XAVIER RAYNAUD

ON A SHALLOW WATER WAVE EQUATION

Doctoral thesis
for the degree of doctor philosophiae

Trondheim, June 2006

Norwegian University of Science and Technology
Faculty for Information Technology, Mathematics
and Electrical Engineering
Department of Mathematical Sciences

NTNU
Innovation and Creativity

NTNU
Norwegian University of Science and Technology

Doctoral thesis
for the degree of doctor philosophiae

Faculty for Information Technology, Mathematics
and Electrical Engineering
Department of Mathematical Sciences

ISBN 82-471-7963-6 (printed version)
ISBN 82-471-7962-8 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2006:103

Printed by NTNU-trykk

To my parents,
Pierre and Colette.

Acknowledgements

This thesis contains my research as a PhD student at the Department of Mathematical Sciences at the Norwegian University of Sciences and Technology. My PhD was founded by the department. The working and research conditions at the department are really excellent and gratefully acknowledged.

First, I would like to thank my advisor, prof. Helge Holden. His genuine interest for this work has been for me one of the most important sources of motivation. I am very grateful to him for his constant support and for the many discussions we have had and that were so determinant for this thesis.

I also would like to thank Henrik Kalisch for having suggested new problems and initiated an interesting and fruitful collaboration.

As part of my PhD I spent four months in fall 2005 at the Mittag Leffler Institute in Djursholm, Sweden. I am very grateful to the institute for the grant they gave me and for the wonderful environment for research they provided.

Finally I would like to thank Mette. The last four years have been hectic – for both of us – but they have also been happy years thanks to her and to our very lively little boy, Espen.

INTRODUCTION

The main topic of this thesis is the study of a nonlinear partial differential equation, the Camassa–Holm (CH) equation:

$$u_t - u_{txx} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0. \tag{1}$$

The origin of the Camassa-Holm equation can by traced back to an article from Fuchssteiner and Fokas ([18]) from 1981 where it appears as one member of a whole family of bi-hamiltonian equations generated by the method of recursion operator. However some coefficients were not correctly computed. This may be the reason why no special attention was given to it until its rediscovery in 1993 by Camassa and Holm in the context of water wave ([7, 8]). They derived equation (1) as a model for unidirectional water wave propagation in shallow water with u representing the height of the water's free surface above a flat bottom. The relevance of the equation as a model for shallow water wave has been further investigated by Johnson in [21].

In 1998 a similar equation, namely

$$u_t - u_{txx} + 3uu_x - \gamma(2u_x u_{xx} + uu_{xxx}) = 0, \tag{2}$$

was discovered independently by Dai as a model for nonlinear waves in cylindrical axially symmetric hyperelastic-rod. In this case, $u(x, t)$ represents the radial stretch and γ a constant depending on the property of the material.

A RICH MATHEMATICAL STRUCTURE

The Camassa–Holm enjoys many remarkable mathematical properties. It is bi-Hamitonian, that is, it possesses two distinct but compatible hamiltonians. Following the methodology described in [24] for general bi-Hamiltonian systems, it is possible to derive an infinite number of conserved quantities for the solutions of (1); the computation is carried out in detail in [22]. The equation admits a Lax-pair and is also formally integrable by means of scattering and inverse scattering techniques. The scattering problem consists of computing the scattered far-field over an obstacle whose "shape" is determined in some way by u (for t fixed). In practice, it means finding the eigenvalues of a linear operator depending on $u(t, x)$. The remarkable fact is that as time evolves, if $u(t, x)$ satisfies (1), then these eigenvalues satisfy trivial linear ordinary differential equations which can be solved explicitly and the far-field can be determined for any time. The inverse scattering problem consists of retrieving the "shape" of the obstacle, that is u, from the knowledge of the scattered far-field. This is also a nontrivial but nevertheless linear problem so that one can think of the scattering-inverse

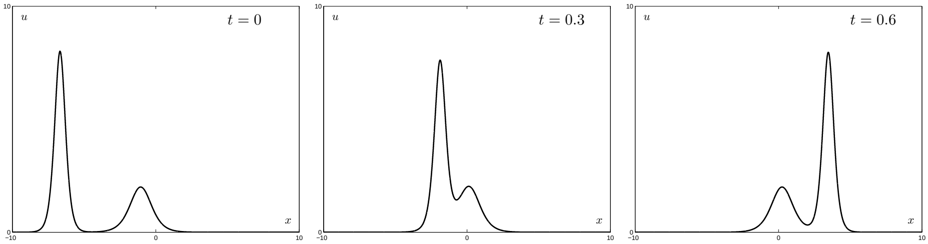


FIGURE 1. Interaction of two solitons for the KdV equation.

scattering method as a way of linearizing the equation. If this approach can be (formally) followed then the system is (formally) integrable and solutions of the Camassa–Holm equation exist and can be computed. For a large class of initial data, it is indeed possible, see [10, 16].

It turns out that many physical relevant equations share the same structure (Lax pair, complete integrability via scattering and inverse scattering techniques), the paramount example being the KdV equation

$$u_t + uu_x - u_{xxx} = 0 \tag{3}$$

which is also used as a model for shallow water wave. These equations exhibit a special type of solutions, the so-called solitons. A single soliton is a traveling wave whose speed is proportional to the height. What makes solitons so special is that when one combines several of them they interact nicely and retain their shape after interaction, see Figure 1 for a two soliton interaction in the case of the KdV equation. The Camassa-Holm equation also possesses solutions of a soliton type, which, because of their shape, have been given the name of peakons. A single peakon is given by

$$u(t,x) = ce^{|x-ct|}. \tag{4}$$

The traveling speed is then equal to the height of the peak. By taking a linear combination of peakons one obtains what is called a multipeakon solution. The multipeakons have the following form

$$u(x,t) = \sum_{i=1}^{n} p_i(t)e^{-|x-q_i(t)|} \tag{5}$$

where p_i and q_i are solutions of the following system of ordinary differential equations

$$\dot{q}_i = \sum_{j=1}^{n} p_j e^{-|q_i-q_j|}, \ \dot{p}_i = \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i-q_j|}. \tag{6}$$

In Figure 2, we show a simple interaction between two peakons. At the peaks, the derivative is discontinuous and the multipeakons can only be solutions of (1)

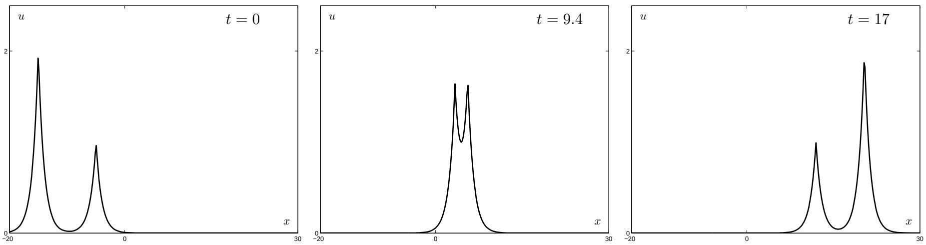


FIGURE 2. Interaction of two peakons for the CH equation.

in a weak sense, see [20] and below. The system of equations (6) is hamiltonian: For H given by $H = \frac{1}{2}\sum_{i,j=1}^{n} p_i p_j e^{-|q_i - q_j|}$, it can be rewritten as

$$\dot{q}_i = \frac{\partial H}{\partial p_i},\ \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{7}$$

The multipeakon solutions as given by (7) can then be seen as a discrete version of the Camassa–Holm equation, see [6].

The Camassa–Holm equation has a geometric interpretation: It is the geodesic equation in the group of diffeomorphism with respect to a right-invariant metric. Using the formalism presented in [1], this geometrical property can in turn be given a physical meaning. For a mechanical system constituted of n distinct particles, the evolution of the system is naturally given by the position of the particles at each time, say $\{y_i(t)\}_{i=1}^{n}$. When we consider a continuous medium, like a fluid, the system is correspondingly described by a function $y(t,\xi)$ which gives the trajectory of the particle labeled by ξ. This is the Lagrangian description. In a fluid, we may assume that vacuum is not created and particles do not accumulate so that, for any time t, $\xi \mapsto y(t,\xi)$ remains a bijection between the labeling space and the physical space. Taking one step further, we may as well assume that $y(t,\cdot)$ remains a smooth diffeomorphism so that $y\colon t \mapsto y(t,\cdot)$ can be seen as a path in the group of diffeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$ (n is the dimension of the system). Formally, the group of smooth diffeomorphism, which we denote G, can be given the structure of a Riemannian Lie group, the Riemannian metric then representing the energy of the system. The physically relevant path is then determined by the least action principle which says that $y : t \mapsto y(t,\cdot)$ is a geodesic in G. We consider a homogeneous fluid for which the particles are undistinguishable. In this case the initial labeling is arbitrary and, at each given time, it must be possible to relabel the particles in a arbitrarily way without changing the evolution of the system. The evolution of the system depends only on the velocity distribution and the actual position of the particles should not matter as they are undistinguishable. A pure Eulerian description of the system is possible: Instead of looking at the trajectory of each individual particle, one consider, for a fixed point x in space, the velocity $u(t,x)$ of the particle that at time t goes through x, that is,

$$u(t,x) = y_t(t, y^{-1}(t,x)).$$

The system enjoys what is called a *relabeling symmetry.* In the topological framework of Arnold and Khesin, the relabeling symmetry corresponds to the fact that the metric of G is right-invariant. Then, by Noether's theorem, one derives the existence of a conserved quantity holding point-wise in space and which, by analogy with the rigid body problem, is called *angular momentum.* Furthermore, this framework provides a generic way of deriving the Euler equation for hydrodynamical systems. For the Camassa–Holm equation, the right-invariant metric is given by

$$\int_{\mathbb{R}} \left[\left(y_t \circ y^{-1}\right)^2 + \left(\partial_x (y_t \circ y^{-1})\right)^2 \right] \, dx = \int_{\mathbb{R}} (u^2 + u_x^2) \, dx$$

and the preservation of the angular momentum writes

$$\big((u - u_{xx}) \circ y(t,\xi)\big) y_\xi(t,\xi)^2 = \big((u - u_{xx}) \circ y(0,\xi)\big) y_\xi(0,\xi)^2 \tag{8}$$

for all time and $\xi \in \mathbb{R}$, see [13, 14, 15].

Local well-posedness and blow-up of the solutions

Local existence and well-posedness of solutions to (1) have been studied in [25, 11] with the help of Kato's semi-group theory and in [23] using a regularization technique. It is shown that, for $u_0 \in H^s(\mathbb{R})$ with $s > \frac{3}{2}$, there exists a unique solution u with

$$u \in C([0,T), H^s(\mathbb{R})) \cap C^1([0,T), H^{\frac{1}{2}}(\mathbb{R}))$$

where $T > 0$ only depends on $\|u_0\|_{H^s(\mathbb{R})}$. Solutions have to be understood in the weak sense or in the sense of distribution. Equation (1) can be rewritten as

$$u_t + uu_x + P_x = 0, \tag{9a}$$

$$P - P_{xx} = u^2 + \frac{1}{2}u_x^2. \tag{9b}$$

The operator $1 - \partial_{xx}$ is a bijection from $\mathcal{S}'$ into $\mathcal{S}'$ where $\mathcal{S}'$ denotes the class of tempered distribution (see [19]) and a sufficient condition for (9) to hold in the sense of distribution is for example that $u \in L^1_{\text{loc}}(\mathbb{R}, H^1(\mathbb{R}))$.

These results hold only for a finite time interval and the CH equation, in contrast with the KdV equation, has smooth solutions that blow up in finite time. Due to their bi-hamiltonian structure, the KdV and CH equations possess infinitely many conserved quantities. In the case of the KdV equation these quantities provide some apriori control on the regularity of the solution and yield global existence and uniqueness of smooth solutions. However, this argument does not apply to the CH equation where only one such conserved quantity, the $H^1(\mathbb{R})$ norm, can be used that way. The solution blows up in the following manner. Let T be the time where a smooth solution eventually loses its regularity, i.e., $\lim_{t\to T}\|u(t,\cdot)\|_{H^s}=\infty$ for all $s>1$. Then,

$$\lim_{t\to T}\inf_{x\in\mathbb{R}} u_x(t,x)=-\infty. \tag{10}$$

There appears a point where the profile of u steepens gradually and ultimately the slope becomes vertical. In the context of water waves, this corresponds to the breaking of a wave. This fact was already noted in the seminal papers of Camassa and Holm ([7, 8]) and was subsequently proved by Constantin and Escher ([11, 12]). Wave breaking is an important physical phenomenon which is not captured by the other standard shallow water equations, as for example the KdV equation, and therefore makes the CH equation particularly interesting in that context.

The peakon and multipeakon as defined in (4) and (5) belong to $H^s(\mathbb{R})$ (for t fixed) only when $s<\frac{3}{2}$ and therefore are not included in the existence theorems mentioned above. The $H^1(\mathbb{R})$ norm is preserved by the equation, it plays a special role in the geometrical interpretation of the equation and $H^1(\mathbb{R})$ can be seen as the natural space for the equation. These facts motivate the investigation of an $H^1(\mathbb{R})$ theory for the CH equation.

Global existence of solutions

The first major step in this direction was accomplished by Constantin and Escher in [17]. They prove that, for $u_0\in H^1(\mathbb{R})$ and $u_0-u_{0,xx}\in\mathcal{M}^+(\mathbb{R})$, the space of positive Radon measure, equation (9) admit a unique global solution u in $C^1(\mathbb{R},L^2(\mathbb{R}))\cap C(\mathbb{R},H^1(\mathbb{R}))$. They proceed as follows. They consider a smooth approximation of the initial data, which for simplicity we also denote u_0, satisfying the sign condition $u_0-u_{0,xx}\geq 0$ and the corresponding solution u given by the local existence theory. Formally, it follows directly from (8) that

$$(u-u_{xx})(t,x)\geq 0 \tag{11}$$

for $t\in[0,T)$ and $x\in\mathbb{R}$; a rigorous proof of (11) is given in [11]. The fact that the sign of $u-u_{xx}$ is preserved leads to an apriori estimate for the total variation of u_x as the following simple (formal) computation shows. We have

$$\mathrm{TV}(u_x)=\|u_{xx}\|_{\mathcal{M}(\mathbb{R})}\leq 2\|u\|_{L^1(\mathbb{R})}, \tag{12}$$

because $\|u_{xx}\|_{\mathcal{M}(\mathbb{R})}=\int_{\mathbb{R}}|u_{xx}|\,dx\leq\int_{\mathbb{R}}|u-u_{xx}|\,dx+\int_{\mathbb{R}}|u|\,dx=2\int_{\mathbb{R}}|u|\,dx$. Since the operator $(1-\partial_{xx})^{-1}$ preserves positivity, we have $u\geq 0$ and $\int_{\mathbb{R}}|u|\,dx=\int_{\mathbb{R}}u\,dx=\int_{\mathbb{R}}u_0\,dx$ (we see directly from (9a) that $\int_{\mathbb{R}}u\,dx$ is a conserved quantity). Therefore, $\mathrm{TV}(u_x)\leq 2\|u_0\|_{L^1(\mathbb{R})}$. This apriori bound implies that u_x remains bounded and the blow-up situation given by (10) cannot occur. There is no wave breaking and the solution exists globally in time. Moreover, it gives enough control on the approximated solutions to prove by compactness the existence of solutions in $C(\mathbb{R},H^1(\mathbb{R}))$. In the two first papers, we present numerical schemes for the same class of initial data based on a finite difference scheme (**Paper I**)

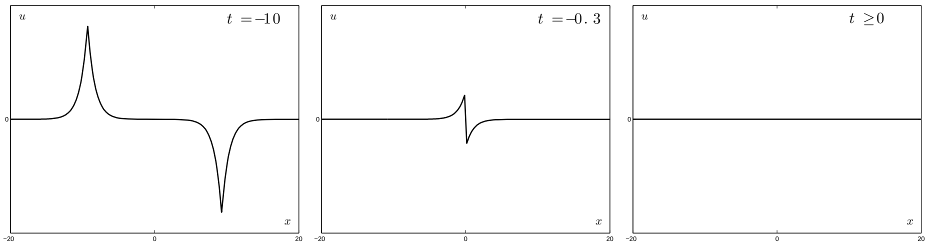


FIGURE 3. Peakon-antipeakon collision. First scenario: The dissipative solution.

and on multipeakons (**Paper II**). Both schemes preserve the positivity of $u - u_{xx}$. This is the key property that enables us to derive an apriori bound of the same type of (12) for our approximated solutions, and the convergence of the schemes is proved by a compactness argument.

In the case of arbitrary initial data in $H^1(\mathbb{R})$, solutions are no longer unique. To illustrate this fact we look at the following multipeakon configuration where a peakon traveling from the left to the right collides with another peakon going in the opposite direction (since this peakon has its peak pointing downwards, it is called antipeakon), see Figure 3. In the antisymmetric case, that is $p_1 = -p_2$ and $q_1 = -q_2$, the solution at collision time is identically zero. Then to prolong the solutions, two scenarios at least are possible. The first one consists of letting u remain identically zero after collusion. It can be checked directly that this is indeed a weak solution of (9). The second scenario is provided by Beals, Sattinger and Szmigielski in [2, 3] where they derive analytical solutions for the multipeakons by using scattering and inverse scattering techniques. For the solution they obtain in the antisymmetric peakon-antipeakon case, the peakons re-emerge after the collision in such a way that the transformation $t \mapsto -t$, $x \mapsto -x$, which let equation (1) invariant, also lets the solution invariant (here we assume that $t = 0$ at collision), see Figure 4. The solution is time reversible. Let $E(t) = \|u(t, \cdot)\|_{H^1(\mathbb{R})}$ denote the energy of the system. In the time reversible case illustrated in Figure 4, for any t different from collision time ($t \neq 0$), $E(t)$ remains equal to the same strictly positive constant, say $E(t) = 1$, while at collision time, we have $E(0) = 0$. In the other case illustrated in Figure 3 we have $E(t) = 1$ for $t < 0$ and $E(t) = 0$ for $t \geq 0$. One can prolong the solution after collision in infinitely many ways but the two scenarios we mentioned are really the only reasonable ones and the question is how they can be characterized and what is the selection principle that can be used to capture each of them.

By using viscous approximations of the equation, Xin and Zhang in [26] obtain the existence of a solution $u \in C([0, \infty) \times \mathbb{R}^1) \cap L^\infty((0, \infty), H^1(\mathbb{R}))$ to the CH equation for any initial data in $u_0 \in H^1(\mathbb{R})$. In particular their solution satisfies

$$E(t) = \|u(t, \cdot)\|_{H^1(\mathbb{R})} \leq E(t') = \|u(t', \cdot)\|_{H^1(\mathbb{R})} \tag{13}$$

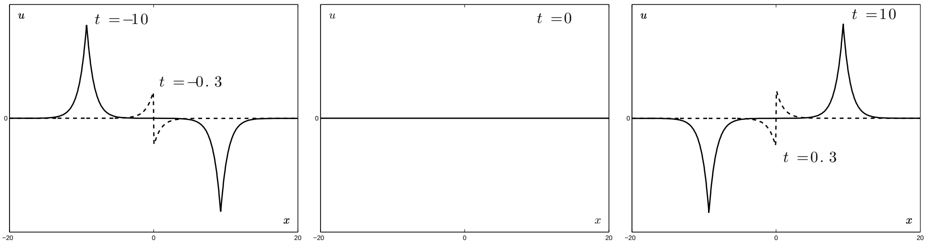


FIGURE 4. Peakon-antipeakon collision. Second scenario: The conservative solution.

for all $t < t'$ and the following one-sided super-norm estimate on u_x holds

$$u_x(t,x) \leq \frac{1}{t} + C, \quad t > 0,\ x \in \mathbb{R}. \tag{14}$$

Because of (13), we call these solutions dissipative solutions: the energy can only decay. In [9] it is proven that dissipative solutions are unique. In the symmetric peakon-antipeakon case, it is clear that the dissipative solution is the one corresponding to Figure 3.

In [4], Bressan and Constantin introduce a new set of variable,

$$w = u(t,y),\ v = 2\arctan u_x,\ q = (1+u_x^2)y_\xi \tag{15}$$

where $y(t,\xi)$ denotes the characteristics, i.e., $y_t(t,\xi) = u(t,y(t,\xi))$. They rewrite equation (1) uniquely in terms of these new variables and the system of equations they obtain, turns out to be a well-posed system of ordinary differential equation in a Banach space. Well-posed ordinary differential equation are time reversible and indeed this change of variable selects the time reversible solution in the antisymmetric peakon-antipeakon problem, which is given in Figure 4. The approach adopted in [5] by Bressan and Fonte is substantially different. They start by considering a system of multipeakons and describe the dynamic of the system, in particular how the multipeakons evolve throughout the collisions. At this stage, they select the conservative solution. Then, they introduce a distance functional inspired by optimal transport theory which satisfies

$$\frac{d}{dt}J(u(t),v(t)) \leq \kappa J(u(t),v(t)) \tag{16}$$

for any conservative multipeakons solutions u and v. Identity (16) is precisely the one needed to use Gronwall's Lemma and obtain stability results. General solutions to the CH equation are finally constructed from the multipeakons by a density argument.

Our approach in **Paper IV** is similar to the one of Bressan and Constantin. We reformulate the equation by using a new set of variables. The variables (y,U,H) we use have a natural interpretation from the Lagrangian point of view.

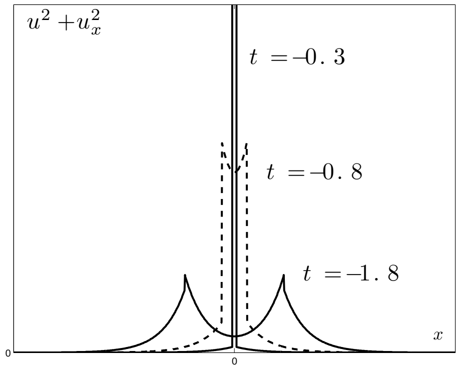


FIGURE 5. The energy density in the peakon-antipeakon case

As before, y denotes the characteristics and is given by

$$y_t(t,\xi) = u(t, y(t,\xi)) \tag{17}$$

while

$$U(t,\xi) = u(t, y(t,\xi)) \text{ and } H(t,\xi) = \int_{-\infty}^{y(t,\xi)} (u^2(t,x) + u_x^2(t,x))\, dx \tag{18}$$

correspond to the Lagrangian velocity and the cumulative energy distribution, respectively. Equation (9) can be rewritten as a system of ordinary differential equation in a Banach space involving uniquely (y, U, H). The system is well-posed and we obtain the global existence of solution.

The original equation (1) which corresponds to the Euler formulation of the problem contains only one unknown function, the velocity field u. The Lagrangian description as we introduced it in the first section contains two unknown functions: the position and the velocity of the particles, y and U. In order to explain why the extra variable H describing the energy distribution is needed, we look again at the peakon-antipeakon problem. At collision time, say $t = 0$, u is identically zero. Since zero is a global solution of (1), it is necessary, in order to select the conservative solution after collision, to take into account what happened before collision, to keep track in some way of the history of the system. The $H^1(\mathbb{R})$ norm of u is a preserved quantity and $\int_{\mathbb{R}} \left(u^2(t,x) + u_x^2(t,x)\right) dx$ remains equal to a constant, say 1, up to collision. In Figure 5, we plot the function $u^2(t,x) + u_x^2(t,x)$ at different times. As it can also be seen from Figure 4, $\lim_{t\to 0} \|u(t,\cdot)\|_{L^\infty(\mathbb{R})} = 0$ and $\lim_{t\to 0}\left\{\sup_{x\in\mathbb{R}\setminus[q_1,q_2]} |u_x(t,x)|\right\} = 0$ so that all the mass of $u^2 + u_x^2$ concentrates at the origin. We have

$$\lim_{t\to 0} (u^2(t,x) + u_x^2(t,x))\, dx = \delta(x)$$

where δ denotes the Dirac function. It is clear now that if we want to prolong the solution after collision while conserving the energy, we have to take into account the fact that the energy at collision time is concentrated at one point. We define the set $\mathcal{D}$ consisting of pairs (u,μ) such that $u\in H^1(\mathbb{R})$, μ is a Radon measure whose absolute continuous part satisfies

$$\mu_{\mathrm{ac}}=(u^2+u_x^2)\,dx. \tag{19}$$

The measure μ represents the energy density. It is strongly related to $(u^2+u_x^2)\,dx$ by (19) but at the same time it also allows the energy to concentrate on singular sets. In the case of peakon-antipeakon, we have $u(0)=0$ and $\mu(0)=\delta$ and it is clear that $(u(0),\mu(0))\in\mathcal{D}$. The question is now whether, knowing both u and μ, we are able to construct a solution in a unique way. For smooth solutions, a simple calculation shows that the energy density $u^2+u_x^2$ satisfies the following transport equation

$$(u^2+u_x^2)_t+(u(u^2+u_x^2))_x=(u^3-2Pu)_x. \tag{20}$$

Rewritten in terms of the Lagrangian variables, equation (20) takes the following simpler form

$$H_t=U^3-2P\circ y\;U. \tag{21}$$

From (9a), we obtain that $U_t=u_t(t,y)+y_tu_x(t,y)=u_t(t,y)+u(t,y)u_x(t,y)$ and therefore

$$U_t=-P_x\circ y. \tag{22}$$

Equations (17), (21) and (22) can be rewritten only in terms of (y,U,H) and they constitute a well-posed system of ordinary differential equation in a Banach space which admits global solutions in time. Going back to the Eulerian variable (u,μ), we prove that u is a weak solution of (9). Thus we have established the global existence of conservative solutions to the CH equation.

In the second part of the paper we address the question of stability and determine the topology on $\mathcal{D}$ which makes the conservative solutions stable. The set of Lagrangian coordinates, that we denote $\mathcal{G}$, can be given the topology induced by the Banach space in which they are embedded and stability with respect to initial data follows directly from the general theory of ordinary differential equation. Furthermore, we obtain the existence a continuous semigroup $S\colon\mathbb{R}_+\times\mathcal{G}\to\mathcal{G}$ of solutions in Lagrangian coordinates. If we assume for a moment that the Lagrangian and Eulerian coordinates are in bijection, that is, that there exists an invertible mapping f from $\mathcal{G}$ to $\mathcal{D}$, then we can define the semigroup of conservative solutions of (9) $T\colon\mathbb{R}_+\times\mathcal{D}\to\mathcal{D}$ by

$$T_t=f\circ S_t\circ f^{-1}.$$

The topology on $\mathcal{D}$ can simply be defined by transporting the metric on $\mathcal{G}$ into $\mathcal{D}$ by the mapping f, i.e.,

$$d_{\mathcal{D}}((u,\mu),(\bar u,\bar\mu))=d_{\mathcal{G}}(f(u,\mu),f(\bar u,\bar\mu)).$$

Since the semigroup S_t is continuous, this metric per definition makes also T_t continuous. As noted in [5], distances defined in terms of convex norms perform well in connection with linear problems, but occasionally fail when nonlinear features become dominant. This is the case here, the set $\mathcal{D}$ is not a vector space and the distance on $\mathcal{D}$ is not derived from a norm of some vector space containing $\mathcal{D}$. At the same time, the metric d_D is perfectly well-suited to the conservative solutions. With three variables in Lagrangian coordinates (y, U, H) versus two in Eulerian coordinates (u, μ), it is clear that the set $\mathcal{G}$ and $\mathcal{D}$ can not be in bijection. However, using the relabeling symmetry, it is possible to identify the Lagrangian variables that correspond to a same Eulerian configuration. Then, a bijection between the two coordinates systems can be established and we obtain the existence of a continuous semigroup in $\mathcal{D}$.

The Lagrangian variables are particularly well-suited to the study of the multipeakon solutions. From (5) and (6) we infer that

$$\dot{q}_i(t) = u(t, q_i(t)).$$

Hence, the positions of the peaks are given by the characteristics. By definition, we have that, for a multipeakon u, $u - u_{xx} = 0$ everywhere between the peaks. Furthermore the conservation of angular momentum (8), which is given in Lagrangian coordinates, tells us that this quantity remains zero. In **Paper V**, we prove that the conservative solutions preserve the multipeakon structure, i.e., multipeakons are conservative solutions in the sense defined in Paper IV. Moreover, we derive a system of ordinary differential equation, globally defined, for the conservative multipeakon solutions.

The Lagrangian approach is sufficiently robust to handle a larger class of equation. In **Paper VI**, we prove the existence of a global continuous semigroup of conservative solutions for

$$u_t - u_{xxt} + f(u)_x - f(u)_{xxx} + (g(u) + \frac{1}{2}f''(u)(u_x)^2)_x = 0 \tag{23}$$

with $f \in W^{3,\infty}_{\mathrm{loc}}(\mathbb{R})$, f strictly convex or concave, $g \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$. For $f(u) = \frac{u^2}{2}$ and $g(u) = \kappa u + u^2$, (23) gives the Camassa–Holm equation while, for $f(u) = \frac{\gamma u^2}{2}$ and $g(u) = \frac{3-\gamma}{2}u^2$, it gives the hyperelastic rod equation (2).

In **Paper III** we look at the smooth-solutions of (1) in contrast with the other papers where the focus was set on solutions with low spatial regularity. We prove the spectral convergence of the Fourier-Galerkin and a de-aliased Fourier-collocation for the Camassa–Holm equation.

References

[1] V. I. Arnold and B. A. Khesin. *Topological methods in hydrodynamics*, volume 125 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1998.

[2] R. Beals, D. H. Sattinger, and J. Szmigielski. Multi-peakons and a theorem of Stieltjes. *Inverse Problems*, 15(1):L1–L4, 1999.

[3] R. Beals, D. H. Sattinger, and J. Szmigielski. Multipeakons and the classical moment problem. *Adv. Math.*, 154(2):229–257, 2000.

[4] A. Bressan and A. Constantin. Global conservative solutions of the Camassa–Holm equation. *Preprint, Submitted*, 2005.

[5] A. Bressan and M. Fonte. An optimal transportation metric for solutions of the Camassa–Holm equation. *Preprint, Submitted*, 2005.
[6] R. Camassa. Characteristics and the initial value problem of a completely integrable shallow water equation. *Discrete Contin. Dyn. Syst. Ser. B*, 3(1):115–139, 2003.
[7] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.
[8] R. Camassa, D. D. Holm, and J. Hyman. A new integrable shallow water equation. *Adv. Appl. Mech.*, 31:1–33, 1994.
[9] G. M. Coclite, H. Holden, and K. H. Karlsen. Global weak solutions to a generalized hyperelastic-rod wave equation. *SIAM J. Math. Anal.*, 37(4):1044–1069 (electronic), 2005.
[10] A. Constantin. On the scattering problem for the Camassa-Holm equation. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 457(2008):953–970, 2001.
[11] A. Constantin and J. Escher. Global existence and blow-up for a shallow water equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 26(2):303–328, 1998.
[12] A. Constantin and J. Escher. Wave breaking for nonlinear nonlocal shallow water equations. *Acta Math.*, 181(2):229–243, 1998.
[13] A. Constantin and B. Kolev. Least action principle for an integrable shallow water equation. *J. Nonlinear Math. Phys.*, 8(4):471–474, 2001.
[14] A. Constantin and B. Kolev. On the geometric approach to the motion of inertial mechanical systems. *J. Phys. A*, 35(32):R51–R79, 2002.
[15] A. Constantin and B. Kolev. Geodesic flow on the diffeomorphism group of the circle. *Comment. Math. Helv.*, 78(4):787–804, 2003.
[16] A. Constantin and J. Lenells. On the inverse scattering approach to the Camassa-Holm equation. *J. Nonlinear Math. Phys.*, 10(3):252–255, 2003.
[17] A. Constantin and L. Molinet. Global weak solutions for a shallow water equation. *Comm. Math. Phys.*, 211(1):45–61, 2000.
[18] B. Fuchssteiner and A. S. Fokas. Symplectic structures, their Bäcklund transformations and hereditary symmetries. *Phys. D*, 4(1):47–66, 1981/82.
[19] H. Holden and X. Raynaud. Convergence of a finite difference scheme for the camassa–holm equation. *SIAM J. Numer. Anal. 2006*, to appear. **paper I** in this thesis.
[20] H. Holden and X. Raynaud. A convergent numerical scheme for the camassa–holm equation based on multipeakons. *Discrete Contin. Dyn. Syst.*, 14(3), 2006. **paper II** in this thesis.
[21] R. S. Johnson. Camassa–Holm, Korteweg-de Vries and related models for water waves. *J. Fluid Mech.*, 455:63–82, 2002.
[22] J. Lenells. Conservation laws of the Camassa-Holm equation. *J. Phys. A*, 38(4):869–880, 2005.
[23] Y. A. Li and P. J. Olver. Well-posedness and blow-up solutions for an integrable nonlinearly dispersive model wave equation. *J. Differential Equations*, 162(1):27–63, 2000.
[24] P. J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.
[25] G. Rodríguez-Blanco. On the Cauchy problem for the Camassa-Holm equation. *Nonlinear Anal.*, 46(3, Ser. A: Theory Methods):309–327, 2001.
[26] Z. Xin and P. Zhang. On the weak solutions to a shallow water equation. *Comm. Pure Appl. Math.*, 53(11):1411–1433, 2000.

Paper I

Convergence of a finite different scheme for the Camassa-Holm equation.

H. Holden and X. Raynaud

To appear in *SIAM J. Numer. Anal.*

CONVERGENCE OF A FINITE DIFFERENCE SCHEME FOR THE CAMASSA–HOLM EQUATION

HELGE HOLDEN AND XAVIER RAYNAUD

ABSTRACT. We prove that a certain finite difference scheme converges to the weak solution of the Cauchy problem on a finite interval with periodic boundary conditions for the Camassa–Holm equation $u_t - u_{xxt} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0$ with initial data $u|_{t=0} = u_0 \in H^1([0,1])$. Here it is assumed that $u_0 - u_0'' \geq 0$ and in this case, the solution is unique, globally defined, and energy preserving.

1. INTRODUCTION

The Camassa–Holm equation (CH) [3]

$$u_t - u_{xxt} + 2\kappa u_x + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0 \tag{1.1}$$

has received considerable attention the last decade. With κ positive it models, see [4, 16, 12], propagation of unidirectional gravitational waves in a shallow water approximation, with u representing the fluid velocity. The Camassa–Holm equation possesses many intriguing properties: It is, for instance, completely integrable and experiences wave breaking in finite time for a large class of initial data. Most attention has been given to the case with $\kappa = 0$ on the full line, that is,

$$u_t - u_{xxt} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0, \tag{1.2}$$

which has so-called peakon solutions, i.e., solutions of the form $u(x,t) = ce^{-|x-ct|}$ for real constants c. Local and global well-posedness results as well as results concerning breakdown are proved in [9, 14, 17, 20].

In this paper we study the Camassa–Holm equation (1.1) on a finite interval with periodic boundary conditions. It is known that certain initial data give global solutions, while other classes of initial data experience wave breaking in the sense that u_x becomes unbounded while the solution itself remains bounded. It suffices to treat the case $\kappa = 0$, since solutions with nonzero κ are obtained from solutions with zero κ by the transformation $v(x,t) = u(x + \kappa t, t) - \kappa$. More precisely, the fundamental existence theorem, due to Constantin and Escher [10], reads as follows: If $u_0 \in H^3([0,1])$ and $m_0 := u_0 - u_0'' \in H^1([0,1])$ is non-negative, then equation (1.2) has a unique global solution

1991 *Mathematics Subject Classification.* Primary: 65M06, 65M12; Secondary: 35B10, 35Q53.

Key words and phrases. Camassa–Holm equation, convergence of finite difference schemes.

$u \in C([0,T), H^3([0,1])) \cap C^1([0,T), H^2([0,1]))$ for any T positive. However, if $m_0 \in H^1([0,1])$, u_0 not identically zero but $\int m_0\,dx = 0$, then the maximal time interval of existence is finite. Furthermore, if $u_0 \in H^1([0,1])$ and $m_0 = u_0 - u_0''$ is a positive Radon measure on $[0,1]$, then (1.2) has a unique global weak solution. Additional results in the periodic case can be found in [7, 10, 8, 11, 18]. Numerical results can be found in [4] where Camassa, Holm, and Hyman study (1.2) using a pseudospectral method. Numerical schemes based on multipeakons are examined in [2, 6, 5, 15].

In this paper, we prove convergence of a particular finite difference scheme for the equation, thereby giving a constructive approach to the actual determination of the solution). We work in the case where one has global solutions, that is, when $m_0 \geq 0$. The scheme is semi-discrete: Time is not discretized, and we have to solve a system of ordinary differential equations. We reformulate (1.1) to give meaning in $C([0,T]; H^1[0,1])$ to solutions such as peakons, and we prove that our scheme converges in $C([0,T]; H^1[0,1])$.

More precisely, we prove the following: Assume that v^n is a sequence of continuous, periodic and piecewise linear functions on intervals $[(i-1)/n, i/n]$, $i = 1, \dots, n$, that converges to the initial data v in $H^1([0,1])$ as $n \to \infty$. Let $u^n = u^n(x,t)$ be the solution of the following system of equations

$$\begin{aligned} m_t^n &= -D_-(m^n u^n) - m^n D u^n \\ m^n &= u^n - D_- D_+ u^n \end{aligned} \tag{1.3}$$

with initial condition $u^n|_{t=0} = v^n$. Here $D_\pm$ denotes forward and backward difference operators relative to the lattice with spacing $1/n$, and $D = (D_+ + D_-)/2$. Extrapolate u^n from its lattice values at points i/n to obtain a continuous, periodic, and piecewise linear function also denoted u^n. Assume that $v^n - D_- D_+ v^n \geq 0$. Then u^n converges in $C([0,T]; H^1([0,1]))$ as $n \to \infty$ to the solution u of the Camassa–Holm equation with initial condition $u|_{t=0} = v$. The result includes the case when the initial data $v \in H^1$ is such that $v - v_{xx}$ is a positive Radon measure, see Corollary 2.5. For the actual computations we discretize (1.3) using the forward Euler method. We prove convergence of that method, see Theorem 3.1.

The numerical scheme (1.3) is tested on various initial data. In addition, we study experimentally the convergence of other numerical schemes for the Camassa–Holm equation. The numerical results are surprisingly sensitive in the explicit form of the scheme, and, among the various schemes we have implemented, only the scheme (1.3) converges to the unique solution.

2. Convergence of the numerical scheme

We consider periodic boundary conditions and solve the equation on the interval $[0,1]$. We are looking for solutions that belong to $H^1([0,1])$ which is the natural space for the equation. Introduce the partition of $[0,1]$ in points separated by a distance $h = 1/n$ denoted $x_i = hi$ for $i = 0, \dots, n-1$. For any $(u_0, \dots, u_{n-1})$ in $\mathbb{R}^n$, we can define a continuous, periodic, piecewise linear function u by

$$u(x_i) = u_i, \tag{2.1}$$

in other words, the periodic polygon that passes through the points (x_i, u_i) for $i = 0, \dots, n-1$. It defines a bijection between $\mathbb{R}^n$ and the set of continuous, periodic, piecewise linear function with possible break points at x_i, and we will use this bijection throughout this paper.

Given $u = (u_0, \dots, u_{n-1})$, the quantity $D_\pm u$ given by

$$(D_\pm u)_i = \frac{\pm 1}{h}(u_{i\pm 1} - u_i)$$

gives the right and left derivatives, respectively, of u at x_i. In these expressions, u_{-1} and u_n are derived from the periodicity conditions: $u_{-1} = u_{n-1}$ and $u_n = u_0$. The average Du between the left and right derivative is given by

$$(Du)_i = \frac{1}{2}\big((D_+u)_i + (D_-u)_i\big) = \frac{1}{2h}(u_{i+1} - u_{i-1}).$$

The Camassa–Holm equation preserves the H^1-norm. In order to see that, we rewrite (1.2) in its Hamiltonian form, see [3]

$$m_t = -(mu)_x - mu_x \tag{2.2}$$

with

$$m = u - u_{xx}. \tag{2.3}$$

Assuming that u is smooth enough so that the integration by parts can be carried out, we get

$$\begin{aligned}
\frac{d}{dt}\left\|u\right\|_{H^1}^2 &= 2\int_0^1 (u_t - u_{xxt})u\,dx = 2\int_0^1 um_t\,dx \\
&= -2\int_0^1 u(mu)_x\,dx - 2\int_0^1 umu_x\,dx \\
&= 2\int_0^1 u_x mu\,dx - 2\int_0^1 umu_x\,dx = 0,
\end{aligned}$$

and the H^1 norm of u is preserved.

From (2.3) and (2.2), we derive a finite difference approximation scheme for the Camassa–Holm equation, and prove that it converges to the right solution. This is our main result.

Theorem 2.1. *Let v^n be a sequence of continuous, periodic and piecewise linear functions on $[0,1]$ that converges to v in $H^1([0,1])$ as $n \to \infty$ and such that $v^n - D_-D_+v^n \geq 0$. Then, for any given $T > 0$, the sequence $u^n = u^n(x,t)$ of continuous, periodic and piecewise linear functions determined by the system of ordinary differential equations*

$$\begin{aligned}
m_t^n &= -D_-(m^n u^n) - m^n Du^n \\
m^n &= u^n - D_-D_+u^n
\end{aligned} \tag{2.4}$$

with initial condition $u^n|_{t=0} = v^n$, converges in $C([0,T]; H^1([0,1]))$ as $n \to \infty$ to the solution u of the Camassa–Holm equation (1.2) with initial condition $u|_{t=0} = v$.

If we interpret the functions as vectors in (2.4), cf. (2.1), the multiplications are term-by-term multiplications of vectors. We also have to rewrite equation (1.2) in order to make it well-defined in the sense of distributions for functions that at least belong to $C([0,T];H^1([0,1]))$, more precisely,

$$u_t - u_{xxt} = -\frac{3}{2}(u^2)_x - \frac{1}{2}(u_x^2)_x + \frac{1}{2}(u^2)_{xxx}. \tag{2.5}$$

A function u in $L^\infty([0,T];H^1)$ is said to be solution of the periodic Camassa–Holm equation if it is periodic and satisfies (2.5) in the sense of distributions. In [11], a different definition of weak solutions for the Camassa–Holm equation is presented. After proving our main theorem at the end of this section, we also prove that these two definitions are equivalent.

In order to solve equation (2.4), we need to compute u^n from m^n. It is simpler first to consider sequences that are defined in $\mathbb{R}^{\mathbb{Z}}$, the set of all sequences, and then discuss the periodic case. Let L denote the linear operator from $\mathbb{R}^{\mathbb{Z}}$ to $\mathbb{R}^{\mathbb{Z}}$ given, for all $u \in \mathbb{R}^{\mathbb{Z}}$, by

$$Lu = u - D_-D_+u.$$

We want to find an expression for L^{-1}. Introduce the Kronecker delta by $\delta_i = 1$ if $i = 0$ and zero otherwise. It is enough to find a solution g of

$$Lg = \delta$$

which decays sufficently fast at infinity because $L^{-1}m$ is then given, for any bounded $m \in \mathbb{R}^{\mathbb{Z}}$, by the discrete convolution product of g and m:

$$L^{-1}m_i = \sum_{j\in\mathbb{Z}} g_{i-j}m_j.$$

The function g satisfies for i nonzero

$$g_i - n^2(g_{i+1} - 2g_i + g_{i-1}) = 0. \tag{2.6}$$

The general solution of (2.6) for all $i \in \mathbb{Z}$ is given by

$$g_i = Ae^{\kappa_1 i} + Be^{\kappa_2 i}$$

where A, B are constants, $\kappa_1 = \ln x_1$, $\kappa_2 = \ln x_2$, and x_1 and x_2 are the solutions of

$$-n^2x^2 + (1+2n^2)x - n^2 = 0.$$

Here x_1 and x_2 are real and positive, and $x_1x_2 = 1$ implies that $\kappa_2 = -\kappa_1$. We set $\kappa = \kappa_1 = -\kappa_2$. After some calculations, we get

$$\kappa = \ln\left(\frac{1+2n^2+\sqrt{1+4n^2}}{2n^2}\right). \tag{2.7}$$

We take g of the form

$$g_i = c\,e^{-\kappa|i|}$$

so that g satisfies (2.6) for all $i \neq 0$ and decays at infinity. The constant c is determined by the condition that $(Lg)_0 = 1$ which yields

$$c = \frac{1}{1 + 2n^2(1 - e^{-\kappa})}.$$

We periodize g in the following manner:

$$g_i^p \equiv \sum_{k\in\mathbb{Z}} g_{i+kn} = c\frac{e^{-\kappa i} + e^{\kappa(i-n)}}{1 - e^{-\kappa n}}$$

for $i \in \{0, \dots, n-1\}$. The inverse of L on the set of periodic sequences is then given by

$$u_i = L^{-1}m_i = \sum_{j=0}^{n-1} g_{i-j}^p m_i = \frac{c}{1 - e^{-\kappa n}} \sum_{j=0}^{n-1}(e^{-\kappa(i-j)} + e^{\kappa(i-j-n)})m_j. \tag{2.8}$$

Hence,

$$L\big(\sum_{j=0}^{n-1} g_{i-j}^p m_j\big)_i = L\big(\sum_{l\in\mathbb{Z}} g_{i-l}m_l\big)_i = m_i.$$

For sufficiently smooth initial data ($u_0 \in H^3$ and $m_0 \in H^1$) which satisfies $m_0 \geq 0$, Constantin and Escher [9] proved that there exists a unique global solution of the Camassa–Holm equation belonging to $C(\mathbb{R}_+; H^3) \cap C^1(\mathbb{R}_+; H^2)$. The proof of this result relies heavily on the fact that if m is non-negative at $t = 0$, then m remains non-negative for all $t > 0$. An important feature of our scheme is that it preserves this property. (For simplicity we have here dropped the superscript n appearing on u and m.)

