} = \begin{pmatrix} 1 & 0 \\ \frac{r(k)}{1-|r(k)|^2}\mathrm{e}^{2\mathrm{i}t\theta} & 1 \end{pmatrix}\begin{pmatrix} 1-|r(k)|^2 & 0 \\ 0 & \frac{1}{1-|r(k)|^2} \end{pmatrix}\begin{pmatrix} 1 & \frac{\bar{r}(\bar{k})}{1-|r(k)|^2}\mathrm{e}^{-2\mathrm{i}t\theta} \\ 0 & 1 \end{pmatrix} \tag{5.9}$$

Similarly to the previous case, an appropriate sequence of deformations of the RH problem is the same as in the case of the SW equation, so we will follow it giving details mainly for items specific to the considered equation.

The deformations involve the removal of the diagonal factor in (5.9) and the consequent absorption of the triangular factors, leading, after an appropriate rescaling, to a model RH problem on a contour consisting of two crosses centered at $k=\pm\hat\kappa$, see [2, 4], which finally leads to the asymptotics in the form of modulated decaying (of the order $\mathrm{O}(t^{-1/2})$) oscillations. The diagonal term is removed introducing $\hat M^{(1)}=\hat M\delta^{-\sigma_3}$, where

$$\delta(k;\hat\zeta)=\exp\left\{\frac{1}{2\pi\mathrm{i}}\left(\int_{-\infty}^{-\hat\kappa}+\int_{\hat\kappa}^{\infty}\right)\log(1+|r(s)|^2)\frac{\mathrm{d}s}{s-k}\right\}. \tag{5.10}$$

solves the scalar RH problem whose jump condition is

$$\delta_+=\delta_-(1+|r(k)|^2)$$

across the contour $(-\infty,-\hat\kappa)\cup(\hat\kappa,\infty)$.

The triangular factors are absorbed into the RH problem for $\hat M^{(2)}$:

$$\hat M^{(2)}=\begin{cases}\hat M^{(1)}\begin{pmatrix}1&0\\-r\delta^{-2}\mathrm{e}^{2\mathrm{i}t\theta}&1\end{pmatrix}, & \operatorname{Im}k>0,\ k\text{ near }(-\hat\kappa,\hat\kappa),\\ \hat M^{(1)}\begin{pmatrix}1&-\frac{\bar r}{1-|r|^2}\delta_+^2\mathrm{e}^{-2\mathrm{i}t\theta}\\0&1\end{pmatrix}, & \operatorname{Im}k>0,\ k\text{ near }\mathbb{R}\setminus[-\hat\kappa,\hat\kappa],\\ \hat M^{(1)}\begin{pmatrix}1&\bar r\delta^2\mathrm{e}^{-2\mathrm{i}t\theta}\\0&1\end{pmatrix}, & \operatorname{Im}k<0,\ k\text{ near }(-\hat\kappa,\hat\kappa),\\ \hat M^{(1)}\begin{pmatrix}1&0\\\frac{r}{1-|r|^2}\delta_-^{-2}\mathrm{e}^{2\mathrm{i}t\theta}&1\end{pmatrix}, & \operatorname{Im}k<0,\ k\text{ near }\mathbb{R}\setminus[-\hat\kappa,\hat\kappa].\end{cases} \tag{5.11}$$

Now, in order to reduce the RH problem for $\hat M^{(2)}$, as $t\to\infty$, to a model problem whose solution can be given explicitly in terms of parabolic cylinder functions, see [15, 4, 22, 21], the leading term of the factor $\delta(k)\mathrm{e}^{-\mathrm{i}t\theta(k)}$ as $k\to\pm\hat\kappa$ is to be evaluated. One has

$$\delta(k)=\left(\frac{\hat\kappa-k}{\hat\kappa+k}\right)^{-\mathrm{i}h}\mathrm{e}^{\chi(k)} \tag{5.12}$$

with

$$h\equiv h(\hat\kappa)=-\frac{1}{2\pi}\log\bigl(1+|r(\hat\kappa)|^2\bigr), \tag{5.13}$$

$$\chi(k)=\frac{1}{2\pi\mathrm{i}}\int_{\mathbb{R}\setminus[-\hat\kappa,\hat\kappa]}\frac{\log(1+|r(s)|^2)}{\log(1+|r(\hat\kappa)|^2)}\frac{\mathrm{d}s}{s-k}. \tag{5.14}$$