Lemma 2.2. *Assume that $m_i(0) \geq 0$ for all $i = 0, \dots, n-1$. For any solution $u(t)$ of the system* (2.4)*, we have that $m_i(t) \geq 0$ for all $t \geq 0$ and for all $i = 0, \dots, n-1$.*

Proof. Let us assume that there exist $t > 0$ and $i \in \{0, \dots, n-1\}$ such that

$$m_i(t) < 0. \tag{2.9}$$

We consider the time interval F in which m remains positive:

$$F = \{t \geq 0 \mid m_i(\tilde{t}) \geq 0, \text{for all } \tilde{t} \leq t \text{ and } i \in \{0, \dots, n-1\}\}.$$

Because of assumption (2.9), F is bounded and we define

$$T = \sup F.$$

By definition of T, for any integer $j > 0$, there exists a $\tilde{t}_j$ and an i_j such that $T < \tilde{t}_j < T + \frac{1}{j}$ and $m_{i_j}(\tilde{t}_j) < 0$. The function $m_{i_j}(t)$ is a continuously differentiable function of t. Hence, $m_{i_j}(T) \geq 0$ and there exists a t_j such that

$$m_{i_j}(t_j) = 0,$$

with $T \leq t_j < T + \frac{1}{j}$.

Since i_j can only take a finite number of values ($i_j \in \{0, \dots, n-1\}$), there exists a $p \in \{0, \dots, n-1\}$ and a subsequence j_k such that $i_{j_k} = p$. The function $m_p(t)$ belongs to C^1 and, since $t_{j_k} \to T$, we have

$$m_p(T) = 0. \tag{2.10}$$

We denote by G the set of indices for which (2.10) holds:

$$G = \{k \in \{0, \dots, n-1\} \mid m_k(T) = 0\}.$$

G is non-empty because it contains p. If $G = \{0, \dots, n-1\}$, then $m_k(T) = 0$ for all k and m must be the zero solution because we know from Picard's theorem that the solution of (2.4) is unique.

If $G \neq \{0, \dots, n-1\}$, then there exists an $l \in \{0, \dots, n-1\}$ such that

$$m_{l-1}(T) > 0, \quad m_l(T) = 0, \quad \frac{dm_l}{dt}(T) \leq 0. \tag{2.11}$$

The last condition, $\frac{dm_l}{dt}(T) \leq 0$, comes from the definition of T that would be contradicted if we had $\frac{dm_l}{dt}(T) > 0$. Note that we also use the periodicity of m which in particular means that if $l = 0$, then $m_{l-1}(T) = m_{-1}(T) = m_{n-1}(T)$.

In (2.4), for $i = l$ and $t = T$, the terms involving $m_l(T)$ cancel and

$$\frac{dm_l}{dt}(T) = \frac{m_{l-1}(T) u_{l-1}(T)}{h}.$$

The fact that all the $m_i(T)$ are positive with one of them, $m_{l-1}(T)$, strictly positive, implies that u_i is strictly positive for all indices i, see (2.8). Since, in addition, $m_{l-1}(T) > 0$, we get

$$\frac{dm_l}{dt}(T) > 0$$

which contradicts the last inequality in (2.11) and therefore our primary assumption (2.9) does not hold. The lemma is proved. $\square$

We want to establish a uniform bound on the H^1 norm of the sequence u^n. Recall that u^n is a continuous piecewise linear function (with respect to the space variable), and its L^2 norm can be computed exactly. We find

$$\|u^n\|_{L^2}^2 = \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{3}((u_{i+1}^n)^2 + u_i^n u_{i+1}^n + (u_i^n)^2). \tag{2.12}$$

The derivative u_x^n of u^n is piecewise constant and therefore we have

$$\|u_x^n\|_{L^2}^2 = \frac{1}{n} \sum_{i=0}^{n-1} (D_+ u^n)_i^2. \tag{2.13}$$

We define a renormalized norm $\|\cdot\|_{l^2}$ and the corresponding scalar product on $\mathbb{R}^n$ by

$$\|u^n\|_{l^2} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (u_i^n)^2}, \quad \langle u^n, v^n \rangle_{l^2} = \frac{1}{n} \sum_{i=0}^{n-1} u_i^n v_i^n.$$

The following inequalities hold

$$\frac{1}{2}\left\|u^n\right\|_{l^2} \leq \left\|u^n\right\|_{L^2} \leq \left\|u^n\right\|_{l^2} \tag{2.14}$$

which make the two norms $\|\cdot\|_{l^2}$ and $\|\cdot\|_{L^2}$ uniformly equivalent independently of n. In (2.14), u^n either denotes an element of $\mathbb{R}^n$ or the corresponding continuous piecewise linear function as defined previously. By using the Cauchy–Schwarz inequality and the periodicity of u^n, it is not hard to prove that

$$\left\|u^n\right\|_{L^2} \leq \left\|u^n\right\|_{l^2}.$$

For the other equality, it suffices to see that (2.12) can be rewritten as

$$\|u^n\|_{L^2}^2 = \frac{1}{3n}\sum_{i=0}^{n-1}\left[(u_{i+1}^n + \frac{1}{2}u_i^n)^2 + \frac{3}{4}(u_i^n)^2\right]$$

which implies

$$\frac{1}{2}\left\|u^n\right\|_{l^2} \leq \left\|u^n\right\|_{L^2}.$$

We are now in position to establish a uniform bound on the H^1-norm of u^n. Let $E_n(t)$ denote

$$E_n(t) = \left(\|u^n(t)\|_{l^2}^2 + \|D_+u^n(t)\|_{l^2}^2\right)^{\frac{1}{2}} \tag{2.15}$$

which provides an approximation of the H^1-norm of $u^n(t)$. We have, from (2.14) and (2.13),

$$\frac{1}{2}\left\|u^n(t)\right\|_{H^1} \leq E_n(t) \leq \left\|u^n(t)\right\|_{H^1}. \tag{2.16}$$

The derivative of $E_n(t)^2$ reads

$$\begin{aligned}
\frac{dE_n(t)^2}{dt} &= \frac{2}{n}\sum_{i=0}^{n-1}\left[u_i^n u_{i,t}^n + D_+u_i^n D_+u_{i,t}^n\right] \\
&= \frac{2}{n}\sum_{i=0}^{n-1}(u_i^n - D_-D_+u_i^n)_t u_i^n \quad \text{(summation by parts)} \\
&= -\frac{2}{n}\sum_{i=0}^{n-1}\left[D_-(m^n u^n)_i u_i^n + m_i^n D u_i^n u_i^n\right] \quad \text{by (2.4)} \\
&= \frac{2}{n}\sum_{i=0}^{n-1}\left[m_i^n u_i^n (D_+u_i^n - Du_i^n)\right].
\end{aligned}$$

Since

$$D_+u_i^n - Du_i^n = \frac{1}{2}\left[D_+u_i^n - D_+u_{i-1}^n\right] = \frac{1}{2n}D_-D_+u_i^n,$$

we get

$$\frac{dE_n(t)^2}{dt} = \frac{1}{n}\sum_{i=0}^{n-1}\left[m_i^n u_i^n \frac{1}{n} D_-D_+u_i^n\right] = \frac{1}{n^2}\sum_{i=0}^{n-1}\left[m_i^n u_i^n(-m_i^n + u_i^n)\right], \tag{2.17}$$

and, because u_i^n is positive (see (2.8)),

$$\frac{dE_n^2(t)}{dt} \leq \frac{1}{n^2}\sum_{i=0}^{n-1} m_i^n (u_i^n)^2. \tag{2.18}$$

A summation by parts gives us that

$$\frac{1}{n}\sum_{i=0}^{n-1} m_i^n u_i^n = E_n(t)^2.$$

Since L^∞ is continuously embedded in H^1, there exists a constant $\mathcal{O}(1)$, independent of n, such that

$$\max_i u_i^n \leq \mathcal{O}(1)\left\|u^n\right\|_{H^1} \leq \mathcal{O}(1)E_n(t).$$

Hence, (2.18) implies

$$E_n'(t) \leq \frac{\mathcal{O}(1)}{n}E_n(t)^2$$

and, after integration,

$$\frac{1}{E_n(t)} \geq \frac{1}{E_n(0)} - \frac{\mathcal{O}(1)}{n}t.$$

Since $u^n(0) = v^n$ tends to v in H^1, $\left\|u^n(0)\right\|_{H^1}$ and therefore $E_n(0)$ are bounded. It implies that $E_n(0)^{-1}$ is bounded from below by a strictly positive constant and, for any given $T > 0$, there exists $N \geq 0$ and constant $C' > 0$ such that for all $n \geq N$ and all $t \in [0, T]$, we have $E_n(0)^{-1} - \mathcal{O}(1)t/n \geq 1/C'$. Hence,

$$\left\|u^n\right\|_{H^1} \leq 2E_n(t) \leq 2C' \tag{2.19}$$

and, by (2.16), the H^1-norm of $u^n(t)$ is uniformly bounded in $[0, T]$. This result also guarantees the existence of solutions to (2.4) in $[0, T]$ (at least, for n big enough) because, on $[0, T]$, we have that $\max_i |u_i^n(t)| = \left\|u^n(\,\cdot\,, t)\right\|_{L^\infty} \leq \mathcal{O}(1)\left\|u^n(t)\right\|_{H^1}$ remains bounded.

To prove that we can extract a converging subsequence of u^n, we need some estimates on the derivative of u^n.

Lemma 2.3. *We have the following properties:*
(i) u_x^n is uniformly bounded in $L^\infty([0,1])$.
(ii) u_x^n has a uniformly bounded total variation.
(iii) u_t^n is uniformly bounded in $L^2([0,1])$.

Proof. (i) From (2.8), we get

$$D_+u_i^n = \frac{c}{1-e^{-\kappa n}}\sum_{j=0}^{n-1}\left[m_j^n e^{-\kappa(i-j)}\left(\frac{e^{-\kappa}-1}{h}\right) + m_j^n e^{\kappa(i-j-n)}\left(\frac{e^{\kappa}-1}{h}\right)\right]$$

where κ is given by (2.7).

One easily gets the following expansion for κ as h tends to 0

$$\kappa = h + o(h^2),$$

which implies that for all $i \in \{0, \dots, n-1\}$,

$$
\begin{aligned}
|D_+ u_i^n| &\le (1 + \mathcal{O}(h)) \frac{c}{1 - e^{-\kappa n}} \sum_{j=0}^{n-1} \left(\left|m_j^n\right| e^{-\kappa(i-j)} + \left|m_j^n\right| e^{\kappa(i-j-n)} \right) \\
&\le (1 + \mathcal{O}(h)) \frac{c}{1 - e^{-\kappa n}} \sum_{j=0}^{n-1} \left(m_j^n e^{-\kappa(i-j)} + m_j^n e^{\kappa(i-j-n)} \right) \\
&\le (1 + \mathcal{O}(h))\, u_i^n, \qquad (2.20)
\end{aligned}
$$

where we have used the positivity of m^n and relation (2.8). Hence, since $\|u^n\|_{L^\infty}$ is uniformly bounded, we get a uniform bound on $\|u_x^n\|_{L^\infty}$.

(ii) For each t the total variation of $u_x^n(\,\cdot\,, t)$ is given by

$$
\mathrm{TV}(u_x^n) = \sup_{\phi \in C^1, \|\phi\|_{L^\infty} \le 1} \int_0^1 u_x^n(x) \phi_x(x)\, dx.
$$

On the interval (x_i, x_{i+1}), the function u_x^n is constant and equal to $D_+ u_i^n$. Therefore,

$$
\begin{aligned}
\int_0^1 u_x^n(x) \phi_x(x)\, dx &= \sum_{i=0}^{n-1} D_+ u_i^n \int_{x_i}^{x_{i+1}} \phi_x(x)\, dx = \sum_{i=0}^{n-1} D_+ u_i^n (\phi(x_{i+1}) - \phi(x_i)) \\
&= \sum_{i=0}^{n-1} \frac{1}{n} D_+ u_i^n D_+ \phi(x_i) = - \sum_{i=0}^{n-1} \frac{1}{n} (D_- D_+ u_i^n) \phi(x_i)
\end{aligned}
$$

and

$$
\mathrm{TV}(u_x^n) \le \frac{1}{n} \sum_{i=0}^{n-1} |D_- D_+ u_i^n|\,.
$$

Since m_i^n and u_i^n are positive for all i,

$$
|D_- D_+ u_i^n| = |m_i^n - u_i^n| \le m_i^n + u_i^n \le 2u_i^n - D_- D_+ u_i^n.
$$

When summing over i on the right-hand side of the last inequality, the term $D_- D_+ u_i^n$ disappears and we get

$$
\mathrm{TV}(u_x^n) \le 2 \max_i u_i^n \le \mathcal{O}(1) \|u^n\|_{H^1} \le \mathcal{O}(1)
$$

for all t.

(iii) In order to make the ideas clearer, we first sketch the proof directly on equation (2.2). Assuming that m is positive and u is in H^1, we see how, from (2.2), u_t can be defined as an element of $L^2([0,1])$. This will be useful when we afterwards derive a uniform bound for u_t^n in $L^2([0,1])$.

For all smooth v, we have

$$
\int_0^1 u_t\, v\, dx = \int_0^1 (\mathcal{L}^{-1} m_t)\, v\, dx
$$

where $\mathcal{L}$ denotes the operator $\mathcal{L}u = u - u_{xx}$, which is a self-adjoint homeomorphism from H^2 to L^2. If we let $w = \mathcal{L}^{-1}v$, the continuity of $\mathcal{L}^{-1}$ implies

$$\|w\|_{H^2} \leq \mathcal{O}(1)\,\|v\|_{L^2} \tag{2.21}$$

for some constant $\mathcal{O}(1)$ independent of v.

We find

$$\begin{aligned}\int_0^1 u_t\,v\,dx &= \int_0^1 \left(\mathcal{L}^{-1}m_t\right)v\,dx = \int_0^1 m_t\,\mathcal{L}^{-1}v\,dx \quad (\mathcal{L}^{-1} \text{ is self-adjoint})\\ &= -\int_0^1((mu)_x + mu_x)w\,dx = \int_0^1 (muw_x - mu_xw)\,dx.\end{aligned}$$

The integrals here must be understood as distributions. Even so, some terms (like mu_x) are not well-defined as distributions. However, we get the same results rigorously by considering the equation written as a distribution (2.5). We have:

$$\begin{aligned}\left|\int_0^1 u_t\,v\,dx\right| &\leq \int_0^1 (|muw_x| + |mu_xw|)\,dx\\ &\leq \left(\|u\|_{L^\infty}\|w_x\|_{L^\infty} + \|u_x\|_{L^\infty}\|w\|_{L^\infty}\right)\int_0^1 |m|\;dx.\end{aligned}$$

Recall that $\|u\|_{L^\infty}$ and $\|u_x\|_{L^\infty}$ are uniformly bounded. Furthermore, m positive implies $\int_0^1 |m| = \int_0^1 m = \int_0^1 u \leq \|u\|_{L^\infty}$ and therefore m is also uniformly bounded. From (2.21) and the fact that H^1 is continuously embedded in L^∞, we get

$$\|w_x\|_{L^\infty} \leq \mathcal{O}(1)\,\|w_x\|_{H^1} \leq \mathcal{O}(1)\,\|w\|_{H^2} \leq \mathcal{O}(1)\,\|v\|_{L^2}\,,$$

and similarly

$$\|w\|_{L^\infty} \leq \mathcal{O}(1)\,\|v\|_{L^2}\,.$$

Finally,

$$\left|\int_0^1 u_t\,v\,dx\right| \leq \mathcal{O}(1)\,\|v\|_{L^2}$$

which implies, by Riesz's representation theorem, that u_t is in L^2 and

$$\|u_t\|_{L^2} \leq \mathcal{O}(1).$$

We now turn to the analogous derivations in the discrete case. Consider the sequence u^n. The aim is to derive a uniform bound for u^n_t in L^2. We take a continuous piecewise linear function v^n,

$$\langle u^n_t, v^n\rangle_{l^2} = \left\langle L^{-1}m^n_t, v^n\right\rangle_{l^2} = \left\langle m^n_t, L^{-1}v^n\right\rangle_{l^2} \tag{2.22}$$

because L and therefore L^{-1} are self-adjoint.

Let w^n denote

$$w^n = L^{-1}v^n.$$

We have

$$\langle v^n, w^n\rangle_{l^2} = \langle Lw^n, w^n\rangle_{l^2} = \frac{1}{n}\sum_{i=0}^{n-1}\left(w^n_i - D_-D_+w^n_i\right)w^n_i = \frac{1}{n}\sum_{i=0}^{n-1}\left[(w^n_i)^2 + (D_+w^n_i)^2\right].$$

Then, after using (2.16) and Cauchy–Schwarz, we get

$$\|w^n\|_{H^1}^2 \leq 4\,\|v^n\|_{l^2}\,\|w^n\|_{l^2}\,.$$

By (2.14), (2.16) we find

$$\|w^n\|_{H^1}^2 \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,\|w^n\|_{H^1}$$

and

$$\|w^n\|_{H^1} \leq \mathcal{O}(1)\,\|v^n\|_{l^2} \tag{2.23}$$

where $\mathcal{O}(1)$ is a constant independent of n. Since H^1 is continuously embedded in L^∞, we get

$$\max_i |w_i^n| \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,. \tag{2.24}$$

Let us define y^n as follows

$$y_i^n = (D_+ w^n)_{i-1}.$$

We want to find a bound on y^n. From (2.14) and (2.23), we get

$$\|y^n\|_{l^2} \leq \|w^n\|_{H^1} \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,. \tag{2.25}$$

We also have, using the definition of y^n and w^n,

$$D_+ y^n = D_- D_+ w^n = w^n - v^n$$

which gives

$$\|D_+ y^n\|_{l^2} \leq \mathcal{O}(1)\,\|v^n\|_{l^2} \tag{2.26}$$

because, by (2.23),

$$\|w^n\|_{l^2} \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,.$$

Equations (2.25), (2.26), and (2.16) give us a uniform bound on the H^1 norm of y^n:

$$\|y^n\|_{H^1} \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,.$$

Since H^1 is continuously embedded in L^∞, we get

$$\max_i |D_+ w_i^n| = \max_i |y_i^n| = \|y^n\|_{L^\infty} \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,. \tag{2.27}$$

Going back to (2.22), we have

$$\begin{aligned}\langle u_t^n, v^n\rangle_{l^2} &= \langle m_t^n, w^n\rangle_{l^2} = \langle -D_-(m^n u^n) - m^n D u^n, w^n\rangle_{l^2}\\ &= \langle m^n u^n, D_+ w^n\rangle_{l^2} - \langle m^n D u^n, w^n\rangle_{l^2}\,.\end{aligned}$$

Hence,

$$|\langle u_t^n, v^n\rangle_{l^2}| \leq \frac{1}{n}\left(\max_i |u_i^n| \max_i |D_+ w_i^n| + \max_i |D_+ u_i^n| \max_i |w_i^n|\right) \sum_{i=0}^{n-1} |m_i^n|\,.$$

The functions u_i^n and $D_+u_i^n$ are uniformly bounded with respect to n. and

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}|m_i^n| &= \frac{1}{n}\sum_{i=0}^{n-1}m_i^n \quad &&(m^n \text{ is positive})\\ &= \frac{1}{n}\sum_{i=0}^{n-1}u_i^n \quad &&(\text{cancellation of } \sum_{i=0}^{n-1}D_-D_+u_i^n)\\ &\leq \mathcal{O}(1). &&(u_i^n \text{ is bounded})\end{aligned}$$

Finally, using the bounds we have derived on w^n, see (2.24), and D_+w^n, see (2.27), we get

$$|\langle u_t^n, v^n\rangle_{l^2}| \leq \mathcal{O}(1)\,\|v^n\|_{l^2}\,.$$

Taking $v^n = u_t^n$ yields

$$\|u_t^n\|_{l^2} \leq \mathcal{O}(1)$$

which, since the l^2 and L^2 norm are uniformly equivalent, gives us a uniform bound on $\|u_t^n\|_{L^2}$. □

To prove the existence of a converging subsequence of u^n in $C([0,T],H^1)$ we recall the following compactness theorem given by Simon [21, Corollary 4].

Theorem 2.4 (Simon)**.** *Let X, B, Y be three continuously embedded Banach spaces*

$$X \subset B \subset Y$$

with the first inclusion, $X \subset B$, compact. We consider a set $\mathcal{F}$ of functions mapping $[0,T]$ into X. If $\mathcal{F}$ is bounded in $L^\infty([0,T],X)$ and $\frac{\partial\mathcal{F}}{\partial t} = \left\{\frac{\partial f}{\partial t} \mid f \in \mathcal{F}\right\}$ is bounded in $L^r([0,T],Y)$ where $r > 1$, then $\mathcal{F}$ is relatively compact in $C([0,T],B)$.

We now turn to the proof of our main theorem.

Proof of Theorem 2.1. (i) First we establish that there exists a subsequence of u^n that converges in $C([0,T],H^1)$ to an element $u \in H^1$. To apply Theorem 2.4, we have to determine the Banach spaces with the required properties. In our case, we take X as the set of functions of H^1 which have derivatives of bounded variation:

$$X = \left\{v \in H^1 \mid v_x \in \mathrm{BV}\right\}.$$

X endowed with the norm

$$\|v\|_X = \|v\|_{H^1} + \|v_x\|_{BV} = \|v\|_{H^1} + \|v_x\|_{L^\infty} + \mathrm{TV}(v_x)$$

is a Banach space. Let us prove that the injection $X \subset H^1$ is compact. We consider a sequence v_n which is bounded in X. Since $\|v_n\|_{L^\infty}$ is bounded ($H^1 \subset L^\infty$ continuously), there exists a point x_0 such that $v_n(x_0)$ is bounded and we can extract a subsequence (that we still denote v_n) such that $v_n(x_0)$ converges to some $l \in \mathbb{R}$. By Helly's theorem, we can also extract a subsequence such that

$$v_{n,x} \to w \quad \text{a.e.} \tag{2.28}$$

for some $w \in L^\infty$. By Lebesgue's dominated convergence theorem, it implies that $v_{n,x} \to w$ in L^2. We set

$$v(x) = l + \int_{x_0}^{x} w(s)\, ds.$$

We have that $v_x = w$ almost everywhere. We also have

$$v_n(x) = v_n(x_0) + \int_{x_0}^{x} v_{n,x}(s)\, ds$$

which together with (2.28) implies that v_n converges to v in L^∞. Therefore v_n converges to v in H^1 and X is compactly embedded in H^1.

The estimates we have derived previously give us that u^n and u^n_t are uniformly bounded in $L^\infty([0,T], X)$ and $L^\infty([0,T], L^2)$, respectively. Since $X \subset H^1 \subset L^2$ with the first inclusion compact, Simon's theorem gives us the existence of a subsequence of u^n that converges in $C([0,T], H^1)$ to some $u \in H^1$.

(ii) Next we show that the limit we get is a solution of the Camassa–Holm equation (1.2).

Let us now take φ in $C^\infty([0,1] \times [0,T])$ and multiply, for each i, the first equation in (2.4) by $h\varphi(x_i, t)$. We denote φ^n the continuous piecewise linear function given by $\varphi^n(x_i, t) = \varphi(x_i, t)$. We sum over i and get, after one summation by parts,

$$\begin{aligned}\sum_{i=0}^{n-1} h\left(u^n_{i,t} - (D_-D_+u^n_i)_t\right)\varphi^n_i = &\underbrace{\sum_{i=0}^{n-1} h(u^n_i)^2 D_+\varphi_i}_{A} - \underbrace{\sum_{i=0}^{n-1} hu^n_i D_-D_+u^n_i D_+\varphi^n_i}_{B} \\ &- \underbrace{\sum_{i=0}^{n-1} hu^n_i Du^n_i \varphi^n_i}_{C} + \underbrace{\sum_{i=0}^{n-1} hD_-D_+u^n_i Du^n_i \varphi^n_i}_{D}.\end{aligned} \tag{2.29}$$

We are now going to prove that each term in this equality converges to the corresponding terms in (2.5).

Term A: We want to prove that

$$\left\langle (u^n)^2 D_+\varphi^n \right\rangle \to \int_0^1 u^2 \varphi_x\, dx, \tag{2.30}$$

where we have introduced the following notation

$$\langle u \rangle = h \sum_{i=0}^{n-1} u_i$$

to denote the average of a quantity u. We have

$$\begin{aligned}\left|\int_0^1 u^2\varphi_x\,dx - \big\langle (u^n)^2 D_+\varphi^n\big\rangle\right| \leq{}& \left|\int_0^1 (u^2-(u^n)^2)\varphi_x\,dx\right|\\ &+\left|\int_0^1 (u^n)^2(\varphi_x - D_+\varphi^n)\,dx\right|\\ &+\left|\int_0^1 (u^n)^2 D_+\varphi^n\,dx - \big\langle (u^n)^2 D_+\varphi^n\big\rangle\right|.\end{aligned}$$

The first term tends to zero because $u^n\to u$ in L^2 for all $t\in[0,T]$. The second tends to zero by Lebesgue's dominated convergence theorem. It remains to prove that the last term tends to zero.

The integral of a product between two continuous piecewise linear function, v and w, and a piecewise constant function z can be computed explicitly. We skip the details of the calculation and give directly the result:

$$\int_0^1 zvw\,dx = \frac{1}{3}\left\langle zS_+vS_+w\right\rangle + \frac{1}{6}\left\langle zS_+vw\right\rangle + \frac{1}{6}\left\langle zvS_+w\right\rangle + \frac{1}{3}\left\langle zvw\right\rangle. \tag{2.31}$$

Here S_+ and S_- denote shift operators

$$(S_\pm u)_i = u_{i\pm 1}.$$

After using (2.31) with $v=w=u^n$ and $z=D_+\varphi^n$, we get

$$\begin{aligned}\int_0^1 (u^n)^2D_+\varphi^n - \big\langle (u^n)^2 D_+\varphi^n\big\rangle ={}& \frac{1}{3}\left\langle (S_+u^n-u^n)D_+\varphi^n u^n\right\rangle\\ &+\frac{1}{3}\big\langle (u^n)^2 D_+(S_-\varphi^n-\varphi^n)\big\rangle.\end{aligned}$$

We use the uniform equivalence of the l^2 and L^2 norm to get the following estimate

$$\begin{aligned}\left\langle (S_+u^n-u^n)D_+\varphi^n u^n\right\rangle &\leq \left\|S_+u^n-u^n\right\|_{l^2}\left\|D_+\varphi^nu^n\right\|_{l^2}\quad\text{(Cauchy–Schwarz)}\\ &\leq \mathcal{O}(1)\left\|u^n(\,\cdot\,+h)-u^n(\,\cdot\,)\right\|_{L^2}.\end{aligned}\tag{2.32}$$

Since $u_n\in H^1$, we have (see, for example, [1]):

$$\left\|u^n(\,\cdot\,+h)-u^n(\,\cdot\,)\right\|_{L^2}\leq h\left\|u^n_x\right\|_{L^2}\leq\mathcal{O}(1)h$$

because $\left\|u^n_x\right\|_{L^\infty}$ is uniformly bounded. Hence $\left|\left\langle (S_+u^n-u^n)D_+\varphi^nu^n\right\rangle\right|$ tends to zero. The quantity $\big\langle (u^n)^2D_+(S_-\varphi^n-\varphi^n)\big\rangle$ tends to zero because φ is C^∞ and u^n uniformly bounded. We have proved (2.30).

Term B: We want to prove

$$\left\langle u^nD_-D_+u^nD_+\varphi^n\right\rangle \to \frac{1}{2}\int_0^1 u^2\varphi_{xxx}\,dx - \int_0^1 u_x^2\varphi_x. \tag{2.33}$$

We rewrite $u^nD_-D_+u^n$ in such a way that the discrete double derivative D_-D_+ does not appear in a product (so that we can later sum by parts). We have

$$u^nD_-D_+u^n = \frac{1}{2}(D_-D_+((u^n)^2) - D_+u^nD_+u^n - D_-u^nD_-u^n).$$

We can prove in the same way as we did for term A that

$$\begin{aligned}\left\langle D_-D_+((u^n)^2)D_+\varphi^n\right\rangle &= \left\langle (u^n)^2 D_-D_+D_+\varphi^n\right\rangle \quad \text{(summation by parts)}\\ &\to \int_0^1 u^2\varphi_{xxx}\,dx.\end{aligned}$$

The quantity $(u_x^n)^2\varphi_x^n$ is a piecewise constant function. Therefore,

$$\int_0^1 (u_x^n)^2\varphi_x^n\,dx = \langle D_+u^nD_+u^nD_+\varphi^n\rangle\,.$$

Since $u_x^n \to$ in L^2 for all $t\in[0,T]$, and

$$\int_0^1 u_x^2\varphi_x\,dx - \langle D_+u^nD_+u^nD_+\varphi^n\rangle = \int_0^1 (u_x^2-(u_x^n)^2)\varphi_x\,dx + \int_0^1 (u_x^n)^2(\varphi_x-\varphi_x^n)\,dx,$$

we have

$$\langle D_+u^nD_+u^nD_+\varphi^n\rangle \to \int_0^1 u_x^2\varphi_x\,dx.$$

In the same way, we get

$$\langle D_-u_i^nD_-u_i^nD_+\varphi^n\rangle \to \int_0^1 u_x^2\varphi_x$$

and (2.33) is proved.

Term C: We want to prove

$$\langle u^nDu^n\varphi^n\rangle \to \int_0^1 uu_x\varphi\,dx. \tag{2.34}$$

We have

$$\begin{aligned}\int_0^1 uu_x\varphi\,dx - \langle u^nD_+u^n\varphi^n\rangle &= \int_0^1 (u-u^n)u_x\varphi\,dx + \int_0^1 u^n(u_x-u_x^n)\varphi\,dx\\ &\quad + \int_0^1 u^nu_x^n(\varphi-\varphi^n)\,dx + \int_0^1 u^nu_x^n\varphi^n\,dx\\ &\quad - \langle u^nD_+u^n\varphi^n\rangle\,.\end{aligned}$$

The first two terms converge to zero because $u^n\to u$ in H^1 for all $t\in[0,T]$. The third term converges to zero by Lebesgue's dominated convergence theorem. We use formula (2.31) to evaluate the last integral:

$$\begin{aligned}\int_0^1 u^nu_x^n\varphi^n\,dx &= \frac{1}{3}\langle D_+u^nS_+u^nS_+\varphi^n\rangle + \frac{1}{6}\langle D_+u^nS_+u^n\varphi^n\rangle\\ &\quad + \frac{1}{6}\langle D_+u^nu^nS_+\varphi^n\rangle + \frac{1}{3}\langle D_+u^nu^n\varphi^n\rangle\,.\end{aligned}$$

Using the same type of arguments as those we have just used for term A, one can show that

$$\int_0^1 u^nu_x^n\varphi^n\,dx \to \langle D_+u^nu^n\varphi^n\rangle\,.$$

Thus, in order to prove (2.34), it remains to prove that

$$\langle D_+u^n u^n \varphi^n\rangle - \langle Du^n u^n \varphi^n\rangle \to 0. \tag{2.35}$$

Since $D = \frac{1}{2}(D_+ + D_-)$, we have:

$$\langle D_+u^n u^n \varphi^n\rangle - \langle Du^n u^n \varphi^n\rangle = \frac{1}{2}\langle (D_+u^n - D_-u^n)u^n\varphi^n\rangle$$

and

$$\begin{aligned}
|\langle (D_+u^n - D_-u^n)u^n\varphi^n\rangle| &\leq C\sum_{i=0}^{n-1} h\left|D_+u_i^n - D_+u_{i-1}^n\right| \\
&\leq \mathcal{O}(1)\int_0^1 |u_x^n(x) - u_x^n(x-h)|\,dx \\
&\leq \mathcal{O}(1)h\ \mathrm{TV}(u_x^n).
\end{aligned}$$

Since $\mathrm{TV}(u_x^n)$ is uniformly bounded, (2.35) holds and we have proved (2.34).

Term D: We want to prove that

$$\langle D_-D_+u^n Du^n\varphi^n\rangle \to -\frac{1}{2}\int_0^1 u_x^2\varphi_x\,dx. \tag{2.36}$$

We have

$$\frac{1}{2}\int_0^1 u_x^2\varphi_x\,dx + \langle D_-D_+u^n Du^n\varphi^n\rangle \tag{2.37}$$

$$= \frac{1}{2}\int_0^1 (u_x^2 - (u_x^n)^2)\varphi_x\,dx + \frac{1}{2}\int_0^1 (u_x^n)^2(\varphi_x - D_-\varphi^n)\,dx \tag{2.38}$$

$$-\frac{1}{2}\langle D_+(D_+u^n D_+u^n)\varphi^n\rangle + \langle D_-D_+u^n Du^n\varphi^n\rangle. \tag{2.39}$$

The two first terms on the right-hand side tend to zero. After using the following identity

$$D_+(D_+u^n D_+u^n) = D_+D_+u^n D_+u^n + D_+D_+u^n D_+S_+u^n,$$

we can rewrite the two last terms in (2.37) as

$$\begin{aligned}
&-\frac{1}{2}\langle D_+(D_+u^n D_+u^n)\varphi^n\rangle + \langle D_-D_+u^n Du^n\varphi^n\rangle \\
&\quad= -\frac{1}{2}\langle D_-D_+S_+u^n D_+u^n\varphi^n\rangle - \frac{1}{2}\langle D_-D_+S_+u^n D_+S_+u^n\varphi^n\rangle \\
&\qquad\quad+\frac{1}{2}\langle D_-D_+u^n D_+S_-u^n\varphi^n\rangle + \frac{1}{2}\langle D_-D_+u^n D_+u^n\varphi^n\rangle \\
&\quad= \frac{1}{2}\langle D_-D_+u^n D_-u^n(\varphi^n - S_-\varphi^n)\rangle + \frac{1}{2}\langle D_-D_+u^n D_+u^n(\varphi^n - S_-\varphi^n)\rangle
\end{aligned}$$

which tends to zero because, as we have seen before, due to the positivity of m, $\langle |D_-D_+u_i^n D_+u_i^n|\rangle$ is uniformly bounded. We have proved (2.36).

Up to now we have not really considered the time variable. We integrate (2.29) with respect to time and integrate by part the left-hand side:

$$\int_0^T \sum_{i=0}^{n-1} h\left(u_{i,t}^n - D_-D_+u_{i,t}^n\right)\varphi(x_i,t)\,dt = -\int_0^T \sum_{i=0}^{n-1} h\left(u_i^n - D_-D_+u_i^n\right)\varphi_t(x_i,t)\,dt + \left[\sum_{i=0}^{n-1} h\left(u_i^n - D_-D_+u_i^n\right)\varphi(x_i,t)\right]_{t=0}^{t=T}$$

and, after summing by parts, the limit of this expression is (we use Lebesgue's dominated convergence theorem with respect to x and t)

$$-\int_0^T\int_0^1 u(\varphi_t - \varphi_{txx})\,dxdt + \left[\int_0^1 u(\varphi - \varphi_{xx})\,dx\right]_{t=0}^{t=T}.$$

It is not hard to see that the right-hand side of (2.29) is uniformly bounded by a constant and we can integrate over time and use the Lebesgue dominated convergence theorem to conclude that u is indeed a solution of (2.5) in the sense of distribution.

The analysis in [11] shows that the weak solution of the Camassa–Holm with initial conditions satisfying $m(x,0) \geq 0$ is unique. This implies that in our algorithm not only a subsequence but the whole sequence u^n converges to the solution. However, in [11], a solution of the Camassa–Holm equation is defined as an element u of H^1 satisfying

$$u_t + uu_x + \left[\int_{-\infty}^{\infty} p(x-y)[u^2(y,t) + \frac{1}{2}u_x^2(y,t)]\,dy\right]_x = 0 \tag{2.40}$$

where p is the solution of

$$\mathcal{A}p \equiv (I - \partial_x^2)p = \delta.$$

We want to prove that weak solutions of (2.40) and (2.5) are the same. Periodic distributions belong to the class of tempered distribution $\mathcal{S}'$ (see for example [13]). The operator $\mathcal{A}$ defines a homeomorphism on the Schwartz class $\mathcal{S}$ (or class of rapidly decreasing function): The Fourier transform is a homeomorphism on $\mathcal{S}$ and $\mathcal{A}$ restricted to $\mathcal{S}$ can be written as

$$\mathcal{A} = \mathcal{F}^{-1}(1+\xi^2)\mathcal{F} \tag{2.41}$$

where ξ denotes the frequency variable. It is clear from (2.41) that the inverse of $\mathcal{A}$ in $\mathcal{S}$ is

$$\mathcal{A}^{-1} = \mathcal{F}^{-1}\frac{1}{1+\xi^2}\mathcal{F}.$$

Hence $\mathcal{A}$ is a homeomorphism on $\mathcal{S}$.

We can now define the inverse $\mathcal{A}^{-1}$ of $\mathcal{A}$ in $\mathcal{S}'$. Given T in $\mathcal{S}'$, $\mathcal{A}^{-1}T$ is given by

$$\left\langle \mathcal{A}^{-1}T, \phi\right\rangle = \left\langle T, \mathcal{A}^{-1}\phi\right\rangle, \quad \phi \in \mathcal{S}.$$

It is easy to check that $\mathcal{A}^{-1}$ indeed satisfies

$$\mathcal{A}^{-1}\mathcal{A} = \mathcal{A}\mathcal{A}^{-1} = \mathrm{Id},$$

and that $\mathcal{A}^{-1}$ is continuous on $\mathcal{S}'$. The operator $\mathcal{A}$ is therefore a homeomorphism on $\mathcal{S}'$.

Let u be a solution of (2.40). Then we have

$$u_t + \partial_x(\frac{u^2}{2}) + \partial_x \mathcal{A}^{-1}[u^2 + \frac{1}{2}u_x^2] = 0. \tag{2.42}$$

The operators ∂_x and $\mathcal{A}^{-1}$ commute because ∂_x and $\mathcal{A}$ commute. We apply $\mathcal{A}$ on both sides of (2.42) and get:

$$u_t - u_{xxt} + \mathcal{A}\partial_x(\frac{1}{2}u^2) + \partial_x[u^2 + \frac{1}{2}u_x^2] = 0, \tag{2.43}$$

which is exactly (2.5). Since $\mathcal{A}$ is a bijection, (2.43) also implies (2.42) and we have proved that the weak solutions of (2.5) are the same as the weak solutions given by (2.40). □

In Theorem 2.1, some restrictions on the initial data v are implicitly imposed by the condition $v^n - D_-D_+v^n \geq 0$. We are going to prove that if $v \in H^1([0,1])$ is periodic with $v - v_{xx} \in \mathcal{M}^+$, where $\mathcal{M}^+$ denotes the space of positive Radon measures, then there exists a sequence of piecewise linear, continuous, periodic functions v^n that converges to v in H^1 and satisfies $v^n - D_-D_+v^n \geq 0$ for all n.

We can then apply Theorem 2.1 and get the existence result contained in the following corollary which coincides with results obtained in [11] by a different method.

Corollary 2.5. *If $u_0 \in H^1$ is such that $u_0 - u_{0,xx} \in \mathcal{M}^+$ then the Camassa–Holm equation has a global solution in $C(\mathbb{R}_+, H^1)$. The solution is obtained as a limit of the numerical scheme defined by* (2.4)*.*

To apply Theorem 2.1, we need to prove that, given $u \in H^1([0,1])$ such that $u - u_{xx} \in \mathcal{M}^+$, there exists a sequence u^n of piecewise linear, continuous and periodic functions such that

$$\begin{aligned} &u^n \to u \text{ in } H^1, \\ &u^n - D_-D_+u^n \geq 0. \end{aligned}$$

Let $\{\psi_i^n\}$ be a partition of unity associated with the covering $\cup_{i=0}^{n-1}(x_{i-1}, x_{i+1})$. For all $i \in \{0, , \dots n-1\}$, the functions ψ_i^n are non-negative with $\operatorname{supp} \psi_i^n \subset (x_{i-1}, x_{i+1})$, and $\sum_{i=0}^{n-1} \psi_i^n = 1$. Define

$$v_i^n = \frac{1}{h}\left\langle u - u_{xx}, \psi_i^n \right\rangle$$

and

$$u_i^n - D_-D_+u_i^n = v_i^n. \tag{2.44}$$

Recall that the operator $u^n - D_-D_+u^n$ is invertible, see (2.8), so that u^n is well-defined by (2.44). Since $u - u_{xx}$ belongs to $\mathcal{M}^+$ and $\psi_i^n \geq 0$, we have $v_i^n = u_i^n - D_-D_+u_i^n \geq 0$ and it only remains to prove that u^n converges to

u in H^1. Since the application $\mathcal{L} : H^1 \to H^{-1}$ given by $\mathcal{L}u = u - u_{xx}$ is an homeomorphism, it is equivalent to prove that

$$u^n - u^n_{xx} \to u - u_{xx} \text{ in } H^{-1}.$$

The homeomorphism $\mathcal{L}$ is also an isometry, so that

$$\|\mathcal{L}u\|_{H^{-1}} = \|u\|_{H^1}.$$

We can find a bound on $\|u^n\|_{H^1}$. Let E_n be defined, as before, by

$$E_n = \left(h \sum_{i=0}^{n-1} \left[(u^n_i)^2 + (D_+ u^n)^2_i \right] \right)^{\frac{1}{2}}.$$

The inequality (2.16) still holds. We have

$$\begin{aligned}
E_n^2 &= h \sum_{i=0}^{n-1} (u^n_i - D_- D_+ u^n_i) u^n_i \\
&= h \sum_{i=0}^{n-1} v^n_i u^n_i \\
&\leq \|u^n\|_{L^\infty} \sum_{i=0}^{n-1} h v^n_i \\
&\leq \|u^n\|_{L^\infty} \langle u - u_{xx}, \sum_{i=0}^{n-1} \psi^n_i \rangle \\
&\leq \|u^n\|_{L^\infty} \|u - u_{xx}\|_{\mathcal{M}^+} \quad (\text{since } \sum_{i=0}^{n-1} \psi^n_i = 1).
\end{aligned}$$

Hence, since L^∞ is continuously embedded in H^1, there exists a constant C (independent of n) such that

$$E_n^2 \leq C \|u^n\|_{H^1} \|u - u_{xx}\|_{\mathcal{M}^+}.$$

We use inequality (2.16) to get the bound on $\|u^n\|_{H^1}$ we were looking for:

$$\|u^n\|_{H^1} \leq 4C \|u - u_{xx}\|_{\mathcal{M}^+}.$$

To prove that $u^n - u^n_{xx} \to u - u_{xx}$ in H^{-1}, since $\|u^n - u^n_{xx}\|_{H^{-1}} = \|u^n\|_{H^1}$ is uniformly bounded, we just need to prove that

$$\langle u^n - u^n_{xx}, \varphi \rangle \to \langle u - u_{xx}, \varphi \rangle$$

for all φ belonging to a dense subset of H^1 (for example C^∞).

The function u^n is continuous and piecewise linear. Its second derivative u^n_{xx} is therefore a sum of Dirac functions:

$$u^n_{xx} = \sum_{i=0}^{n-1} h D_- D_+ u^n_i \delta_{x_i}$$

and, for any φ in C^∞, we have

$$\begin{aligned}\langle u^n - u^n_{xx}, \varphi\rangle &= \int_0^1 u^n(x)\varphi(x)\,dx - h\sum_{i=0}^{n-1} D_-D_+u_i^n\varphi(x_i)\\ &= \int_0^1 u^n(x)(\varphi(x)-\varphi^n(x))\,dx + \int_0^1 u^n(x)\varphi^n(x)\,dx \\ &\quad - h\sum_{i=0}^{n-1} u_i^n\varphi_i^n + h\sum_{i=0}^{n-1} v_i\varphi_i^n\end{aligned} \tag{2.45}$$

where φ^n denotes the piecewise linear, continuous function that coincides with φ on x_i, $i = 0, \dots, n-1$.

The first integral in (2.45) tends to zero by the Lebesgue dominated convergence theorem. We use formula (2.31) to compute the second integral:

$$\int_0^1 u^n(x)\varphi^n(x)\,dx = \frac{2}{3}\left\langle u^n\varphi^n\right\rangle + \frac{1}{6}\left\langle S_+u^n\varphi^n\right\rangle + \frac{1}{6}\left\langle u^nS^+\varphi^n\right\rangle.$$

One can prove that this term tends to $\langle u^n\varphi^n\rangle$ (see the proof of the convergence of term A in the proof of Theorem 2.1). The last sum equals

$$\sum_{i=0}^{n-1} hv_i^n\varphi(x_i) = \left\langle u - u_{xx}, \sum_{i=0}^{n-1}\varphi_i^n\psi_i^n(x)\right\rangle.$$

For all $x \in [0,1]$, there exists a k such that $x \in [x_k, x_{k+1}]$. Then,

$$\begin{aligned}\left|\varphi(x) - \sum_{i=0}^{n-1}\varphi_i^n\psi_i^n(x)\right| &= \left|\sum_{i=0}^{n-1}(\varphi(x)-\varphi(x_i))\psi_i^n(x)\right| \\ &\leq |\varphi(x)-\varphi(x_k)| + |\varphi(x)-\varphi(x_{k+1})| \\ &\leq 2\sup_{|z-y|\leq h}|\varphi(y)-\varphi(z)|\end{aligned}$$

and therefore, by the uniform continuity of φ,

$$\sum_{i=0}^{n-1}\varphi(x_i)\psi_i^n(x) \to \varphi(x) \text{ in } L^\infty.$$

Thus,

$$\sum_{i=0}^{n-1} hv_i^n\varphi(x_i) = \left\langle u - u_{xx}, \sum_{i=0}^{n-1}\varphi(x_i)\psi_i^n\right\rangle \to \langle u - u_{xx}, \varphi\rangle$$

and, from (2.45), we get

$$\langle u^n - u^n_{xx}, \varphi\rangle \to \langle u - u_{xx}, \varphi\rangle.$$

As already explained, it implies that

$$u^n \to u \text{ in } H^1.$$

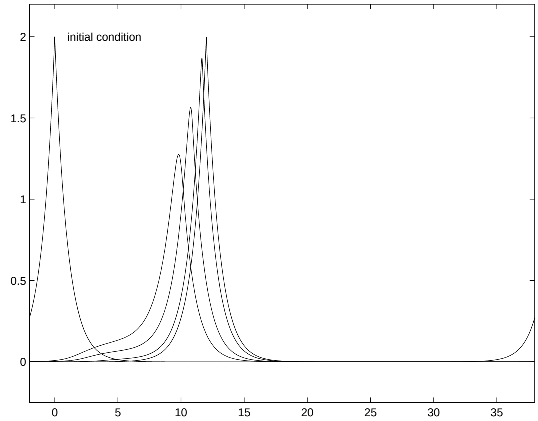


FIGURE 1. Periodic single peakon. The initial condition is given by $u(x, 0) = 2e^{-|x|}$ and period $a = 40$. The computed solutions are shown at time $t = 6$ for (from left to right) $n = 2^{10}$, $n = 2^{12}$, $n = 2^{14}$ together with the exact solution (at the far right).

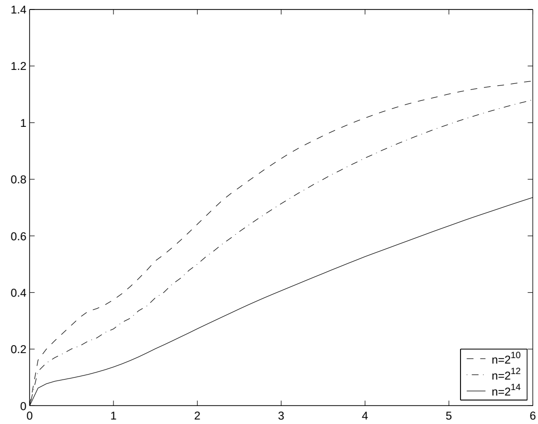


FIGURE 2. Plot of $\|u(t) - u^n(t)\|_{H^1} / \|u(t)\|_{H^1}$ in the one peakon case of Figure 1.

3. Numerical results

The numerical scheme (2.4) is semi-discrete: The time derivative has not been discretized, and hence we work with a system of ordinary differential equations. However, for numerical computations we integrate in time by using an explicit Euler method. Given a positive time T and $l \in \mathbb{N}$, we consider the time step $\Delta t = T/l$. We compute $m_j^{n,l}$, the approximate value of m^n at time $t_j = j\Delta t$, by taking

$$m_{j+1}^{n,l} = m_j^{n,l} + \Delta t\left(-D_-(m_j^{n,l}u_j^{n,l}) - m_j^{n,l}Du_j^{n,l}\right), \tag{3.1}$$

where

$$m_j^{n,l} = u_j^{n,l} - D_-D_+u_j^{n,l}. \tag{3.2}$$

Here $m_j^{n,l} = (m_{0,j}^{n,l}, \dots, m_{n-1,j}^{n,l})$ and $u_j^{n,l} = (u_{0,j}^{n,l}, \dots, u_{n-1,j}^{n,l})$. Given $m_j^{n,l}$, one can still recompute $u_j^{n,l}$ using (2.8), that is,

$$u_{i,j}^{n,l} = L^{-1}m_{i,j}^{n,l} = \frac{c}{1-e^{-\kappa n}}\sum_{k=0}^{n-1}(e^{-\kappa(i-k)} + e^{\kappa(i-k-n)})m_{k,j}^{n,l}. \tag{3.3}$$

Lemma 2.2 does not apply in this setting, and the proof of convergence for the fully discrete scheme proceeds differently. Writing (2.4) as

$$m_t^n = f(m^n),$$

where $f\colon \mathbb{R}^n \to \mathbb{R}^n$, we observe (cf. (2.4) and (2.8)) that each component of $f(x)$ is a polynomial in the components $x_0, \dots, x_{n-1}$ of x. Hence, f is continuously differentiable. From (2.19) and (2.4), we obtain that, when n is large enough, there exists a constant C which is independent of n such that

$$|m_i^n(t)| \leq 5n^2 \max_i |u_i^n(t)| \leq Cn^2$$

for all $t \in [0,T]$. Hence, $m^n(t)$ is bounded in $[0,T]$ and therefore the Euler method converges, see, for example, [19], that is,

$$\lim_{l\to\infty} \max_{j=1,\dots,l} \left\| m_j^{n,l} - m^n(t_j) \right\| = 0. \tag{3.4}$$

All norms are equivalent in finite dimensional vector spaces, and therefore (3.4) holds for any norm in $\mathbb{R}^n$. We denote by $m^{n,l}(t)$ the piecewise linear function in $C([0,T],\mathbb{R}^n)$ satisfying $m^{n,l}(t_j) = m_j^{n,l}$. It is given by

$$m^{n,l}(t) = \frac{1}{\Delta t}(t_{j+1}-t)m_j^{n,l} + \frac{1}{\Delta t}(t-t_j)m_{j+1}^{n,l}$$

for $t \in [t_j, t_{j+1}]$. Let us prove that

$$\lim_{l\to\infty} \left\| m^{n,l} - m^n \right\|_{C([0,T],\mathbb{R}^n)} = 0. \tag{3.5}$$

We have, for $t \in [t_j, t_{j+1}]$,

$$\begin{aligned} m^{n,l}(t) - m^n(t) &= \frac{1}{\Delta t}(t_{j+1}-t)(m_j^{n,l} - m^n(t_j)) + \frac{1}{\Delta t}(t-t_j)(m_{j+1}^{n,l} - m^n(t_{j+1})) \\ &\quad + \frac{1}{\Delta t}(t_{j+1}-t)(m^n(t_j) - m^n(t)) + \frac{1}{\Delta t}(t-t_j)(m^n(t_{j+1}) - m^n(t)). \end{aligned} \tag{3.6}$$

Let $\varepsilon > 0$. Since $m^n \in C([0,T],\mathbb{R}^n)$, m^n is uniformly continuous and there exists $\delta > 0$ such that $\|m^n(t_1) - m^n(t_2)\| < \varepsilon/2$ for all $t_1, t_2 \in [0,T]$ with $|t_2 - t_1| < \delta$.

We can choose l large enough so that $\Delta t = T/l < \delta$. Then, for $t \in [t_j, t_{j+1}]$, we have $t - t_j < \delta$ and $t_{j+1} - t < \delta$, and

$$\begin{aligned}\left\| \frac{1}{\Delta t}(t_{j+1} - t)(m^n(t_j) - m^n(t)) + \frac{1}{\Delta t}(t - t_j)(m^n(t_{j+1}) - m^n(t_{j+1})) \right\| \\ < \frac{1}{\Delta t}(t_{j+1} - t)\frac{\varepsilon}{2} + \frac{1}{\Delta t}(t - t_j)\frac{\varepsilon}{2} \\ < \frac{\varepsilon}{2}.\end{aligned} \tag{3.7}$$

By (3.4), we can choose l large enough so that $\max_{j=1,\dots,l} \left\| m_j^{n,l} - m^n(t_j) \right\| < \varepsilon/2$. Hence,

$$\left\| \frac{1}{\Delta t}(t_{j+1} - t)(m_j^{n,l} - m^n(t_j)) + \frac{1}{\Delta t}(t - t_j)(m_{j+1}^{n,l} - m^n(t_{j+1})) \right\| < \frac{\varepsilon}{2}. \tag{3.8}$$

Comparing (3.6), (3.7) and (3.8), we obtain

$$\left\| m^{n,l}(t) - m^n(t) \right\| < \varepsilon$$

for l large enough and any $t \in [0, T]$. Hence, (3.5) is proved. The mapping $L^{-1} \colon \mathbb{R}^n \to \mathbb{R}^n$, $L^{-1} m^n = u^n$, is continuous, and therefore we have

$$\lim_{l \to \infty} \left\| u^{n,l} - u^n \right\|_{C([0,T],\mathbb{R}^n)} = 0.$$

Finally, after using the identification of $\mathbb{R}^n$ with the set of continuous, periodic, piecewise linear functions, we get that

$$\lim_{l \to \infty} u^{n,l} = u^n,$$

and, from Theorem 2.1,

$$\lim_{n \to \infty} \lim_{l \to \infty} u^{n,l} = u$$

in $C([0, T], H^1)$. We summarize the result in the following theorem.

Theorem 3.1. *Let $\Delta t = T/l$, and define the function $u_{i,j}^{n,l}$ by (3.1)–(3.3). Define the corresponding interpolating function $u^{n,l}$ in $C([0, T], H^1)$ by*

$$\begin{aligned}u^{n,l}(x, t) = \frac{n}{\Delta t}\Big((t_{j+1} - t)\big[(x_{i+1} - x)u_{i,j}^{n,l} + (x - x_i)u_{i+1,j}^{n,l}\big] \\ + (t - t_j)\big[(x_{i+1} - x)u_{i,j+1}^{n,l} + (x - x_i)u_{i+1,j+1}^{n,l}\big]\Big)\end{aligned}$$

for $x \in [x_i, x_{i+1}]$ and $t \in [t_j, t_{j+1}]$. Then

$$\lim_{n \to \infty} \lim_{l \to \infty} u^{n,l} = u \tag{3.9}$$

in $C([0, T], H^1)$ where u is the solution of the Camassa–Holm equation (1.2).

To compute the discrete spatial derivative, we need at each step to compute u from m. The function u is given by a discrete convolution product

$$u_i = h \sum_{j=0}^{n-1} g_{i-j}^p m_j.$$

It is advantageous to apply the Fast Fourier Transform (FFT), see [13]. In the frequency space, a convolution product becomes a multiplication which is cheap to evaluate. Going back and forth to the frequency space is not very expensive due to the efficiency of the FFT. We use a formula of the form (see [13] for more details):

$$u = \mathcal{F}_N^{-1}(\mathcal{F}_N[g] \cdot \mathcal{F}_N[m])$$

where $\mathcal{F}_N$ denotes the FFT.

We have tested algorithm (3.1) with single and double peakons. In the single peakon case, the initial condition is given by

$$u(x,0) = c\frac{\cosh(d - \frac{a}{2})}{\sinh \frac{a}{2}}, \tag{3.10}$$

which is the periodized version of $u(x,0) = ce^{-|x|}$. The period is denoted by a and $d = \min(x, a-x)$ is the distance from x to the boundary of the interval $[0,a]$. The peakons travel at a speed equal to their height, that is

$$u(x,t) = ce^{-|x-ct|}.$$

If u satisfies the initial condition $u(x,0) = e^{-|x|}$, then $m = 2\delta$ at $t = 0$ and we take

$$m_i(0) = \begin{cases} \frac{2}{h} & \text{if } i = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.11}$$

as initial discrete condition. The function m_i gives a discrete approximation of 2δ. Figure 1 shows the result of the computation for different refinements. Figure 2 indicates that the computed solution converges to the exact solution.

The sharp increase of the error $\|u(t) - u^n(t)\|_{H^1}$ at time $t = 0$ can be predicted by looking at (2.17) which gives a first-order approximation of the time derivative of $\|u(t)\|_{H^1}^2$:

$$\frac{dE_n(t)^2}{dt} = -\sum_{i=0}^{n-1} u_i(hm_i)^2 + \mathcal{O}(h).$$

Hence,

$$\frac{d\|u\|_{H^1}^2}{dt} \approx \frac{dE_n(t)^2}{dt} \approx -4 \quad \text{at } t = 0.$$

At the beginning of the computation, we can therefore expect a sharp decrease of the H^1 norm. To get convergence in H^1, it is therefore necessary that the solution becomes smooth enough so that $\frac{d\|u\|_{H^1}^2}{dt} \to 0$. In any case, we cannot hope for high accuracy and convergence rate in this case. Figure 3 shows the same plots in the two peakon case.

We have tested our algorithm with smooth initial conditions. In this case, the H^1 norm remains constant in a much more accurate manner. The convergence is probably much better but we have no analytical solution to compare with.

Other time integration methods (second-order Runge–Kutta method, variable order Adams–Bashforth–Moulton) have also been tried and the results do not differ significantly from those given by (3.1). It follows that the CH equation is

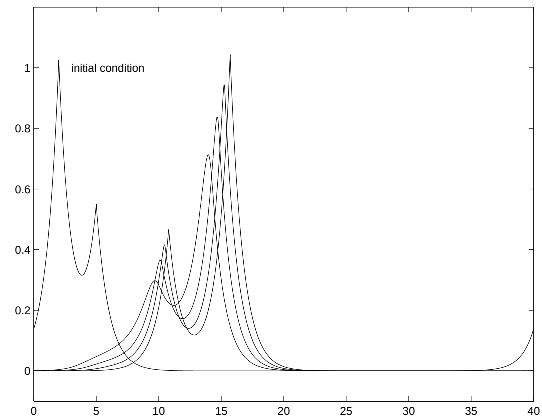


FIGURE 3. Two peakon case. The initial condition is the periodized version of $2e^{-|x-2|} + e^{-|x-5|}$. The computed solutions are shown at time $t = 12$ for (from left to right) $n = 2^{10}$, $n = 2^{12}$, $n = 2^{14}$ together with the exact solution (at the far right).

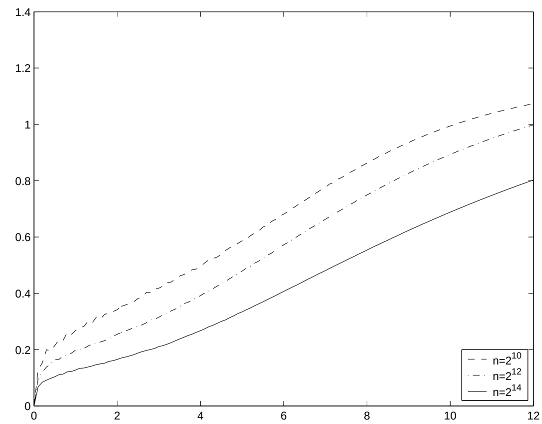


FIGURE 4. Plot of $\|u(t) - u^n(t)\|_{H^1} / \|u(t)\|_{H^1}$ in the two peakon case of Figure 3.

not very sensitive to the way time is discretized. But the situation is completely different when we consider different space discretizations. The following schemes

$$m_t = -D_-(mu)_i - m_i D_+ u_i, \tag{3.12}$$

$$m_t = -D(mu)_i - m_i D u_i, \tag{3.13}$$

$$m_t = -D_+(mu)_i - m_i D_- u_i \tag{3.14}$$

are all at first glance good candidates for solving the CH equation. They preserve the H^1 norm, are finite difference approximations of (2.2) and finally look very similar to (2.4). But, tested on a single peakon, (3.12) produces a peakon that grows, (3.13) produces oscillations, and (3.14) behaves in a completely unexpected manner (at the first time step, m becomes a negative Dirac function and starts traveling backward!).

Let us have a closer look at the scheme (3.12). We compute $\frac{dE_n^2}{dt}$:

$$\frac{1}{2}\frac{dE_n^2}{dt} = \sum_{i=0}^{n-1} m_{i,t}^n u_i^n = \sum_{i=0}^{n-1} \big(- D_-(m^n u^n)_i u_i - m_i^n D_+ u_i u_i \big) = 0.$$

Thus, E_n is exactly preserved. Lemma 2.2 still holds since the same proof applies to (3.12). It allows us to derive the bounds of Lemma 2.3 and, after applying Simon's theorem, we get the existence of a converging subsequence. The problem is that, in general, this subsequence *does not* converge to the solution of the Camassa–Holm equation. In order to see that, we compare how our original algorithm (3.12) and algorithm (3.13) handle a peakon solution $u = ce^{-|x-ct|}$. The only terms that differ are $m^n Du^n$ and $m^n D_+ u^n$. We have proved earlier that, for any smooth function φ,

$$\sum_{i=0}^{n-1} m_i^n Du_i^n \varphi(x_i) \to \frac{1}{2}\int_0^1 (u^2 - u_x^2)\varphi(x)\,dx$$

as $n \to \infty$. In the peakon case, $u^2 = u_x^2$ and this term tends to zero. Roughly speaking, we can say that m^n converges to a Dirac function, see (3.11), but at the same time it is multiplied by Du^n which is the average of the left and right derivatives and which tends to zero at the top of the peak. Eventually the whole product $m^n Du^n$ tends to zero. We follow the same heuristic approach with the term $m^n D_+ u^n$ in (3.13). This time, m^n is multiplied by the right derivative $D_+ u^n$ of u^n which tends, at the top of the peak, to $-c$. Hence, $-m^n D_+ u^n$ tends to $c\delta$ and not zero as it would if (3.13) converged to the correct solution. This example shows how sensitive the numerical approximation is, regarding the explicit form of the finite difference scheme, for the Camassa–Holm equation.

Acknowledgements. HH acknowledges helpful discussions with Nils Henrik Risebro and Kenneth H. Karlsen on discretizations of the Camassa–Holm equation.

References

[1] Haïm Brezis, *Analyse Fonctionnelle*, Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree], Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[2] Roberto Camassa, *Characteristics and the initial value problem of a completely integrable shallow water equation*, Discrete Contin. Dyn. Syst. Ser. B, 3 (2003), pp. 115–139.

[3] Roberto Camassa and Darryl D. Holm, *An integrable shallow water equation with peaked solitons*, Phys. Rev. Lett., 71 (1993), pp. 1661–1664.

[4] Roberto Camassa, Darryl D. Holm, and J. Hyman, *A new integrable shallow water equation*, Adv. Appl. Mech., 31 (1994), pp. 1–33.

[5] Roberto Camassa, J. Huang, and L. Lee, *Integral and integrable algorithms for a nonlinear shallow-water wave equation*, Preprint. Submitted, (2004).

[6] Roberto Camassa, Jingfang Huang, and Long Lee, *On a completely integrable numerical scheme for a nonlinear shallow-water wave equation*, J. Nonlinear Math. Phys., 12 (2005), pp. 146–162.

[7] Adrian Constantin, *On the Cauchy problem for the periodic Camassa-Holm equation*, J. Differential Equations, 141 (1997), pp. 218–235.

[8] ———, *Existence of permanent and breaking waves for a shallow water equation: a geometric approach*, Ann. Inst. Fourier (Grenoble), 50 (2000), pp. 321–362.

[9] Adrian Constantin and Joachim Escher, *Global existence and blow-up for a shallow water equation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 26 (1998), pp. 303–328.

[10] ———, *Well-posedness, global existence, and blowup phenomena for a periodic quasi-linear hyperbolic equation*, Comm. Pure Appl. Math., 51 (1998), pp. 475–504.

[11] Adrian Constantin and Luc Molinet, *Global weak solutions for a shallow water equation*, Comm. Math. Phys., 211 (2000), pp. 45–61.

[12] Holger R. Dullin, Georg A. Gottwald, and Darryl D. Holm, *Camassa-Holm, Korteweg-de Vries-5 and other asymptotically equivalent equations for shallow water waves*, Fluid Dynam. Res., 33 (2003), pp. 73–95. In memoriam Prof. Philip Gerald Drazin 1934–2002.

[13] C. Gasquet and P. Witomski, *Fourier Analysis and Applications*, vol. 30 of Texts in Applied Mathematics, Springer-Verlag, New York, 1999. Filtering, numerical computation, wavelets, Translated from the French and with a preface by R. Ryan.

[14] A. Alexandrou Himonas and Gerard Misiołek, *The Cauchy problem for an integrable shallow-water equation*, Differential Integral Equations, 14 (2001), pp. 821–831.

[15] Helge Holden and Xavier Raynaud, *A convergent numerical scheme for the camassa–holm equation based on multipeakons*, Discrete Contin. Dyn. Syst., 14 (2006), pp. 505–523.

[16] R. S. Johnson, *Camassa-Holm, Korteweg-de Vries and related models for water waves*, J. Fluid Mech., 455 (2002), pp. 63–82.

[17] Yi A. Li and Peter J. Olver, *Well-posedness and blow-up solutions for an integrable nonlinearly dispersive model wave equation*, J. Differential Equations, 162 (2000), pp. 27–63.

[18] G. Misiołek, *Classical solutions of the periodic Camassa-Holm equation*, Geom. Funct. Anal., 12 (2002), pp. 1080–1104.

[19] Robert Plato, *Concise numerical mathematics*, vol. 57 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003. Translated from the 2000 German original by Richard Le Borne and Sabine Le Borne.

[20] Guillermo Rodríguez-Blanco, *On the Cauchy problem for the Camassa-Holm equation*, Nonlinear Anal., 46 (2001), pp. 309–327.

[21] Jacques Simon, *Compact sets in the space $L^p(0,T;B)$*, Ann. Mat. Pura Appl. (4), 146 (1987), pp. 65–96.

Paper II

A convergent numerical scheme for the Camassa-Holm equation based on multipeakons.

H. Holden and X. Raynaud

Published in *Discrete Contin. Dyn. Syst.* 14 (2006), no. 3, 505–523.

A CONVERGENT NUMERICAL SCHEME FOR THE CAMASSA–HOLM EQUATION BASED ON MULTIPEAKONS

HELGE HOLDEN AND XAVIER RAYNAUD

Abstract. The Camassa–Holm equation $u_t - u_{xxt} + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0$ enjoys special solutions of the form $u(x,t) = \sum_{i=1}^n p_i(t)e^{-|x-q_i(t)|}$, denoted multipeakons, that interact in a way similar to that of solitons. We show that given initial data $u|_{t=0} = u_0$ in $H^1(\mathbb{R})$ such that $u - u_{xx}$ is a positive Radon measure, one can construct a sequence of multipeakons that converges in $L^\infty_{\text{loc}}(\mathbb{R}, H^1_{\text{loc}}(\mathbb{R}))$ to the unique global solution of the Camassa–Holm equation. The approach also provides a convergent, energy preserving nondissipative numerical method which is illustrated on several examples.

1. Introduction

The Camassa–Holm equation (CH) [4, 5]

$$u_t - u_{xxt} + 2\kappa u_x + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0 \tag{1.1}$$

has received considerable attention the last decade. With κ positive it models, see [14], propagation of unidirectional gravitational waves in a shallow water approximation, with u representing the fluid velocity. The Camassa–Holm equation possesses many intriguing properties: It is, for instance, completely integrable and experiences wave breaking in finite time for a large class of initial data. In this article we consider the case $\kappa = 0$ on the real line, that is,

$$u_t - u_{xxt} + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0, \tag{1.2}$$

and henceforth we refer to (1.2) as the Camassa–Holm equation.

Local and global well-posedness results as well as results concerning breakdown are proved in [8, 12, 16, 17]. It is known that certain initial data give global solutions, while other classes of initial data experience wave breaking in the sense that u_x becomes unbounded while the solution itself remains bounded. More precisely, the fundamental existence theorem, due to Constantin, Escher, and Molinet [8, 9], reads as follows: If $u_0 \in H^1(\mathbb{R})$ and $m_0 := u_0 - u_0''$ is a positive Radon measure, then equation (1.2) has a unique global weak solution $u \in C([0,T), H^1(\mathbb{R}))$, for any T positive, with initial data u_0. However, any

1991 *Mathematics Subject Classification.* Primary: 65M06, 65M12; Secondary: 35B10, 35Q53.

Key words and phrases. Camassa–Holm equation, peakons, numerical methods.

Supported in part by the BeMatA program of the Research Council of Norway.

solution with odd initial data u_0 in $H^3(\mathbb{R})$ such that $u_{0,x}(0) < 0$ blows up in a finite time ([8]).

The Camassa–Holm equation (1.2) exhibits so-called multipeakon solutions (see [5]), i.e., solutions of the form

$$u(x,t) = \sum_{i=1}^{n} p_i(t) e^{-|x-q_i(t)|}$$

where p_i and q_i are solutions of the following system of ordinary differential equations

$$\begin{aligned} \dot{q}_i &= \sum_{j=1}^{n} p_j e^{-|q_i-q_j|}, \\ \dot{p}_i &= \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i-q_j|}. \end{aligned} \tag{1.3}$$

The main idea in this article is to use multipeakons to approximate solutions of the Camassa–Holm equation. This gives rise to a numerical scheme for which we prove convergence.

In [5], Camassa, Holm, and Hyman use a pseudospectral method to solve (1.2) numerically but they do not study convergence of the method. We have shown in [13] how a particular finite difference scheme converges to the unique global solution in the case with periodic initial data.

The idea of using multipeakons has also been used by Camassa, Huang, and Lee in [3, 7, 6]. In [3], Camassa reformulates equation (1.2) in term of characteristics. The characteristics $q(\xi,t)$ are defined as solutions of the equation

$$q_t(\xi,t) = u(q(\xi,t),t)$$

with initial condition $q(\xi,0) = \xi$. After introducing the auxiliary variable p, which is directly related to the momentum $m = u - u_{xx}$ of the system,

$$p(\xi,t) = m(q(\xi,t),t)\frac{\partial q}{\partial \xi}(\xi,t),$$

Camassa shows that (1.2) reduces to the following system of partial differential equations

$$\begin{aligned} q_t(\xi,t) &= \frac{1}{2}\int_{-\infty}^{\infty} \exp\left(-\left|q(\xi,t) - q(\eta,t)\right|\right) p(\eta,t)\, d\eta, \\ p_t(\xi,t) &= \frac{1}{2} p(\xi,t) \int_{-\infty}^{\infty} \operatorname{sgn}(\xi-\eta) \exp\left(-\left|q(\xi,t) - q(\eta,t)\right|\right) p(\eta,t)\, d\eta. \end{aligned} \tag{1.4}$$

In [7, 6], Camassa, Huang, and Lee discretize system (1.4) by considering a finite number n of "particles" whose positions and momenta are given by

$$q_i(t) = q(\xi_i,t) \quad p_i = p(\xi_i,t)$$

for some equidistributed ξ_i. By approximating the integrals in (1.4) by their Riemann sums, equation (1.4) reduces to the system of ordinary differential equations given by (1.3). For initial data such that $p_i > 0$, (1.3) has global solutions in time. In this case they show that the scheme is convergent in the following sense. Let p and q be solutions of (1.4) and $\{p_i(t)\}_{i=1}^n$ and $\{q_i(t)\}_{i=1}^n$ be solutions of (1.3) with initial conditions $p_i(0) = p(\xi_i, 0)$ and $q_i(0) = q(\xi_i, 0)$. Then, when the number of particles n increases, $\{p_i(t)\}_{i=1}^n$ and $\{q_i(t)\}_{i=1}^n$ converge uniformly for any time interval $[0, T]$ to $\{p(\xi_i, t)\}_{i=1}^n$ and $\{q(\xi_i, t)\}_{i=1}^n$ in some discrete l_1 norm.

The approach we adopt here is different, and we obtain a more general convergence result, see Theorem 3.1. However, the numerical method, which is based on solving (1.3), is the same. We consider distributional solutions of (1.2), and show first that multipeakons are indeed distributional solutions. Given general initial data for (1.2), we construct a sequence of multipeakons and prove that it converges to the exact solution of the equation when the number of peakons is increased appropriately. More precisely, we prove that, given $u_0 \in H^1(\mathbb{R})$ such that $u_0 - u_{0,xx}$ is a positive Radon measure, there exists a sequence of multipeakons that converges in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1_{\mathrm{loc}}(\mathbb{R}))$ to the solution of the Camassa–Holm equation with initial data u_0. The proofs extend to the periodic case as well. Our proofs are constructive in the sense that we provide an explicit method, either by a collocation method or by a minimization technique (see Proposition 3.2 and Remark 3.4) to construct the multipeakon approximation. This gives a constructive proof of existence of solutions to the Camassa–Holm equation and shows that the multipeakons span the set of solutions at least in the case where the initial data satisfy the condition mentioned above. Furthermore, this leads to a numerical method which, in contrast to the finite difference scheme presented in [13], does not contain any dissipation and preserves the $H^1(\mathbb{R})$ norm exactly. In the last section we illustrate the method on two numerical examples.

2. Global existence of multipeakon solutions

The Camassa–Holm equation may be rewritten as

$$m_t + um_x + 2mu_x = 0 \tag{2.1}$$

where the momentum m equals $u - u_{xx}$.

Definition 2.1. We say that u in $L^1_{\mathrm{loc}}([0, T), H^1_{\mathrm{loc}})$ is a weak solution of the Camassa–Holm equation if it satisfies

$$u_t - u_{xxt} + \frac{3}{2}(u^2)_x + \frac{1}{2}(u_x^2)_x - \frac{1}{2}(u^2)_{xxx} = 0 \tag{2.2}$$

in the sense of distributions.

When u is smooth, (2.1) and (2.2) are equivalent. Multipeakons are solutions of the form

$$u(x,t) = \sum_{i=1}^n p_i(t) e^{-|x - q_i(t)|}, \tag{2.3}$$

which are continuous and piecewise C^∞ functions in $H^1(\mathbb{R})$ for any given t. But since they have discontinuous first derivative, they cannot satisfy the Camassa–Holm equation in the classical sense. For functions with these properties the left-hand side of (2.2) is a distribution which consists of regular terms (piecewise C^∞

functions) and singular terms (Dirac functions or their derivatives at the points q_i) that we can compute explicitly. We only give the details of the computation of the last term, $(u^2)_{xxx}$, in (2.2), the other terms being obtained similarly. For each $i \in \{0, \dots, n+1\}$ we introduce the function

$$u_i(x,t) = \sum_{j=1}^{i} p_j(t) e^{-(x-q_j(t))} + \sum_{j=i+1}^{n} p_j(t) e^{(x-q_j(t))}$$

which is C^∞ in the space variable. Then (2.3) can be rewritten as

$$u(x,t) = \sum_{i=0}^{n} u_i(x,t)\chi_i(x)$$

where χ_i denotes the characteristic function of the interval $[q_i, q_{i+1})$ with the convention that $q_0 = -\infty$ and $q_{n+1} = \infty$. Since the χ_i have disjoint supports, we have

$$u^2 = \sum_{i=0}^{n} u_i^2 \chi_i \tag{2.4}$$

and, after differentiating (2.4),

$$\begin{aligned}
(u^2)_x &= \sum_{i=0}^{n} (u_i^2)_x \chi_i + \sum_{i=1}^{n} u_i^2(q_i)\delta_{q_i} - \sum_{i=0}^{n-1} u_i^2(q_{i+1})\delta_{q_{i+1}} \\
&= \sum_{i=0}^{n} (u_i^2)_x \chi_i + \sum_{i=1}^{n} \left(u_i^2(q_i) - u_{i-1}^2(q_i)\right)\delta_{q_i} \\
&= \sum_{i=0}^{n} (u_i^2)_x \chi_i + \sum_{i=1}^{n} \left[u^2\right]_{q_i} \delta_{q_i} \qquad (2.5)
\end{aligned}$$

where the bracket $[v]_{q_i}$ denotes the jump of v across q_i, that is, $[v]_{q_i} = v(q_i^+) - v(q_i^-)$. Since u is continuous, $\left[u^2\right]_{q_i} = 0$, and the last term in (2.5) vanishes. We differentiate (2.5) and get

$$\begin{aligned}
(u^2)_{xx} &= \sum_{i=0}^{n} (u_i^2)_{xx} \chi_i + \sum_{i=1}^{n} (u_i^2)_x(q_i)\delta_{q_i} - \sum_{i=0}^{n-1} (u_i^2)_x(q_{i+1})\delta_{q_{i+1}} \\
&= \sum_{i=0}^{n} (u_i^2)_{xx} \chi_i + \sum_{i=1}^{n} \left[(u^2)_x\right]_{q_i} \delta_{q_i}. \qquad (2.6)
\end{aligned}$$

On every interval (q_i, q_{i+1}), since $u = u_i$, u is differentiable and every derivative of u admits a limit when x tends to q_i from one side. It follows that the jump $\left[(u^2)_x\right]_{q_i}$ is a well-defined quantity and justifies its use in (2.6). Finally, after differentiating (2.6) once more, we get

$$(u^2)_{xxx} = \sum_{i=0}^{n} (u_i^2)_{xxx} \chi_i + \sum_{i=1}^{n} \left[(u^2)_{xx}\right]_{q_i} \delta_{q_i} + \sum_{i=1}^{n} \left[(u^2)_x\right]_{q_i} \delta'_{q_i}.$$

In a similar way we can compute the other terms in (2.2) and we end up with

$$\begin{aligned} u_t - u_{xxt} + \frac{3}{2}(u^2)_x + \frac{1}{2}(u_x^2)_x - \frac{1}{2}(u^2)_{xxx} & \\ = \sum_{i=0}^{n} \Big(u_{i,t} - u_{i,xxt} + \frac{3}{2}(u_i^2)_x + \frac{1}{2}(u_{i,x}^2)_x - \frac{1}{2}(u_i^2)_{xxx} \Big) \chi_i & \\ + \sum_{i=1}^{n} \Big(- [u_{xt}]_{q_i} + \frac{1}{2}\left[u_x^2\right]_{q_i} - \frac{1}{2}\left[(u^2)_{xx}\right]_{q_i} \Big) \delta_{q_i} & \\ + \sum_{i=1}^{n} \Big(- [u_t]_{q_i} - \frac{1}{2}\left[(u^2)_x\right]_{q_i} \Big) \delta'_{q_i}. & \end{aligned} \tag{2.7}$$

We already noted the equivalence between (2.2) and (2.1) when u is smooth. The same equivalence obviously holds for u_i and, after introducing m_i to denote $u_i - u_{i,xx}$, we have

$$u_{i,t} - u_{i,xxt} + \frac{3}{2}(u_i^2)_x + \frac{1}{2}(u_{i,x}^2)_x - \frac{1}{2}(u_i^2)_{xxx} = m_{i,t} + u_i m_{i,x} + 2m_i u_{i,x} = 0$$

because, from the definition of u_i as a linear combination of e^{-x} and e^x, it is clear that $m_i = 0$. Thus, the first sum on the right-hand side of (2.7) vanishes. The values of the jumps in (2.7) can be computed from (2.3). We have

$$[u_x]_{q_i} = -2p_i \tag{2.8}$$

and, after some calculation,

$$\begin{aligned} &\left[(u^2)_{xx}\right]_{q_i} = 0, \quad [u_t]_{q_i} = 2p_i\dot{q}_i, \quad [u_{xt}]_{q_i} = -2\dot{p}_i, \\ &\left[u_x^2\right]_{q_i} = [u_x]_{q_i}\left(u_x(q_i^+) + u_x(q_i^-)\right) = 4p_i \sum_{j=1}^{n} p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i - q_j|}, \\ &\left[(u^2)_x\right]_{q_i} = 2u(q_i)\,[u_x]_{q_i} = -4p_i \sum_{j=1}^{n} p_j e^{-|q_i - q_j|}. \end{aligned} \tag{2.9}$$

Assume that the q_i are all distinct. Then (2.2) holds if and only if the coefficients multiplying δ_{q_i} and δ'_{q_i} in (2.7) all vanish. Hence, after using (2.9), (2.7) and (2.2), we end up with the system

$$\begin{cases} \dot{q}_i = \sum_{j=1}^{n} p_j e^{-|q_i - q_j|}, \\ \dot{p}_i = \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i - q_j|} \end{cases} \tag{2.10}$$

with the convention that $\operatorname{sgn}(x) = 0$ if $x = 0$. We summarize the discussion in the following lemma.

Lemma 2.2. *The function* (2.3) *is a weak solution of the Camassa–Holm equation if and only if* p_i, q_i *satisfy the system* (2.10) *of ordinary differential equations.*

The system (2.10) is Hamiltonian with Hamiltonian H given by

$$H = \frac{1}{2}\sum_{i,j=1}^{n} p_i p_j e^{-|q_i-q_j|}.$$

It means that (2.10) can be rewritten as

$$\dot{q}_i = \frac{\partial H}{\partial p_i},\ \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{2.11}$$

From (2.3), the momentum is given by

$$m = 2\sum_{i=1}^{n} p_i \delta_{q_i}.$$

Hence,

$$\|u\|^2_{H^1(\mathbb{R})} = \langle u - u_{xx}, u\rangle_{H^{-1}} = \sum_{i=1}^{n} 2p_i u(q_i) = 2\sum_{i,j=1}^{n} p_i p_j e^{-|q_i-q_j|}.$$

and the Hamiltonian H and the H^1 norm of u satisfy

$$H = \frac{1}{2}\sum_{i,j=1}^{n} p_i p_j e^{-|q_i-q_j|} = \frac{1}{4}\|u\|^2_{H^1(\mathbb{R})}. \tag{2.12}$$

Because of the sign function, the right-hand side in (2.10) is not Lipchitz, and we cannot apply Picard's theorem to get existence and uniqueness of solutions of (2.10). However, the Lipschitz condition would hold if we knew in advance that $q_i - q_j$ does not change sign. We are going to prove that the peaks do not cross and that the sign of $q_i - q_j$ is indeed preserved.

Let us first assume, without loss of generality, that the positions of the peaks at time $t = 0$, $\{q_i\}_{i=1}^n$, are distinct and ordered as follows

$$q_i(0) < q_j(0) \text{ for all } i < j. \tag{2.13}$$

We consider the system of ordinary differential equations

$$\begin{cases} \dot{q}_i = \displaystyle\sum_{j=1}^{n} p_j e^{-|q_i-q_j|}, \\ \dot{p}_i = \displaystyle\sum_{j=1}^{n} p_i p_j \operatorname{sgn}(i-j) e^{-|q_i-q_j|}. \end{cases} \tag{2.14}$$

This system is equivalent to (2.10) as long as the positions of the peaks q_i satisfy the ordering defined in (2.13). In contrast to (2.10), the system (2.14) fulfills the Lipchitz condition of Picard's theorem, and therefore there exists a unique maximal solution. If, in addition, the p_i are strictly positive initially then the solution exists for all time.

Lemma 2.3. *Let* $\{p_i, q_i\}$ *be the maximal solutions of* (2.14). *If we have*

$$q_i(t) < q_j(t) \quad \textit{for all } i < j, \tag{2.15}$$

$$p_i(t) > 0 \quad \textit{for all } i, \tag{2.16}$$

when $t = 0$, *then* $\{p_i(t), q_i(t)\}$ *are globally defined on* $[0, \infty)$ *and inequalities* (2.15) *and* (2.16) *remain true for all* t.

Proof. We call T the maximal time of existence. Let us assume that (2.15) and (2.16) do not hold for all $t \in [0, T)$. Then, since p_i and q_i are continuous, there exist t_0 in $[0, T)$ such that (2.15) and (2.16) hold in $[0, t_0)$ and either

$$q_i(t_0) = q_j(t_0) \quad \text{for some } i \text{ and } j \text{ with } i < j$$

or

$$p_i(t_0) = 0 \quad \text{for some } i.$$

In the first case when $q_i(t_0) = q_j(t_0) = \alpha$, we have that q_i and q_j are both solutions of the ordinary differential equation

$$\dot{q} = \sum_{k=1}^{n} p_k e^{-|q - q_k|}$$

with initial condition $q(t_0) = \alpha$. The function q plays the role of the unknown while p_k and q_k are given (they are the solutions of (2.14)). By Picard's theorem, we know that, given some initial condition, the solution is unique and therefore $q_i = q_j$ at least in a small interval centered around t_0. This contradicts the assumption that $q_i(t) < q_j(t)$ in $[0, t_0)$. In the second case when $p_i(t_0) = 0$, the function p_i is solution of

$$\dot{p} = p \sum_{j=1}^{n} p_j \operatorname{sgn}(i - j) e^{-|q_i - q_j|}$$

with initial condition $p(t_0) = 0$. Zero is an obvious solution and since the solution is unique, we must have $p_i = 0$ on $[0, T)$. This contradicts our assumption, and hence (2.15) and (2.16) hold for all time t in $[0, T)$. We denote by M the following quantity

$$M = 2 \sum_{i=1}^{n} p_i. \tag{2.17}$$

As we will see in the next section, M corresponds to the total mass of the system and it is a conserved quantity of the governing equation (1.1). We can directly check this statement here since we have

$$\frac{dM}{dt} = 2 \sum_{i,j=1}^{n} p_i p_j \operatorname{sgn}(i - j) e^{-|q_i - q_j|} = 0. \tag{2.18}$$

We have proved that the p_i are positive for all t in $[0, T)$. Therefore we have

$$0 < p_i(t) < \frac{M}{2},$$

for all i and all $t \in [0, T)$, which implies the p_i are bounded. It follows that $\dot{p}_i$ and $\dot{q}_i$ in (2.14) are bounded and the maximum solution is therefore defined for all time, i.e., $T = \infty$. $\square$

Lemma 2.3 tells us that the ordering of the positions of the peaks is preserved, and in this case, as we already mentioned, (2.10) and (2.14) are equivalent. Thus we have established the following result.

Lemma 2.4. *If $q_i < q_j$ for $i < j$ and $p_i > 0$ at $t = 0$, then the system* (2.10) *has a unique, globally defined solution on* $[0, \infty)$.

Remark 2.5. A similar result is proved by other means in [6].

3. Convergence of multipeakon sequences

Multipeakon solutions can be used to prove the existence of solutions for the Camassa–Holm equation.

Theorem 3.1. *Given u_0 in $H^1(\mathbb{R})$ such that $m_0 = u_0 - u_{0,xx}$ is in $\mathcal{M}^+$, the space of positive finite Radon measures, there exists a sequence of multipeakons that converges in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1_{\mathrm{loc}}(\mathbb{R}))$ to the unique solution of the Camassa–Holm equation with initial condition u_0.*

The proof of Theorem 3.1 is presented at the end of the section. The sequence of multipeakons mentioned in the theorem is denoted by $u^n(x,t) = \sum_{i=1}^n p_i^n(t)e^{-|x-q_i^n(t)|}$. We require that the initial conditions $u_0^n(x) = u^n(x,0)$ satisfy the following properties

$$u_0^n \to u_0 \text{ in } H^1(\mathbb{R}), \tag{3.1a}$$

$$u_0^n \text{ is uniformly bounded in } L^1(\mathbb{R}), \tag{3.1b}$$

$$p_i^n \geq 0 \text{ for all } i \text{ and } n. \tag{3.1c}$$

In the next proposition we give a constructive proof that such sequences exist. The sequence u_0^n is defined by collocation: It coincides with the given initial function u_0 at a given number of points.

Proposition 3.2. *Given $u_0 \in H^1(\mathbb{R})$ such that $u_0 - u_{0,xx} \in \mathcal{M}^+$. For each n, let $q_{i,n} = i/n$. Then, there exists a unique $(p_{i,n})_{i=-n^2}^{n^2}$ such that $u_0^n(x) = \sum_{i=-n^2}^{n^2} p_{i,n}e^{-|x-q_{i,n}|}$ coincides with u_0 at the $q_{i,n}$, that is,*

$$u_0^n(q_{i,n}) = u_0(q_{i,n}) \tag{3.2}$$

for all $i \in \{-n^2, \dots, n^2\}$. The initial multipeakon sequence u_0^n satisfies condition (3.1).