As $k\to-\hat\kappa$,

$$\theta(k)=\frac{1}{2\hat\kappa}+\frac{1}{4\hat\kappa^3}(k+\hat\kappa)^2+\mathrm{O}((k+\hat\kappa)^3). \tag{5.15}$$

Therefore, introducing the scaled spectral variable $\hat k$ by

$$k+\hat\kappa=\frac{\hat k}{\sqrt{\hat\kappa^{-3}t}}, \tag{5.16}$$

the factor $\delta^2(k)\mathrm{e}^{-2\mathrm{i}t\theta(k)}$ can be approximated as

$$\delta^2(k)\mathrm{e}^{-2\mathrm{i}t\theta(k)}\approx\tilde\delta^2\hat k^{2\mathrm{i}h}\mathrm{e}^{-\mathrm{i}\hat k^2/2}, \tag{5.17}$$

where

$$\tilde\delta^2=\left(\frac{4t}{\hat\kappa}\right)^{-\mathrm{i}h}\mathrm{e}^{-\frac{\mathrm{i}t}{\hat\kappa}}\mathrm{e}^{-2\chi(\hat\kappa)}. \tag{5.18}$$

Similarly for k near $\hat\kappa$.

Following [22], the solution of the RH problem for $\hat{M}^{(2)}$ formulated on two crosses centered at $k=\pm\hat{\kappa}$ with the jump matrix $\hat{J}^{(2)}=(\hat{M}_-^{(2)})^{-1}\hat{M}_+^{(2)}$ that follows from (5.11), can be approximated, for large t, in terms of the solution m^X of the model problem formulated in the $\hat{k}$-plane on a cross centered at $\hat{k}=0$ and evaluated for large $\hat{k}$. In our case, the evaluation of the model problem is given by (5.4), (5.5) with $q=r(-\hat{k})$ and h as in (5.13). Taking into account that we are interested in the expansion of $\hat{M}^{(2)}(k)$ as $k\to 0$, a reasoning similar to [22, Eqs (2.36)–(2-39)] leads to an approximation for $\hat{M}^{(2)}(k)$ in terms of m^X:

$$\begin{aligned}\hat{M}^{(2)}(k)=I&+\frac{1}{\pi}\operatorname{Im}\int_{|k+\hat{\kappa}|=\rho}(m_0^{-1}(s)-I)\frac{\mathrm{d}s}{s}\\&+\frac{k}{\mathrm{i}\pi}\operatorname{Re}\int_{|k+\hat{\kappa}|=\rho}(m_0^{-1}(s)-I)\frac{\mathrm{d}s}{s^2}+\mathrm{O}(k^2t^{-1/2-\varepsilon})\end{aligned}\tag{5.19}$$

with some $\rho>0$ and $\varepsilon>0$, where

$$m_0(k)=\tilde{\delta}^{\sigma_3}m^X\left(\hat{\zeta},\sqrt{\frac{t}{\hat{\kappa}^3}}(k+\hat{\kappa})\right)\tilde{\delta}^{-\sigma_3}.$$

In view of (5.4) and (5.16), and for large t,

$$m_0(k)=I+\frac{\mathrm{i}\sqrt{\hat{\kappa}^3}}{\sqrt{t}(k+\hat{\kappa})}\begin{pmatrix}0&-\beta^X\tilde{\delta}^2\\-\bar{\beta}^X\tilde{\delta}^{-2}&0\end{pmatrix}+\mathrm{O}(t^{-1/2-\varepsilon}).\tag{5.20}$$

Now recall that $\hat{M}=\hat{M}^{(1)}\delta^{\sigma_3}$ and that $\hat{M}^{(1)}$ is related to $\hat{M}^{(2)}$ by (5.11), where $r(k)=\mathrm{O}(k^2)$. Evaluating $\delta(k;\hat{\zeta})$ as $k\to 0$ gives

$$\delta(k;\hat{\zeta})=1-\mathrm{i}kQ+\mathrm{O}(k^2),$$

where

$$Q=\frac{1}{\pi}\int_{\hat{\kappa}}^{\infty}\frac{\log(1+|r(s)|^2)}{s^2}\,\mathrm{d}s.$$