Proof. In order to simplify the notation, we write u and u^n instead of u_0 and u_0^n. First we show that (3.2) defines a unique $p_{i,n}$. The equation (3.2) is equivalent to the following system

$$A\overline{p} = \overline{u} \tag{3.3}$$

where $\overline{p}$ and $\overline{u}$ are vectors of $\mathbb{R}^{2n^2+1}$ given by $(p_{i,n})_{i=-n^2}^{n^2}$ and $(u(q_{i,n}))_{i=-n^2}^{n^2}$, respectively, and A equals the matrix

$$A = (A_{i,j})_{i,j=-n^2}^{n^2}, \quad A_{i,j} = e^{-|q_{i,n}-q_{j,n}|}.$$

The method is well-posed if A is invertible. In fact, A is symmetric and positive definite. Symmetry is obvious. To prove the positivity of A, we associate to any vector $\overline{r}$ in $\mathbb{R}^{2n^2+1}$ the function v in $H^1(\mathbb{R})$ given by $v(x) = \sum_{i=-n^2}^{n^2} r_{i,n} e^{-|x-q_{i,n}|}$. The H^1 norm of v has already been calculated, see (2.12), and we have

$$\overline{r}^t A \overline{r} = \frac{1}{2} \|v\|_{H^1(\mathbb{R})}^2 \geq 0. \tag{3.4}$$

Hence, A is positive. Let us prove that A is invertible. Assume $A\overline{r} = 0$. From (3.4), we have $v = 0$. Thus, since $v - v_{xx} = 2\sum_{i=-n^2}^{n^2} r_{i,n}\delta_{q_{i,n}}$, we have

$$2 \sum_{i=-n^2}^{n^2} r_{i,n}\delta_{q_{i,n}} = 0. \tag{3.5}$$

Since the $q_{i,n}$ are all distinct, it follows that $\overline{r} = 0$. Hence, A is invertible, and thus there exists a unique $\overline{p}$ solving (3.3) for any given $\overline{u}$.

Let us prove (3.1a). Let f and v^n denote $u - u_{xx}$ and $u - u^n$, respectively. We want to prove that v^n tends to zero in $H^1(\mathbb{R})$. Note that $v^n - v_{xx}^n = f - 2\sum_{i=-n^2}^{n^2} p_{i,n}\delta_{q_{i,n}}$ is a Radon measure and we have

$$\|v^n\|_{H^1(\mathbb{R})}^2 = \langle v^n - v_{xx}^n, v^n \rangle = \langle f, v^n \rangle - 2p_{i,n} v^n(q_{i,n})$$

where the bracket $\langle \mu, g \rangle$ denotes the integration of g with respect to the Radon measure μ. By assumption (3.2), we have $v(q_{i,n}) = 0$, and it follows that

$$\|v^n\|_{H^1(\mathbb{R})}^2 = \langle f, v^n \rangle. \tag{3.6}$$

We consider a partition of unity of $\mathbb{R}$ that we denote $\{\phi_{i,n}\}_{i=-\infty}^{\infty}$ and which corresponds to the decomposition $\mathbb{R} = \cup_{i=-\infty}^{\infty} \left(\frac{i-1}{n}, \frac{i+1}{n}\right)$. The functions $\phi_{i,n}$ satisfy $0 \leq \phi_{i,n} \leq 1$, $\sum_{i=-\infty}^{\infty} \phi_{i,n} = 1$ and $\operatorname{supp}\phi_{i,n} \subset (\frac{i-1}{n}, \frac{i+1}{n})$. Then we have

$$\langle f, v^n \rangle = \langle f, \psi_n v^n \rangle + \sum_{i=-n^2}^{n^2} \langle f, \phi_i v^n \rangle \tag{3.7}$$

where $\psi_n = 1 - \sum_{i=-n^2}^{n^2} \phi_i$. We estimate separately the two terms on the right-hand side of (3.7). Since the support of ϕ_i is contained in $(q_{i-1,n}, q_{i+1,n})$, we have

$$\phi_i(x)v^n(x) \leq \sup_{x\in(q_{i-1,n},q_{i+1,n})} |v^n(x)|\, \phi_i(x). \tag{3.8}$$

Since

$$v^n(x) = v^n(q_{i,n}) + \int_{q_{i,n}}^x v_x^n(t)\, dt$$

and $v^n(q_{i,n}) = 0$, we have

$$\begin{aligned}\sup_{x\in(q_{i-1,n},q_{i+1,n})} |v^n(x)| &\leq \int_{q_{i-1,n}}^{q_{i+1,n}} |v^n_x(t)|\; dt \\ &\leq \sqrt{\frac{2}{n}}\, \|v^n\|_{H^1(\mathbb{R})} \quad \text{(Cauchy–Schwarz).} \end{aligned} \tag{3.9}$$

The positivity of f directly implies that f is monotone: If $u \leq v$, then $\langle f, u\rangle \leq \langle f, v\rangle$. Hence, from (3.8), (3.9) and the monotonicity of f, we get

$$\sum_{i=-n^2}^{n^2} \langle f, \phi_i v^n\rangle \leq \sum_{i=-n^2}^{n^2} \sqrt{\frac{2}{n}}\, \|v^n\|_{H^1(\mathbb{R})} \langle f, \phi_i\rangle \leq \sqrt{\frac{2}{n}}\, \|v^n\|_{H^1(\mathbb{R})} \|f\|_{\mathcal{M}}\,. \tag{3.10}$$

Since $H^1(\mathbb{R})$ is continuously embedded in $L^\infty(\mathbb{R})$, we have, for some constant C independent of n,

$$\psi_n(x)v^n(x) \leq \|v^n\|_{L^\infty} \psi_n(x) \leq C\, \|v^n\|_{H^1(\mathbb{R})} \psi_n(x)$$

and, after using the monotonicity of f,

$$\langle f, \psi_n v^n\rangle \leq C\, \|v^n\|_{H^1(\mathbb{R})} \langle f, \psi_n\rangle \tag{3.11}$$

Gathering (3.6), (3.7), (3.10) and (3.11), we get

$$\|v^n\|^2_{H^1(\mathbb{R})} \leq \sqrt{\frac{2}{n}}\, \|v^n\|_{H^1(\mathbb{R})} \|f\|_{\mathcal{M}} + C\, \|v^n\|_{H^1(\mathbb{R})} \langle f, \psi_n\rangle$$

which, after dividing both terms by $\|v^n\|_{H^1(\mathbb{R})}$,

$$\|v^n\|_{H^1(\mathbb{R})} \leq \sqrt{\frac{2}{n}}\, \|f\|_{\mathcal{M}} + C\, \langle f, \psi_n\rangle\,. \tag{3.12}$$

It remains to prove that $\langle f, \psi_n\rangle$ tends to zero. The space of Radon measures and C_0^*, the dual of C_0, where C_0 denotes the closure of C_c in $L^\infty(\mathbb{R})$, are isometrically isomorphic (see, e.g., [11, Chapter 7]), and we have

$$\|f\|_{\mathcal{M}} = \sup_{\substack{\varphi\in C_c \\ \|\varphi\|_{L^\infty}\leq 1}} \langle f, \varphi\rangle\,.$$

Therefore, for all $\varepsilon > 0$, there exists $\tilde\varphi \in C_c$ with $\|\tilde\varphi\|_{L^\infty} \leq 1$ and such that

$$\|f\|_{\mathcal{M}} \leq \langle f, \tilde\varphi\rangle + \varepsilon.$$

For n big enough, the supports of ψ_n and $\tilde\varphi$ do not intersect and therefore we have $\|\psi_n + \tilde\varphi\|_{L^\infty} \leq 1$. Hence,

$$\begin{aligned}\langle f, \psi_n + \tilde\varphi\rangle &\leq \|f\|_{\mathcal{M}} \|\psi_n + \tilde\varphi\|_{L^\infty} \\ &\leq \|f\|_{\mathcal{M}} \\ &\leq \langle f, \tilde\varphi\rangle + \varepsilon\end{aligned}$$

which implies

$$\langle f, \psi_n\rangle \leq \varepsilon,$$

and this proves that $\langle f, \psi_n \rangle \to 0$. Then, by (3.12), we get that v^n tends to zero, and (3.1a) is proved.

Let us prove (3.1b), namely that $p_{i,n} \geq 0$ for all $-n^2 \leq i \leq n^2$. Let f again denote $u - u_{xx}$. By assumption, f is positive. In a first step, we assume that f belongs to $C^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$. We will remove this smoothness assumption afterwards. A notable property of u^n is that it is always bounded by u, i.e.,

$$u^n(x) \leq u(x) \text{ for all } x. \tag{3.13}$$

We see this as follows. Let $v = u - u^n$. Since we have $u^n - u^n_{xx} = 0$ everywhere except at the $q_{i,n}$ and (3.2) holds, v satisfies, for every $i \in \{-n^2, \ldots, n^2 - 1\}$, the Dirichlet problem

$$\begin{aligned} &v - v_{xx} = f \quad \text{on } (q_{i,n}, q_{i+1,n}), \\ &v(q_{i,n}) = v(q_{i+1,n}) = 0. \end{aligned} \tag{3.14}$$

The Green's function $G(x, \xi)$ is defined as the solution of (3.14) with $f = \delta(x - \xi)$. We can compute G (see for example [1]), and we get

$$G(x,\xi) = \frac{1}{\sinh(q_{i+1,n} - q_{i,n})} \begin{cases} \sinh(x - q_{i,n}) \sinh(q_{i+1,n} - \xi) & \text{for } q_{i,n} \leq x \leq \xi, \\ \sinh(\xi - q_{i,n}) \sinh(q_{i+1,n} - x) & \text{for } \xi < x \leq q_{i+1,n}. \end{cases} \tag{3.15}$$

The general solution of (3.14) is then given by

$$v(x) = \int_{q_{i,n}}^{q_{i+1,n}} G(x, \xi) f(\xi)\, d\xi. \tag{3.16}$$

Since $G(x, \xi)$ is positive, it follows from (3.16) that $v \geq 0$ on every interval $[q_{i,n}, q_{i+1,n}]$. On the intervals $(-\infty, q_{-n^2,n}]$ and $[q_{n^2,n}, \infty)$, v solves a Dirichlet problem similar to (3.14) and the Green's functions are obtained from (3.15) by letting $q_{-n^2-1,n}$ tend to $-\infty$ and $q_{n^2+1,n}$ to $+\infty$, respectively. The Green's functions are still positive and that implies, as before, that $v \geq 0$ on $(-\infty, q_{-n^2,n}] \cup [q_{n^2,n}, \infty)$. This concludes the proof of (3.13). From (2.8), we have

$$\begin{aligned} p_{i,n} &= -\frac{1}{2} \left[u^n_x\right]_{q_{i,n}} \\ &= -\frac{1}{2} \lim_{h \downarrow 0} \left[\frac{u^n(q_{i,n} + h) - u^n(q_{i,n})}{h} - \frac{u^n(q_{i,n}) - u^n(q_{i,n} - h)}{h} \right] \end{aligned}$$

and, after using (3.13) and (3.2),

$$\begin{aligned} p_{i,n} &\geq -\frac{1}{2} \lim_{h \downarrow 0} \left[\frac{u(q_{i,n} + h) - u(q_{i,n})}{h} - \frac{u(q_{i,n}) - u(q_{i,n} - h)}{h} \right] \\ &\geq -\frac{1}{2} \left[u_x\right]_{q_{i,n}}. \end{aligned}$$

Since f is smooth, u is smooth and therefore $\left[u_x\right]_{q_{i,n}} = 0$. Hence,

$$p_{i,n} \geq 0. \tag{3.17}$$

We want to prove (3.17) without any extra smoothness assumption on f. Let ρ be a positive, C^∞ and even function which satisfies $\int_{-\infty}^{\infty} \rho(x)\, dx = 1$. We denote by ρ_ε the mollifier $\rho_\varepsilon = \frac{1}{\varepsilon}\rho(x/\varepsilon)$. Let $f_\varepsilon = \rho_\varepsilon * f$ and $u_\varepsilon = \rho_\varepsilon * u$. The mollified function u_ε tends to u in $H^1(\mathbb{R})$ and therefore in $L^\infty(\mathbb{R})$. Hence, for all i in $\{-n^2, \ldots, n^2\}$, $u_\varepsilon(q_{i,n})$ tends to $u(q_{i,n})$ or, using the previous notations,

$$\overline{u}_\varepsilon \to \overline{u}. \tag{3.18}$$

We can construct multipeakons u_ε^n from the regularized function u_ε whose coefficients $p_{\varepsilon,i,n}$ are determined by

$$A\overline{p}_\varepsilon = \overline{u}_\varepsilon, \quad \overline{p}_\varepsilon = (p_{\varepsilon,i,n})_{i=-n^2}^{n^2}. \tag{3.19}$$

Since f is positive, f_ε is positive and, since it also belongs to $C^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$, we have already established, see (3.17), that $\overline{p}_\varepsilon \geq 0$. Thus, by (3.18),

$$\overline{p} = A\overline{u} = \lim_{\varepsilon\to 0} A\overline{u}_\varepsilon = \lim_{\varepsilon\to 0} \overline{p}_\varepsilon, \tag{3.20}$$

implying that $\overline{p}$ is positive and (3.1b) is proved.

It remains to prove (3.1c), namely that u^n is bounded in $L^1(\mathbb{R})$. The regularized f_ε of f belongs to $C^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ and is positive. Hence, (3.13) holds when u^n and u are replaced by u_ε^n and u_ε:

$$u_\varepsilon^n \leq u_\varepsilon. \tag{3.21}$$

From (3.20), we have $\overline{p}_\varepsilon \to \overline{p}$ when $\varepsilon \to 0$. Then by looking at the definitions of u_ε^n and u^n it is clear that u_ε^n tends to u^n in $L^\infty(\mathbb{R})$. We have already seen that u_ε tends to u_ε in $L^\infty(\mathbb{R})$. Hence, after letting ε tend to zero in (3.21), we get that (3.13) holds for all f without any further smoothness assumption. Moreover, u is positive since the positivity of $\overline{p}$ implies the positivity of u^n. From (3.13), we get

$$\int_{-\infty}^{\infty} u^n(x)\,dx \leq \int_{-\infty}^{\infty} u(x)\,dx. \tag{3.22}$$

If u belongs to $L^1(\mathbb{R})$, then a bound on $\|u^n\|_{L^1}$ follows directly from (3.22). Again, we consider the regularized f_ε of f. Since u_ε satisfies $u_\varepsilon - u_{\varepsilon,xx} = f_\varepsilon$, it is known that u_ε can be expressed as

$$u_\varepsilon(x) = \int_{-\infty}^{\infty} e^{-|x-y|} f_\varepsilon(y)\,dy.$$

Hence,

$$\begin{aligned}\int_{-\infty}^{\infty} u_\varepsilon\,dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-|x-y|} f_\varepsilon(y)\,dydx \\ &= 2\|f_\varepsilon\|_{L^1} \quad \text{(after applying Fubini's theorem)} \\ &= 2\|f_\varepsilon\|_{\mathcal{M}}.\end{aligned}$$

Since u_ε is positive and converges to u in $L^\infty(\mathbb{R})$, by Fatou's lemma, we get

$$\begin{aligned}\int_{-\infty}^{\infty} u(x)\,dx &\leq \liminf \int_{-\infty}^{\infty} u_\varepsilon(x)\,dx \\ &\leq 2\liminf \|f_\varepsilon\|_{\mathcal{M}}.\end{aligned} \tag{3.23}$$

Let us estimate $\|f_\varepsilon\|_{\mathcal{M}}$. For any continuous function ϕ with compact support, we have

$$\langle f_\varepsilon, \phi\rangle = \langle \rho_\varepsilon * f, \phi\rangle = \langle f, \phi * \rho_\varepsilon\rangle . \tag{3.24}$$

Note that the last equality in (3.24) holds because of the parity of ρ_ε (see, e.g., [11, Chapter 9] for general formulas on convolutions of distributions). Hence,

$$\begin{aligned} |\langle f_\varepsilon, \phi\rangle| &\leq \|f\|_{\mathcal{M}} \|\phi * \rho_\varepsilon\|_{L^\infty} \\ &\leq \|f\|_{\mathcal{M}} \|\phi\|_{L^\infty} \|\rho_\varepsilon\|_{L^1} \quad \text{(Young's inequality)} \end{aligned}$$

and, since $\|\rho_\varepsilon\|_{L^1} = 1$, it implies

$$\|f_\varepsilon\|_{\mathcal{M}} \leq \|f\|_{\mathcal{M}} .$$

Inequality (3.23) now gives

$$\int_{-\infty}^{\infty} u(x)\, dx \leq 2\|f\|_{\mathcal{M}}$$

which implies that u belongs to $L^1(\mathbb{R})$. From (3.22), we get that $\|u^n\|_{L^1}$ is bounded. This concludes the proof of the proposition. $\square$

Remark 3.3. The initial multipeakon sequence $u_0^n(x) = \sum_{i=-n^2}^{n^2} p_{i,n} e^{-|x-q_{i,n}|}$ defined by setting

$$p_{i,n} = \frac{1}{2}\langle m_0, \phi_{i,n}\rangle \quad \text{for } i \in \{-n^2, \dots, n^2\}$$

where $\{\phi_i\}_{i=-\infty}^{\infty}$ denotes the partition of unity used in (3.7), also satisfies the condition (3.1). The proof of that result is much shorter than the proof of Proposition 3.2. However, the method is not directly applicable numerically (we would have to construct the ϕ_i and compute $2n^2+1$ integrals), which makes Proposition 3.2 more interesting.

Remark 3.4. Another natural way to construct a sequence of multipeakons from the set of points $q_{i,n}$, is to choose $\overline{p}$ so that it minimizes $\|u_0 - u_0^n\|_{H^1(\mathbb{R})}$, that is,

$$\overline{p} = \underset{p_{i,n}}{\operatorname{Argmin}} \left\| u_0 - \sum_{i=-n^2}^{n^2} p_{i,n} e^{-|x-q_{i,n}|} \right\|_{H^1(\mathbb{R})} . \tag{3.25}$$

It turns out that the sequence that this minimization method produces and the one of Proposition 3.2 are the same. One can prove this as follows. We have

$$\begin{aligned} \|u_0 - u_0^n\|_{H^1(\mathbb{R})}^2 &= \|u_0\|_{H^1(\mathbb{R})}^2 - 2\langle u_0^n, u_0\rangle_{H^1} + \|u_0^n\|_{H^1(\mathbb{R})}^2 \\ &= \|u_0\|_{H^1(\mathbb{R})}^2 - 2\left\langle u_0^n - u_{0,xx}^n, u_0\right\rangle_{\mathcal{M}} + 2\overline{p}^t A \overline{p} \end{aligned}$$

and, since

$$\left\langle u_0^n - u_{0,xx}^n, u_0\right\rangle_{\mathcal{M}} = \left\langle \sum_{i=-n^2}^{n^2} 2p_{i,n}\delta_{q_{i,n}}, u_0 \right\rangle = 2\sum_{i=-n^2}^{n^2} p_{i,n} u_0(q_{i,n}) = 2\overline{p}^t \overline{u},$$

we get

$$\|u_0 - u_0^n\|_{H^1(\mathbb{R})}^2 = \|u_0\|_{H^1(\mathbb{R})}^2 - 4\overline{p}^t\overline{u} + 2\overline{p}^t A\overline{p}. \tag{3.26}$$

By differentiating (3.26) with respect to $\overline{p}$, we can easily check that the minimizer $\overline{p}$ of (3.26) satisfies (3.3). In addition, since A is positive definite, $\overline{p}$ is the unique strict minimizer of (3.26).

The estimates contained in the following lemma will be needed to derive the existence of a converging subsequence.

Lemma 3.5. *Let $u^n(x,t)$ be a sequence of multipeakons with initial data satisfying* (3.1)*. The following properties hold:*
(i) u^n *is uniformly bounded in* $H^1(\mathbb{R})$*,*
(ii) u_x^n *is uniformly bounded in* $L^\infty(\mathbb{R})$*,*
(iii) u_x^n *has a uniformly bounded total variation,*
(iv) u_t^n *is uniformly bounded in* $L^2(\mathbb{R})$*.*

Proof. From Lemma 2.4 we know, using assumption (3.1c), that the system (2.10) has a unique global solution, and hence we have a globally defined sequence of multipeakons denoted $u^n(x,t)$. In order to simplify the notation, we drop the superscript n on p_i^n and q_i^n and write

$$u^n = \sum_{i=1}^n p_i e^{-|x-q_i|}.$$

Property (i) is obvious because of (3.1a) and the fact that the H^1 norm is automatically preserved due to the Hamiltonian structure of (2.10). We have

$$u_x^n(t,x) = \sum_{i=1}^n -p_i(t)\operatorname{sgn}(x-q_i(t))e^{-|x-q_i(t)|} \quad \text{a.e.}$$

Hence,

$$\begin{aligned} |u_x^n(x,t)| &\le \sum_{i=1}^n p_i(t)e^{-|x-q_i(t)|} \quad && (p_i \ge 0) \\ &\le \|u^n\|_{L^\infty} \\ &\le C\,\|u^n\|_{H^1(\mathbb{R})} && (H^1(\mathbb{R}) \text{ is continuously embedded in } L^\infty(\mathbb{R})) \end{aligned}$$

and (ii) follows from (i). The total variation of u_x^n equals $\|u_{xx}^n\|_{\mathcal{M}}$ (see, e.g., [10, Chapter 6]). We have

$$\|u_{xx}^n\|_{\mathcal{M}} \le \|u^n\|_{\mathcal{M}} + \|u^n - u_{xx}^n\|_{\mathcal{M}}\,. \tag{3.27}$$

Since $u^n \in L^1(\mathbb{R})$, we have $\|u^n\|_{\mathcal{M}} = \|u^n\|_{L^1}$. At the same time, due to the positivity of the p_i, the L^1-norm of u is equal to M, the total mass of the system,

$$M = \int u^n(x,t)\,dx = 2\sum_{i=1}^n p_i(t). \tag{3.28}$$

As we have seen in the previous section, M is a conserved quantity. Hence,

$$M = \|u^n\|_{L^1} = \|u_0^n\|_{L^1}.$$

Since the p_i are positive, the fact that $m^n = u^n - u_{xx}^n = \sum_{i=1}^n 2p_i\delta_{q_i}$ and (3.28) imply that

$$\|u^n - u_{xx}^n\|_{\mathcal{M}} = 2\sum_{i=1}^n p_i = M. \tag{3.29}$$

Hence, from (3.27),

$$\|u_{xx}^n\|_{\mathcal{M}} \leq 2M \ = 2\,\|u_0^n\|_{L^1}$$

and (iii) follows from (3.1b) .

The derivative u_t^n is given by

$$u_t^n = \sum_{i=1}^n \left(\dot{p}_i e^{-|x-q_i|} + p_i\dot{q}_i \operatorname{sgn}(x-q_i)e^{-|x-q_i|}\right),$$

or, after using (2.10),

$$u_t^n = \sum_{i,j=1}^n p_ip_je^{-|x-q_i|}e^{-|q_i-q_j|}\left(\operatorname{sgn}(q_i-q_j)+\operatorname{sgn}(x-q_i)\right).$$

Hence, since the p_i are all positive,

$$\begin{aligned}\|u_t^n\|_{L^2} &\leq 2\sum_{i,j=1}^n p_ip_je^{-|q_i-q_j|}\left\|e^{-|x-q_i|}\right\|_{L^2}\\ &\leq 2\sum_{i,j=1}^n p_ip_je^{-|q_i-q_j|}\\ &\leq \|u^n\|_{H^1(\mathbb{R})}^2\end{aligned}$$

and assertion (iv) follows from (i). □

To prove the existence of a converging subsequence of u^n in $C([0,T],H^1_{\mathrm{loc}}(\mathbb{R}))$ we recall the following compactness theorem adapted from Simon [18, Corollary 4].

Theorem 3.6 (Simon)**.** *Let X, B, Y be three continuously embedded Banach spaces*

$$X \subset B \subset Y$$

with the first inclusion, $X \subset B$, compact. We consider a set $\mathcal{F}$ of continuous functions mapping $[0,T]$ into X. If $\mathcal{F}$ is bounded in $L^\infty([0,T],X)$ and $\frac{\partial\mathcal{F}}{\partial t} = \left\{\frac{\partial f}{\partial t} \mid f\in\mathcal{F}\right\}$ is bounded in $L^r([0,T],Y)$ where $r>1$, then $\mathcal{F}$ is relatively compact in $C([0,T],B)$.

Proof of Theorem 3.1. Given initial data $u_0 \in H^1(\mathbb{R})$ with $u_0 - u_{0,xx} \in \mathcal{M}^+$ we know from Proposition 3.2 that there exists a sequence u_0^n satisfying condition (3.1). Furthermore, by using Lemma 2.4, we infer that there exists a sequence of multipeakons $u^n(x,t)$ such that $u^n|_{t=0} = u_0^n$. The sequence then possesses the properties stated in Lemma 3.5.

To apply Theorem 3.6, we have to determine the Banach spaces with the required properties. Let K be a compact subset of $\mathbb{R}$. We define $X = X(K)$ as the set of functions of $H^1(K)$ which have derivatives of bounded variation, that is,

$$X(K) = \left\{ v \in H^1(K) \mid v_x \in \mathrm{BV}(K) \right\}$$

endowed with the norm

$$\|v\|_{X(K)} = \|v\|_{H^1(K)} + \|v_x\|_{BV(K)} = \|v\|_{H^1(K)} + \|v_x\|_{L^\infty(K)} + \mathrm{TV}_K(v_x).$$

It follows that $X(K)$ is a Banach space. Let us prove that the injection $X(K) \subset H^1(K)$ is compact. We consider a sequence v_n which is bounded in $X(K)$. By the Rellich–Kondrachov theorem, since $\|v_n\|_{H^1(K)}$ is bounded, there exists a subsequence (that we still denote v_n) which converges to some v in $L^2(\mathbb{R})$. Since $\mathrm{TV}_K(v_{n,x})$ is bounded, Helly's theorem allow us to extract another subsequence such that

$$v_{n,x} \to w \quad \text{a.e. in } K \tag{3.30}$$

for some $w \in L^\infty(K)$. We have $\|v_{n,x}\|_{L^\infty(K)}$ bounded. From (3.30) we get, by Lebesgue's dominated convergence theorem, that $v_{n,x} \to w$ in $L^2(K)$. Using the distributional definition of a derivative, it is not hard to check that w must coincide with v_x. Therefore v_n converges to v in $H^1(K)$ and $X(K)$ is compactly embedded in $H^1(K)$.

The estimates we have derived previously imply that u^n and u_t^n are uniformly bounded in $L^\infty([0,T], X(K))$ and $L^\infty([0,T], L^2(K))$, respectively. Since $X(K) \subset H^1(K) \subset L^2(K)$ with the first inclusion compact, Simon's theorem gives us the existence of a subsequence of u^n that converges to some $u \in H^1(K)$ in $C([0,T], H^1(K))$. We consider a sequence of compact sets K_m such that $\mathbb{R} = \cup_{m\in\mathbb{N}} K_m$ and a sequence of time T_m such that $\lim_{m\to\infty} T_m = \infty$. By a diagonal argument, we can find a subsequence (that we still denote u^n) that converges to some $u \in C([0,T_m], H^1(K_m))$ in $L^\infty([0,T_m], H^1(K_m))$ for all m. Therefore u belongs to $C(\mathbb{R}, H^1_{\mathrm{loc}}(\mathbb{R}))$ and u^n converges to u in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1_{\mathrm{loc}}(\mathbb{R}))$.

It remains to prove that u is solution of the Camassa–Holm equation. This simply comes from the fact that the u^n are all weak solutions of (2.2), and since they converge to u in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1_{\mathrm{loc}}(\mathbb{R}))$, u is a weak solution of (2.2). The solutions of the Camassa–Holm equation for the class of initial data we are considering in the theorem are unique, see [8]. It implies that not only a subsequence, but the whole sequence of multipeakons converges to the solution. □

4. Numerical results

Multipeakons can be used in a numerical scheme to solve the Camassa–Holm equation with initial data satisfying $u_0 - u_{0,xx} \in \mathcal{M}^+$. The scheme consists of solving the system of ordinary differential equations (2.10) where the initial conditions are computed as in Proposition 3.2.

In the numerical experiments that follow, we solve (2.10) by using the explicit Runge–Kutta solver `ode45` for ordinary differential equation from Matlab. In the case where u_0 is sufficiently smooth, an initial multipeakon sequence can be obtained without having to solve (3.2). This is the aim of the following proposition.

Proposition 4.1. *Let u_0 be such that $u_0 - u_{0,xx}$ is a positive function in $H^1(\mathbb{R}) \cap L^1(\mathbb{R})$. We set*

$$\begin{aligned} q_{i,n} &= \frac{i}{n}, \\ p_{i,n} &= \frac{1}{2n}[u_0 - u_{0,xx}](q_{i,n}) = \frac{1}{2n} m_0(q_{i,n}). \end{aligned} \tag{4.1}$$

Then the sequence $u_0^n = \sum_{i=-n^2}^{n^2} p_{i,n} e^{-|x-q_{i,n}|}$ of multipeakons satisfies the conditions given in (3.1).

Proof. Condition (3.1c) follows directly from the definition of $p_{i,n}$ and the positivite of m_0. Let us prove (3.1a), i.e., that $u_0^n \to u_0$ in $H^1(\mathbb{R})$. It is enough to show that m_0^n tends to m_0 in H^{-1} because the mapping $v \mapsto v - v_{xx}$ is an homeomorphism from H^1 to H^{-1} (see [2, chapter 8]). For any function ϕ in $H^1(\mathbb{R})$, we have to prove that

$$\langle m_0^n, \phi \rangle = \sum_{i=-n^2}^{n^2} 2 p_{i,n} \phi(q_{i,n}) = \frac{1}{n} \sum_{i=-n^2}^{n^2} m_0(q_{i,n}) \phi(q_{i,n})$$

converges to

$$\langle m_0, \phi \rangle = \int_{\mathbb{R}} m_0(x) \phi(x) \, dx.$$

If ϕ is continuous with compact support, the above convergence simply follows from the fact that for continuous functions, the Riemann sums converge to the integral. To prove that $\langle m_0^n, \phi \rangle \to \langle m_0, \phi \rangle$ for any $\phi \in H^1(\mathbb{R})$, it is then enough to show that $\|m_0^n\|_{H^{-1}}$ is uniformly bounded. In fact, m_0^n is uniformly bounded in $\mathcal{M}$ and

$$\|m_0^n\|_{\mathcal{M}} = \frac{1}{n} \sum_{i=-n^2}^{n^2} m_0(q_{i,n}) \to \|m_0\|_{L^1}. \tag{4.2}$$

Let us prove (4.2). We have

$$\int_{\mathbb{R}} m_0(x) \, dx = \int_{-\infty}^{-n} m_0(x) \, dx + \int_{-n}^{n+\frac{1}{n}} m_0(x) \, dx + \int_{n+\frac{1}{n}}^{\infty} m_0(x) \, dx.$$

The first and the last integral tend to zero because m_0 belongs to $L^1(\mathbb{R})$. Then we have

$$\left|\int_{-n}^{n+\frac{1}{n}} m_0(x)\,dx - \frac{1}{n}\sum_{i=-n^2}^{n^2} m_0(q_{i,n})\right| \leq \sum_{i=-n^2}^{n^2}\int_{q_{i,n}}^{q_{i+1,n}} |m_0(x)-m_0(q_{i,n})|\;dx$$
$$\leq \sum_{i=-n^2}^{n^2}\int_{q_{i,n}}^{q_{i+1,n}}\int_{q_{i,n}}^{x} |m_0'(\xi)|\;d\xi dx.$$

We change the order of integration, introduce $\chi_{i,n}$ to denote the characteristic function of the interval $(q_{i,n}, q_{i+1,n})$, and get

$$\begin{aligned}
|\int_{-n}^{n+\frac{1}{n}} m_0(x)\,dx - \frac{1}{n}\sum_{i=-n^2}^{n^2} m_0(q_{i,n})| &\leq \sum_{i=-n^2}^{n^2}\int_{q_{i,n}}^{q_{i+1,n}}\int_{\xi}^{q_{i+1,n}} |m_0'(\xi)|\;dxd\xi \\
&= \int_{-\infty}^{\infty} |m_0'(\xi)| \sum_{i=-n^2}^{n^2}\chi_{i,n}(\xi)(q_{i+1,n}-\xi)\,d\xi \\
&\leq \|m_0'\|_{L^2}\Big[\int_{-\infty}^{\infty}\Big(\sum_{i=-n^2}^{n^2}\chi_{i,n}(\xi)(q_{i+1,n}-\xi)\Big)^2 d\xi\;\Big]^{1/2} \\
&\leq \|m_0\|_{H^1(\mathbb{R})}\Big[\int_{-\infty}^{\infty}\sum_{i=-n^2}^{n^2}\chi_{i,n}(\xi)(q_{i+1,n}-\xi)^2\,d\xi\;\Big]^{1/2} \\
&\leq \|m_0\|_{H^1(\mathbb{R})}\Big[\sum_{i=-n^2}^{n^2}\int_{q_{i,n}}^{q_{i+1,n}}(q_{i+1,n}-\xi)^2\,d\xi\;\Big]^{1/2} \\
&\leq \|m_0\|_{H^1(\mathbb{R})}\Big[\sum_{i=-n^2}^{n^2}\frac{1}{3n^3}\Big]^{1/2} \\
&\leq \frac{1}{\sqrt{n}}\|m_0\|_{H^1(\mathbb{R})}
\end{aligned}$$

which tends to zero. This concludes the proof of (3.1a) and condition (3.1b) follows from (4.2) since we have, see (3.29),

$$\|u_0^n\|_{L^1} = 2\sum_{i=1}^{n} p_i = \|m_0^n\|_{\mathcal{M}}\,.$$

□

We tested our algorithm with smooth traveling waves. Smooth traveling waves are solutions of the form

$$u(x,t) = f(x-ct)$$

where f is solution of the second-order ordinary differential equation

$$f_{xx} = f - \frac{\alpha}{(f-c)^2}. \tag{4.3}$$

In order to give rise to a smooth traveling wave, the constants c and α cannot be chosen arbitrarily, see [15]. Here we consider periodic smooth traveling waves. The approach, based on functions in $H^1(\mathbb{R})$, which was developed in the previous sections, can be adapted to handle solutions with periodic boundary conditions. We then have to consider periodic multipeakons which are solutions of the form

$$u(x,t) = \sum_{i=1}^{n} p_i(t) G(x, q_i(t)) \tag{4.4}$$

where G is given by

$$G(x,y) = \frac{\cosh(d(x,y) - \frac{a}{2})}{\sinh \frac{a}{2}}.$$

In the expression above, a is the period and $d(x,y) = \min\left(|x-y|, a - |x-y|\right)$ is the distance in the interval $[0,a]$, identifying the end points 0 and a of the interval. The function $G(x,y)$ can be interpreted as the periodized version of $e^{-|x-y|}$ as we have $G(x,y) = \sum_{k=-\infty}^{\infty} e^{-|x-y+ka|}$. The coefficients p_i and q_i satisfy equation (2.11) when H is replaced by the Hamiltonian

$$H_{\text{per}} = \frac{1}{2} \sum_{i,j=1}^{n} p_i p_j G(q_i, q_j).$$

For periodic functions, we have $H_{\text{per}} = 4\|u\|_{H^1((0,a))}$ for u given by (4.4). It is not hard to prove that, with the necessary amendments, Theorem 3.1 and Proposition 3.2 hold also for periodic functions in $H^1([0,a])$.

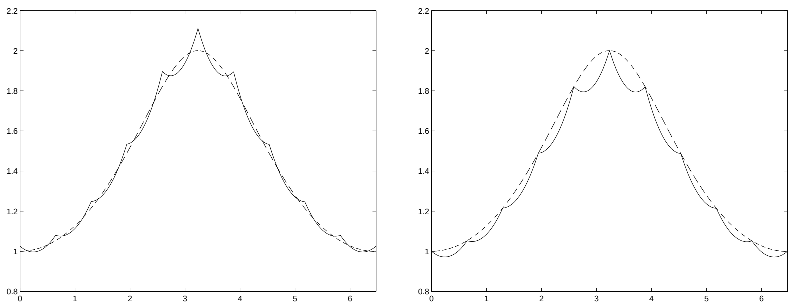


FIGURE 1. Approximation of a smooth traveling wave (dashed curve) by (4.4) with $n = 10$. On the left, the coefficients p_i are computed by using the method of Proposition 4.1 designed for smooth functions. On the right, they are computed by using the collocation method of Proposition 3.2.

TABLE 1. Convergence rate in the case of a smooth traveling wave.

Number of peakons	5	10	20	40
$\|u - u_{\text{exact}}\|_{H^1(\mathbb{R})}$ at $t = 0$	1.48	0.76	0.38	0.20
Ratio		1.95	2	1.9
$\|u - u_{\text{exact}}\|_{H^1(\mathbb{R})}$ at $t = 2$	1.31	0.68	0.34	0.17
Ratio		1.93	2	2

A high precision solution of equation (4.3) is used as a reference solution for the smooth traveling wave. We take $\alpha = c = 3$. With initial condition $f(0) = 1, f_x(0) = 0$, it gives rise to a smooth traveling wave of period $a \approx 6.4723$. In our multipeakon scheme, we approximate initial data by using (4.1) because the initial data is smooth. In Figure 1, we show the result of such approximation in the case of 10 multipeakons. In Table 1 we give the error in the H^1 norm between the computed and the exact solutions at time $t = 0$ and $t = 2$ (at $t = 2$, the wave has approximately traveled over a distance equal to one period). We can see that the computed solution converges to the exact solution at a linear rate. It is to be noted that the error does not grow in time and is apparently only due to the error which is made in approximating the initial data.

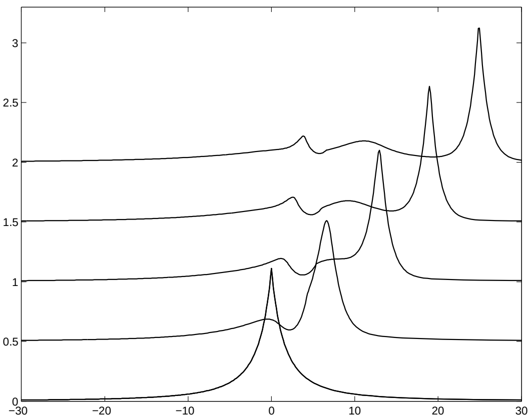


FIGURE 2. Solution with initial data $u_0(x) = 10(3 + |x|)^{-2}$ at $t = 0, 5, 10, 15, 20$ (from the bottom to the top).

Our next example deals with a initial data function u_0 which has discontinuous derivative. We take

$$u_0(x) = \frac{10}{(3 + |x|)^2}.$$

The function u_0 satisfies $u_0 - u_{0,xx} \geq 0$ and it is plotted in Figure 2. In our multipeakon scheme, we use Proposition 3.2 to set the initial sequence of multipeakons.

TABLE 2. Convergence rate for an initial data given by $u_0(x) = 10(3+|x|)^{-2}$.

Number of peakons	61	127	251	501	1001
$\Vert u - u_{\text{exact}}\Vert_{H^1(\mathbb{R})}$ at $t=0$	0.27	0.14	0.079	0.053	0.045
Ratio	–	1.93	1.77	1.49	1.18
$\Vert u - u_{\text{exact}}\Vert_{H^1(\mathbb{R})}$ at $t=10$	0.58	0.18	0.074	0.028	–
Ratio	–	3.22	2.43	1.95	–

In Figure 2 the solution is computed with very high resolution ($n = 1000$ peakons spread over the interval $[-30, 30]$) and in Table 2, the error is evaluated by taking this numerical solution as an approximation of the exact solution (except a time $t = 0$ where we can use u_0). The convergence rate at time $t = 0$ is not linear, as in the previous case. This is due to the fact that we only took peakons on the interval $[-30, 30]$. We have considered the error in $H^1([-30, 30])$, and in that case the convergence is linear. As in the case of smooth traveling waves, the error does not grow in time showing the robustness of the algorithm.

REFERENCES

[1] G. Birkhoff and G.-C. Rota, Ordinary Differential Equations, Wiley, New York, fourth edition, 1989.
[2] H. Brezis, Analyse fonctionnelle, Masson, Paris, 1983.
[3] R. Camassa, *Characteristics and the initial value problem of a completely integrable shallow water equation,* Discrete Contin. Dyn. Syst. Ser. B, **3** (2003), 115–139.
[4] R. Camassa and D. D. Holm, *An integrable shallow water equation with peaked solitons,* Phys. Rev. Lett., **71** (1993), 1661–1664.
[5] R. Camassa and D. D. Holm, and J. Hyman, *A new integrable shallow water equation,* Adv. Appl. Mech., **31** (1994), 1–33.
[6] R. Camassa, J. Huang, and L. Lee, *Integral and integrable algorithms for a nonlinear shallow-water wave equation,* Preprint Submitted (2004).
[7] R. Camassa, J. Huang, and L. Lee, *On a completely integrable numerical scheme for a nonlinear shallow-water equation,* Preprint Submitted (2004).
[8] A. Constantin and J. Escher, *Global existence and blow-up for a shallow water equation,* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), **26** (1998), 303–328.
[9] A. Constantin and L. Molinet, *Global weak solutions for a shallow water equation,* Comm. Math. Phys., **211** (2000), 45–61.
[10] L. C. Evans and R. F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton, FL, 1992.
[11] G. B. Folland, Real Analysis, Wiley, New York, second edition, 1999.
[12] A. A. Himonas and G. Misiołek, *The Cauchy problem for an integrable shallow-water equation,* Differential Integral Equations, **14** (2001), 821–831.
[13] H. Holden and X. Raynaud, *Convergence of a finite difference scheme for the Camassa–Holm equation,* SIAM J. Numer. Anal. 2006, to appear. **paper I** in this thesis.
[14] R. S. Johnson, *Camassa–Holm, Korteweg–de Vries and related models for water,* J. Fluid Mech, **455** (2002), 63–82.
[15] H. Kalisch, *Stability of solitary waves for a nonlinearly dispersive equation,* Discrete Contin. Dyn. Syst., **10** (2004), 709–717.

[16] Yi A. Li and P. J. Olver, *Well-posedness and blow-up solutions for an integrable nonlinearly dispersive model wave equation,* J. Differential Equations, **162** (2000), 27–63.

[17] G. Rodríguez-Blanco, *On the Cauchy problem for the Camassa–Holm equation,* Nonlinear Anal., **46** (2001), 309–327.

[18] J. Simon, *Compact sets in the space $L^p(0, T; B)$,* Ann. Mat. Pura Appl. (4), **146** (1987), 65–96.

Paper III

Convergence of a Spectral Projection of the Camassa-Holm equation.

H. Kalisch and X. Raynaud

To appear in *Numerical Methods for Partial Differential Equations.*

CONVERGENCE OF A SPECTRAL PROJECTION OF THE CAMASSA-HOLM EQUATION

HENRIK KALISCH AND XAVIER RAYNAUD

Abstract. A spectral semi-discretization of the Camassa-Holm equation is defined. The Fourier-Galerkin and a de-aliased Fourier-collocation method are proved to be spectrally convergent. The proof is supplemented with numerical explorations which illustrate the convergence rates and the use of the dealiasing method.

1. Introduction

In this article, consideration is given to the error analysis of a spectral projection of the periodic Camassa–Holm equation

$$u_t - u_{xxt} + \omega u_x + 3uu_x - \gamma(2u_x u_{xx} + uu_{xxx}) = 0 \tag{1.1}$$

on the interval $[0, 2\pi]$. Spectral discretizations of this equation have been in use ever since the work of Camassa and Holm [2] and Camassa, Holm and Hyman [3]. However, to the knowledge of the authors, no proof that such a discretization actually converges has appeared heretofore. Therefore, this issue is taken up here. Our method of proof is related to the work of Maday and Quarteroni on the convergence of a Fourier-Galerkin and collocation method for the Korteweg-de Vries equation [22]. While they were able to treat the unfiltered collocation approximation, we resort to proving the convergence of a de-aliased collocation projection which turns out to be equivalent to a Galerkin scheme. Before we get to the heart of the subject, a few words about the range of applicability of the equation are in order. The validity of the Camassa-Holm equation as a model for water waves in a channel of uniform width and depth has been a somewhat controversial subject. The discussion seems to have finally been settled in the recent articles of Johnson [16] and Kunze and Schneider [20]. One merit of the equation is the fact that it allows wave breaking typical of hyperbolic systems. Such wave breaking is observed in fluid flows, and under this aspect, the Camassa-Holm equation could be seen as a more suitable model than the well known Korteweg-de Vries equation, for instance. On the other hand, in the derivation of the Camassa-Holm equation, it is assumed that the solutions are more regular than breaking waves [20]. In this respect, smooth solutions are more closely related to the fluid flow problem than are irregular solutions. From

This research was supported by the BeMatA program of the Research Council of Norway.

this point of view, a spectral approximation seems a natural choice for a spatial discretization.

Another application of the Camassa-Holm equation arises when $\omega = 0$. In this case, the equation can be derived as a model equation for mechanical vibrations in a compressible elastic rod. As explained by Dai and Huo [8], the range of the parameter γ is roughly from -29.5 to 3.4. The equation has even found its place in the context of differential geometry, where it can be seen as a re-expression for geodesic flow on an infinite-dimensional Lie group [6, 13, 23].

Notwithstanding its importance as a model equation, one reason for the interest in the Camassa-Holm equation is its vast supply of novel mathematical issues, such as its integrable bi-Hamiltonian structure. This property alone has led to many interesting developments, a sample of which can be found in [2, 3, 4, 9, 10, 11], and the references contained therein. One aspect of the integrability of the equation in case $\gamma = 1$ is that the solitary-wave solutions are solitons [2, 4], similar to the solitary-wave solutions of the Korteweg-de Vries equation. However, the Camassa-Holm equation also admits solitary waves which are not smooth, but rather have a peak or even a cusp. These peaked solitary waves are well known, and owing to their soliton-like properties they have been termed peakons. In the case that $\omega = 0$ and $\gamma = 1$, they are of the form

$$u(x,t) = de^{-|x-dt|},$$

where $d \in \mathbb{R}$ is the wavespeed. For general ω and γ, a similar formula was found by one of the authors in [17]. Even more general shapes have been described in [21], where a classification of traveling-wave solutions is given. For the numerical approximation of these peaked or cusped waves, spectral methods may not be the best choice. Other methods based on finite-difference approximations have been used for instance in [1, 14, 15].

For the purpose of numerical study, it is important to have a satisfactory theory of existence of solutions, as well as uniqueness and continuous dependence with respect to the initial data. For the periodic case, an example of such well posedness results has been provided by Constantin and Escher in [5]. However, for our purposes, the available results are not quite strong enough. In particular, it appears that it is possible for solutions emanating from smooth initial data to form singularities in finite time. These singularities manifest themselves in the form of steepening up to the point where the first derivative becomes close to $-\infty$. In the context of one-dimensional water-wave theory, this may be understood as wave breaking which we have alluded to earlier. The idea of the proof of this phenomenon is to make use of a differential inequality which goes back at least to the work of Seliger [25] (see also Whitham [26]). It is clear that such singularity formation will prevent the spectral, or super-polynomial convergence usually exhibited by spectral discretizations. To circumvent this problem, we assume that the solution has at least four derivatives in the space of square-summable functions. This requirement turns out to be sufficient to obtain convergence of the spectral projection, with a convergence rate dependent upon the regularity of the solution to be approximated. Of course, this requirement also restricts the pool of possible solutions, and thus limits the applicability of our method. A spectral discretization may still be used for solutions that have lower regularity, but the convergence is then not known.

For the sake of simplicity, we will only give proofs in the case where $\omega = 0$ and $\gamma = 1$ but all the proofs extend to the general case with only small changes in the constants. To prepare the equation for the discretization, it is convenient to write (1.1) in the form

$$u_t + \tfrac{1}{2}(u^2)_x + K(u^2 + \tfrac{1}{2}(u_x)^2) = 0, \tag{1.2}$$

where K is a Fourier multiplier operator with the symbol

$$\widehat{Kf}(k) = \frac{ik}{1+k^2}\hat{f}(k),$$

which formally corresponds to

$$K = \frac{\partial_x}{1-\partial_x^2}.$$

This formulation reveals that it is natural to use a spectral Fourier discretization, as the symbol of the Green's function K is already known. In section 3, we define a Fourier-Galerkin approximation of the Camassa-Holm equation, and prove that this approximation converges if appropriate assumptions are made on the initial data and the solution. Indeed, it will transpire that for smooth solutions, the convergence is indeed spectral, i.e. super-polynomial. In section 4, a similar result will be proved for a de-aliased Fourier-collocation scheme. Finally, the last section contains some numerical computations, which illustrate the results obtained in sections 3 and 4, and which show that the de-aliased collocation scheme is preferable to an unfiltered collocation approximation, especially when approximating solutions which are not smooth.

2. Notation

In order to facilitate our study, we start by introducing some mathematical notation. Denote the inner product in $L^2(0, 2\pi)$ by

$$(f, g) = \int_0^{2\pi} f(x)\,\overline{g(x)}\,dx.$$

The Fourier coefficients $\hat{f}(k)$ of a function $f \in L^2(0, 2\pi)$ are defined by

$$\hat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx} f(x)\,dx.$$

Recall the inversion formula

$$f(x) = \sum_{k\in\mathbb{Z}} e^{ikx}\hat{f}(k),$$

and the convolution formula

$$(\hat{f} * \hat{g})(k) = \widehat{fg}(k),$$

where the convolution of two functions $\hat{f}$ and $\hat{g}$ on $\mathbb{Z}$ is formally defined by

$$(\hat{f} * \hat{g})(k) = \sum_{m+n=k} \hat{f}(m)\, \hat{g}(n).$$

Denote by $\|\cdot\|_{H^m}$ the Sobolev norm, given by

$$\|f\|_{H^m}^2 = \sum_{k\in\mathbb{Z}} (1+|k|^2)^m |\hat{f}(k)|^2.$$

The space of periodic Sobolev functions on the interval $[0, 2\pi]$ is defined as the closure of the space of smooth periodic functions with respect to the H^m-norm, and will be simply denoted by H^m. In particular, for $m = 0$, we recover the space $L^2(0, 2\pi)$ whose norm will be denoted by $\|\cdot\|_{L^2}$. The subspace of $L^2(0, 2\pi)$ spanned by the set

$$\left\{ e^{ikx} \;\middle|\; k \in \mathbb{Z},\ -\frac{N}{2} \le k \le \frac{N}{2} - 1 \right\}$$

for N even is denoted by S_N. In the following, it will always be assumed that N is even. The operator P_N denotes the orthogonal, self-adjoint, projection from L^2 onto S_N, defined by

$$P_N f(x) = \sum_{-N/2 \le k \le N/2-1} e^{ikx} \hat{f}(k).$$

For $f \in H^m$, the estimates

$$\|f - P_N f\|_{L^2} \le C_P N^{-m} \|\partial_x^m f\|_{L^2}\,, \tag{2.1}$$

$$\|f - P_N f\|_{H^n} \le C_P N^{n-m} \|\partial_x^m f\|_{L^2} \tag{2.2}$$

hold for an appropriate constant C_P and a positive integer n. For the proof of these inequalities, the reader is referred to [7].

The space of continuous functions from the interval $[0, T]$ into the space H^n is denoted by $C([0, T], H^n)$. Similarly, we also consider the space $C([0, T], S_N)$, where the topology on the finite-dimensional space S_N can be given by any norm. Finally note the inverse inequality

$$\|\partial_x^m \phi\|_{L^2} \le N^m \|\phi\|_{L^2}\,, \tag{2.3}$$

which holds for integers $m > 0$ and $\phi \in S_N$. A proof of this estimate can also be found in [7]. We will make use of the Sobolev lemma, which guarantees the existence of a constant c, such that

$$\sup_x |f(x)| \le c\,\|f\|_{H^1}\,. \tag{2.4}$$

Another standard result is that the assignment $(f, g) \mapsto fg$ is a continuous bilinear map from $H^1 \times H^{-1}$ to H^{-1}, as shown by the estimate

$$\|fg\|_{H^{-1}} \le c\,\|f\|_{H^1}\,\|g\|_{H^{-1}}\,, \tag{2.5}$$

where the same constant c has been used for simplicity.

In order to obtain a unique solution, equation (1.1) has to be supplemented by appropriate boundary and initial conditions. For the purpose of numerical approximation, the problem will be studied on a finite interval with periodic boundary conditions. The periodic initial value problem associated to equation (1.1) is

$$\begin{cases} u_t - u_{xxt} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0 & x \in [0, 2\pi]\ , t \geq 0, \\ \qquad\qquad u(0,t) = u(2\pi, t), & t \geq 0, \\ \qquad\qquad u(x,0) = u_0(x). & \end{cases} \tag{2.6}$$

In the following, it will always be assumed that a solution of this problem exists on some time interval $[0,T]$, and with a certain amount of spatial regularity. In particular, we suppose that a solution exists in the space $C([0,T], H^4)$ for some $T > 0$. With these preliminaries in place, we are set to attack the problem of defining a suitable spectral projection of (2.6) and proving the convergence of such a projection. First, the Fourier-Galerkin method is presented and a proof of convergence given. Then in section 4, a de-aliased collocation scheme will be treated.

3. The Fourier-Galerkin method.

A space-discretization of (2.6) is defined by utilizing the equivalent formulation (1.2). Thus the problem is to find a function u_N from $[0,T]$ to S_N which satisfies

$$\begin{cases} \left(\partial_t u_N + \frac{1}{2}\partial_x (u_N)^2 + K\left(u_N^2 + \frac{1}{2}(\partial_x u_N)^2\right), \phi\right) = 0, & t \in [0,T], \\ \qquad\qquad u_N(0) = P_N u_0, & \end{cases} \tag{3.1}$$

for all $\phi \in S_N$. Since for each t, $u_N(\cdot, t) \in S_N$, u_N has the form

$$u_N(x,t) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{u}_N(k,t) e^{ikx},$$

where $\hat{u}_N(k,t)$ are the Fourier coefficients of $u_N(\cdot, t)$.

Taking $\phi = e^{ikx}$ for $-N/2 \leq k \leq N/2 - 1$ in (3.1) yields the following system of equations for the Fourier coefficients of u_N.

$$\begin{cases} \dfrac{d}{dt}\hat{u}_N(k,t) = -\dfrac{1}{2} ik (\hat{u}_N * \hat{u}_N)(k,t) - \dfrac{ik}{1+k^2}\Big[(\hat{u}_N * \hat{u}_N)(k,t) \\ \qquad\qquad + \ \frac{1}{2}((ik\hat{u}_N) * (ik\hat{u}_N))(k,t)\Big], \\ \hat{u}_N(k,0) = \hat{u}_0(k), \end{cases} \tag{3.2}$$

for $-N/2 \leq k < N/2 - 1$..

The short-time existence of a maximal solution of (3.2) is proved using the contraction mapping principle, and the solution is unique on its maximal interval of definition, $[0, t_N^m)$, where t_N^m is possibly equal to T. Since the argument is standard, the proof is omitted here. The main result of this paper is the fact that the Galerkin approximation u_N converges to the exact solution u when u is smooth enough. This is stated in the next theorem.

Theorem 3.1. *Suppose that a solution u of the Camassa-Holm equation* (2.6) *exists in the space $C([0,T],H^m)$ for $m \geq 4$, and for some time $T > 0$. Then for N large enough, there exists a unique solution u_N of the finite-dimensional problem* (3.1)*. Moreover, there exists a constant λ such that*

$$\sup_{t\in[0,T]} \|u(\cdot,t) - u_N(\cdot,t)\|_{L^2} \leq \lambda N^{1-m}.$$

Before the proof is given, note that the assumptions of the theorem encompass the existence of a constant κ, such that

$$\sup_{t\in[0,T]} \|u(\cdot,t)\|_{H^m} \leq \kappa.$$

In particular, it follows then that there is another constant Λ, such that

$$\sup_{t\in[0,T]} \|u(\cdot,t)\|_{H^2} \leq \Lambda.$$

The main ingredient in the proof of the theorem is a local error estimate which will be established by the following lemma.

Lemma 3.2. *Suppose that a solution u_N of* (3.1) *exists on the time interval $[0,t_N^*]$, and that $\sup_{t\in[0,t_N^*]} \|u_N(\cdot,t)\|_{H^2} \leq 2\Lambda$. Then the error estimate*

$$\sup_{t\in[0,t_N^*]} \|u(\cdot,t) - u_N(\cdot,t)\|_{L^2} \leq \lambda N^{1-m} \tag{3.3}$$

holds for some constant λ which only depends on T, Λ and κ.

Proof. Let $h = P_N u - u_N$. We apply P_N to both sides of (1.2) and, since P_N commutes with derivation, we obtain

$$\partial_t P_N u + \tfrac{1}{2} P_N \partial_x (u^2) + K P_N (u^2 + \tfrac{1}{2} u_x^2) = 0.$$

We multiply this equation by h, integrate over $[0,2\pi]$ and subtract the result from (3.1) where we have used h, which belongs to S_N, as a test function. We get

$$\begin{aligned}(h_t,h) = -\frac{1}{2}\left(P_N\partial_x u^2 - \partial_x u_N^2, h\right) - \left(KP_N\left[u^2 + \frac{1}{2}(\partial_x u)^2\right], h\right) \\ + \left(K\left[u_N^2 + \frac{1}{2}(\partial_x u_N^2)^2\right], h\right)\end{aligned}$$

Using the fact that P_N is self-adjoint on L^2, and $h \in S_N$, this may be rewritten as

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|h\|_{L^2}^2 = -\frac{1}{2}\left(\partial_x u^2 - \partial_x u_N^2, h\right) - \left(K\left[u^2 - u_N^2\right], h\right) \\ - \frac{1}{2}\left(K\left[(\partial_x u)^2 - (\partial_x u_N)^2\right], h\right).\end{aligned} \tag{3.4}$$

Let's estimate the three terms on the right-hand side of (3.4) in the time interval $[0, t_N^*]$ where the H^2-norm of u_N is bounded by 2Λ. We have

$$\begin{aligned}
\big(\partial_x u^2 - \partial_x u_N^2, h\big) &= \big(\partial_x \left\{(u+u_N)(u-u_N)\right\}, h\big) \\
&= \big(\partial_x (u+u_N)\,(u-u_N), h\big) + \big((u+u_N)\,\partial_x (u-u_N), h\big) \\
&= \big(\partial_x (u+u_N)\,(u-u_N), h\big) \\
&\quad + \big((u+u_N)\partial_x (u-P_N u), h\big) \\
&\quad + \big((u+u_N)\partial_x (P_N u-u_N), h\big).
\end{aligned}$$

Consequently, there appears the estimate

$$\begin{aligned}
\left|\big(\partial_x u^2 - \partial_x u_N^2, h\big)\right| &\leq \sup_x |\partial_x (u+u_N)| \, \|u-u_N\|_{L^2} \, \|h\|_{L^2} \\
&\quad + \sup_x |u+u_N| \, \|\partial_x (u-P_N u)\|_{L^2} \, \|h\|_{L^2} \\
&\quad + \left| \int_0^{2\pi} (u+u_N) h_x h \, dx \right| \\
&\leq c \, \|u+u_N\|_{H^2} \left(\|u-P_N u\|_{L^2} + \|P_N u - u_N\|_{L^2} \right) \|h\|_{L^2} \\
&\quad + c \, \|u+u_N\|_{H^1} \, \|u-P_N u\|_{H^1} \, \|h\|_{L^2} \\
&\quad + \frac{1}{2} \int_0^{2\pi} h^2 |\partial_x (u+u_N)| \, dx \\
&\leq 3c\Lambda \left(C_P N^{-m} \|u\|_{H^m} + \|h\|_{L^2} \right) \|h\|_{L^2} \\
&\quad + 3c\Lambda C_P N^{1-m} \, \|u\|_{H^m} \, \|h\|_{L^2} \\
&\quad + \frac{1}{2} \sup_x |\partial_x (u+u_N)| \int_0^{2\pi} h^2 \, dx.
\end{aligned}$$

Noting that the last integral is bounded by $\frac{1}{2} 3c\Lambda \|h\|_{L^2}^2$, there appears the estimate

$$\left|\big(\partial_x u^2 - \partial_x u_N^2, h\big)\right| \leq 3c\Lambda \|h\|_{L^2} \left(\frac{3}{2} \|h\|_{L^2} + C_P \|u\|_{H^m} \left(N^{-m} + N^{1-m}\right) \right). \tag{3.5}$$

We turn to the second term in (3.4). The operator K is a continuous operator from H^{-1} to L^2. Therefore, after using the Cauchy-Schwartz inequality and (2.5), there appears

$$\begin{aligned}
\big(K(u^2 - u_N^2), h\big) &\leq \left\|u^2 - u_N^2\right\|_{H^{-1}} \|h\|_{L^2} \\
&\leq \|(u-u_N)(u+u_N)\|_{H^{-1}} \|h\|_{L^2} \\
&\leq c \, \|u+u_N\|_{H^1} \, \|u-u_N\|_{H^{-1}} \, \|h\|_{L^2} .
\end{aligned} \tag{3.6}$$

Then, since

$$\begin{aligned}\|u-u_N\|_{H^{-1}} &\leq \|u-u_N\|_{L^2}\\ &\leq \|u-P_N u\|_{L^2}+\|h\|_{L^2}\\ &\leq C_P N^{-m}\|u\|_{H^m}+\|h\|_{L^2},\end{aligned}$$

and $\|u+u_N\|_{H^1}$ is bounded (recall that the estimates are established on $[0,t_N^*]$ where the H^2-norm of u_N is bounded by 2Λ), we get from (3.6)

$$\big(K(u^2-u_N^2),h\big) \leq 3c\Lambda\|h\|_{L^2}\left(C_P N^{-m}\|u\|_{H^m}+\|h\|_{L^2}\right). \tag{3.7}$$

Similarly for the remaining term in (3.4) we have

$$\begin{aligned}\big(K((\partial_x u)^2-(\partial_x u_N)^2),h\big) &\leq \|h\|_{L^2}\left\|(\partial_x u)^2-(\partial_x u_N)^2\right\|_{H^{-1}}\\ &\leq c\|\partial_x(u+u_N)\|_{H^1}\|\partial_x(u-u_N)\|_{H^{-1}}\|h\|_{L^2}\\ &\leq 3c\Lambda\|h\|_{L^2}\left(C_P\|u\|_{H^m}N^{-m}+\|h\|_{L^2}\right). \end{aligned}\tag{3.8}$$

Gathering the estimates (3.5), (3.7) and (3.8), it transpires that

$$\frac{d}{dt}\|h\|_{L^2} \leq \frac{27}{4}c\Lambda\|h\|_{L^2}+\frac{15}{2}c\Lambda C_P\kappa N^{1-m}.$$

Consequently, Gronwall's inequality gives

$$\sup_{t\in[0,t_N^*]}\|h(\cdot,t)\|_{L^2} \leq \lambda N^{1-m} \tag{3.9}$$

for an appropriate constant λ which depends on T, Λ and κ. After decomposing $u-u_N$ as the sum $u-P_N u+h$ and using (2.1) and the triangle inequality, (3.9) yields

$$\sup_{t\in[0,t_N^*]}\|u(\cdot,t)-u_N(\cdot,t)\|_{L^2} \leq \lambda N^{1-m}$$

for another constant λ which again only depends on T, Λ and κ. □

Lemma 3.3. *Suppose that a solution u_N of* (3.1) *exists on the time interval $[0,t_N^*]$, and that $\sup_{t\in[0,t_N^*]}\|u_N(\cdot,t)\|_{H^2}\leq 2\Lambda$. Then the error estimate*

$$\sup_{t\in[0,t_N^*]}\|u(\cdot,t)-u_N(\cdot,t)\|_{H^2} \leq \lambda N^{3-m} \tag{3.10}$$

holds for some constant λ which only depends on T, Λ and κ.

The proof of this lemma follows from (3.9) after application of the triangle inequality and the inverse inequality (2.3).

Proof of Theorem 3.1. We want to extend the estimate (3.3) to the time interval $[0,T]$. Note that the time t_N^* appearing in Lemma 3.2 has so far been unspecified. We now define t_N^* by

$$t_N^* = \sup\left\{t\in[0,T] \mid \text{ for all } t'\leq t, \|u_N(\cdot,t')\|_{H^2}\leq 2\Lambda\right\}. \tag{3.11}$$

Thus the time t_N^* corresponds to the largest time in $[0,T]$ for which the H^2-norm of u_N is uniformly bounded by 2Λ. Since $\|u_N(\cdot,0)\|_{H^2}=\|P_N u(\cdot,0)\|_{H^2}$, we have

$$\|u_N(\cdot,0)\|_{H^2} \leq \|u(\cdot,0)\|_{H^2} \leq \Lambda.$$

Hence, $t_N^* > 0$ for all N. Note that t_N^* is necessarily smaller than the maximum time of existence t_N^m. On the other hand, we are going to prove that there exists N_* such that

$$t_N^* = T \quad \text{for all } N \geq N_*, \tag{3.12}$$

and therefore the supremum in (3.3) holds on $[0,T]$. By definition (3.11), we either have $t_N^* = T$ or $t_N^* < T$ and in this case, since $\|u_N(t)\|_{H^2}$ is a continuous function in time, $\|u_N(t_N^*)\|_{H^2} = 2\Lambda$. Suppose that $t_N^* < T$. Then using the triangle inequality yields

$$\begin{aligned}
2\Lambda &= \|u_N(\cdot, t_N^*)\|_{H^2} \\
&\leq \|(u_N(\cdot, t_N^*) - u(\cdot, t)\|_{H^2} + \sup_{t\in[0,T]} \|u(\cdot, t)\|_{H^2} \\
&= \|(u_N(\cdot, t_N^*) - u(\cdot, t_N^*)\|_{H^2} + \Lambda,
\end{aligned}$$

by the definition of Λ. Hence,

$$\Lambda \leq \|(u_N(\cdot, t_N^*) - u(\cdot, t_N^*)\|_{H^2}\,.$$

By Lemma 3.3, it follows that

$$\Lambda \leq \lambda N^{3-m}$$

or

$$N \leq \left(\frac{\lambda}{\Lambda}\right)^{\frac{1}{m-3}}.$$

In conclusion, for $N^* > \left(\frac{\lambda}{\Lambda}\right)^{\frac{1}{m-3}}$, we cannot have $t_N^* < T$ and the claim (3.12) holds. It follows that for $N \geq N_*$ the solution u_N of (3.2) is defined on $[0,T]$ because, as we noticed earlier, $t_N^* < t_N^m$ and, from (3.3), we get

$$\sup_{t\in[0,T]} \|u(\cdot, t) - u_N(\cdot, t)\|_{L^2} \leq \lambda N^{1-m}.$$

The following corollary is immediate from the estimate (3.9) and the inequalities (2.2) and (2.3).

Corollary 3.4. *Suppose that a solution u of the Camassa-Holm equation* (2.6) *exists in the space $C([0,T], H^m)$ for $m \geq 4$, and for some time $T > 0$. Then for N large enough, there exists a unique solution u_N of the finite-dimensional problem* (3.1)*. Moreover, there exists a constant λ such that*

$$\sup_{t\in[0,T]} \|u(\cdot, t) - u_N(\cdot, t)\|_{H^2} \leq \lambda N^{3-m}.$$

4. The Fourier-collocation method

The Galerkin method is not very attractive from a computational point of view, because the computation of the convolution sums in (3.2) is very expensive. If the convolution is computed by means of the Fast Fourier Transform (FFT), the computational time is minimized, but an additional error known as *aliasing* is introduced. This means that high wavenumbers are projected back into low wavenumber modes, causing spurious oscillations. We refer to [7, 12] for more details about the aliasing phenomenon. Methods that use the FFT are often called *collocation* methods since they are sometimes algebraically equivalent to a collocation scheme in the case of a Fourier basis. The problem of aliasing can be somewhat alleviated if enough modes are used to resolve all frequencies. However, this remedy is mostly applicable to the study of one-dimensional problems, and it supposes that the amplitudes decay in a reasonable fashion. In the case of the Camassa-Holm equation, one may want to use the spectral discretization to study the peakon solutions mentioned in the introduction. In the case that $\omega = 0$ and $\gamma = 1$, the representation

$$u(x,t) = de^{-|x-dt|},$$

reveals that u has an elementary Fourier transform, given for instance at time $t = 0$ by

$$\hat{u}(k,0) = d\sqrt{\frac{2}{\pi}}\frac{1}{1+k^2}. \tag{4.1}$$

As this expression shows, the Fourier amplitudes decay only quadratically, and therefore, it will be nearly impossible to avoid aliasing, even when using a large number of modes. In light of this problem, we have chosen to treat the case of a de-aliased scheme. In fact, it will be shown that the dealiasing we choose yields a scheme which is equivalent to the Galerkin scheme treated in the previous section.

The collocation operator in S_N denoted I_N is defined as follows. Let the collocation points be $x_j = \frac{2\pi j}{N}$ for $j = 0, 1, ..., N-1$. Then, given a continuous and periodic function f, $I_N f$ is the unique element in S_N, such that

$$I_N f(x_j) = f(x_j),$$

for $j = 0, 1, ..., N-1$. $I_N f$ is also called the N-th trigonometric interpolant of f. When restricted to S_N, the collocation operator reduces to the identity operator, as highlighted by the identity

$$I_N \phi = \phi \quad \text{for all } \phi \in S_N. \tag{4.2}$$

It has been proved in [19, 24] that when $f \in H^m$ with $m \geq 1$, there exists a constant C_I, such that

$$\|f - I_N f\|_{L^2} \leq C_I N^{-m} \|\partial_x^m f\|_{L^2}, \tag{4.3}$$

and more generally

$$\|f - I_N f\|_{H^n} \leq C_I N^{n-m} \|\partial_x^m f\|_{L^2}. \tag{4.4}$$

The collocation approximation to (2.6) is given by a function u_N from $[0,T]$ to S_N, such that

$$\begin{cases} \partial_t u_N + \frac{1}{2}\partial_x I_N[u_N^2] + K I_N\left[u_N^2 + \frac{1}{2}(\partial_x u_N)^2\right] = 0, & t \in [0,T], \\ & u_N(0) = I_N u_0. \end{cases} \tag{4.5}$$

Note that (4.5) reduces to (3.1) if the interpolation operator I_N is replaced by the projection operator P_N. In order to apply the FFT, we take the *discrete Fourier transform*, denoted here by $\mathcal{F}_N$. We again refer to [7, 12] for more details on this algorithm and its properties. The discrete Fourier transform $\mathcal{F}_N$ of a continuous function u is the vector in $\mathbb{C}^N$ defined as

$$\mathcal{F}_N(u)(k) = \widehat{I_N(u)}(k), \text{ for } -N/2 \le k < N/2 - 1. \tag{4.6}$$

Applying the discrete Fourier transform to each term in (4.5), and using the definition (4.6) of $\mathcal{F}_N$, there appears the equation

$$\left(\partial_t \mathcal{F}_N u_N + \frac{ik}{2}\mathcal{F}_N(u_N^2) + \frac{ik}{1+k^2}\mathcal{F}_N\left[u_N^2 + \tfrac{1}{2}(\partial_x u_N)^2\right]\right)(k) = 0. \tag{4.7}$$

Let $\tilde{u}_N \in \mathbb{C}^N$ denote $\mathcal{F}_N(u_N)$. The solutions of (4.5), or equivalently of (4.7), are obtained by solving the following system of ordinary differential equations:

$$\begin{cases} \dfrac{d}{dt}\tilde{u}_N(k) + \dfrac{ik}{2}\mathcal{F}_N((\mathcal{F}_N^{-1}\tilde{u}_N)^2)(k) + \dfrac{ik}{1+k^2}\mathcal{F}_N\big[(\mathcal{F}_N^{-1}\tilde{u}_N)^2 \\ \qquad\qquad\qquad\qquad\qquad\qquad + \tfrac{1}{2}(\mathcal{F}_N^{-1}(ik\tilde{u}_N))^2\big](k) = 0, \\ \tilde{u}_N(k,0) = \mathcal{F}_N u(k,0) \end{cases}$$

for $-N/2 \le k < N/2 - 1$. This method is appealing because of the efficiency of the FFT which allows us to rapidly compute the discrete Fourier transform and its inverse. It has to be compared to (3.2) where the computation of the convolution is extremely expensive. Nevertheless the FFT introduces errors due to aliasing. In order to avoid aliasing, we apply the well known 2/3-rule. Thus we consider instead the following initial value problem: Find $u_N \in C([0,T], S_N)$ such that

$$\begin{cases} \partial_t u_N + \tfrac{1}{2}\partial_x I_N[(P_{\frac{2N}{3}}u_N)^2] + K I_N\left[(P_{\frac{2N}{3}}u_N)^2 + \tfrac{1}{2}(\partial_x(P_{\frac{2N}{3}}u_N))^2\right] = 0, \\ u_N(0) = I_N u_0. \end{cases} \tag{4.8}$$

The corresponding system of ordinary differential equation satisfied by $\tilde{u}_N = \mathcal{F}_N u_N$ is

$$\begin{cases} \dfrac{d}{dt}\tilde{u}_N(k) + \dfrac{ik}{2}\mathcal{F}_N((P_{\frac{2N}{3}}\mathcal{F}_N^{-1}\tilde{u}_N)^2)(k) \\ \qquad + \dfrac{ik}{1+k^2}\mathcal{F}_N\left[(P_{\frac{2N}{3}}\mathcal{F}_N^{-1}\tilde{u}_N)^2 + \tfrac{1}{2}(P_{\frac{2N}{3}}\mathcal{F}_N^{-1}(ik\tilde{u}_N))^2\right](k) = 0, \\ \tilde{u}_N(k,0) = \mathcal{F}_N u(k,0) \end{cases} \tag{4.9}$$

which again can be solved efficiently by the use of the FFT. When implementing (4.9) numerically, it is important to note that $P_{\frac{2N}{3}}\mathcal{F}_N^{-1}\tilde{u}_N$ (and similarly $P_{\frac{2N}{3}}\mathcal{F}_N^{-1}(ik\tilde{u}_N)$) can be rewritten according to

$$P_{\frac{2N}{3}}\mathcal{F}_N^{-1}\tilde{u}_N = \mathcal{F}_N^{-1}\tilde{u}_N^c,$$

where $\tilde{u}_N^c$ is obtained by cutting off the frequencies higher than M, i.e.

$$\tilde{u}_N^c(k) = \begin{cases} \tilde{u}_N(k) & \text{if } \dfrac{2N}{3} - 1 \le k \le \dfrac{2N}{3}, \\ 0 & \text{otherwise,} \end{cases}$$

thus making these quantities easy to compute. The use of the projection $P_{\frac{2N}{3}}$ is justified by the following identity which is derived in [7],

$$P_M I_N \left(P_M f \cdot P_M g\right) = P_M \left(P_M f \cdot P_M g\right), \tag{4.10}$$

and which holds for any continuous functions f, g and any $M \leq \frac{2N}{3}$. The identity (4.10) essentially means that the interpolation operator I_N, which generally introduces aliasing, becomes harmless when we apply an M-filter, that is when we cut off the frequencies higher than M.

In order to prove convergence of the scheme, we introduce $v_N = P_{\frac{2N}{3}} u_N$ and $h = v_N - P_{\frac{2N}{3}} u$. Note that $v_N, \partial_x v_N, h \in S_{\frac{2N}{3}}$. After applying $P_{\frac{2N}{3}}$ in (4.8) and taking the scalar product with h, we get

$$(\partial_t v_N, h) + \frac{1}{2}\Big(P_{\frac{2N}{3}} \partial_x I_N(v_N^2), h\Big) + \Big(P_{\frac{2N}{3}} K I_N\left[(v_N)^2 + \tfrac{1}{2}(\partial_x v_N)^2\right], h\Big) = 0. \tag{4.11}$$

The projection operator $P_{\frac{2N}{3}}$ commutes with ∂_x and K. Using (4.10) with $M = \frac{2N}{3}$, we get $P_{\frac{2N}{3}} I_N(v_N^2) = P_{\frac{2N}{3}}(v_N^2)$ and $P_{\frac{2N}{3}} I_N(\partial_x v_N^2) = P_{\frac{2N}{3}}(\partial_x v_N^2)$ since $v_N, \partial_x v_N \in S_{\frac{2N}{3}}$. Hence, from (4.11), we get

$$(\partial_t v_N, h) + \frac{1}{2}\Big(P_{\frac{2N}{3}} \partial_x (v_N^2), h\Big) + \Big(P_{\frac{2N}{3}} K\left[(v_N)^2 + \tfrac{1}{2}(\partial_x v_N)^2\right], h\Big) = 0.$$

Then using the fact that $P_{\frac{2N}{3}}$ is self-adjoint and $P_{\frac{2N}{3}} h = h$, there obtains

$$(\partial_t v_N, h) + \frac{1}{2}\left(\partial_x (v_N^2), h\right) + \left(K\left[(v_N)^2 + \tfrac{1}{2}(\partial_x v_N)^2\right], h\right) = 0. \tag{4.12}$$

After applying $P_{\frac{2N}{3}}$ to (1.2) and taking the scalar product with h, we get

$$\Big(P_{\frac{2N}{3}} \partial_t u, h\Big) + \frac{1}{2}\left(\partial_x(u^2), h\right) + \left(K\left[u^2 + \tfrac{1}{2}(\partial_x u)^2\right], h\right) = 0, \tag{4.13}$$

where we have used again the fact that $P_{\frac{2N}{3}}$ is self-adjoint and $P_{\frac{2N}{3}} h = h$. Subtracting (4.13) from (4.12), we obtain

$$\begin{aligned}(\partial_t h, h) + \tfrac{1}{2}\left(\partial_x(v_N^2) - \partial_x u^2, h\right) + \left(K\left[(v_N)^2 - u^2\right], h\right)& \\ + \tfrac{1}{2}\left(K\left[(\partial_x v_N)^2 - (\partial_x u)^2\right], h\right) &= 0. \end{aligned} \tag{4.14}$$

We now proceed exactly in the same way as for the Galerkin method in the previous section. After introducing

$$t_N^* = \sup\left\{t \in [0,T] \mid \text{ for all } t' \leq t, \|v_N(t',\cdot)\|_{H^2} \leq 2\Lambda\right\}$$

is appears that the estimate

$$\sup_{t \in [0, t_N^*]} \|v_N - v\|_{L^2} \leq \lambda \left(\frac{2N}{3}\right)^{1-m} \tag{4.15}$$

holds for some λ depending only on T, Λ and κ. The factor $2N/3$ in (4.15) comes from the fact that we used the projection on $S_{\frac{2N}{3}}$ instead of the projection on S_N that was used to derive (3.3). As in the previous section, we can prove that for N large enough the inequality (4.15) holds when taking the supremum over all t in $[0,T]$. Thus we are led to the following theorem.

Theorem 4.1. *Suppose that a solution u of the Camassa-Holm equation* (1.1) *exists in the space $C([0,T],H^m)$ for $m \geq 4$ and for some time $T > 0$. Then for N large enough, there exists a unique solution u_N to the finite-dimensional problem* (4.8)*. Moreover, there exists a constant λ such that,*

$$\sup_{t\in[0,T]} \|u(\cdot,t) - v_N(\cdot,t)\|_{L^2} \leq \lambda \left(\frac{2N}{3}\right)^{1-m},$$

where $v_N = P_{\frac{2N}{3}} u_N$.

5. Numerical experiments

It appears that spectral discretizations have been widely used to study mathematical properties of the Camassa-Holm equation. In particular, the interaction of two or more peakon solution has been a topic of intense interest. Here, we restrict ourselves to the computation of single traveling waves in order to validate the results of the previous sections. Traveling waves have the form

$$u(x,t) = \phi(x - dt),$$

were ϕ is either known exactly, or can be approximated, and where d is the wave speed. On the real line $\mathbb{R}$, the so-called peakons are well known. In the case that $\omega = 0$ and $\gamma = 1$, they are given by

$$\phi(x) = de^{-|x|}. \tag{5.1}$$

Since we consider periodic boundary conditions, it is convenient to know that there are also periodic peaked traveling waves. These are given by

$$\phi(x) = d\frac{\cosh(\frac{1}{2} - x)}{\cosh(\frac{1}{2})} \tag{5.2}$$

on the interval $[0,1]$, and are periodic with period 1.

The Camassa-Holm equation (1.1) also admits smooth periodic traveling-wave solutions. These are not known in closed form, but they can be approximated. In this case, ϕ is given implicitly by

$$|x - x_0| = \int_{\phi_0}^{\phi} \frac{\sqrt{d-y}}{\sqrt{(M-y)(y-m)(y-z)}} dy, \tag{5.3}$$

where $\phi(x_0) = \phi_0$, and $z = d - M - m$. If $z < m < M < d$, then ϕ is a smooth function with $m = \min_{x\in\mathbb{R}} \varphi(x)$ and $M = \max_{x\in\mathbb{R}} \varphi(x)$. Once this integral is evaluated for a sufficient number of values, the function is inverted, and an approximation is found via a spline interpolation. This procedure for finding smooth traveling waves is explained in more detail in [18, 21].

For the purpose of numerical integration, the Fourier discretization is supplemented with the well known explicit four-stage Runge-Kutta scheme. Since the equation is only mildly stiff, an explicit method appears to be more advantageous than an implicit method. The scheme is explained as follows. If the wave profile $v_N(\cdot, t_i)$ is known at a particular time t_i, the four-stage Runge-Kutta method consists of letting

$$\begin{aligned} V_1 &= \hat{v}_N(\cdot, t_i), & \Gamma_1 &= F(V_1), \\ V_2 &= V_1 + \frac{\Delta t}{2}\Gamma_1, & \Gamma_2 &= F(V_2), \\ V_3 &= V_2 + \frac{\Delta t}{2}\Gamma_2, & \Gamma_3 &= F(V_3), \\ V_4 &= V_3 + \Delta t \Gamma_3, & \Gamma_4 &= F(V_4). \end{aligned}$$

Finally, these functions are combined to compute $v_N(\cdot, t_{i+1}) = v_N(\cdot, t_i + \Delta t)$, according to

$$\hat{v}_N(\cdot, t_i + \Delta t) = \hat{v}_N(\cdot, t_i) + \frac{\Delta t}{6}\left(\Gamma_1 + 2\Gamma_2 + 2\Gamma_3 + \Gamma_4\right).$$

This scheme is formally fourth-order convergent, meaning that if the time step Δt is halved, the error should decrease by a factor of 16.

The traveling-wave solutions can be used to test the numerical algorithm, because their time evolution is simply given by translation. If $v_N(\cdot, T)$ is the result of a numerical computation with initial data $\phi(x)$, it can be compared with the translated function $\phi(x - dT)$. In this way, the error produced by the discretization can be calculated. In particular, the smooth traveling wave shown in Figure 1 can be used to exhibit the convergence rate of the numerical approximation. In Table 1, it can be seen that the 4-th order convergence of the temporal integration scheme is approximately achieved. The spectral convergence of the spatial discretization is visible in Table 2. Note however, that the error is limited below because the smooth traveling waves are not known in closed form.

For the Camassa-Holm equation, the peakon solutions are of special importance. For these, we cannot expect spectral convergence. However, as shown in Table 3, the spectral discretization converges nevertheless, albeit at a lower rate. The Fourier representation (4.1) reveals that the peakon is in the Sobolev space $H^s(\mathbb{R})$ for any $s < \frac{3}{2}$. Thus it appears that the convergence at rate of N^{1-m} is probably not an optimal result. We hasten to mention however that this result as it stands is not applicable to peakons, because the proof makes use of H^4-regularity. In order to analyze the advantage of the dealiasing scheme, we compared our calculations with computations performed using an unfiltered scheme. For both the smooth approximate traveling waves, and the peaked waves, the differences in convergence were minute on small time scales. In fact the observed rate of convergence was almost exactly the same. However, as shown in Figure 3, the dealiasing did reduce some spurious oscillations in the approximations. The advantage of the dealiasing became apparent when integrating over intermediate time scales. The periodic peaked traveling waves then suffered serious destabilization when computed with an aliased scheme. As shown in Figure 4, this problem was completely avoided by the use of an appropriate filter. This shows in fact that it is preferable to perform de-aliased computations, especially for solutions with low regularity.

FIGURE 1. A smooth periodic traveling wave. The solid curve shows the initial data, while the dashed-dotted curve shows the computed solution at $T = 2.255$.

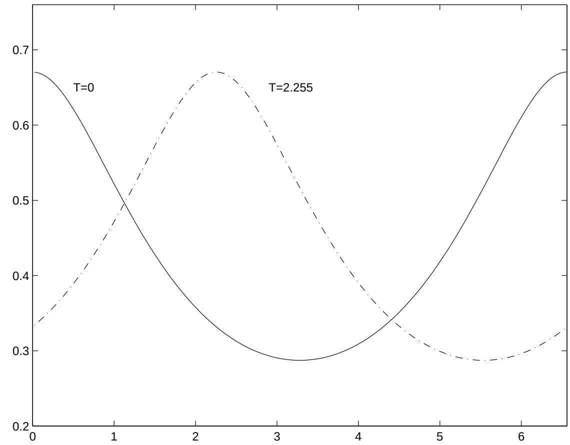


TABLE 1. Temporal discretization error for a smooth periodic traveling wave over the time domain $[0, 2.255]$. The number of grid points is $N = 256$.

Δt	L^2-Error	Ratio
0.036	1.50e-07	
0.018	9.30e-09	16.08
0.009	5.58e-10	16.66
0.0045	1.82e-11	30.74
0.00225	2.52e-11	7.20

TABLE 2. Spatial discretization error for a smooth periodic traveling wave over the time domain $[0, 3.28]$. The time step is $\Delta t = 0.0002$.

N	L^2-Error	Ratio
4	8.67e-02	
8	8.12e-03	10.67
16	1.11e-04	73.14
32	7.98e-08	1391.0
64	1.49e-11	5354.1
128	1.49e-11	0.9945

FIGURE 2. A peakon. The solid curve shows the initial data, while the dashed-dotted curve shows the computed solution at $T = 3.2$. The size of the domain is $L = 50$.

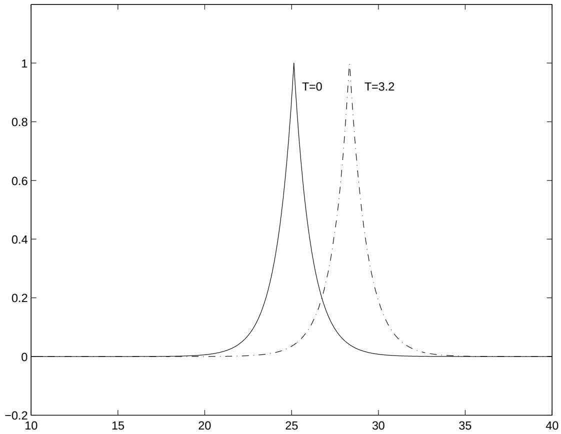


TABLE 3. Discretization error for a peaked traveling wave on the real line over the time domain $[0, 3.2]$. The time step is $\Delta t = 0.0002$, and the size of the spatial domain is 50.

N	L^2-Error	Ratio
512	7.26e-02	
1024	3.29e-02	2.20
2048	1.45e-02	2.27
4096	6.35e-03	2.28
8192	2.81e-03	2.26
16384	1.28e-03	2.18
32768	6.20e-04	2.06

TABLE 4. Discretization error for a peaked periodic traveling wave over the time domain $[0, 3.2]$. The time step is $\Delta t = 0.003$.

N	L^2-Error	Ratio
32	8.89e-01	
64	3.95e-01	2.25
128	1.86e-01	2.12
512	9.36e-02	1.98
1024	4.66e-02	2.01
2048	2.31e-02	2.02
4096	1.15e-02	2.01

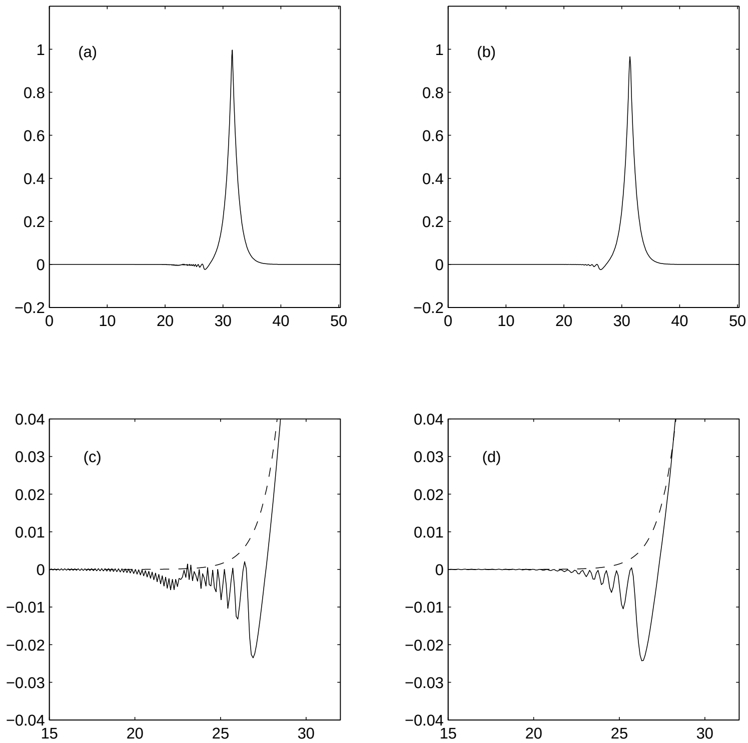


FIGURE 3. A peakon, computed with and without dealiasing. The number of grid points is $N = 512$. The left column shows the aliased calculations. (a) $T = 6.4$, (c) $T = 6.4$, close-up. The right column shows the de-aliased calculations. (b) $T = 6.4$, (d) $T = 6.4$, close-up. The dashed curve shows the exact peakon, translated by an appropriate amount according to the time $T = 6.4$ and the speed $d = 1$. It appears that the de-aliased computation shown on the right is more advantageous.