Taking this into account and substituting (5.20) into (5.19), one obtains

$$\hat{M}(k)=I+\frac{c_1}{\sqrt{t}}\begin{pmatrix}0&1\\-1&0\end{pmatrix}+\mathrm{i}k\left(-Q\sigma_3+\frac{Qc_1+c_2}{\sqrt{t}}\begin{pmatrix}0&1\\1&0\end{pmatrix}\right)+\mathrm{O}(k^2t^{-1/2-\varepsilon}),$$

where

$$c_1=2\sqrt{\hat{\kappa}}\operatorname{Im}\{\beta^X\tilde{\delta}^2\},\qquad c_2=\frac{2}{\sqrt{\hat{\kappa}}}\operatorname{Re}\{\beta^X\tilde{\delta}^2\}.$$

Finally, using (3.24b)–(3.24d) one arrives at the asymptotic formulas

$$\hat{u}=\frac{c_2}{\sqrt{t}}(1+\mathrm{o}(1)),\qquad x-\hat{x}=Q(1+\mathrm{o}(1)),\qquad t\to\infty,$$

which imply the asymptotics for u in the original variables:

$$u(x,t)=2\sqrt{\frac{|h(\varkappa)|}{\varkappa t}}\cos\left\{\frac{t}{\varkappa}+h(\varkappa)\log t+\phi_0(\varkappa)\right\},\tag{5.21}$$

where

$$\varkappa=\frac{1}{2}\sqrt{\frac{t}{|x|}}\tag{5.22}$$

and

$$\begin{aligned}\phi_0(\varkappa)=-\frac{\pi}{4}-\arg(r(\varkappa))-\arg\Gamma(\mathrm{i}h(\varkappa))+\frac{1}{\pi}\int_{\mathbb{R}\setminus[-\varkappa,\varkappa]}\log|k-s|\,\mathrm{d}\log(1+|r(s)|^2)\\+\frac{2\varkappa}{\pi}\int_{\varkappa}^{\infty}\frac{\log(1+|r(s)|^2)}{s^2}\mathrm{d}s+h(\varkappa)\log\frac{4}{\varkappa}.\end{aligned}\tag{5.23}$$

Theorem 5.1 (solitonless asymptotics)**.** *Let $u(x,t)$ be the solution of the Cauchy problem* (1.7)*. Assume that the spectral function $a(k)$ constructed from $u_0(x)$ has no zeros in the upper half-plane. Then the behavior of u as $t \to \infty$ is described as follows. Let ε be any small positive number.*

(i) *In the domain $\zeta \equiv x/t > \varepsilon$, $u(x,t)$ tends to* 0 *with fast decay.*

(ii) *In the domain $\zeta \equiv x/t < -\varepsilon$, $u(x,t)$ exhibits decaying (of the order $\mathrm{O}(t^{-1/2})$) modulated oscillations given by* (5.21)*, where $h(\varkappa)$ and $\phi_0(\varkappa)$ are functions of $\varkappa = 1/(2\sqrt{|\zeta|})$ given in terms of the associated reflection coefficient $r(k)$; in particular, by* (5.13) *and* (5.23)

$$h(\varkappa) = -\frac{1}{2\pi}\log\bigl(1+|r(\varkappa)|^2\bigr). \tag{5.24}$$

For completeness, we present also the asymptotics in the soliton case.

Theorem 5.2 (soliton asymptotics)**.** *Assume that $a(k)$ has $N = 2n+m$ simple zeros*

$$\{k_j\}_1^n \cup \{-\bar{k}_j\}_1^n \cup \{i\nu_j\}_1^m,$$

where $\mu_j := \operatorname{Re} k_j > 0$, $\nu_j := \operatorname{Im} k_j > 0$. Assume also that if $\mu_j^2+\nu_j^2 \neq \mu_l^2+\nu_l^2$ if $j \neq l$. Then:

(i) *For $\varepsilon > 0$ sufficiently small, the asymptotics of u in each sector $\left|\frac{x}{t} + \frac{1}{4(\mu_j^2+\nu_j^2)}\right| < \varepsilon$ is given by*

$$u(x,t) = u_j(x,t) + \mathrm{O}(t^{-1/2}),$$

where u_j is given, parametrically, as follows:

a) *If $\mu_j = 0$, then u_j is given by* (4.5) *with ν replaced by ν_j and ϕ replaced by*

$$\phi_j = 2\nu_j\hat{x} + \frac{t}{2\nu_j} + \phi_j^0.$$

If $\mu_j \neq 0$, then u_j is given by (4.9)*, with μ, ν, ϕ, and ψ replaced respectively by μ_j, ν_j,*

$$\phi_j = 2\nu_j\hat{x} + \frac{\nu_j t}{2(\mu_j^2+\nu_j^2)} + \phi_j^0, \qquad \psi_j = 2\mu_j\hat{x} - \frac{\mu_j t}{2(\mu_j^2+\nu_j^2)} + \psi_j^0.$$

Here ϕ_j^0 and ψ_j^0 are constants determined by the scattering data $\{r(k), k \in \mathbb{R}; \{k_j,\gamma_j\}_1^N\}$.

(ii) *Outside these sectors, $u(x,t) = \mathrm{O}(t^{-1/2})$.*

Remark 5.3. The asymptotic results presented above imply the following.

a) In the solitonless case as well as in the case when all the zeros k_j of $a(k)$ are located outside the sector $\left|\frac{\operatorname{Im} k_j}{\operatorname{Re} k_j}\right| \geq \tan\frac{\pi}{8}$, there exists $T > 0$ such that for all $t > T$, the solution of the Cauchy problem (1.7) is a smooth classical solution (possibly after passing through wave breakings).

b) A sufficient condition for wave breaking: If $a(k)$ has a zero k^* in the sector $\left|\frac{\operatorname{Im} k^*}{\operatorname{Re} k^*}\right| \geq \tan\frac{\pi}{8}$, then wave breaking occurs at a certain finite time.

Notice that another sufficient condition for finite time wave breaking has been obtained in [23] using the method of characteristics and conserved quantities.

Remark 5.4. The asymptotic formula presented in Theorem 5.1 improves the asymptotics obtained in [20] and [26] by using different methods (not relying on the integrability of the SP equation).

References

[1] R. Beals and R. R. Coifman. Scattering and inverse scattering for first order systems. *Comm. Pure Appl. Math.*, 37(1):39–90, 1984.

[2] A. Boutet de Monvel, A. Kostenko, D. Shepelsky, and G. Teschl. Long-time asymptotics for the Camassa–Holm equation. *SIAM J. Math. Anal.*, 41(4):1559–1588, 2009.

[3] A. Boutet de Monvel and D. Shepelsky. Riemann-Hilbert approach for the Camassa–Holm equation on the line. *C. R. Math. Acad. Sci. Paris*, 343(10):627–632, 2006.

[4] A. Boutet de Monvel and D. Shepelsky. Long-time asymptotics of the Camassa–Holm equation on the line. In *Integrable systems and random matrices*, volume 458 of *Contemp. Math.*, pages 99–116. Amer. Math. Soc., Providence, RI, 2008.

[5] A. Boutet de Monvel and D. Shepelsky. Riemann–Hilbert problem in the inverse scattering for the Camassa–Holm equation on the line. In *Probability, geometry and integrable systems*, volume 55 of *Math. Sci. Res. Inst. Publ.*, pages 53–75. Cambridge Univ. Press, Cambridge, 2008.

[6] A. Boutet de Monvel and D. Shepelsky. A Riemann–Hilbert approach for the Degasperis–Procesi equation. *Nonlinearity*, 26:2081–2107, 2013.

[7] A. Boutet de Monvel and D. Shepelsky. The Ostrovsky–Vakhnenko equation by a Riemann–Hilbert approach. *J. Phys. A: Math. Theor.*, 48:035204, 2015.

[8] A. Boutet de Monvel, D. Shepelsky, and L. Zielinski. The short-wave model for the Camassa–Holm equation: a Riemann–Hilbert approach. *Inverse Problems*, 27:105006, 2011.

[9] J. C. Brunelli. The short pulse hierarchy. *J. Math. Phys.*, 46:123507, 2005.

[10] J. C. Brunelli. The bi-Hamiltonian structure of the short pulse equation. *Phys. Lett. A*, 353(6):475–478, 2006.

[11] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.