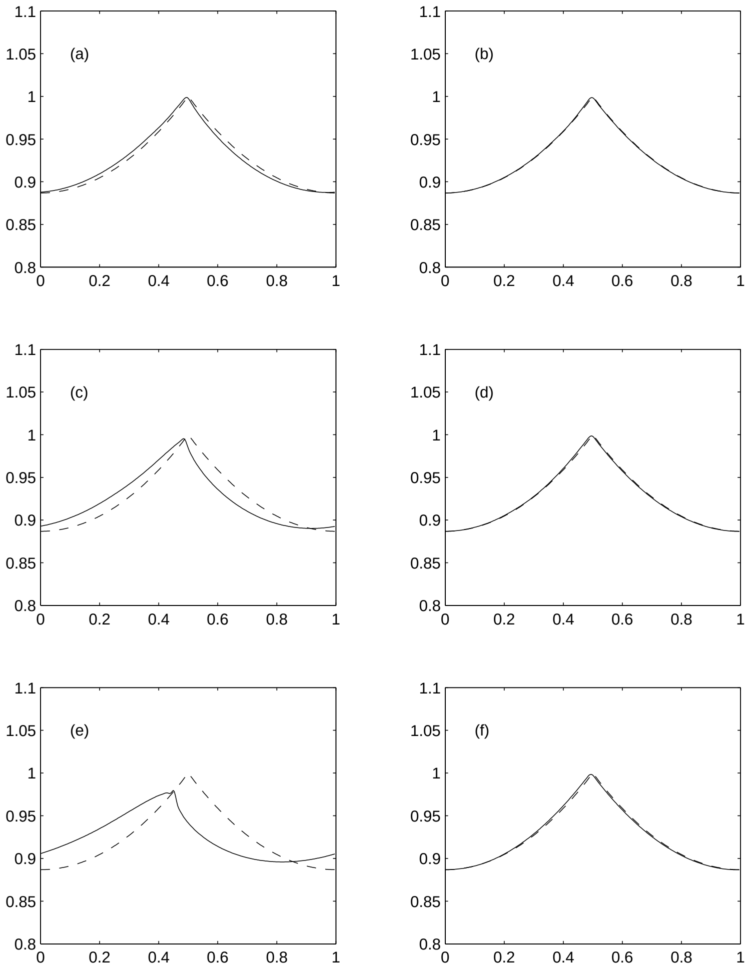


FIGURE 4. Comparison of aliased and de-aliased calculations of a periodic peakon with $N = 256$ and $\Delta t = 0.003$. The left column shows the aliased calculations. (a) $T = 4$, (c) $T = 6$, (e) $T = 8$. The right column shows the de-aliased calculations. (b) $T = 4$, (d) $T = 6$, (f) $T = 8$. Solid curves are computed waves after 4, 6 and 8 periods. Dashed curves are the initial data, translated by an appropriate amount. In the de-aliased computation, a 2/3 filter was used, so that the effective number of modes is only 170.

References

[1] R. Artebrant and H.J. Schroll, Numerical simulation of Camassa-Holm peakons by adaptive upwinding, to appear in Appl. Num. Math.

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

[3] R. Camassa, D. Holm and J. Hyman, A new integrable shallow water equation, Adv. Appl. Mech. 31 (1994), 1-33.

[4] A. Constantin, On the scattering problem for the Camassa-Holm equation, Proc. Roy. Soc. London A 457 (2001), 953-970.

[5] A. Constantin and J. Escher, Well-posedness, global existence, and blowup phenomena for a periodic quasi-linear hyperbolic equation, Comm. Pure Appl. Math. 5 (1998), 475-504.

[6] A. Constantin and B. Kolev, On the geometric approach to the motion of inertial mechanical systems, J. Phys. A 35 (2002), R51-R79.

[7] C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, Spectral Methods in Fluid Dynamics, Springer, 1988.

[8] H.-H. Dai and Y. Huo, Solitary shock waves and other travelling waves in a general compressible hyperelastic rod, Proc. Roy. Soc. London A 456 (2000), 331-363.

[9] A. S. Fokas, On a class of physically important integrable equations, Physica D 87 (1995), 145-150.

[10] B. Fuchssteiner and A. S. Fokas, Symplectic structures, their Bächlund transformation and hereditary symmetries, Physica D 4 (1981), 47-66.

[11] B. Fuchssteiner, Some tricks from the symmetry-toolbox for nonlinear equations: generalizations of the Camassa-Holm equation, Physica D 95 (1996), 229-243.

[12] C. Gasquet and P. Witomski, Fourier Analysis and Applications, Springer, 1999

[13] D.D. Holm, J. Marsden and T. Ratiu, The Euler-Poincaré equations and semidirect products with applications to continuum theories, Adv. Math. 137 (1998), 1-81.

[14] H. Holden and X. Raynaud, Convergence of a finite difference scheme for the Camassa-Holm equation, SIAM J. Numer. Anal. 2006, to appear. **paper I** in this thesis.

[15] H. Holden and X. Raynaud, A convergent numerical scheme for the Camassa–Holm equation based on multipeakons, Disc. Cont. Dyn. Sys. 14 (2006) 505-523. **paper II** in this thesis.

[16] R. Johnson, Camassa-Holm, Korteweg-de Vries and related models for water waves, J. Fluid Mech. **455** (2002), 63-82.

[17] H. Kalisch, Stability of solitary waves for a nonlinearly dispersive equation, Disc. Cont. Dyn. Sys. **10** (2004), 709-716.

[18] H. Kalisch and J. Lenells, Numerical study of traveling-wave solutions for the Camassa-Holm equation, Chaos, Solitons & Fractals **25** (2005) 287-298.

[19] H.-O. Kreiss and J. Oliger, Stability of the Fourier method, SIAM J. Numer. Anal. **16** (1979), 421-433.

[20] M. Kunze and G. Schneider, Estimates for the KdV-limit of the Camassa-Holm equation, Lett. Math. Phys. 72 (2005), 17-26.

[21] J. Lenells, Traveling wave solutions of the Camassa-Holm equation, J. Diff. Eq. 217 (2005), 393-430.

[22] Y. Maday and A. Quarteroni, Error analysis for spectral approximation of the Korteweg-de Vries equation, RAIRO Modl. Math. Anal. Numr. 22 (1988), 499-529.

[23] G. Misiolek, A shallow water equation as a geodesic flow on the Bott-Viroasoro group, J. Geom. Phys. 24 (1998), 203-208.

[24] J.E. Pasciak, Spectral and pseudospectral methods for advection equations, Math. Comput. 35 (1980), 1081-1092.

[25] R.L. Seliger, A note on breaking waves, Proc. Roy. Soc. A 303 (1968), 493-496.

[26] G.B. Whitham, Linear and Nonlinear Waves, Wiley, 1974.

Paper IV

Global conservative solutions of the Camassa–Holm equation – A Lagrangian point of view.

H. Holden and X. Raynaud

Submitted for publication.

GLOBAL CONSERVATIVE SOLUTIONS OF THE CAMASSA–HOLM EQUATION – A LAGRANGIAN POINT OF VIEW

HELGE HOLDEN AND XAVIER RAYNAUD

Abstract. We show that the Camassa–Holm equation $u_t - u_{xxt} + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0$ possesses a global continuous semigroup of weak conservative solutions for initial data $u|_{t=0}$ in H^1. The result is obtained by introducing a coordinate transformation into Lagrangian coordinates. To characterize conservative solutions it is necessary to include the energy density given by the positive Radon measure μ with $\mu_{\text{ac}} = (u^2 + u_x^2)\,dx$. The total energy is preserved by the solution.

1. Introduction

The Cauchy problem for the Camassa–Holm equation [7, 8]

$$u_t - u_{xxt} + 2\kappa u_x + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0, \quad u|_{t=0} = \bar{u}, \tag{1.1}$$

has received considerable attention the last decade. With κ positive it models, see [25], propagation of unidirectional gravitational waves in a shallow water approximation, with u representing the fluid velocity. The Camassa–Holm equation has a bi-Hamiltonian structure and is completely integrable. It has infinitely many conserved quantities. In particular, for smooth solutions the quantities

$$\int u\,dx, \quad \int (u^2 + u_x^2)\,dx, \quad \int (u^3 + uu_x^2)\,dx \tag{1.2}$$

are all time independent.

In this article we consider the case $\kappa = 0$ on the real line, that is,

$$u_t - u_{xxt} + 3uu_x - 2u_xu_{xx} - uu_{xxx} = 0, \tag{1.3}$$

and henceforth we refer to (1.3) as the Camassa–Holm equation. The equation can be rewritten as the following system

$$u_t + uu_x + P_x = 0, \tag{1.4a}$$

$$P - P_{xx} = u^2 + \frac{1}{2}u_x^2. \tag{1.4b}$$

2000 *Mathematics Subject Classification.* Primary: 65M06, 65M12; Secondary: 35B10, 35Q53.

Key words and phrases. Camassa–Holm equation, conservative solutions.

Supported in part by the Research Council of Norway.

A highly interesting property of the equation is that for a wide class of initial data the solution experiences wave breaking in finite time in the sense that the solution u remains bounded pointwise while the spatial derivative u_x becomes unbounded pointwise. However, the H^1 norm of u remains finite. More precisely, Constantin, Escher, and Molinet [12, 14] showed the following result: If the initial data $u|_{t=0} = \bar{u} \in H^1(\mathbb{R})$ and $\bar{m} := \bar{u} - \bar{u}''$ is a positive Radon measure, then equation (1.3) has a unique global weak solution $u \in C([0,T], H^1(\mathbb{R}))$, for any T positive, with initial data $\bar{u}$. However, any solution with odd initial data $\bar{u}$ in $H^3(\mathbb{R})$ such that $\bar{u}_x(0) < 0$ blows up in a finite time.

The problem how to extend the solution beyond wave breaking can nicely be illustrated by studying an explicit class of solutions. The Camassa–Holm equation possesses solutions, denoted (multi)peakons, of the form

$$u(t,x) = \sum_{i=1}^{n} p_i(t) e^{-|x-q_i(t)|}, \tag{1.5}$$

where the $(p_i(t), q_i(t))$ satisfy the explicit system of ordinary differential equations

$$\dot{q}_i = \sum_{j=1}^{n} p_j e^{-|q_i - q_j|}, \quad \dot{p}_i = \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i - q_j|}.$$

Observe that the solution (1.5) is not smooth even with continuous functions $(p_i(t), q_i(t))$; one possible way to interpret (1.5) as a weak solution of (1.3) is to rewrite the equation (1.3) as

$$u_t + \big(\frac{1}{2}u^2 + (1-\partial_x^2)^{-1}(u^2 + \frac{1}{2}u_x^2)\big)_x = 0.$$

Peakons interact in a way similar to that of solitons of the Korteweg–de Vries equation, and wave breaking may appear when at least two of the q_i's coincide. If all the $p_i(0)$ have the same sign, the peakons move in the same direction. Furthermore, in that case the solution experiences no wave breaking, and one has a global solution. Higher peakons move faster than the smaller ones, and when a higher peakon overtakes a smaller, there is an exchange of mass, but no wave breaking takes place. Furthermore, the $q_i(t)$ remain distinct. However, if some of $p_i(0)$ have opposite sign, wave breaking may incur, see, e.g., [3, 26]. For simplicity, consider the case with $n = 2$ and one peakon $p_1(0) > 0$ (moving to the right) and one antipeakon $p_2(0) < 0$ (moving to the left). In the symmetric case ($p_1(0) = -p_2(0)$ and $q_1(0) = -q_2(0) < 0$) the solution will vanish pointwise at the collision time t^* when $q_1(t^*) = q_2(t^*)$, that is, $u(t^*, x) = 0$ for all $x \in \mathbb{R}$. Clearly, at least two scenarios are possible; one is to let $u(t,x)$ vanish identically for $t > t^*$, and the other possibility is to let the peakon and antipeakon "pass through" each other in a way that is consistent with the Camassa–Holm equation. In the first case the energy $\int (u^2 + u_x^2)\, dx$ decreases to zero at t^*, while in the second case, the energy remains constant except at t^*. Clearly, the well-posedness of the equation is a delicate matter in this case. The first solution could be denoted a dissipative solution, while the second one could be called conservative. Other solutions are also possible. Global dissipative solutions of a more general class of equations were recently derived by Coclite, Holden, and Karlsen [9, 10]. In their approach the solution was obtained by first regularizing the equation by adding a small diffusion term ϵu_{xx} to the equation, and subsequently analyzing the vanishing viscosity limit $\epsilon \to 0$. Multipeakons are fundamental building blocks for general solutions. Indeed, if the initial data $\bar{u}$ is in H^1 and $\bar{m} := \bar{u} - \bar{u}''$ is a positive Radon measure, then it can proved, see [23], that one can construct a sequence of multipeakons that converges in $L^\infty_{\rm loc}(\mathbb{R}; H^1_{\rm loc}(\mathbb{R}))$ to the unique global solution of the Camassa–Holm equation.

The problem of continuation beyond wave breaking was recently considered by Bressan and Constantin [4]. They reformulated the Camassa–Holm equation as a semilinear system of ordinary differential equations taking values in a Banach space. This formulation allowed them to continue the solution beyond collision time, giving a global conservative solution where the energy is conserved for almost all times. Thus in the context of peakon-antipeakon collisions they considered the solution where the peakons and antipeakons "passed through" each other. Local existence of the semilinear system is obtained by a contraction argument. Furthermore, the clever reformulation allows for a global solution where all singularities disappear. Going back to the original function u, one obtains a global solution of the Camassa–Holm equation. The well-posedness, i.e., the uniqueness and stability of the solution, is resolved as follows. In addition to the solution u, one includes a family of non-negative Radon measures μ_t with density $u_x^2\, dx$ with respect to the Lebesgue measure. The pair (u, μ_t) constitutes a continuous semigroup, in particular, one has uniqueness and stability.

Very recently, Bressan and Fonte [5, 20] presented another approach to the Camassa–Holm equation. The flow map $\bar{u} \mapsto u(t)$ is, as we have seen, neither a continuous map on H^1 nor on L^2. However, they introduced a new distance $J(u, v)$ with the property

$$c_1 \left\| u - v \right\|_{L^1} \leq J(u, v) \leq c_2 \left\| u - v \right\|_{H^1}.$$

Furthermore, it satisfies

$$J(u(t), \bar{u}) \leq c_3 \left| t \right|, \quad J(u(t), v(t)) \leq J(\bar{u}, \bar{v}) e^{c_4 |t|},$$

where $u(t), v(t)$ are solutions with initial data $\bar{u}, \bar{v}$, respectively. The distance is introduced by first defining it for multipeakons, using the global, conservative solution described above. Subsequently it is shown that multipeakons are dense in the space H^1. This enables them to construct a semi-group of conservative solutions for the Camassa–Holm equation which is continuous with respect to the distance J.

In this paper, as Bressan and Constantin [4], we reformulate the equation using a different set of variables and obtain a semilinear system of ordinary differential equations. However, the change of variables we use is distinct from that of Bressan and Constantin and simply corresponds to the transformation between Eulerian and Lagrangian coordinates. Let $u = u(t, x)$ denote the solution, and $y(t, \xi)$ the corresponding characteristics, thus $y_t(t, \xi) = u(t, y(t, \xi))$. Our new variables are

$y(t,\xi)$,

$$U(t,\xi)=u(t,y(t,\xi)),\quad H(t,\xi)=\int_{-\infty}^{y(t,\xi)}(u^2+u_x^2)\,dx \tag{1.6}$$

where U corresponds to the Lagrangian velocity while H could be interpreted as the Lagrangian cumulative energy distribution. Furthermore, let

$$Q(t,\xi)=-\frac{1}{4}\int_{\mathbb{R}}\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi+H_\xi\big)(\eta)\,d\eta,$$
$$P(t,\xi)=\frac{1}{4}\int_{\mathbb{R}}\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi+H_\xi\big)\,d\eta.$$

Then one can show that

$$\begin{cases} y_t=U,\\ U_t=-Q,\\ H_t=U^3-2PU,\end{cases} \tag{1.7}$$

is equivalent to the Camassa–Holm equation. Global existence of solutions of (1.7) is obtained starting from a contraction argument, see Theorem 2.8. The uniqueness issue is resolved by considering the set $\mathcal{D}$ (see Definition 3.1) which consists of pairs (u,μ) such that $(u,\mu)\in\mathcal{D}$ if $u\in H^1(\mathbb{R})$ and μ is a positive Radon measure whose absolutely continuous part satisfies $\mu_{\text{ac}}=(u^2+u_x^2)\,dx$. With three Lagrangian variables (y,U,H) versus two Eulerian variables (u,μ), it is clear that there can be no bijection between the two coordinates systems. However, we define a group of relabeling transformations which acts on the Lagrangian variables and let the system of equations (1.7) invariant. Using this group, we are able to establish a bijection between the space of Eulerian variables and the space of Lagrangian variables when we identify variables that are invariant under relabeling. This bijection allows us to transform the results obtained in the Lagrangian framework (in which the equation is well-posed) into the Eulerian framework (in which the situation is much more subtle). In particular, and this constitutes the main result of this paper, we obtain a metric $d_{\mathcal{D}}$ on $\mathcal{D}$ and a continuous semi-group of solutions on $(\mathcal{D},d_{\mathcal{D}})$. The distance $d_{\mathcal{D}}$ gives $\mathcal{D}$ the structure of a complete metric space. This metric is compared with some more standard topologies, and we obtain that convergence in $H^1(\mathbb{R})$ implies convergence in $(\mathcal{D},d_{\mathcal{D}})$ which itself implies convergence in $L^\infty(\mathbb{R})$, see Propositions 5.1 and 5.2. The properties of the spaces as well as the various mappings between them are described in great detail, see Section 3. Our main result, Theorem 4.2, states that for given initial data in $\mathcal{D}$ there exists a unique weak solution of the Camassa–Holm equation. The associated measure $\mu(t)$ has constant total mass, i.e., $\mu(t)(\mathbb{R})=\mu(0)(\mathbb{R})$ for all t, which corresponds to the total energy of the system. This is the reason why our solutions are called conservative.

The method described here can be studied in detail for multipeakons, see [24] for details. By suitably modifying the techniques described in this paper, the results can be extended to show global existence of conservative solutions for the generalized hyperelastic-rod equation

$$\begin{cases} u_t + f(u)_x + P_x = 0 \\ P - P_{xx} = g(u) + \dfrac{1}{2} f''(u) u_x^2. \end{cases} \tag{1.8}$$

where $f, g \in C^\infty(\mathbb{R})$ and f is strictly convex. Observe that if $g(u) = \kappa u + u^2$ and $f(u) = \frac{u^2}{2}$, then (1.8) is the classical Camassa–Holm equation (1.1). With $g(u) = \frac{3-\gamma}{2} u^2$ and $f(u) = \frac{\gamma}{2} u^2$, Dai [15, 16, 17] derived (1.8) as an equation describing finite length, small amplitude radial deformation waves in cylindrical compressible hyperelastic rods, and the equation is often referred to as the hyperelastic-rod wave equation. See [9, 10] for a recent proof of existence of dissipative solutions of (1.8). The details will be described in a forthcoming paper.

Furthermore, the methods presented in this paper can be used to derive numerical methods that converge to conservative solutions rather than dissipative solutions. This contrasts finite difference methods that normally converge to dissipative solutions, see [22] and [21] for the related Hunter–Saxton equation. See also [23]. Results will be presented separately.

2. Global solutions in Lagrangian coordinates

2.1. **Equivalent system.** Assuming that u is smooth, it is not hard to check that

$$(u^2 + u_x^2)_t + (u(u^2 + u_x^2))_x = (u^3 - 2Pu)_x. \tag{2.1}$$

Let us introduce the characteristics $y(t, \xi)$ defined as the solutions of

$$y_t(t, \xi) = u(t, y(t, \xi)) \tag{2.2}$$

for a given $y(0, \xi)$. Equation (2.1) gives us information about the evolution of the amount of energy contained between two characteristics. Indeed, given ξ_1, ξ_2 in $\mathbb{R}$, let $H(t) = \int_{y(t,\xi_1)}^{y(t,\xi_2)} \left(u^2 + u_x^2\right) dx$ be the energy contained between the two characteristic curves $y(t, \xi_1)$ and $y(t, \xi_2)$. Then, using (2.1) and (2.2), we obtain

$$\frac{dH}{dt} = \left[(u^3 - 2Pu) \circ y\right]_{\xi_1}^{\xi_2}. \tag{2.3}$$

Solutions of the Camassa–Holm blow up when characteristics arising from different points collide. It is important to notice that we do not get shocks as the Camassa–Holm preserves the H^1 norm and therefore solutions remain continuous. However, it is not obvious how to continue the solution after collision time. It turns out that, when two characteristics collide, the energy contained between these two characteristics has a limit which can be computed from (2.3). As we will see, knowing this energy enables us to prolong the characteristics and thereby the solution, after collisions.

We now derive a system equivalent to (1.4). All the derivations in this section are formal and will be justified later. Let y still denote the characteristics. We introduce two other variables, the Lagrangian velocity and cumulative energy distribution, U and H, defined as $U(t,\xi) = u(t,y(t,\xi))$ and

$$H(t,\xi) = \int_{-\infty}^{y(t,\xi)} \left(u^2 + u_x^2\right)\,dx. \tag{2.4}$$

From the definition of the characteristics, it follows that

$$U_t(t,\xi) = u_t(t,y) + y_t(t,\xi)u_x(t,y) = -P_x\circ y\,(t,\xi). \tag{2.5}$$

This last term can be expressed uniquely in term of U, y, and H. From (1.4b), we obtain the following explicit expression for P,

$$P(t,x) = \frac{1}{2}\int_{\mathbb{R}} e^{-|x-z|}(u^2(t,z) + \frac{1}{2}u_x^2(t,z))\,dz. \tag{2.6}$$

Thus we have

$$P_x\circ y\,(t,\xi) = -\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(y(t,\xi)-z)e^{-|y(t,\xi)-z|}(u^2(t,z) + \frac{1}{2}u_x^2(t,z))\,dz$$

and, after the change of variables $z = y(t,\eta)$,

$$P_x\circ y(t,\xi) = -\frac{1}{2}\int_{\mathbb{R}} \Big[\operatorname{sgn}(y(t,\xi)-y(t,\eta))e^{-|y(t,\xi)-y(t,\eta)|} \\ \times\left(u^2(t,y(t,\eta)) + \frac{1}{2}u_x^2(t,y(t,\eta))\right)y_\xi(t,\eta)\Big]\,d\eta.$$

Finally, since $H_\xi = (u^2+u_x^2)\circ y\,y_\xi$,

$$P_x\circ y\,(\xi) = -\frac{1}{4}\int_{\mathbb{R}} \operatorname{sgn}(y(\xi)-y(\eta))\exp(-|y(\xi)-y(\eta)|)\big(U^2y_\xi + H_\xi\big)(\eta)\,d\eta \tag{2.7}$$

where the t variable has been dropped to simplify the notation. Later we will prove that y is an increasing function for any fixed time t. If, for the moment, we take this for granted, then $P_x\circ y$ is equivalent to Q where

$$Q(t,\xi) = -\frac{1}{4}\int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi+H_\xi\big)(\eta)\,d\eta, \tag{2.8}$$

and, slightly abusing the notation, we write

$$P(t,\xi) = \frac{1}{4}\int_{\mathbb{R}} \exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi + H_\xi\big)(\eta)\,d\eta. \tag{2.9}$$

Thus $P_x\circ y$ and $P\circ y$ can be replaced by equivalent expressions given by (2.8) and (2.9) which only depend on our new variables U, H, and y. We introduce yet another variable, $\zeta(t,\xi)$, simply defined as $\zeta(t,\xi) = y(t,\xi) - \xi$. It will turn out that $\zeta \in L^\infty(\mathbb{R})$. We now derive a new system of equations, formally equivalent to the Camassa–Holm equation. Equations (2.5), (2.3) and (2.2) give us

$$\begin{cases} \zeta_t = U, \\ U_t = -Q, \\ H_t = U^3 - 2PU. \end{cases} \tag{2.10}$$

As we will see, the system (2.10) of ordinary differential equations for (ζ, U, H) from $[0,T]$ to E is well-posed, where E is Banach space to be defined in the next section. We have

$$Q_\xi = -\frac{1}{2}H_\xi - \left(\frac{1}{2}U^2 - P\right) y_\xi \quad \text{and} \quad P_\xi = Q y_\xi. \tag{2.11}$$

Hence, differentiating (2.10) yields

$$\begin{cases} \zeta_{\xi t} = U_\xi \ (\text{or } y_{\xi t} = U_\xi), \\ U_{\xi t} = \dfrac{1}{2}H_\xi + \left(\dfrac{1}{2}U^2 - P\right) y_\xi, \\ H_{\xi t} = -2Q\, U y_\xi + \left(3U^2 - 2P\right) U_\xi. \end{cases} \tag{2.12}$$

The system (2.12) is semilinear with respect to the variables y_ξ, U_ξ and H_ξ.

2.2. **Existence and uniqueness of solutions of the equivalent system.** In this section, we focus our attention on the system of equations (2.10) and prove, by a contraction argument, that it admits a unique solution. Let V be the Banach space defined by

$$V = \{f \in C_b(\mathbb{R}) \mid f_\xi \in L^2(\mathbb{R})\}$$

where $C_b(\mathbb{R}) = C(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and the norm of V is given by $\|f\|_V = \|f\|_{L^\infty(\mathbb{R})} + \|f_\xi\|_{L^2(\mathbb{R})}$. Of course $H^1(\mathbb{R}) \subset V$ but the converse is not true as V contains functions that do not vanish at infinity. We will employ the Banach space E defined by

$$E = V \times H^1(\mathbb{R}) \times V$$

to carry out the contraction mapping argument. For any $X = (\zeta, U, H) \in E$, the norm on E is given by $\|X\|_E = \|\zeta\|_V + \|U\|_{H^1(\mathbb{R})} + \|H\|_V$. The following lemma gives the Lipschitz bounds we need on Q and P.

Lemma 2.1. *For any $X = (\zeta, U, H)$ in E, we define the maps $\mathcal{Q}$ and $\mathcal{P}$ as $\mathcal{Q}(X) = Q$ and $\mathcal{P}(X) = P$ where Q and P are given by* (2.8) *and* (2.9)*, respectively. Then, $\mathcal{P}$ and $\mathcal{Q}$ are Lipschitz maps on bounded sets from E to $H^1(\mathbb{R})$. Moreover, we have*

$$Q_\xi = -\frac{1}{2}H_\xi - \left(\frac{1}{2}U^2 - P\right)(1 + \zeta_\xi), \tag{2.13}$$

$$P_\xi = Q(1 + \zeta_\xi). \tag{2.14}$$

Proof. We rewrite $\mathcal{Q}$ as

$$\begin{aligned} \mathcal{Q}(X)(\xi) = &-\frac{e^{-\zeta(\xi)}}{4}\int_{\mathbb{R}} \chi_{\{\eta<\xi\}} e^{-|\xi-\eta|} e^{\zeta(\eta)} \left[U(\eta)^2(1+\zeta_\xi(\eta)) + H_\xi(\eta)\right] d\eta \\ &+ \frac{e^{\zeta(\xi)}}{4}\int_{\mathbb{R}} \chi_{\{\eta>\xi\}} e^{-|\xi-\eta|} e^{-\zeta(\eta)} \left[U(\eta)^2(1+\zeta_\xi(\eta)) + H_\xi(\eta)\right] d\eta, \end{aligned} \tag{2.15}$$

where χ_B denotes the indicator function of a given set B. We decompose $\mathcal{Q}$ into the sum $\mathcal{Q}_1 + \mathcal{Q}_2$ where $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are the operators corresponding to the two terms on the right-hand side of (2.15). Let $h(\xi) = \chi_{\{\xi>0\}}(\xi)e^{-\xi}$ and A be the map defined by $A\colon v \mapsto h \star v$. Then, Q_1 can be rewritten as

$$\mathcal{Q}_1(X)(\xi) = -\frac{e^{-\zeta(\xi)}}{4} A \circ R(\zeta, U, H)(\xi) \tag{2.16}$$

where R is the operator from E to $L^2(\mathbb{R})$ given by $R(\zeta,U,H)(t,\xi) = \chi_{\{\eta<\xi\}}e^{\zeta}(U^2 + U^2\zeta_\xi + H_\xi)$. We claim that A is continuous from $L^2(\mathbb{R})$ into $H^1(\mathbb{R})$. The Fourier transform of h can easily be computed, and we obtain

$$\hat{h}(\eta) = \int_{\mathbb{R}} h(\xi)e^{-2i\pi\eta\xi}\,d\xi = \frac{1}{1+2i\pi\eta}. \tag{2.17}$$

The $H^1(\mathbb{R})$ norm can be expressed in term of the Fourier transform as follows, see, e.g., [19],

$$\|h \star v\|_{H^1(\mathbb{R})} = \left\|(1+\eta^2)^{\frac{1}{2}}\widehat{h \star v}\right\|_{L^2(\mathbb{R})}.$$

Since $\widehat{h \star v} = \hat{h}\hat{v}$, we have

$$\begin{aligned}
\|h \star v\|_{H^1(\mathbb{R})} &= \left\|(1+\eta^2)^{\frac{1}{2}}\hat{h}\hat{v}\right\|_{L^2(\mathbb{R})} \\
&\leq C\,\|\hat{v}\|_{L^2(\mathbb{R})} \qquad \text{by (2.17)} \\
&= C\,\|v\|_{L^2(\mathbb{R})} \qquad \text{by Plancherel equality}
\end{aligned}$$

for some constant C. Hence, $A\colon L^2(\mathbb{R}) \to H^1(\mathbb{R})$ is continuous. We prove that $R(\zeta,U,H)$ belongs to $L^2(\mathbb{R})$ by using the assumption that $g(0) = 0$. Then, $A \circ R(\zeta,U,H)$ belongs to H^1. We say that an operator is B-Lipschitz when it is Lipschitz on bounded sets. Let us prove that $\mathcal{Q}_1\colon E \to H^1(\mathbb{R})$ is B-Lipschitz. It is not hard to prove that R is B-Lipschitz from E into $L^2(\mathbb{R})$ and therefore from E into $H^{-1}(\mathbb{R})$. Since $A\colon H^{-1}(\mathbb{R}) \to H^1(\mathbb{R})$ is linear and continuous, $A \circ R$ is B-Lipschitz from E to $H^1(\mathbb{R})$. Then, we use the following lemma whose proof is left to the reader.

Lemma 2.2. *Let $\mathcal{R}_1\colon E \to V$ and $\mathcal{R}_2\colon E \to H^1(\mathbb{R})$, or $\mathcal{R}_2\colon E \to V$, be two B-Lipschitz maps. Then, the product $X \mapsto R_1(X)R_2(X)$ is also a B-Lipschitz map from E to $H^1(\mathbb{R})$, or from E to V.*

Since the mapping $X \mapsto e^{-\zeta}$ is B-Lipschitz from E to V, Q_1 is the product of two B-Lipschitz maps, one from E to $H^1(\mathbb{R})$ and the other from E to V, it is B-Lipschitz map from E to $H^1(\mathbb{R})$. Similarly, one proves that $\mathcal{Q}_2$ is B-Lipschitz and therefore $\mathcal{Q}$ is B-Lipschitz. Furthermore, $\mathcal{P}$ is B-Lipschitz. The formulas (2.13) and (2.14) are obtained by direct computation using the product rule, see [18, p. 129]. ☐

In the next theorem, by using a contraction argument, we prove the short-time existence of solutions to (2.10).

Theorem 2.3. *Given $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in E, there exists a time T depending only on $\|\bar{X}\|_E$ such that the system (2.10) admits a unique solution in $C^1([0,T],E)$ with initial data $\bar{X}$.*

Proof. Solutions of (2.10) can be rewritten as

$$X(t) = \bar{X} + \int_0^t F(X(\tau))\,d\tau \tag{2.18}$$

where $F\colon E \to E$ is given by $F(X) = (U, -\mathcal{Q}(X), U^3 - 2\mathcal{P}(X)U)$ where $X = (\zeta, U, H)$. The integrals are defined as Riemann integrals of continuous functions on the Banach space E. Using Lemma 2.1, we can check that each component of $F(X)$ is a product of functions that satisfy one of the assumptions of Lemma 2.2 and using this same lemma, we obtain that $F(X)$ is a Lipchitz function on any bounded set of E. Since E is a Banach space, we use the standard contraction argument to prove the theorem. □

We now turn to the proof of existence of global solutions of (2.10). We are interested in a particular class of initial data that we are going to make precise later, see Definition 2.6. In particular, we will only consider initial data that belong to $E \cap \left[W^{1,\infty}(\mathbb{R})\right]^3$ where $W^{1,\infty}(\mathbb{R}) = \{f \in C_b(\mathbb{R}) \mid f_\xi \in L^\infty(\mathbb{R})\}$. Given $(\bar{\zeta}, \bar{U}, \bar{H}) \in E \cap [W^{1,\infty}(\mathbb{R})]^3$, we consider the short-time solution $(\zeta, U, H) \in C([0,T], E)$ of (2.10) given by Theorem 2.3. Using the fact that $\mathcal{Q}$ and $\mathcal{P}$ are Lipschitz on bounded sets (Lemma 2.1) and, since $X \in C([0,T], E)$, we can prove that P and Q belongs to $C([0,T], H^1(\mathbb{R}))$. We now consider U, P and Q as given function in $C([0,T], H^1(\mathbb{R}))$. Then, for any *fixed* $\xi \in \mathbb{R}$, we can solve the system of ordinary differential equations

$$\begin{cases} \dfrac{d}{dt}\alpha(t,\xi) = \beta(t,\xi), \\ \dfrac{d}{dt}\beta(t,\xi) = \dfrac{1}{2}\gamma(t,\xi) + \left[(\dfrac{1}{2}U^2 - P)(t,\xi)\right](1 + \alpha(t,\xi)), \\ \dfrac{d}{dt}\gamma(t,\xi) = -\left[2(Q\,U)(t,\xi)\right](1 + \alpha(t,\xi)) + \left[(3U^2 - 2P)(t,\xi)\right]\beta(t,\xi), \end{cases} \tag{2.19}$$

which is obtained by substituting ζ_ξ, U_ξ and H_ξ in (2.12) by the unknowns α, β, and γ, respectively. We have to specify the initial conditions for (2.19). Let $\mathcal{A}$ be the following set

$$\mathcal{A} = \{\xi \in \mathbb{R} \mid \left|\bar{U}_\xi(\xi)\right| \leq \left\|\bar{U}_\xi\right\|_{L^\infty(\mathbb{R})},\ \left|\bar{H}_\xi(\xi)\right| \leq \left\|\bar{H}_\xi\right\|_{L^\infty(\mathbb{R})},\ \left|\bar{\zeta}_\xi(\xi)\right| \leq \left\|\bar{\zeta}_\xi\right\|_{L^\infty(\mathbb{R})}\},$$

We have that $\mathcal{A}$ has full measure, that is, $\operatorname{meas}(\mathcal{A}^c) = 0$. For $\xi \in \mathcal{A}$ we define $(\alpha(0,\xi), \beta(0,\xi), \gamma(0,\xi)) = (\bar{U}_\xi(\xi), \bar{H}_\xi(\xi), \bar{\zeta}_\xi(\xi))$. However, for $\xi \in \mathcal{A}^c$ we take $(\alpha(0,\xi), \beta(0,\xi), \gamma(0,\xi)) = (0,0,0)$.

Lemma 2.4. *Given initial condition $\bar{X} = (\bar{U}, \bar{H}, \bar{\zeta}) \in E \cap [W^{1,\infty}(\mathbb{R})]^3$, we consider the solution $X = (\zeta, U, H) \in C^1([0,T], E)$ of* (2.19) *given by Theorem 2.3. Then, $X \in C^1([0,T], E \cap [W^{1,\infty}(\mathbb{R})]^3)$. The functions $\alpha(t,\xi)$, $\beta(t,\xi)$ and $\gamma(t,\xi)$ which are obtained by solving* (2.19) *for any fixed given ξ with the initial condition specified above, coincide for almost every ξ and for all time t with ζ_ξ, U_ξ and H_ξ, respectively, that is, for all $t \in [0,T]$, we have*

$$(\alpha(t,\xi), \beta(t,\xi), \gamma(t,\xi)) = (\zeta_\xi(t,\xi), U_\xi(t,\xi), H_\xi(t,\xi)) \tag{2.20}$$

for almost every $\xi \in \mathbb{R}$.

Thus, this lemma allows us to pick up a special representative for $(\zeta_\xi, U_\xi, H_\xi)$ given by (α, β, γ), which is defined for all $\xi \in \mathbb{R}$ and which, for any given ξ, satisfies the ordinary differential equation (2.19). In the remaining we will of course identify the two and set $(\zeta_\xi, U_\xi, H_\xi)$ equal to (α, β, γ). To prove this lemma, we will need the following proposition which is adapted from [27, p. 134, Corollary 2].

Proposition 2.5. *Let R be a bounded linear operator on a Banach space X into a Banach space Y. Let f be in $C([0,T], X)$. Then, Rf belongs to $C([0,T], Y)$ and therefore is Riemann integrable, and $\int_{[0,T]} Rf(t)\,dt = R\int_{[0,T]} f(t)\,dt$.*

Proof of Lemma 2.4. We introduce the Banach space of *everywhere* bounded function $B^\infty(\mathbb{R})$ whose norm is naturally given by $\|f\|_{B^\infty(\mathbb{R})} = \sup_{\xi\in\mathbb{R}} |f(\xi)|$. Obviously, $C_b(\mathbb{R})$ is included in $B^\infty(\mathbb{R})$. We define (α, β, γ) as the solution of (2.19) in $[B^\infty(\mathbb{R})]^3 \cap \left[L^2(\mathbb{R})\right]^3$ with initial data as given above. Thus, strictly speaking, this is a different definition than the one given in the lemma but we will see that they are in fact equivalent. We note that the system (2.19) is affine (it consists of a sum of a linear transformation and a constant) and therefore it is not hard to prove, by using a contraction argument in $[B^\infty(\mathbb{R})]^3 \cap \left[L^2(\mathbb{R})\right]^3$, the short-time existence of solutions. Moreover, the solution exists on $[0,T]$, the interval on which (ζ, U, H) is defined. Let us assume the opposite. Then, $Z_1(t) = \|\alpha(t,\,\cdot\,)\|_{B^\infty(\mathbb{R})\cap L^2(\mathbb{R})} + \|\beta(t,\,\cdot\,)\|_{B^\infty(\mathbb{R})\cap L^2(\mathbb{R})} + \|\gamma(t,\,\cdot\,)\|_{B^\infty(\mathbb{R})\cap L^2(\mathbb{R})}$ has to blow up when t approaches some time strictly smaller than T. We rewrite (2.19) in integral form:

$$\begin{cases} \alpha(t) = \alpha(0) + \displaystyle\int_0^t \beta(\tau)\,d\tau, \\ \beta(t) = \beta(0) + \displaystyle\int_0^t \Big(\frac{1}{2}\gamma + (\frac{1}{2}U^2 - P)(1+\alpha)\Big)(\tau)\,d\tau, \\ \gamma(t) = \gamma(0) + \displaystyle\int_0^t \big(-2Q\,U(1+\alpha) + (3U^2 - 2P)\beta\big)(\tau)\,d\tau. \end{cases} \tag{2.21}$$

Note that in (2.21) all the terms belong to $B^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$ and the equalities hold in this space. After taking the norms on both sides of the three equations in (2.21) and adding them term by term, we obtain the following inequality

$$Z_1(t) \le Z_1(0) + CT + C\int_0^t Z_1(\tau)\,d\tau$$

where C is a constant which depends on the $C([0,T], H^1(\mathbb{R}))$-norms of U, P and Q, which, by assumption, are bounded. From Gronwall's lemma, we get $Z_1(t) \le (Z_1(0) + CT)e^{CT}$ and therefore $Z_1(t)$ cannot blow up and α, β and γ belong to $C^1([0,T], B^\infty(\mathbb{R}) \cap L^2(\mathbb{R}))$. For any given ξ, the map $f \mapsto f(\xi)$ from $B^\infty(\mathbb{R})$ to $\mathbb{R}$ is linear and continuous. Hence, after applying this map to each term in (2.21) and using Proposition 2.5, we recover the original definition of

α, β and γ as solutions, for any given $\xi \in \mathbb{R}$, of the system (2.19) of ordinary differential equations in $\mathbb{R}^3$. The derivation map ∂_ξ is continuous from V and $H^1(\mathbb{R})$ into $L^2(\mathbb{R})$. We can apply it to each term in (2.10) written in integral from and, by Proposition 2.5, this map commutes with the integral. We end up with, after using (2.13) and (2.14),

$$\begin{cases} \zeta_\xi(t) = \bar{\zeta}_\xi + \int_0^t U_\xi(\tau)\,d\tau, \\ U_\xi(t) = \bar{U}_\xi + \int_0^t \left(\frac{1}{2}H_\xi + (\frac{1}{2}U^2 - P)(1+\zeta_\xi)\right)(\tau)\,d\tau, \\ H_\xi(t) = \bar{H}_\xi + \int_0^t \left(-2Q\,U(1+\zeta_\xi) + (3U^2 - 2P)U_\xi\right)(\tau)\,d\tau. \end{cases} \tag{2.22}$$

The injection map from $B^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$ to $L^2(\mathbb{R})$ is of course continuous, we can apply it to (2.21) and again use Proposition 2.5. Then, we can subtract each equation in (2.22) from the corresponding one in (2.21), take the norm and add them. After introducing $Z_2(t) = \|\alpha(t,\,\cdot\,) - \zeta_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|\beta(t,\,\cdot\,) - U_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|\gamma(t,\,\cdot\,) - H_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})}$, we end up with the following equation

$$Z_2(t) \le Z_2(0) + C\int_0^t Z_2(\tau)\,d\tau$$

where C is a constant which, again, only depends on the $C([0,T], H^1(\mathbb{R}))$-norms, of U, P and Q. By assumption on the initial conditions, we have $Z_2(0) = 0$ because $\alpha(0) = \bar{\zeta}_\xi$, $\beta(0) = \bar{U}_\xi$, $\gamma(0) = \bar{H}_\xi$ almost everywhere and therefore, by Gronwall's lemma, we get $Z_2(t) = 0$ for all $t \in [0,T]$. This is just a reformulation of (2.20), and this concludes the proof of the lemma. □

It is possible to carry out the contraction argument of Theorem 2.3 in the Banach space $[W^{1,\infty}(\mathbb{R})]^3$ but the topology on $[W^{1,\infty}(\mathbb{R})]^3$ turns out to be too strong for our purpose and that is why we prefer E whose topology is in some sense weaker. Our goal is to find solutions of (1.4) with initial data $\bar{u}$ in H^1 because H^1 is the natural space for the equation. Theorem 2.3 gives us the existence of solutions to (2.10) for initial data in E. Therefore we have to find initial conditions that match the initial data $\bar{u}$ and belong to E. A natural choice would be to use $\bar{y}(\xi) = y(0,\xi) = \xi$ and $\bar{U}(\xi) = u(\xi)$. Then $y(t,\xi)$ gives the position of the particle which is at ξ at time $t = 0$. But, if we make this choice, then $\bar{H}_\xi = \bar{u}^2 + \bar{u}_x^2$ and H_ξ does not belong to $L^2(\mathbb{R})$ in general. We consider instead $(\bar{y}, \bar{U}, \bar{H})$ given by the relations

$$\xi = \int_{-\infty}^{\bar{y}(\xi)} (\bar{u}^2 + \bar{u}_x^2)\,dx + \bar{y}(\xi),\ \bar{U} = \bar{u}\circ\bar{y}\,,\ \text{and}\ \bar{H} = \int_{-\infty}^{\bar{y}} \left(\bar{u}^2 + \bar{u}_x^2\right)\,dx. \tag{2.23}$$

Later (see Remark 3.10), we will prove that $(\bar{y} - \mathrm{Id}, \bar{U}, \bar{H})$ belongs to $\mathcal{G}$ where $\mathcal{G}$ is defined as follows.

Definition 2.6. The set $\mathcal{G}$ is composed of all $(\zeta, U, H) \in E$ such that

$$(\zeta, U, H) \in \left[W^{1,\infty}(\mathbb{R})\right]^3, \tag{2.24a}$$

$$y_\xi \geq 0, H_\xi \geq 0, y_\xi + H_\xi > 0 \text{ almost everywhere, and } \lim_{\xi \to -\infty} H(\xi) = 0, \tag{2.24b}$$

$$y_\xi H_\xi = y_\xi^2 U^2 + U_\xi^2 \text{ almost everywhere,} \tag{2.24c}$$

where we denote $y(\xi) = \zeta(\xi) + \xi$.

Note that if all functions are smooth and $y_\xi > 0$, we have $u_x \circ y = \frac{U_\xi}{y_\xi}$ and condition (2.24c) is equivalent to (2.4). For initial data in $\mathcal{G}$, the solution of (2.10) exists globally.

Lemma 2.7. *Given initial data $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in $\mathcal{G}$, let $X(t) = (\zeta(t), U(t), H(t))$ be the short-time solution of (2.10) in $C([0,T], E)$ for some $T > 0$ with initial data $(\bar{\zeta}, \bar{U}, \bar{H})$. Then,*
- *(i) $X(t)$ belongs to $\mathcal{G}$ for all $t \in [0,T]$,*
- *(ii) for almost every $t \in [0,T]$, $y_\xi(t,\xi) > 0$ for almost every $\xi \in \mathbb{R}$,*
- *(iii) For all $t \in [0,T]$, $\lim_{\xi \to \pm\infty} H(t,\xi)$ exists and is independent of time.*

We denote by $\mathcal{A}$ the set where the absolute values of $\bar{\zeta}_\xi(\xi)$, $\bar{H}_\xi(\xi)$, and $\bar{U}_\xi(\xi)$ all are smaller than $\left\|\bar{X}\right\|_{[W^{1,\infty}(\mathbb{R})]^3}$ and where the inequalities in (2.24b) and (2.24c) are satisfied for y_ξ, U_ξ and H_ξ. By assumption, we have $\operatorname{meas}(\mathcal{A}^c) = 0$ and we set $(\bar{U}_\xi, \bar{H}_\xi, \bar{\zeta}_\xi)$ equal to zero on $\mathcal{A}^c$. Thus, as allowed by Lemma 2.4, we choose a special representative for $(\zeta(t,\xi), U(t,\xi), H(t,\xi))$ which satisfies (2.12) as an ordinary differential equation, for every $\xi \in \mathbb{R}$.

Proof. (i) We already proved in Lemma 2.4 that the space $[W^{1,\infty}(\mathbb{R})]^3$ is preserved and therefore $X(t)$ satisfies (2.24a) for all $t \in [0,T]$. Let us prove that (2.24c) and the inequalities in (2.24b) hold for any $\xi \in \mathcal{A}$ and therefore almost everywhere. We consider a fixed ξ in $\mathcal{A}$ and drop it in the notations when there is no ambiguity. From (2.12), we have, on the one hand,

$$(y_\xi H_\xi)_t = y_{\xi t} H_\xi + H_{\xi t} y_\xi = U_\xi H_\xi + (3U^2 U_\xi - 2y_\xi QU - 2PU_\xi) y_\xi,$$

and, on the other hand,

$$\begin{aligned}(y_\xi^2 U^2 + U_\xi^2)_t &= 2y_{\xi t} y_\xi U^2 + 2y_\xi^2 U_t U + 2U_{\xi t} U_\xi \\ &= 3U_\xi U^2 y_\xi - 2PU_\xi y_\xi + H_\xi U_\xi - 2y_\xi^2 QU.\end{aligned}$$

Thus, $(y_\xi H_\xi - y_\xi^2 U^2 - U_\xi^2)_t = 0$, and since $y_\xi H_\xi(0) = (y_\xi^2 U^2 + U_\xi^2)(0)$, we have $y_\xi H_\xi(t) = (y_\xi^2 U^2 + U_\xi^2)(t)$ for all $t \in [0,T]$. We have proved (2.24c). Let us introduce t^* given by

$$t^* = \sup\{t \in [0,T] \mid y_\xi(t') \geq 0 \text{ for all } t' \in [0,t]\}.$$

Here we recall that we consider a fixed $\xi \in \mathcal{A}$ and dropped it in the notation. Assume that $t^* < T$. Since $y_\xi(t)$ is continuous with respect to time, we have

$$y_\xi(t^*) = 0. \tag{2.25}$$

Hence, from (2.24c) that we just proved, $U_\xi(t^*)=0$ and, by (2.12),

$$y_{\xi t}(t^*)=U_\xi(t^*)=0. \tag{2.26}$$

From (2.12), since $y_\xi(t^*)=U_\xi(t^*)=0$, we get

$$y_{\xi tt}(t^*)=U_{\xi t}(t^*)=\frac{1}{2}H_\xi(t^*). \tag{2.27}$$

If $H_\xi(t^*)=0$, then $(y_\xi,U_\xi,H_\xi)(t^*)=(0,0,0)$ and, by the uniqueness of the solution of (2.12), seen as a system of ordinary differential equations, we must have $(y_\xi,U_\xi,H_\xi)(t)=0$ for all $t\in[0,T]$. This contradicts the fact that $y_\xi(0)$ and $H_\xi(0)$ cannot vanish at the same time ($\bar{y}_\xi+\bar{H}_\xi>0$ for all $\xi\in\mathcal{A}$). If $H_\xi(t^*)<0$, then $y_{\xi tt}(t^*)<0$ and, because of (2.25) and (2.26), there exists a neighborhood $\mathcal{U}$ of t^* such that $y(t)<0$ for all $t\in\mathcal{U}\setminus\{t^*\}$. This contradicts the definition of t^*. Hence, $H_\xi(t^*)>0$ and, since we now have $y_\xi(t^*)=y_{\xi t}(t^*)=0$ and $y_{\xi tt}(t^*)>0$, there exists a neighborhood of t^* that we again denote $\mathcal{U}$ such that $y_\xi(t)>0$ for all $t\in\mathcal{U}\setminus\{t^*\}$. This contradicts the fact that $t^*<T$, and we have proved the first inequality in (2.24b), namely that $y_\xi(t)\geq 0$ for all $t\in[0,T]$. Let us prove that $H_\xi(t)\geq 0$ for all $t\in[0,T]$. This follows from (2.24c) when $y_\xi(t)>0$. Now, if $y_\xi(t)=0$, then $U_\xi(t)=0$ from (2.24c) and we have seen that $H_\xi(t)<0$ would imply that $y_\xi(t')<0$ for some t' in a punctured neighborhood of t, which is impossible. Hence, $H_\xi(t)\geq 0$ and we have proved the second inequality in (2.24b). Assume that the third inequality in (2.24c) does not hold. Then, by continuity, there exists a time $t\in[0,T]$ such that $(y_\xi+H_\xi)(t)=0$. Since y_ξ and H_ξ are positive, we must have $y_\xi(t)=H_\xi(t)=0$ and, by (2.24c), $U_\xi(t)=0$. Since zero is a solution of (2.12), this implies that $y_\xi(0)=U_\xi(0)=H_\xi(0)$, which contradicts $(y_\xi+H_\xi)(0)>0$. The fact that $\lim_{\xi\to-\infty}H(t,\xi)=0$ will be proved below in (*iii*).

(*ii*) We define the set

$$\mathcal{N}=\{(t,\xi)\in[0,T]\times\mathbb{R}\mid y_\xi(t,\xi)=0\}.$$

Fubini's theorem gives us

$$\operatorname{meas}(\mathcal{N})=\int_{\mathbb{R}}\operatorname{meas}(\mathcal{N}_\xi)\,d\xi=\int_{[0,T]}\operatorname{meas}(\mathcal{N}_t)\,dt \tag{2.28}$$

where $\mathcal{N}_\xi$ and $\mathcal{N}_t$ are the ξ-section and t-section of $\mathcal{N}$, respectively, that is,

$$\mathcal{N}_\xi=\{t\in[0,T]\mid y_\xi(t,\xi)=0\}\text{ and }\mathcal{N}_t=\{\xi\in\mathbb{R}\mid y_\xi(t,\xi)=0\}.$$

Let us prove that, for all $\xi\in\mathcal{A}$, $\operatorname{meas}(\mathcal{N}_\xi)=0$. If we consider the sets $\mathcal{N}_\xi^n$ defined as

$$\mathcal{N}_\xi^n=\{t\in[0,T]\mid y_\xi(t,\xi)=0\text{ and }y_\xi(t',\xi)>0\text{ for all }t'\in[t-1/n,t+1/n]\setminus\{t\}\},$$

then

$$\mathcal{N}_\xi=\bigcup_{n\in\mathbb{N}}\mathcal{N}_\xi^n. \tag{2.29}$$

Indeed, for all $t\in\mathcal{N}_\xi$, we have $y_\xi(t,\xi)=0$, $y_{\xi t}(t,\xi)=0$ from (2.24c) and (2.12) and $y_{\xi tt}(t,\xi)=\frac{1}{2}H_\xi(t,\xi)>0$ from (2.12) and (2.24b) (y_ξ and H_ξ cannot vanish at the same time for $\xi \in \mathcal{A}$). This implies that, on a small punctured neighborhood of t, y_ξ is strictly positive. Hence, t belongs to some $\mathcal{N}_\xi^n$ for n large enough. This proves (2.29). The set $\mathcal{N}_\xi^n$ consists of isolated points that are countable since, by definition, they are separated by a distance larger than $1/n$ from one another. This means that $\mathrm{meas}(\mathcal{N}_\xi^n) = 0$ and, by the subadditivity of the measure, $\mathrm{meas}(\mathcal{N}_\xi) = 0$. It follows from (2.28) and since $\mathrm{meas}(\mathcal{A}^c) = 0$ that

$$\mathrm{meas}(\mathcal{N}_t) = 0 \text{ for almost every } t \in [0, T]. \tag{2.30}$$

We denote by $\mathcal{K}$ the set of times such that $\mathrm{meas}(\mathcal{N}_t) > 0$, i.e.,

$$\mathcal{K} = \{t \in \mathbb{R}_+ \mid \mathrm{meas}(\mathcal{N}_t) > 0\}. \tag{2.31}$$

By (2.30), $\mathrm{meas}(\mathcal{K}) = 0$. For all $t \in \mathcal{K}^c$, $y_\xi > 0$ almost everywhere and, therefore, $y(t, \xi)$ is strictly increasing and invertible (with respect to ξ).

(*iii*) For any given $t \in [0, T]$, since $H_\xi(t, \xi) \geq 0$, $H(t, \xi)$ is an increasing function with respect to ξ and therefore, as $H(t, \cdot\,) \in L^\infty(\mathbb{R})$, $H(t, \xi)$ has a limit when $\xi \to \pm\infty$. We denote those limits $H(t, \pm\infty)$. Since $U(t, \cdot\,) \in H^1(\mathbb{R})$, we have $\lim_{\xi\to\pm\infty} U(t, \xi) = 0$ for all $t \in [0, T]$. We have

$$H(t, \xi) = H(0, \xi) + \int_0^t \left[U^3 - 2PU\right](\tau, \xi)\, d\tau. \tag{2.32}$$

We let ξ tend to $\pm\infty$. Since U and P are bounded in $L^\infty([0, T] \times \mathbb{R})$, we can apply the Lebesgue dominated convergence theorem and it follows from (2.32), as $\lim_{\xi\to\pm\infty} U(t, \xi) = 0$, that $H(t, \pm\infty) = H(0, \pm\infty)$ for all $t \in [0, T]$. Since $\bar{X} \in \mathcal{G}$, $H(0, -\infty) = 0$ and therefore $H(t, -\infty) = 0$ for all $t \in [0, T]$. □

We are now ready to prove global existence of solutions to (2.10).

Theorem 2.8. *For any $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) \in \mathcal{G}$, the system* (2.10) *admits a unique global solution $X(t) = (y(t), U(t), H(t))$ in $C^1(\mathbb{R}_+, E)$ with initial data $\bar{X} = (\bar{y}, \bar{U}, \bar{H})$. We have $X(t) \in \mathcal{G}$ for all times. If we equip $\mathcal{G}$ with the topology inducted by the E-norm, then the mapping $S\colon \mathcal{G} \times \mathbb{R}_+ \to \mathcal{G}$ defined as*

$$S_t(\bar{X}) = X(t)$$

is a continuous semigroup.

Proof. The solution has a finite time of existence T only if $\|(\zeta, U, H)(t, \cdot\,)\|_E$ blows up when t tends to T because, otherwise, by Theorem 2.3, the solution can be prolongated by a small time interval beyond T. Let (ζ, U, H) be a solution of (2.10) in $C([0, T), E)$ with initial data $(\bar{\zeta}, \bar{U}, \bar{H})$. We want to prove that

$$\sup_{t\in[0,T)} \|(\zeta(t, \cdot\,), U(t, \cdot\,), H(t, \cdot\,))\|_E < \infty. \tag{2.33}$$

We have already seen that $H(t, \xi)$ is an increasing function in ξ for all t and, from Lemma 2.7, we have $\lim_{\xi\to\infty} H(t, \xi) = \lim_{\xi\to\infty} H(0, \xi)$. This shows that

$\sup_{t\in[0,T)}\|H(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is bounded by $\|\bar{H}\|_{L^\infty(\mathbb{R})}$ and therefore is finite. To simplify the notation we suppress the dependence in t for the moment. We have

$$U^2(\xi) = 2\int_{-\infty}^{\xi} U(\eta)U_\xi(\eta)\,d\eta = 2\int_{\{\eta\leq\xi\,|\,y_\xi(\eta)>0\}} U(\eta)U_\xi(\eta)\,d\eta \tag{2.34}$$

since, from (2.24c), $U_\xi(\xi) = 0$ when $y_\xi(\xi) = 0$. For almost every ξ such that $y_\xi(\xi) > 0$, we have

$$|U(\xi)U_\xi(\xi)| = \left|\sqrt{y_\xi}U(\xi)\frac{U_\xi(\xi)}{\sqrt{y_\xi(\xi)}}\right| \leq \frac{1}{2}\left(U(\xi)^2 y_\xi(\xi) + \frac{U_\xi^2(\xi)}{y_\xi(\xi)}\right) = \frac{1}{2}H_\xi(\xi),$$

from (2.24c). Inserting this inequality in (2.34), we obtain $U^2(\xi) \leq H(\xi)$ and $\sup_{t\in[0,T)}\|U(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is therefore finite. Then, from the governing equation (2.10), it follows that

$$|\zeta(t,\xi)| \leq |\zeta(0,\xi)| + \sup_{t\in[0,T)}\|U(t,\cdot)\|_{L^\infty(\mathbb{R})}T$$

and $\sup_{t\in[0,T)}\|\zeta(t,\cdot)\|_{L^\infty(\mathbb{R})} < \infty$. Let us prove that $\sup_{t\in[0,T)}\|Q(t,\cdot)\|_{L^\infty(\mathbb{R})} < \infty$. After one integration by parts, Q can be rewritten as

$$Q(t,\xi) = -\frac{1}{4}\int_{\mathbb{R}} e^{-|y(\xi)-y(\eta)|}y_\xi(\eta)\left[\operatorname{sgn}(\xi-\eta)U(\eta)^2 - H(\eta)\right]\,d\eta - \frac{1}{2}H(t,\xi).$$

Hence, we get, after a change of variable,

$$|Q(t,\xi)| \leq C\int_{\mathbb{R}} e^{-|y(\xi)-y(\eta)|}y_\xi(\eta)\,d\eta + C \leq 3C,$$

where the constant C is a constant which depends only on $\sup_{t\in[0,T)}\|H(t,\cdot)\|_{L^\infty(\mathbb{R})}$ and $\sup_{t\in[0,T)}\|U(t,\cdot)\|_{L^\infty(\mathbb{R})}$. Similarly, one proves that that $\|P(t,\cdot)\|_{L^\infty(\mathbb{R})} < \infty$. We denote

$$C_1 = \sup_{t\in[0,T)}\{\|U(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|H(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|\zeta(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|P(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|Q(t,\cdot)\|_{L^\infty(\mathbb{R})}\}.$$

We have just proved that $C_1 < \infty$. Let $t \in [0,T)$. Looking back at (2.16) and the definition of R, we obtain that

$$\|R(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})})$$

for some constant C depending only on C_1. Since A is a continuous linear mapping from $L^2(\mathbb{R})$ to $H^1(\mathbb{R})$, we get

$$\|A\circ R(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})})$$

for another constant C which again only depends on C_1. From now on, we denote generically by C such constants that only depends on C_1. From (2.16), as $\|e^{-\zeta(t,\cdot)}\|_{L^\infty(\mathbb{R})} \leq C$, we obtain that

$$\|Q_1(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})})$$

The same bound holds for Q_2 and therefore

$$\|Q(t,\,\cdot\,)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})}). \quad (2.35)$$

Similarly, one proves

$$\|P(t,\,\cdot\,)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})}). \quad (2.36)$$

Let $Z(t) = \|U(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|U_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\,\cdot\,)\|_{L^2(\mathbb{R})}$, then the theorem will be proved once we have established that $\sup_{t\in[0,T)} Z(t) < \infty$. From the integrated version of (2.10) and (2.22), after taking the $L^2(\mathbb{R})$-norms on both sides, adding the relevant terms and using (2.36), we obtain

$$Z(t) \leq Z(0) + C\int_0^t Z(\tau)\,d\tau.$$

Hence, Gronwall's lemma gives us that $\sup_{t\in[0,T)} Z(t) < \infty$. From standard ordinary differential equation theory, we infer that S_t is a continuous semi-group.

□

3. From Eulerian to Lagrangian coordinates and vice versa

As noted in [4], even if $H^1(\mathbb{R})$ is a natural space for the equation, there is no hope to obtain a semigroup of solutions by only considering $H^1(\mathbb{R})$. Thus, we introduce the following space $\mathcal{D}$, which characterizes the solutions in *Eulerian coordinates*:

Definition 3.1. The set $\mathcal{D}$ is composed of all pairs (u,μ) such that u belongs to $H^1(\mathbb{R})$ and μ is a positive finite Radon measure whose absolute continuous part, μ_{ac}, satisfies

$$\mu_{\mathrm{ac}} = (u^2 + u_x^2)\,dx. \quad (3.1)$$

We derived the equivalent system (2.10) by using characteristics. Since y satisfies (2.2), y, for a given ξ, can also be seen as the position of a particle evolving in the velocity field u, where u is the solution of the Camassa–Holm equation. We are then working in *Lagrangian coordinates*. In [13], the Camassa–Holm equation is derived as a geodesic equation on the group of diffeomorphism equipped with a right-invariant metric. In the present paper, the geodesic curves correspond to $y(t,\,\cdot\,)$. Note that y does not remain a diffeomorphism since it can become non invertible, which agrees with the fact that the solutions of the geodesic equation may break down, see [11]. The right-invariance of the metric can be interpreted as an invariance with respect to relabeling as noted in [2]. This is a property that we also observe in our setting. We denote by G the subgroup of the group of homeomorphisms from $\mathbb{R}$ to $\mathbb{R}$ such that

$$f - \mathrm{Id} \text{ and } f^{-1} - \mathrm{Id} \text{ both belong to } W^{1,\infty}(\mathbb{R}) \quad (3.2)$$

where Id denotes the identity function. The set G can be interpreted as the set of relabeling functions. For any $\alpha > 1$, we introduce the subsets G_α of G defined by

$$G_\alpha = \{f \in G \mid \|f - \mathrm{Id}\|_{W^{1,\infty}(\mathbb{R})} + \left\|f^{-1} - \mathrm{Id}\right\|_{W^{1,\infty}(\mathbb{R})} \leq \alpha\}.$$

The subsets G_α do not possess the group structure of G. The next lemma provides a useful characterization of G_α.

Lemma 3.2. *Let $\alpha \geq 0$. If f belongs to G_α, then $1/(1+\alpha) \leq f_\xi \leq 1+\alpha$ almost everywhere. Conversely, if f is absolutely continuous, $f - \mathrm{Id} \in L^\infty(\mathbb{R})$ and there exists $c \geq 1$ such that $1/c \leq f_\xi \leq c$ almost everywhere, then $f \in G_\alpha$ for some α depending only on c and $\|f - \mathrm{Id}\|_{L^\infty(\mathbb{R})}$.*

Proof. Given $f \in G_\alpha$, let B be the set of points where f^{-1} is differentiable. Rademacher's theorem says that $\mathrm{meas}(B^c) = 0$. For any $\xi \in f^{-1}(B)$, we have

$$\lim_{\xi' \to \xi} \frac{f^{-1}(f(\xi')) - f^{-1}(f(\xi))}{f(\xi') - f(\xi)} = (f^{-1})_\xi(f(\xi))$$

because f is continuous and f^{-1} is differentiable at $f(\xi)$. On the other hand, we have

$$\frac{f^{-1}(f(\xi')) - f^{-1}(f(\xi))}{f(\xi') - f(\xi)} = \frac{\xi' - \xi}{f(\xi') - f(\xi)}.$$

Hence, f is differentiable for any $\xi \in f^{-1}(B)$ and

$$f_\xi(\xi) \geq \frac{1}{\|(f^{-1})_\xi\|_{L^\infty(\mathbb{R})}} \geq \frac{1}{1+\alpha}. \tag{3.3}$$

The estimate (3.3) holds only on $f^{-1}(B)$ but, since $\mathrm{meas}(B^c) = 0$ and f^{-1} is Lipschitz and one-to-one, $\mathrm{meas}(f^{-1}(B^c)) = 0$ (see, e.g., [1, Remark 2.72]), and therefore (3.3) holds almost everywhere. We have $f_\xi \leq 1 + \|f_\xi - 1\|_{L^\infty(\mathbb{R})} \leq 1+\alpha$.

Let us now consider a function f which is absolutely continuous and such that $f - \mathrm{Id} \in L^\infty(\mathbb{R})$ and $1/c \leq f_\xi \leq c$ almost everywhere for some $c \geq 1$. Since f_ξ is bounded, f and therefore $f - \mathrm{Id}$ are Lipschitz and $f - \mathrm{Id} \in W^{1,\infty}(\mathbb{R})$. Since $f_\xi \geq 1/c$ almost everywhere, f is strictly increasing and, since it is also continuous, it is invertible. As f is Lipschitz, we can make the following change of variables (see, for example, [1]) and get that, for all ξ_1, ξ_2 in $\mathbb{R}$ such that $\xi_1 < \xi_2$,

$$f^{-1}(\xi_2) - f^{-1}(\xi_1) = \int_{[f^{-1}(\xi_1), f^{-1}(\xi_2)]} \frac{f_\xi}{f_\xi}\, d\xi \leq c(\xi_2 - \xi_1).$$

Hence, f^{-1} is Lipschitz and $(f^{-1})_\xi \leq c$. We have $f^{-1}(\xi') - \xi' = \xi - f(\xi)$ for $\xi' = f(\xi)$ and therefore $\|f - \mathrm{Id}\|_{L^\infty(\mathbb{R})} = \left\|f^{-1} - \mathrm{Id}\right\|_{L^\infty(\mathbb{R})}$. Finally, we get

$$\begin{aligned}
\|f - \mathrm{Id}\|_{W^{1,\infty}(\mathbb{R})} + \left\|f^{-1} - \mathrm{Id}\right\|_{W^{1,\infty}(\mathbb{R})} &\leq 2\|f - \mathrm{Id}\|_{W^{1,\infty}(\mathbb{R})} + 2 \\
&\quad + \|f_\xi\|_{L^\infty(\mathbb{R})} + \left\|(f^{-1})_\xi\right\|_{L^\infty(\mathbb{R})} \\
&\leq 2\|f - \mathrm{Id}\|_{L^\infty(\mathbb{R})} + 2 + 2c.
\end{aligned}$$

□

We define the subsets $\mathcal{F}_\alpha$ and $\mathcal{F}$ of $\mathcal{G}$ as follows

$$\mathcal{F}_\alpha = \{X = (y, U, H) \in \mathcal{G} \mid y + H \in G_\alpha\},$$

and

$$\mathcal{F} = \{X = (y, U, H) \in \mathcal{G} \mid y + H \in G\}.$$

For $\alpha = 0$, $G_0 = \{\mathrm{Id}\}$. As we will see, the space $\mathcal{F}_0$ will play a special role. These sets are relevant only because they are in some sense preserved by the governing equation (2.10) as the next lemma shows.

Lemma 3.3. *The space $\mathcal{F}$ is preserved by the governing equation* (2.10)*. More precisely, given $\alpha, T \geq 0$ and $\bar{X} \in \mathcal{F}_\alpha$, we have*

$$S_t(\bar{X}) \in \mathcal{F}_{\alpha'}$$

for all $t \in [0, T]$ where α' only depends on T, α and $\|\bar{X}\|_E$.

Proof. Let $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) \in \mathcal{F}_\alpha$, we denote $X(t) = (y(t), U(t), H(t))$ the solution of (2.10) with initial data $\bar{X}$ and set $h(t, \xi) = y(t, \xi) + H(t, \xi)$, $\bar{h}(\xi) = \bar{y}(\xi) + \bar{H}(\xi)$. By definition, we have $\bar{h} \in G_\alpha$ and, from Lemma 3.2, $1/c \leq \bar{h}_\xi \leq c$ almost everywhere, for some constant $c > 1$ depending only α. We consider a fixed ξ and drop it in the notation. Applying Gronwall's inequality backward in time to (2.12), we obtain

$$|y_\xi(0)| + |H_\xi(0)| + |U_\xi(0)| \leq e^{CT} \left(|y_\xi(t)| + |H_\xi(t)| + |U_\xi(t)|\right) \tag{3.4}$$

for some constant C which depends on $\|X(t)\|_{C([0,T],E)}$, which itself depends only on $\|\bar{X}\|_E$ and T. From (2.24c), we have

$$|U_\xi(t)| \leq \sqrt{y_\xi(t) H_\xi(t)} \leq \frac{1}{2}(y_\xi(t) + H_\xi(t)).$$

Hence, since y_ξ and H_ξ are positive, (3.4) gives us

$$\frac{1}{c} \leq \bar{y}_\xi + \bar{H}_\xi \leq \frac{3}{2} e^{CT}(y_\xi(t) + H_\xi(t)),$$

and $h_\xi(t) = y_\xi(t) + H_\xi(t) \geq \frac{2}{3c} e^{-CT}$. Similarly, by applying Gronwall's lemma forward in time, we obtain $y_\xi(t) + H_\xi(t) \leq \frac{3}{2} c e^{CT}$. We have $\|(y + H)(t) - \xi\|_{L^\infty(\mathbb{R})} \leq \|X(t)\|_{C([0,T],E)} \leq C$ for another constant C which also only depends on $\|\bar{X}\|_E$ and T. Hence, applying Lemma 3.2, we obtain that $y(t, \cdot\,) + H(t, \cdot\,) \in G_{\alpha'}$ and therefore $X(t) \in \mathcal{F}_{\alpha'}$ for some α' depending only on α, T and $\|\bar{X}\|_E$. $\square$

For the sake of simplicity, for any $X = (y, U, H) \in \mathcal{F}$ and any function $f \in G$, we denote $(y \circ f, U \circ f, H \circ f)$ by $X \circ f$.

Proposition 3.4. *The map from $G \times \mathcal{F}$ to $\mathcal{F}$ given by $(f, X) \mapsto X \circ f$ defines an action of the group G on $\mathcal{F}$.*

Proof. We have to prove that $X \circ f$ belongs to $\mathcal{F}$ for any $X = (y, U, H) \in \mathcal{F}$ and $f \in G$. We denote $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) = X \circ f$. As compositions of two Lipschitz maps, $\bar{y}$, $\bar{U}$ and $\bar{H}$ are Lipschitz. We have

$$\begin{aligned}
\|\bar{y} - \mathrm{Id}\|_{L^\infty(\mathbb{R})} &\leq \|\bar{y} \circ f - f\|_{L^\infty(\mathbb{R})} + \|f - \mathrm{Id}\|_{L^\infty(\mathbb{R})} \\
&\leq \|\bar{y} - \mathrm{Id}\|_{L^\infty(\mathbb{R})} + \|f - \mathrm{Id}\|_{L^\infty(\mathbb{R})} < +\infty.
\end{aligned}$$

Hence, $(\bar y-\mathrm{Id},\bar U,\bar H)\in[W^{1,\infty}(\mathbb{R})]^3$. Let us prove that

$$\bar y_\xi=y_\xi\circ f\,f_\xi,\ \bar U_\xi=U_\xi\circ f\,f_\xi \text{ and } \bar H_\xi=H_\xi\circ f\,f_\xi \tag{3.5}$$

almost everywhere. Let B_1 be the set where y is differentiable and B_2 the set where $\bar y$ and f are differentiable. Using Radamacher's theorem, we get that $\mathrm{meas}(B_1^c)=\mathrm{meas}(B_2^c)=0$. For $\xi\in B_3=B_2\cap f^{-1}(B_1)$, we consider a sequence ξ_i converging to ξ ($\xi_i\neq\xi$). We have

$$\frac{y(f(\xi_i))-y(f(\xi))}{f(\xi_i)-f(\xi)}\,\frac{f(\xi_i)-f(\xi)}{\xi_i-\xi}=\frac{\bar y(\xi_i)-\bar y(\xi)}{\xi_i-\xi}. \tag{3.6}$$

Since f is continuous, $f(\xi_i)$ converges to $f(\xi)$ and, as y is differentiable at $f(\xi)$, the left-hand side of (3.6) tends to $y_\xi\circ f(\xi)\,f_\xi(\xi)$. The right-hand side of (3.6) tends to $\bar y_\xi(\xi)$, and we get that

$$y_\xi(f(\xi))f_\xi(\xi)=\bar y_\xi(\xi) \tag{3.7}$$

for all $\xi\in B_3$. Since f^{-1} is Lipschitz, one-to-one and $\mathrm{meas}(B_1^c)=0$, we have $\mathrm{meas}(f^{-1}(B_1)^c)=0$ and therefore (3.7) holds everywhere. One proves the two other identities in (3.5) similarly. From Lemma 3.2, we have that $f_\xi>0$ almost everywhere. Then, using (3.5) we easily check that (2.24b) and (2.24c) are fulfilled. Thus, we have proved that $(\bar y-\mathrm{Id},\bar U,\bar H)$ fulfills (2.24). It remains to prove that $(\bar y-\mathrm{Id},\bar U,\bar H)\in E$. Since $f\in G$, $f\in G_\alpha$ for some large enough α and, by Lemma 3.2, there exists a constant $c>0$ such that $1/c\leq f_\xi\leq c$ almost everywhere. We have, after a change of variables,

$$\left\|\bar U\right\|^2_{L^2(\mathbb{R})}=\int_{\mathbb{R}}(U\circ f\,)^2\,d\xi\leq c\int_{\mathbb{R}}(U\circ f\,)^2 f_\xi\,d\xi=c\left\|U\right\|^2_{L^2(\mathbb{R})}.$$

Hence, $\bar U\in L^2(\mathbb{R})$. Similarly, one proves that $y_\xi-1$, U_ξ and H_ξ belong to $L^2(\mathbb{R})$ and therefore $(y,U,H)\in\mathcal{G}$. We have $\bar y+\bar H=(y+H)\circ f$ which implies, since $y+H$ and f belongs to G and G is a group, that $\bar y+\bar H\in G$. Therefore $\bar X\in\mathcal{F}$ and the proposition is proved. □

Since G is acting on $\mathcal{F}$, we can consider the quotient space $\mathcal{F}/G$ of $\mathcal{F}$ with respect to the action of the group G. The equivalence relation on $\mathcal{F}$ is defined as follows: For any $X,X'\in\mathcal{F}$, X and X' are equivalent if there exists $f\in G$ such that $X'=X\circ f$. We denote by $\Pi(X)=[X]$ the projection of $\mathcal{F}$ into the quotient space $\mathcal{F}/G$. We introduce the mapping $\Gamma\colon\mathcal{F}\to\mathcal{F}_0$ given by

$$\Gamma(X)=X\circ(\,y+H)^{-1}$$

for any $X=(y,U,H)\in\mathcal{F}$. We have $\Gamma(X)=X$ when $X\in\mathcal{F}_0$. It is not hard to prove Γ is invariant under the G action, that is, $\Gamma(X\circ f)=\Gamma(X)$ for any $X\in\mathcal{F}$ and $f\in G$. Hence, there corresponds to Γ a mapping $\tilde\Gamma$ from the quotient space $\mathcal{F}/G$ to $\mathcal{F}_0$ given by $\tilde\Gamma([X])=\Gamma(X)$ where $[X]\in\mathcal{F}/G$ denotes the equivalence class of $X\in\mathcal{F}$. For any $X\in\mathcal{F}_0$, we have $\tilde\Gamma\circ\Pi(X)=\Gamma(X)=X$. Hence, $\tilde\Gamma\circ\Pi_{|\mathcal{F}_0}=\mathrm{Id}_{|\mathcal{F}_0}$. Any topology defined on $\mathcal{F}_0$ is naturally transported into $\mathcal{F}/G$ by this isomorphism. We equip $\mathcal{F}_0$ with the metric induced by the E-norm, i.e.,

$d_{\mathcal{F}_0}(X, X') = \|X - X'\|_E$ for all $X, X' \in \mathcal{F}_0$. Since $\mathcal{F}_0$ is closed in E, this metric is complete. We define the metric on $\mathcal{F}/G$ as

$$d_{\mathcal{F}/G}([X], [X']) = \|\Gamma(X) - \Gamma(X')\|_E\,,$$

for any $[X], [X'] \in \mathcal{F}/G$. Then, $\mathcal{F}/G$ is isometrically isomorphic with $\mathcal{F}_0$ and the metric $d_{\mathcal{F}/G}$ is complete.

Lemma 3.5. *Given $\alpha \geq 0$. The restriction of Γ to $\mathcal{F}_\alpha$ is a continuous mapping from $\mathcal{F}_\alpha$ to $\mathcal{F}_0$.*

Remark 3.6. The mapping Γ is not continuous from $\mathcal{F}$ to $\mathcal{F}_0$. The spaces $\mathcal{F}_\alpha$ were precisely introduced in order to make the mapping Γ continuous.

Proof. As for $\mathcal{F}_0$, we equip $\mathcal{F}_\alpha$ with the topology induced by the E-norm. Let $X_n = (y_n, U_n, H_n) \in \mathcal{F}_\alpha$ be a sequence that converges to $X = (y, U, H)$ in $\mathcal{F}_\alpha$. We denote $\bar{X}_n = (\bar{y}_n, \bar{U}_n, \bar{H}_n) = \Gamma(X_n)$ and $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) = \Gamma(X)$. By definition of $\mathcal{F}_0$, we have $\bar{H}_n = -\bar{\zeta}_n$ (recall that $\zeta_n = y_n - \mathrm{Id}$). Let us prove first that $\bar{H}_n$ tends to $\bar{H}$ in $L^\infty(\mathbb{R})$. We denote $f_n = y_n + H_n$, $f = y + H$, and we have $f_n, f \in G_\alpha$. Thus $\bar{H}_n - \bar{H} = (H_n - H) \circ {f_n}^{-1} + \bar{H} \circ f \circ {f_n}^{-1} - \bar{H}$ and we have

$$\left\|\bar{H}_n - \bar{H}\right\|_{L^\infty(\mathbb{R})} \leq \|H_n - H\|_{L^\infty(\mathbb{R})} + \left\|\bar{H} \circ f - \bar{H} \circ f_n\right\|_{L^\infty(\mathbb{R})}. \tag{3.8}$$

From the definition of $\mathcal{F}_0$, we know that $\bar{H}$ is Lipschitz with Lipschitz constant smaller than one. Hence,

$$\left\|\bar{H} \circ f - \bar{H} \circ f_n\right\|_{L^\infty(\mathbb{R})} \leq \|f_n - f\|_{L^\infty(\mathbb{R})}. \tag{3.9}$$

Since H_n and f_n converges to H and f, respectively, in $L^\infty(\mathbb{R})$, from (3.8) and (3.9), we get that $\bar{H}_n$ converges to $\bar{H}$ in $L^\infty(\mathbb{R})$. Let us prove now that $\bar{H}_{n,\xi}$ tend to $\bar{H}_\xi$ in $L^2(\mathbb{R})$. We have $\bar{H}_{n,\xi} - \bar{H}_\xi = \frac{H_{n,\xi}}{f_{n,\xi}} \circ {f_n}^{-1} - \frac{H_\xi}{f_\xi} \circ f^{-1}$ which can be decomposed into

$$\bar{H}_{n,\xi} - \bar{H}_\xi = \left(\frac{H_{n,\xi} - H_\xi}{f_{n,\xi}}\right) \circ {f_n}^{-1} + \frac{H_\xi}{f_{n,\xi}} \circ {f_n}^{-1} - \frac{H_\xi}{f_\xi} \circ f^{-1}. \tag{3.10}$$

Since $f_n \in G_\alpha$, there exists a constant $c > 0$ independent of n such that $1/c \geq f_{n,\xi} \geq c$ almost everywhere, see Lemma 3.2. We have

$$\left\|\left(\frac{H_{n,\xi} - H_\xi}{f_{n,\xi}}\right) \circ {f_n}^{-1}\right\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} (H_{n,\xi} - H_\xi)^2 \frac{1}{f_{n,\xi}}\, d\xi \leq c \left\|H_{n,\xi} - H_\xi\right\|^2_{L^2(\mathbb{R})}, \tag{3.11}$$

where we have made the change of variables $\xi' = {f_n}^{-1}(\xi)$. Hence, the left-hand side of (3.11) converges to zero. If we can prove that $\frac{H_\xi}{f_{n,\xi}} \circ {f_n}^{-1} \to \frac{H_\xi}{f_\xi} \circ f^{-1}$ in $L^2(\mathbb{R})$, then, using (3.10), we get that $\bar{H}_{n,\xi} \to \bar{H}_\xi$ in $L^2(\mathbb{R})$, which is the desired result. We have

$$\frac{H_\xi}{f_{n,\xi}} \circ {f_n}^{-1} = \frac{(\bar{H}_\xi \circ f) f_\xi}{f_{n,\xi}} \circ {f_n}^{-1} = (\bar{H}_\xi \circ g_n) g_{n,\xi}$$

where $g_n = f\circ {f_n}^{-1}$. Let us prove that $\lim_{n\to\infty}\|g_{n,\xi}-1\|_{L^2(\mathbb{R})}=0$. We have, after using a change of variables,

$$\|g_{n,\xi}-1\|^2_{L^2(\mathbb{R})}=\int_{\mathbb{R}}\left(\frac{f_\xi}{f_{n,\xi}}\circ {f_n}^{-1}-1\right)^2 d\xi = c\,\|f_\xi-f_{n,\xi}\|^2_{L^2(\mathbb{R})}. \tag{3.12}$$

Hence, since $f_{n,\xi}\to f_\xi$ in $L^2(\mathbb{R})$, $\lim_{n\to\infty}\|g_{n,\xi}-1\|_{L^2(\mathbb{R})}=0$. We have

$$\left\|\bar H_\xi\circ g_n g_{n,\xi}-\bar H_\xi\right\|_{L^2(\mathbb{R})}\le\left\|\bar H_\xi\circ g_n\right\|_{L^\infty(\mathbb{R})}\|g_{n,\xi}-1\|_{L^2(\mathbb{R})}+\left\|\bar H_\xi\circ g_n-\bar H_\xi\right\|_{L^2(\mathbb{R})}. \tag{3.13}$$

We have $\left\|\bar H_\xi\circ g_n\right\|_{L^\infty(\mathbb{R})}\le 1$ since, as we already noted, $\bar H$ is Lipschitz with Lipschitz constant smaller than one. Hence, the first term in the sum in (3.13) converges to zero. As far as the second term is concerned, one can always approximate $\bar H_\xi$ in $L^2(\mathbb{R})$ by a continuous function h with compact support. After observing that $1/c^2\le g_{n,\xi}\le c^2$ almost everywhere, we can prove, as we have done several times now, that $\|H_\xi\circ g_n-h\circ g_n\|^2_{L^2(\mathbb{R})}\le c^2\|H_\xi-h\|^2_{L^2(\mathbb{R})}$ and $h\circ g_n$ can be chosen arbitrarily close to $H_\xi\circ g_n$ in $L^2(\mathbb{R})$ independently of n, that is, for all $\varepsilon>0$, there exists h such that

$$\|H_\xi\circ g_n-h\circ g_n\|_{L^2(\mathbb{R})}\le\frac{\varepsilon}{3}\ \text{ and }\ \|H_\xi-h\|_{L^2(\mathbb{R})}\le\frac{\varepsilon}{3} \tag{3.14}$$

for all n. Since $f_n\to f$ in $L^\infty(\mathbb{R})$, $g_n\to \mathrm{Id}$ in $L^\infty(\mathbb{R})$ and there exists a compact K independent of n such that $\operatorname{supp}(h\circ g_n)\subset K$. Then, by the Lebesgue dominated convergence theorem, we obtain that $h\circ g_n\to h$ in $L^2(\mathbb{R})$. Hence, for n large enough, we have $\|h\circ g_n-h\|_{L^2(\mathbb{R})}\le\frac{\varepsilon}{3}$ which, together with (3.14), implies $\left\|\bar H_\xi\circ g_n-\bar H_\xi\right\|_{L^2(\mathbb{R})}\le\varepsilon$, and $\bar H_\xi\circ g_n\to\bar H_\xi$ in $L^2(\mathbb{R})$. From (3.10), (3.11), (3.12) and (3.13), we obtain that $\bar H_{n,\xi}\to\bar H_\xi$ in $L^2(\mathbb{R})$. It follows that $\bar\zeta_{n,\xi}\to\bar\zeta_\xi$ in $L^2(\mathbb{R})$ and, similarly, one proves that $\bar U_{n,\xi}\to\bar U_\xi$ in $L^2(\mathbb{R})$. It remains to prove that $U_n\to U$ in $L^2(\mathbb{R})$. We write

$$\bar U_n-\bar U=(U_n-U)\circ {f_n}^{-1}+U\circ {f_n}^{-1}-U\circ f^{-1}. \tag{3.15}$$

We have, after a change of variable,

$$\left\|(U_n-U)\circ {f_n}^{-1}\right\|^2_{L^2(\mathbb{R})}=\int_{\mathbb{R}}(U_n-U)^2 f_{n,\xi}\,d\xi\le c\,\|U_n-U\|^2_{L^2(\mathbb{R})}. \tag{3.16}$$

We also have, after the same change of variable, that

$$\left\|U\circ {f_n}^{-1}-U\circ f^{-1}\right\|^2_{L^2(\mathbb{R})}\le c\int_{\mathbb{R}}(U-U\circ f^{-1}\circ f_n)^2\,d\xi. \tag{3.17}$$

By approximating U by continuous functions with compact support as we did before, we prove that $\int_{\mathbb{R}}(U-U\circ f^{-1}\circ f_n)^2$ tends to zero. Hence, by (3.15), (3.16) and (3.17), we get that $\bar U_n\to U$ in $L^2(\mathbb{R})$, which concludes the proof of the lemma. ∎

3.1. **Continuous semigroup of solutions in $\mathcal{F}/G$.** We denote by $S\colon \mathcal{F}\times\mathbb{R}_+\to\mathcal{F}$ the continuous semigroup which to any initial data $\bar{X}\in\mathcal{F}$ associates the solution $X(t)$ of the system of differential equation (2.10) at time t. As we indicated earlier, the Camassa–Holm equation is invariant with respect to relabeling, more precisely, using our terminology, we have the following result.

Theorem 3.7. *For any $t>0$, the mapping $S_t\colon\mathcal{F}\to\mathcal{F}$ is G-equivariant, that is,*

$$S_t(X\circ f)=S_t(X)\circ f \tag{3.18}$$

for any $X\in\mathcal{F}$ and $f\in G$. Hence, the mapping $\tilde{S}_t$ from $\mathcal{F}/G$ to $\mathcal{F}/G$ given by

$$\tilde{S}_t([X])=[S_tX]$$

is well-defined. It generates a continuous semigroup.