[12] Y. Chung, C. K. R. T. Jones, T. Schäfer, and C. E. Wayne. Ultra-short pulses in linear and nonlinear media *Nonlinearity*, 18(3):1351–1374, 2005.

[13] G. M. Coclite and L. di Ruvo, Well-posedness results for the short pulse equation. *Z. Angew. Math. Phys.*, 66:1529–1557, 2015.

[14] G. M. Coclite and L. di Ruvo, Wellposedness of bounded solutions of the non-homogeneous initial boundary for the short pulse equation. *Boll. Unione Mat. Ital.*, 8(1):31–44, 2015.

[15] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation. *Ann. of Math. (2)*, 137(2):295–368, 1993.

[16] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian methods in the theory of solitons.* Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1987.

[17] B. Fuchssteiner. Some tricks from the symmetry-toolbox for nonlinear equations: generalizations of the Camassa–Holm equation. *Phys. D*, 95:229–243, 1996.

[18] B. Fuchssteiner and A. S. Fokas. Symplectic structures, their Bäcklund transformations and hereditary symmetries. *Phys. D*, 4:47–66, 1981.

[19] G. Gambino, U. Tanriver, P. Guha, A. Choudhury, and S. Choudhury. Regular and singular pulse and front solutions and possible isochronous behavior in the short-pulse equation: phase-plane, multi-infinite series and variational approaches. *Commun. Nonlinear Sci. Numer. Simul.*, 20(2):375–388, 2015.

[20] N. Hayashi and P. Naumkin. Large time asymptotics for the reduced Ostrovsky equation. *Commun. Math. Phys.* 335:713–738, 2015.

[21] J. Lenells. The nonlinear steepest descent method: asymptotics for initial-boundary value problems. *SIAM J. Math. Anal.* 48(3):2076–2118, 2016.

[22] J. Lenells. The nonlinear steepest descent method for Riemann–Hilbert problems of low regularity. arXiv:1501.05329.

[23] Y. Liu, D. Pelinovsky, and A. Sakovich. Wave breaking in the short-pulse equation. *Dynamics of PDE*, 6:291–310, 2009.

[24] Y. Matsuno. Multiloop soliton and multibreather solutions of the short pulse model equation. *J. Phys. Soc. Japan*, 76(8):084003, 6 pp., 2007.

[25] Y. Matsuno. Periodic solutions of the short pulse model equation. *J. Math. Phys.*, 49(7):073508, 18 pp., 2008.

[26] T. Niizato. Asymptotic behavior of solutions to the short pulse equation with critical nonlinearity. *Nonlinear Analysis*, 111:15–32, 2014.

[27] P. J. Olver and P. Rosenau. Tri-Hamiltonian duality between solitons and solitary-wave solutions having compact support. *Phys. Rev. E*, 53:1900–1906, 1996.

[28] D. Pelinovsky and A. Sakovich. Global well-posedness of the short-pulse and sine-Gordon equations in energy space. , *Comm. Partial Differential Equations*, 35(4):613–629, 2010.

[29] Z. Qiao. A new integrable equation with cuspons and W/M-shape-peaks solitons. *J. Math. Phys.*, 47:112701, 2006.

[30] A. Sakovich and S. Sakovich. The short pulse equation is integrable. *J. Phys. Soc. Japan*, 74:239–241, 2005.

[31] A. Sakovich and S. Sakovich. Solitary wave solutions of the short pulse equation. , *J. Phys. A: Math. Gen.*, 39:L361–L367, 2006.

[32] T. Schäfer and C. E. Wayne. Propagation of ultra-short optical pulse in nonlinear media. *Physica D*, 196:90-105, 2004.

*Institut de Mathématiques de Jussieu-PRG, Université Paris Diderot, Avenue de France, Bât. Sophie Germain, case 7012, 75205 Paris Cedex 13, France
E-mail address: anne.boutet-de-monvel@imj-prg.fr

†Mathematical Division, Institute for Low Temperature Physics, 47 Lenin Avenue, 61103 Kharkiv, Ukraine
E-mail address: shepelsky@yahoo.com

‡LMPA, Université du Littoral Côte d'Opale, 50 rue F. Buisson, CS 80699, 62228 Calais, France
E-mail address: Lech.Zielinski@lmpa.univ-littoral.fr