Proof. For any $X_0=(y_0,U_0,H_0)\in\mathcal{F}$ and $f\in G$, we denote $\bar{X}_0=(\bar{y}_0,\bar{U}_0,\bar{H}_0)=X_0\circ f$, $X(t)=S_t(X_0)$ and $\bar{X}(t)=S_t(\bar{X}_0)$. We claim that $X(t)\circ f$ satisfies (2.10) and therefore, since $X(t)\circ f$ and $\bar{X}(t)$ satisfy the same system of differential equation with the same initial data, they are equal. We denote $\hat{X}(t)=(\hat{y}(t),\hat{U}(t),\hat{H}(t))=X(t)\circ f$. We have

$$\hat{U}_t=\frac{1}{4}\int_{\mathbb{R}}\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(\hat{y}(\xi)-y(\eta))\big)\left[U(\eta)^2y_\xi(\eta)+H_\xi(\eta)\right]\,d\eta. \tag{3.19}$$

We have $\hat{y}_\xi(\xi)=y_\xi(f(\xi))f_\xi(\xi)$ and $\hat{H}_\xi(\xi)=H_\xi(f(\xi))f_\xi(\xi)$ for almost every $\xi\in\mathbb{R}$. Hence, after the change of variable $\eta=f(\eta')$, we get from (3.19) that

$$\hat{U}_t=\frac{1}{4}\int_{\mathbb{R}}\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(\hat{y}(\xi)-\hat{y}(\eta))\big)\left[\hat{U}(\eta)^2\hat{y}_\xi(\eta)+\hat{H}_\xi(\eta)\right]\,d\eta.$$

We treat similarly the other terms in (2.10), and it follows that $(\hat{y},\hat{U},\hat{H})$ is a solution of (2.10). Since $(\hat{y},\hat{U},\hat{H})$ and $(\bar{y},\bar{U},\bar{H})$ satisfy the same system of ordinary differential equations with the same initial data, they are equal, i.e., $\bar{X}(t)=X(t)\circ f$ and (3.18) is proved. We have the following diagram:

$$\begin{tikzcd}
\mathcal{F}_0 \arrow[r, "\Pi"] & \mathcal{F}/G \\
\mathcal{F}_\alpha \arrow[u, "\Gamma"] & \\
\mathcal{F}_0 \arrow[u, "S_t"] \arrow[r, "\Pi"] & \mathcal{F}/G \arrow[uu, "\tilde{S}_t"']
\end{tikzcd} \tag{3.20}$$

on a bounded domain of $\mathcal{F}_0$ whose diameter together with t determines the constant α, see Lemma 3.3. By the definition of the metric on $\mathcal{F}/G$, the mapping Π is an isometry from $\mathcal{F}_0$ to $\mathcal{F}/G$. Hence, from the diagram (3.20), we see that $\tilde{S}_t\colon\mathcal{F}/G\to\mathcal{F}/G$ is continuous if and only if $\Gamma\circ S_t\colon\mathcal{F}_0\to\mathcal{F}_0$ is continuous. Let us prove that $\Gamma\circ S_t\colon\mathcal{F}_0\to\mathcal{F}_0$ is sequentially continuous. We consider a sequence $X_n\in\mathcal{F}_0$ that converges to $X\in\mathcal{F}_0$ in $\mathcal{F}_0$, that is, $\lim_{n\to\infty}\|X_n-X\|_E=0$. From

Theorem 2.8, we get that $\lim_{n\to\infty}\|S_t(X_n)-S_t(X)\|_E=0$. Since $X_n\to X$ in E, there exists a constant $C\geq 0$ such that $\|X_n\|\leq C$ for all n. Lemma 3.3 gives us that $S_t(X_n)\in\mathcal{F}_\alpha$ for some α which depends on C and t. Hence, $S_t(X_n)\to S_t(X)$ in $\mathcal{F}_\alpha$. Then, by Lemma 3.5, we obtain that $\Gamma\circ S_t(X_n)\to\Gamma\circ S_t(X)$ in $\mathcal{F}_0$. □

3.2. **Mappings between the two coordinate systems.** Our next task is to derive the correspondence between Eulerian coordinates (functions in $\mathcal{D}$) and Lagrangian coordinates (functions in $\mathcal{F}/G$). Earlier we considered initial data in $\mathcal{D}$ with a special structure: The energy density μ was given by $(u^2+u_x^2)\,dx$ and therefore μ did not have any singular part. The set $\mathcal{D}$ however allows the energy density to have a singular part and a positive amount of energy can concentrate on a set of Lebesgue measure zero. We constructed corresponding initial data in $\mathcal{F}_0$ by the means of (2.23). This construction can be generalized in the following way. Let us denote by $L\colon\mathcal{D}\to\mathcal{F}/G$ the mapping transforming Eulerian coordinates into Lagrangian coordinates whose definition is contained in the following theorem.

Theorem 3.8. *For any (u,μ) in $\mathcal{D}$, let*

$$y(\xi)=\sup\{y\mid\mu((-\infty,y))+y<\xi\}\,,\tag{3.21a}$$

$$H(\xi)=\xi-y(\xi),\tag{3.21b}$$

$$U(\xi)=u\circ y(\xi)\,.\tag{3.21c}$$

Then $(y,U,H)\in\mathcal{F}_0$. We define $L(u,\mu)\in\mathcal{F}/G$ to be the equivalence class of (y,U,H).

Proof. Clearly (3.21a) implies that y is increasing and $\lim_{\xi\to\pm\infty}y(\xi)=\pm\infty$. For any $z>y(\xi)$, we have $\xi\leq z+\mu((-\infty,z))$. Hence, $\xi-z\leq\mu(\mathbb{R})$ and, since we can choose z arbitrarily close to $y(\xi)$, we get $\xi-y(\xi)\leq\mu(\mathbb{R})$. It is not hard to check that $y(\xi)\leq\xi$. Hence,

$$|y(\xi)-\xi|\leq\mu(\mathbb{R})\tag{3.22}$$

and $\|y-\mathrm{Id}\|_{L^\infty(\mathbb{R})}\leq\mu(\mathbb{R})$ and $y-\mathrm{Id}\in L^\infty(\mathbb{R})$. Let us prove that y is Lipschitz with Lipschitz constant at most one. We consider ξ, ξ' in $\mathbb{R}$ such that $\xi<\xi'$ and $y(\xi)<y(\xi')$ (the case $y(\xi)=y(\xi')$ is straightforward). It follows from the definition that there exists an increasing sequence, x_i', and a decreasing one, x_i such that $\lim_{i\to\infty}x_i=y(\xi)$, $\lim_{i\to\infty}x_i'=y(\xi')$ with $\mu((-\infty,x_i'))+x_i'<\xi'$ and $\mu((-\infty,x_i))+x_i\geq\xi$. Subtracting the these two inequalities one to the other, we obtain

$$\mu((-\infty,x_i'))-\mu((-\infty,x_i))+x_i'-x_i<\xi'-\xi.\tag{3.23}$$

For i large enough, since by assumption $y(\xi)<y(\xi')$, we have $x_i<x_i'$ and therefore $\mu((-\infty,x_i'))-\mu((-\infty,x_i))=\mu([x_i,x_i'))\geq 0$. Hence, $x_i'-x_i<\xi'-\xi$. Letting i tend to infinity, we get $y(\xi')-y(\xi)\leq\xi'-\xi$. Hence, y is Lipschitz with Lipschitz constant bounded by one and, by Rademacher's theorem, differentiable almost everywhere. Following [19], we decompose μ into its absolute continuous, singular continuous and singular part, denoted μ_{ac}, μ_{sc} and μ_{s}, respectively. Here, since $(u,\mu)\in\mathcal{D}$, we have $\mu_{\mathrm{ac}}=(u^2+u_x^2)\,dx$. The support of μ_{s} consists of a countable set of points. Let $F(x)=\mu((-\infty,x))$, then F is lower semi-continuous and its points of continuity exactly coincide with the support of μ_{s} (see [19]). Let A denote the complement of $y^{-1}(\mathrm{supp}(\mu_{\mathrm{s}}))$. We claim that for any $\xi\in A$, we have

$$\mu((-\infty,y(\xi)))+y(\xi)=\xi. \tag{3.24}$$

From the definition of $y(\xi)$ follows the existence of an increasing sequence x_i which converges to $y(\xi)$ and such that $F(x_i)+x_i<\xi$. Since F is lower semi-continuous, $\lim_{i\to\infty}F(x_i)=F(y(\xi))$ and therefore

$$F(y(\xi))+y(\xi)\leq\xi. \tag{3.25}$$

Let us assume that $F(y(\xi))+y(\xi)<\xi$. Since $y(\xi)$ is a point of continuity of F, we can then find an x such that $x>y(\xi)$ and $F(x)+x<\xi$. This contradicts the definition of $y(\xi)$ and proves our claim (3.24). In order to check that (2.24c) is satisfied, we have to compute y_ξ and U_ξ. We define the set B_1 as

$$B_1=\left\{x\in\mathbb{R}\mid\lim_{\rho\downarrow0}\frac{1}{2\rho}\mu((x-\rho,x+\rho))=(u^2+u_x^2)(x)\right\}.$$

Since $(u^2+u_x^2)\,dx$ is the absolutely continuous part of μ, we have, from Besicovitch's derivation theorem (see [1]), that $\mathrm{meas}(B_1^c)=0$. Given $\xi\in y^{-1}(B_1)$, we denote $x=y(\xi)$. We claim that for all $i\in\mathbb{N}$, there exists $0<\rho<\frac{1}{i}$ such that $x-\rho$ and $x+\rho$ both belong to $\mathrm{supp}(\mu_{\mathrm{s}})^c$. Assume namely the opposite. Then for any $z\in(x-\frac{1}{i},x+\frac{1}{i})\setminus\mathrm{supp}(\mu_{\mathrm{s}})$, we have that $z'=2x-z$ belongs to $\mathrm{supp}(\mu_{\mathrm{s}})$. Thus we can construct an injection between the uncountable set $(x-\frac{1}{i},x+\frac{1}{i})\setminus\mathrm{supp}(\mu_{\mathrm{s}})$ and the countable set $\mathrm{supp}(\mu_{\mathrm{s}})$. This is impossible, and our claim is proved. Hence, since y is surjective, we can find two sequences ξ_i and ξ_i' in A such that $\frac{1}{2}(y(\xi_i)+y(\xi_i'))=y(\xi)$ and $y(\xi_i')-y(\xi_i)<\frac{1}{i}$. We have, by (3.24), since $y(\xi_i)$ and $y(\xi_i')$ belong to A,

$$\mu([y(\xi_i),y(\xi_i')))+y(\xi_i')-y(\xi_i)=\xi_i'-\xi_i. \tag{3.26}$$

Since $y(\xi_i)\notin\mathrm{supp}(\mu_{\mathrm{s}})$, $\mu(\{y(\xi_i)\})=0$ and $\mu([y(\xi_i),y(\xi_i')))=\mu((y(\xi_i),y(\xi_i')))$. Dividing (3.26) by $\xi_i'-\xi_i$ and letting i tend to ∞, we obtain

$$y_\xi(\xi)(u^2+u_x^2)(y(\xi))+y_\xi(\xi)=1 \tag{3.27}$$

where y is differentiable in $y^{-1}(B_1)$, that is, almost everywhere in $y^{-1}(B_1)$. We now derive a short lemma which will be useful several times in this proof.

Lemma 3.9. *Given a Lipschitz function $f\colon\mathbb{R}\to\mathbb{R}$, for any set B of measure zero, we have $f_\xi=0$ almost everywhere in $f^{-1}(B)$.*

Proof of Lemma 3.9. The Lemma follows directly from the area formula:

$$\int_{f^{-1}(B)}f_\xi(\xi)\,d\xi=\int_{\mathbb{R}}\mathcal{H}^0\left(f^{-1}(B)\cap f^{-1}(\{x\})\right)\,dx \tag{3.28}$$

where $\mathcal{H}^0$ is the multiplicity function, see [1] for the formula and the precise definition of $\mathcal{H}^0$. The function $\mathcal{H}^0\left(f^{-1}(B)\cap f^{-1}(\{x\})\right)$ is Lebesgue measurable (see [1]) and it vanishes on B^c. Hence, $\int_{f^{-1}(B)} f_\xi\,d\xi = 0$ and therefore, since $f_\xi \geq 0$, $f_\xi = 0$ almost everywhere in $f^{-1}(B)$. □

We apply Lemma 3.9 to B_1^c and get, since $\mathrm{meas}(B_1^c) = 0$, that $y_\xi = 0$ almost everywhere on $y^{-1}(B_1^c)$. On $y^{-1}(B_1)$, we proved that y_ξ satisfies (3.27). It follows that $0 \leq y_\xi \leq 1$ almost everywhere, which implies, since $H_\xi = 1 - y_\xi$, that $H_\xi \geq 0$. In the same way as we proved that y was Lipschitz with Lipschitz constant at most one, we can prove that the function $\xi \mapsto \int_{-\infty}^{y(\xi)} u_x^2\,dx$ is also Lipschitz with Lipschitz constant at most one. Indeed, from (3.23), for i large enough, we have

$$\int_{x_i}^{x_i'} u_x^2\,dx \leq \mu_{\mathrm{ac}}([x_i, x_i')) \leq \mu([x_i, x_i')) < \xi' - \xi.$$

Since $\lim_{i\to\infty} x_i' = y(\xi')$ and $\lim_{i\to\infty} x_i = y(\xi)$, letting i tend to infinity, we obtain $\int_{y(\xi)}^{y(\xi')} u_x^2\,dx < \xi' - \xi$ and the function $\xi \mapsto \int_{-\infty}^{y(\xi)} u_x^2\,dx$ is Lipchitz with Lipschitz coefficient at most one. For all $(\xi, \xi') \in \mathbb{R}^2$, we have, after using the Cauchy–Schwarz inequality,

$$\begin{aligned}
|U(\xi') - U(\xi)| &= \int_{y(\xi)}^{y(\xi')} u_x\,dx \\
&\leq \sqrt{y(\xi') - y(\xi)}\sqrt{\int_{y(\xi)}^{y(\xi')} u_x^2\,dx} \\
&\leq |\xi' - \xi|
\end{aligned} \tag{3.29}$$

because y and $\int_{-\infty}^{y(\xi)} u_x^2\,dx$ are Lipschitz with Lipschitz constant at most one. Hence, U is also Lipschitz and therefore differentiable almost everywhere. We denote by B_2 the set of Lebesgue points of u_x in B_1, i.e.,

$$B_2 = \{x \in B_1 \mid \lim_{\rho\to 0} \frac{1}{\rho}\int_{x-\rho}^{x+\rho} u_x(t)\,dt = u_x(x)\}.$$

We have $\mathrm{meas}(B_2^c) = 0$. We choose a sequence ξ_i and ξ_i' such that $\frac{1}{2}(y(\xi_i) + y(\xi_i')) = x$ and $y(\xi_i') - y(\xi_i) \leq \frac{1}{i}$. Thus

$$\frac{U(\xi_i') - U(\xi_i)}{\xi_i' - \xi_i} = \frac{\int_{y(\xi_i)}^{y(\xi_i')} u_x(t)\,dt}{y(\xi_i') - y(\xi_i)}\,\frac{y(\xi_i') - y(\xi_i)}{\xi_i' - \xi_i}.$$

Hence, letting i tend to infinity, we get that for every ξ in $y^{-1}(B_2)$ where U and y are differentiable, that is, almost everywhere on $y^{-1}(B_2)$,

$$U_\xi(\xi) = y_\xi(\xi) u_x(y(\xi)). \tag{3.30}$$

From (3.29) and using the fact that $\int_{-\infty}^{y(\xi)} u_x^2\,dx$ is Lipschitz with Lipschitz constant at most one, we get

$$\left|\frac{U(\xi') - U(\xi)}{\xi' - \xi}\right| \leq \sqrt{\frac{y(\xi') - y(\xi)}{\xi' - \xi}}.$$

Hence, for almost every ξ in $y^{-1}(B_2^c)$, we have

$$|U_\xi(\xi)| \leq \sqrt{y_\xi(\xi)}. \tag{3.31}$$

Since $\operatorname{meas}(B_2^c) = 0$, we have by Lemma 3.9, that $y_\xi = 0$ almost everywhere on $y^{-1}(B_2^c)$. Hence, $U_\xi = 0$ almost everywhere on $y^{-1}(B_2^c)$. Thus, we have computed U_ξ almost everywhere. It remains to verify (2.24c). We have, after using (3.27) and (3.30), that $y_\xi H_\xi = y_\xi(1-y_\xi) = y_\xi^2(u^2+u_x^2)\circ y$ and, finally, $y_\xi H_\xi = y_\xi^2 U^2 + U_\xi^2$ almost everywhere on $y^{-1}(B_2)$. On $y^{-1}(B_2^c)$, we have $y_\xi = U_\xi = 0$ almost everywhere. Therefore (2.24c) is satisfied almost everywhere. Up to now we have proved that $X = (y, U, H)$ satisfies (2.24a), (2.24c), the three inequalities in (2.24b) and, by definition, $y + H = \mathrm{Id}$. It remains to prove that $X \in E$ and $\lim_{\xi\to-\infty} H(\xi) = 0$. From (3.24), we have $H(\xi) = \mu((-\infty, y(\xi)))$ for any $\xi \in A$. We can find a sequence $\xi_i \in A$ such that $\lim_{i\to\infty} \xi_i = -\infty$ and we have $\lim_{i\to\infty} H(\xi_i) = 0$. Since H is monotone, it implies that $\lim_{\xi\to-\infty} H(\xi) = 0$. From (3.22) and (3.21b), we obtain $\|H\|_{L^\infty(\mathbb{R})} \leq \mu(\mathbb{R})$. We have, since $H_\xi \geq 0$,

$$\|H_\xi\|_{L^2(\mathbb{R})}^2 \leq \|H_\xi\|_{L^\infty(\mathbb{R})} \|H_\xi\|_{L^1(\mathbb{R})} \leq \|H\|_{L^\infty(\mathbb{R})}^2 \leq \mu(\mathbb{R})$$

and $H \in V$. Since $\zeta = -H$, we have $\zeta \in V$. From (2.24c) we obtain

$$\|U_\xi\|_{L^2(\mathbb{R})}^2 \leq \|y_\xi H_\xi\|_{L^1(\mathbb{R})} \leq (1 + \|\zeta_\xi\|_{L^\infty(\mathbb{R})}) \|H\|_{L^\infty(\mathbb{R})}.$$

Hence, $U_\xi \in L^2(\mathbb{R})$. Let $B_3 = \{\xi \in \mathbb{R} \mid y_\xi < \frac{1}{2}\}$. Since $\zeta_\xi = y_\xi - 1$ and $y_\xi \geq 0$, $B_3 = \{\xi \in \mathbb{R} \mid |\zeta_\xi| > \frac{1}{2}\}$ and, after using the Chebychev inequality, as $\zeta_\xi \in L^2(\mathbb{R})$, we obtain $\operatorname{meas}(B_3) < \infty$. Hence,

$$\begin{aligned}
\int_{\mathbb{R}} U^2(\xi)\,d\xi &= \int_{B_3} U^2(\xi)\,d\xi + \int_{B_3^c} U^2(\xi)\,d\xi \\
&\leq \operatorname{meas}(B_3)\,\|u\|_{L^\infty(\mathbb{R})}^2 + 2\int_{B_3^c} (u\circ y)^2 y_\xi\,d\xi \\
&\leq \operatorname{meas}(B_3)\,\|U\|_{L^\infty(\mathbb{R})}^2 + 2\,\|u\|_{L^2(\mathbb{R})}^2,
\end{aligned}$$

after a change of variables. Hence, $U \in L^2(\mathbb{R})$ and, finally, we have $(y-\mathrm{Id}, U, H) \in E$. ∎

Remark 3.10. If μ is absolutely continuous, then $\mu = (u^2 + u_x^2)dx$ and, from (3.24), we get

$$\int_{-\infty}^{y(\xi)} (u^2 + u_x^2)\,dx + y(\xi) = \xi$$

for all $\xi \in \mathbb{R}$.

At the very beginning, $H(t, \xi)$ was introduced as the energy contained in a strip between $-\infty$ and $y(t, \xi)$, see (2.4). This interpretation still holds. We obtain μ, the energy density in Eulerian coordinates, by pushing forward by y the energy density in Lagrangian coordinates, $H_\xi\,d\xi$. We recall that the push-forward of a measure ν by a measurable function f is the measure $f_\#\nu$ defined as

$$f_\#\nu(B)=\nu(f^{-1}(B))$$

for all Borel set B. We are led to the mapping M which transforms Lagrangian coordinates into Eulerian coordinates and whose definition is contained in the following theorem.

Theorem 3.11. *Given any element $[X]$ in $\mathcal{F}/G$. Then, (u,μ) defined as follows*

$$u(x)=U(\xi)\text{ for any }\xi\text{ such that }x=y(\xi),\tag{3.32a}$$

$$\mu=y_\#(H_\xi\,d\xi)\tag{3.32b}$$

belongs to $\mathcal{D}$ and is independent of the representative $X=(y,U,H)\in\mathcal{F}$ we choose for $[X]$. We denote by $M\colon\mathcal{F}/G\to\mathcal{D}$ the mapping which to any $[X]$ in $\mathcal{F}/G$ associates (u,μ) as given by (3.32).

Proof. First we have to prove that the definition of u makes sense. Since y is surjective, there exists ξ, which may not be unique, such that $x=y(\xi)$. It remains to prove that, given ξ_1 and ξ_2 such that $x=y(\xi_1)=y(\xi_2)$, we have

$$U(\xi_1)=U(\xi_2).\tag{3.33}$$

Since $y(\xi)$ is an increasing function in ξ, we must have $y(\xi)=x$ for all $\xi\in[\xi_1,\xi_2]$ and therefore $y_\xi(\xi)=0$ in $[\xi_1,\xi_2]$. From (2.24c), we get that $U_\xi(\xi)=0$ for all $\xi\in[\xi_1,\xi_2]$ and (3.33) follows.

Since y is proper and $H_\xi\,d\xi$ is a Radon measure, we have, see [1, Remark 1.71], that μ is also a Radon measure. For any $\bar X=(\bar y,\bar U,\bar H)\in\mathcal{F}$ which is equivalent to X, we denote $(\bar u,\bar\mu)$ the pair given by (3.32) when we replace X by $\bar X$. There exists $f\in G$ such that $X=\bar X\circ f$. For any x, there exists ξ' such that $x=\bar y(\xi')$ and $\bar u(x)=\bar U(\xi')$. Let $\xi=f^{-1}(\xi')$. As $x=\bar y(\xi')=y(\xi)$, by (3.32a), we get $u(x)=U(\xi)$ and, since $U(\xi)=\bar U(\xi')$, we finally obtain $\bar u(x)=u(x)$. For any function $\phi\in C_b(\mathbb{R})$, we have

$$\int_{\mathbb{R}}\phi\,d\bar\mu=\int_{\mathbb{R}}\phi\circ\bar y(\xi')\bar H_\xi(\xi')\,d\xi',$$

see [1]. Hence, after making the change of variables $\xi'=f(\xi)$, we obtain

$$\int_{\mathbb{R}}\phi\,d\bar\mu=\int_{\mathbb{R}}\phi\circ\bar y\circ f(\xi)\,\bar H_\xi\circ f(\xi)\,f_\xi(\xi)\,d\xi$$

and, since $H_\xi=\bar H_\xi\circ f f_\xi$ almost everywhere,

$$\int_{\mathbb{R}}\phi\,d\bar\mu=\int_{\mathbb{R}}\phi\circ y(\xi)H_\xi(\xi)\,d\xi=\int_{\mathbb{R}}\phi\,d\mu.$$

Since ϕ was arbitrary in $C_b(\mathbb{R})$, we get $\bar\mu=\mu$. This proves that X and $\bar X$ give raise to the same pair (u,μ), which therefore does not depend on the representative of $[X]$ we choose.

Let us prove that $u\in H^1(\mathbb{R})$. We start by proving that $u_x\in L^2(\mathbb{R})$. For any smooth function ϕ, we have, using the change of variable $x=y(\xi)$,

$$\int_{\mathbb{R}} u(x)\phi_x(x)\,dx=\int_{\mathbb{R}} U(\xi)\phi_x(y(\xi))y_\xi(\xi)\,d\xi=-\int_{\mathbb{R}} U_\xi(\xi)(\phi\circ y)(\xi)\,d\xi, \quad (3.34)$$

after integrating by parts. Let $B_1=\{\xi\in\mathbb{R}\mid y_\xi(\xi)>0\}$. Because of (2.24c), and since $y_\xi\geq 0$ almost everywhere, we have $U_\xi=0$ almost everywhere on B_1^c. Hence, we can restrict the integration domain in (3.34) to B_1. We divide and multiply by $\sqrt{y_\xi}$ the integrand in (3.34) and obtain, after using the Cauchy–Schwarz inequality,

$$\left|\int_{\mathbb{R}} u\phi_x\,dx\right|=\left|\int_{B_1}\frac{U_\xi}{\sqrt{y_\xi}}(\phi\circ y)\sqrt{y_\xi}\,d\xi\right|\leq\sqrt{\int_{B_1}\frac{U_\xi^2}{y_\xi}\,d\xi}\sqrt{\int_{B_1}(\phi\circ y)^2y_\xi\,d\xi}.$$

By (2.24c), we have $\frac{U_\xi^2}{y_\xi}\leq H_\xi$. Hence, after another change of variables, we get

$$\left|\int_{\mathbb{R}} u\phi_x\,dx\right|\leq\sqrt{H(\infty)}\,\|\phi\|_{L^2(\mathbb{R})}\,,$$

which implies that $u_x\in L^2(\mathbb{R})$. Similarly, taking again a smooth function ϕ, we have

$$\left|\int_{\mathbb{R}} u\phi\,dx\right|=\left|\int_{\mathbb{R}} U(\phi\circ y)y_\xi\,d\xi\right|\leq\|\phi\|_{L^2(\mathbb{R})}\sqrt{\int_{\mathbb{R}} U^2y_\xi\,d\xi}\leq\sqrt{H(\infty)}\,\|\phi\|_{L^2(\mathbb{R})}$$

because $U^2y_\xi\leq H_\xi$ from (2.24c). Hence, $u\in L^2(\mathbb{R})$.

Let us prove that the absolute continuous part of μ is equal to $(u^2+u_x^2)\,dx$. We introduce the sets Z and B defined as follows

$$Z=\Big\{\xi\in\mathbb{R}\mid\ y\text{ is differentiable at }\xi\text{ and }y_\xi(\xi)=0$$
$$\text{or }y\text{ or }U\text{ are not differentiable at }\xi\Big\}$$

and

$$B=\{x\in y(Z)^c\mid u\text{ is differentiable at }x\}\,.$$

Since u belongs to $H^1(\mathbb{R})$, it is differentiable almost everywhere. We have, since y is Lipschitz and by the definition of Z, that $\operatorname{meas}(y(Z))=\int_Z y_\xi(\xi)\,d\xi=0$. Hence, $\operatorname{meas}(B^c)=0$. For any $\xi\in y^{-1}(B)$, we denote $x=y(\xi)$. By necessity, we have $\xi\in Z^c$. Let ξ_i be a sequence converging to ξ such that $\xi_i\neq\xi$ for all i. We write $x_i=y(\xi_i)$. Since $y_\xi(\xi)>0$, for i large enough, $x_i\neq x$. The following quantity is well-defined

$$\frac{U(\xi_i)-U(\xi)}{\xi_i-\xi}=\frac{u(x_i)-u(x)}{x_i-x}\,\frac{x_i-x}{\xi_i-\xi}.$$

Since u is differentiable at x and ξ belongs to Z^c, we obtain, after letting i tend to infinity, that

$$U_\xi(\xi)=u_x(y(\xi))y_\xi(\xi). \quad (3.35)$$

For all subsets B' of B, we have

$$\mu(B')=\int_{y^{-1}(B')}H_\xi\,d\xi=\int_{y^{-1}(B')}\left(U^2+\frac{U_\xi^2}{y_\xi^2}\right)y_\xi\,d\xi.$$

We can divide by y_ξ in the integrand above because y_ξ does not vanish on $y^{-1}(B)$. After a change of variables and using (3.35), we obtain

$$\mu(B')=\int_{B'}(u^2+u_x^2)\,dx. \tag{3.36}$$

Since (3.36) holds for any set $B'\subset B$ and $\mathrm{meas}(B^c)=0$, we have $\mu_{\mathrm{ac}}=(u^2+u_x^2)\,dx$. ☐

The next theorem shows that the transformation from Eulerian to Lagrangian coordinates is a bijection.

Theorem 3.12. *The mapping M and L are invertible. We have*

$$L\circ M=\mathrm{Id}_{\mathcal{F}/G}\ \text{ and } M\circ L=\mathrm{Id}_{\mathcal{D}}\,.$$

Proof. Given $[X]$ in $\mathcal{F}/G$, we choose $X=(y,U,H)=\tilde\Gamma([X])$ as a representative of $[X]$ and consider (u,μ) given by (3.32) for this particular X. Note that, from the definition of $\tilde\Gamma$, we have $X\in\mathcal{F}_0$. Let $\bar X=(\bar y,\bar U,\bar H)$ be the representative of $L(u,\mu)$ in $\mathcal{F}_0$ given by the formulas (3.21). We claim that $(\bar y,\bar U,\bar H)=(y,U,H)$ and therefore $L\circ M=\mathrm{Id}_{\mathcal{F}/G}$. Let

$$g(x)=\sup\{\xi\in\mathbb{R}\mid y(\xi)<x\}. \tag{3.37}$$

It is not hard to prove, using the fact that y is increasing and continuous, that

$$y(g(x))=x \tag{3.38}$$

and $y^{-1}((-\infty,x))=(-\infty,g(x))$. For any $x\in\mathbb{R}$, we have, by (3.32b), that

$$\mu((-\infty,x))=\int_{y^{-1}((-\infty,x))}H_\xi\,d\xi=\int_{-\infty}^{g(x)}H_\xi\,d\xi=H(g(x))$$

because $H(-\infty)=0$. Since $X\in\mathcal{F}_0$, $y+H=\mathrm{Id}$ and we get

$$\mu((-\infty,x))+x=g(x). \tag{3.39}$$

From the definition of $\bar y$, we then obtain that

$$\bar y(\xi)=\sup\{x\in\mathbb{R}\mid g(x)<\xi\}. \tag{3.40}$$

For any given $\xi\in\mathbb{R}$, let us consider an increasing sequence x_i tending to $\bar y(\xi)$ such that $g(x_i)<\xi$; such sequence exists by (3.40). Since y is increasing and using (3.38), it follows that $x_i\le y(\xi)$. Letting i tend to ∞, we obtain $\bar y(\xi)\le y(\xi)$. Assume that $\bar y(\xi)<y(\xi)$. Then, there exists x such that $\bar y(\xi)<x<y(\xi)$ and equation (3.40) then implies that $g(x)\ge\xi$. On the other hand, $x=y(g(x))<y(\xi)$ implies $g(x)<\xi$ because y is increasing, which gives us a contradiction. Hence, we have $\bar y=y$. It follows directly from the definitions, since $y+H=\mathrm{Id}$, that $\bar H=H$ and $\bar U=U$ and we have proved that $L\circ M=\mathrm{Id}_{\mathcal{F}/G}$.

Given (u,μ) in $\mathcal{D}$, we denote by (y,U,H) the representative of $L(u,\mu)$ in $\mathcal{F}_0$ given by (3.21). Then, let $(\bar{u},\bar{\mu}) = M\circ L(u,\mu)$. We claim that $(\bar{u},\bar{\mu}) = (u,\mu)$. Let g be the function defined as before by (3.37). The same computation that leads to (3.39) now gives

$$\bar{\mu}((-\infty,x)) + x = g(x). \tag{3.41}$$

Given $\xi\in\mathbb{R}$, we consider an increasing sequence x_i which converges to $y(\xi)$ and such that $\mu((-\infty,x_i)) + x_i < \xi$. The existence of such sequence is guaranteed by (3.21a). Passing to the limit and since $F(x) = \mu((-\infty,x))$ is lower semi-continuous, we obtain $\mu((-\infty,y(\xi))) + y(\xi)\leq\xi$. We take $\xi = g(x)$ and get

$$\mu((-\infty,x)) + x \leq g(x). \tag{3.42}$$

From the definition of g, there exists an increasing sequence ξ_i which converges to $g(x)$ such that $y(\xi_i) < x$. The definition (3.21a) of y tells us that $\mu((-\infty,x))+x\geq\xi_i$. Letting i tend to infinity, we obtain $\mu((-\infty,x)) + x\geq g(x)$ which, together with (3.42), yields

$$\mu((-\infty,x)) + x = g(x). \tag{3.43}$$

Comparing (3.43) and (3.41) we get that $\mu=\bar{\mu}$. It is clear from the definitions that $\bar{u} = u$. Hence, $(\bar{u},\bar{\mu}) = (u,\mu)$ and $M\circ L = \mathrm{Id}_{\mathcal{D}}$. □

4. Continuous semigroup of solutions on $\mathcal{D}$

Now comes the justification of all the analysis done in the previous section. The fact that we have been able to establish a bijection between the two coordinate systems, $\mathcal{F}/G$ and $\mathcal{D}$, enables us now to transport the topology defined in $\mathcal{F}/G$ into $\mathcal{D}$. On $\mathcal{D}$ we define the distance $d_{\mathcal{D}}$ which makes the bijection L between $\mathcal{D}$ and $\mathcal{F}/G$ into an isometry:

$$d_{\mathcal{D}}((u,\mu),(\bar{u},\bar{\mu})) = d_{\mathcal{F}/G}(L(u,\mu),L(\bar{u},\bar{\mu})).$$

Since $\mathcal{F}/G$ equipped with $d_{\mathcal{F}/G}$ is a complete metric space, we have the following theorem.

Theorem 4.1. *$\mathcal{D}$ equipped with the metric d_D is a complete metric space.*

For each $t\in\mathbb{R}$, we define the mapping T_t from $\mathcal{D}$ to $\mathcal{D}$ as

$$T_t = M\tilde{S}_tL.$$

We have the following commutative diagram:

$$\begin{tikzcd} \mathcal{D} & \mathcal{F}/G \arrow[l, "M"'] \\ \mathcal{D} \arrow[u, "T_t"] \arrow[r, "L"] & \mathcal{F}/G \arrow[u, "\tilde{S}_t"'] \end{tikzcd} \tag{4.1}$$

Our main theorem reads as follows.

Theorem 4.2. *$T\colon \mathcal{D}\times\mathbb{R}_+ \to \mathcal{D}$ (where $\mathcal{D}$ is defined by Definition 3.1) defines a continuous semigroup of solutions of the Camassa–Holm equation, that is, given $(\bar{u},\bar{\mu}) \in \mathcal{D}$, if we denote $t \mapsto (u(t),\mu(t)) = T_t(\bar{u},\bar{\mu})$ the corresponding trajectory, then u is a weak solution of the Camassa–Holm equation* (1.4). *Moreover μ is a weak solution of the following transport equation for the energy density*

$$\mu_t + (u\mu)_x = (u^3 - 2Pu)_x. \tag{4.2}$$

Furthermore, we have that

$$\mu(t)(\mathbb{R}) = \mu(0)(\mathbb{R}) \text{ for all } t \tag{4.3}$$

and

$$\mu(t)(\mathbb{R}) = \mu_{\mathrm{ac}}(t)(\mathbb{R}) = \|u(t)\|_{H^1}^2 = \mu(0)(\mathbb{R}) \text{ for almost all } t. \tag{4.4}$$

Remark 4.3. We denote the unique solution described in the theorem as a *conservative* weak solution of the Camassa–Holm equation.

Proof. We want to prove that, for all $\phi \in C^\infty(\mathbb{R}_+\times\mathbb{R})$ with compact support,

$$\int_{\mathbb{R}_+\times\mathbb{R}} \left[-u(t,x)\phi_t(t,x) + u(t,x)u_x(t,x)\phi(t,x)\right] dxdt = \int_{\mathbb{R}_+\times\mathbb{R}} -P_x(t,x)\phi(t,x)\,dxdt \tag{4.5}$$

where P is given by (1.4b) or equivalently (2.6). Let $(y(t),U(t),H(t))$ be a representative of $L(u(t),\mu(t))$ which is solution of (2.10). Since y is Lipschitz in ξ and invertible for $t \in \mathcal{K}^c$ (see (2.31) for the definition of $\mathcal{K}$, in particular, we have $\operatorname{meas}(\mathcal{K}) = 0$), we can use the change of variables $x = y(t,\xi)$ and, using (3.30), we get

$$\begin{aligned}&\int_{\mathbb{R}_+\times\mathbb{R}} \left[-u(t,x)\phi_t(t,x) + u(t,x)u_x(t,x)\phi(t,x)\right] dxdt\\ &= \int_{\mathbb{R}_+\times\mathbb{R}} \left[-U(t,\xi)y_\xi(t,\xi)\phi_t(t,y(t,\xi)) + U(t,\xi)U_\xi(t,\xi)\phi(t,y(t,\xi))\right] d\xi dt.\end{aligned} \tag{4.6}$$

Using the fact that $y_t = U$ and $y_{\xi t} = U_\xi$, one easily check that

$$(Uy_\xi\phi\circ y)_t - (U^2\phi)_\xi = Uy_\xi\phi_t\circ y - UU_\xi\phi\circ y + U_t y_\xi\phi\circ y. \tag{4.7}$$

After integrating (4.7) over $\mathbb{R}_+\times\mathbb{R}$, the left-hand side of (4.7) vanishes and we obtain

$$\begin{aligned}&\int_{\mathbb{R}_+\times\mathbb{R}} \left[-Uy_\xi\,\phi_t\circ y + UU_\xi\,\phi\circ y\right] d\xi dt\\ &= \frac{1}{4}\int_{\mathbb{R}_+\times\mathbb{R}^2} \Big[\operatorname{sgn}(\xi-\eta)e^{-\{\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\}}\times\big(U^2y_\xi + H_\xi\big)(\eta)y_\xi(\xi)\phi\circ y(\xi)\Big]\, d\eta d\xi dt\end{aligned} \tag{4.8}$$

by (2.10). Again, to simplify the notation, we deliberately omitted the t variable. On the other hand, by using the change of variables $x = y(t,\xi)$ and $z = y(t,\eta)$

when $t\in\mathcal{K}^c$, we have

$$-\int_{\mathbb{R}_+\times\mathbb{R}} P_x(t,x)\phi(t,x)\,dxdt=\frac{1}{2}\int_{\mathbb{R}_+\times\mathbb{R}^2}\Big[\operatorname{sgn}(y(\xi)-y(\eta))e^{-|y(\xi)-y(\eta)|}\\ \times\big(u^2(t,y(\eta))+\frac{1}{2}u_x^2(t,y(\eta))\big)\phi(t,y(\xi))y_\xi(\eta)y_\xi(\xi)\Big]\,d\eta d\xi dt.$$

Since, from Lemma 2.7, y_ξ is strictly positive for $t\in\mathcal{K}^c$ and almost every ξ, we can replace $u_x(t,y(t,\eta))$ by $U_\xi(t,\eta)/y_\xi(t,\eta)$, see (3.30), in the equation above and, using the fact that y is an increasing function and the identity (2.24c), we obtain

$$-\int_{\mathbb{R}_+\times\mathbb{R}} P_x(t,x)\phi(t,x)\,dxdt=\frac{1}{4}\int_{\mathbb{R}_+\times\mathbb{R}^2}\Big[\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta)\big)\\ \times\big(U^2y_\xi+H_\xi\big)(\eta)y_\xi(\xi)\phi(t,y(\xi))\Big]\,d\eta d\xi dt. \quad (4.9)$$

Thus, comparing (4.8) and (4.9), we get

$$\int_{\mathbb{R}_+\times\mathbb{R}}\left[-Uy_\xi\,\phi_t(t,y)+UU_\xi\,\phi\right]\,d\xi dt=-\int_{\mathbb{R}_+\times\mathbb{R}}P_x(t,x)\phi(t,x)\,dxdt$$

and (4.5) follows from (4.6). Similarly, one proves that $\mu(t)$ is solution of (4.2). From (3.32a), we obtain

$$\mu(t)(\mathbb{R})=\int_{\mathbb{R}}H_\xi\,d\xi=H(t,\infty)$$

which is constant in time, see Lemma 2.7 (*iii*). Hence, (4.3) is proved. We know from Lemma 2.7 (*ii*) that, for $t\in\mathcal{K}^c$, $y_\xi(t,\xi)>0$ for almost every $\xi\in\mathbb{R}$. Given $t\in\mathcal{K}^c$ (the time variable is suppressed in the notation when there is no ambiguity), we have, for any Borel set B,

$$\mu(t)(B)=\int_{y^{-1}(B)}H_\xi\,d\xi=\int_{y^{-1}(B)}\left(U^2+\frac{U_\xi^2}{y_\xi^2}\right)y_\xi\,d\xi \quad (4.10)$$

from (2.24c) and because $y_\xi(t,\xi)>0$ almost everywhere for $t\in\mathcal{K}^c$. Since y is one-to-one when $t\in\mathcal{K}^c$ and $u_x\circ yy_\xi=U_\xi$ almost everywhere, we obtain from (4.10) that

$$\mu(t)(B)=\int_B(u^2+u_x^2)(t,x)\,dx.$$

Hence, as $\operatorname{meas}(\mathcal{K})=0$, (4.4) is proved. $\square$

5. The topology on $\mathcal{D}$

The metric $d_{\mathcal{D}}$ gives to $\mathcal{D}$ the structure of a complete metric space while it makes continuous the semigroup T_t of conservative solutions for the Camassa–Holm equation as defined in Theorem 4.2. In that respect, it is a suitable metric for the Camassa–Holm equation. However, as the definition of $d_{\mathcal{D}}$ is not straightforward, this metric is not so easy to manipulate and in this section we compare it with more standard topologies. More precisely, we establish that convergence in $H^1(\mathbb{R})$ implies convergence in $(\mathcal{D}, d_{\mathcal{D}})$, which itself implies convergence in $L^\infty(\mathbb{R})$.

Proposition 5.1. *The mapping*

$$u \mapsto (u, (u^2 + u_x^2)dx)$$

is continuous from $H^1(\mathbb{R})$ into $\mathcal{D}$. In other words, given a sequence $u_n \in H^1(\mathbb{R})$ converging to u in $H^1(\mathbb{R})$, then $(u_n, (u_n^2 + u_{nx}^2)dx)$ converges to $(u, (u^2 + u_x^2)dx)$ in $\mathcal{D}$.

Proof. We write $g_n = u_n^2 + u_{n,x}^2$ and $g = u^2 + u_x^2$. Let $X_n = (y_n, U_n, H_n)$ and $X = (y, U, H)$ be the representatives in $\mathcal{F}_0$ given by (3.21) of $L(u_n, (u_n^2 + u_{nx}^2)dx)$ and $L(u, (u^2 + u_x^2)dx)$, respectively. Following Remark 3.10, we have

$$\int_{-\infty}^{y(\xi)} g(x)\, dx + y(\xi) = \xi \ , \quad \int_{-\infty}^{y_n(\xi)} g_n(x)\, dx + y_n(\xi) = \xi \tag{5.1}$$

and, after taking the difference between the two equations, we obtain

$$\int_{-\infty}^{y(\xi)} (g - g_n)(x)\, dx + \int_{y_n(\xi)}^{y(\xi)} g_n(x)\, dx + y(\xi) - y_n(\xi) = 0. \tag{5.2}$$

Since g_n is positive, $\left| y - y_n + \int_{y_n}^{y} g_n(x)\, d\xi \right| = |y - y_n| + \left| \int_{y_n}^{y} g_n(x)\, d\xi \right|$ and (5.2) implies

$$|y(\xi) - y_n(\xi)| \leq \int_{-\infty}^{y(\xi)} |g - g_n|\ dx \leq \|g - g_n\|_{L^1(\mathbb{R})}\,.$$

Since $u_n \to u$ in $H^1(\mathbb{R})$, $g_n \to g$ in $L^1(\mathbb{R})$ and it follows that $\zeta_n \to \zeta$ and $H_n \to H$ in $L^\infty(\mathbb{R})$. We recall that $\zeta(\xi) = y(\xi) - \xi$ and $H = -\zeta$ (as $X, X_n \in \mathcal{F}_0$). The measures $(u^2 + u_x^2)dx$ and $(u_n^2 + u_{n,x}^2)dx$ have, by definition, no singular part and in that case (3.27) holds almost everywhere, that is,

$$y_\xi = \frac{1}{g \circ y + 1} \quad \text{and} \quad y_{n,\xi} = \frac{1}{g_n \circ y + 1} \tag{5.3}$$

almost everywhere. Hence,

$$\begin{aligned} \zeta_{n,\xi} - \zeta_\xi &= (g \circ y - g_n \circ y_n) y_{n,\xi} y_\xi \\ &= (g \circ y - g \circ y_n) y_{n,\xi} y_\xi + (g \circ y_n - g_n \circ y_n) y_{n,\xi} y_\xi. \end{aligned} \tag{5.4}$$

Since $0 \leq y_\xi \leq 1$, we have

$$\int_{\mathbb{R}} |g \circ y_n - g_n \circ y_n|\, y_{n,\xi} y_\xi\, d\xi \leq \int_{\mathbb{R}} |g \circ y_n - g_n \circ y_n|\, y_{n,\xi}\, d\xi = \|g - g_n\|_{L^1(\mathbb{R})}\,. \tag{5.5}$$

For any $\varepsilon > 0$, there exists a continuous function h with compact support such that $\|g - h\|_{L^1(\mathbb{R})} \leq \varepsilon/3$. We can decompose the first term in the right-hand side of (5.4) into

$$\begin{aligned} (g \circ y - g \circ y_n) y_{n,\xi} y_\xi &= (g \circ y - h \circ y) y_{n,\xi} y_\xi \\ &\quad + (h \circ y - h \circ y_n) y_{n,\xi} y_\xi + (h \circ y_n - g \circ y_n) y_{n,\xi} y_\xi. \end{aligned} \tag{5.6}$$

Then, we have

$$\int_{\mathbb{R}} |g\circ y - h\circ y|\, y_{n,\xi} y_\xi \, d\xi \leq \int |g\circ y - h\circ y|\, y_\xi \, d\xi = \|g-h\|_{L^1(\mathbb{R})} \leq \varepsilon/3$$

and, similarly, we obtain $\int_{\mathbb{R}} |g\circ y_n - h\circ y_n|\, y_{n,\xi} y_\xi \, d\xi \leq \varepsilon/3$. Since $y_n \to y$ in $L^\infty(\mathbb{R})$ and h is continuous with compact support, by applying Lebesgue dominated convergence theorem, we obtain $h\circ y_n \to h\circ y$ in $L^1(\mathbb{R})$ and we can choose n big enough so that

$$\int_{\mathbb{R}} |h\circ y - h\circ y_n|\, y_{n,\xi} y_\xi \, d\xi \leq \|h\circ y - h\circ y_n\|_{L^1(\mathbb{R})} \leq \varepsilon/3.$$

Hence, from (5.6), we get that $\int_{\mathbb{R}} |g\circ y - g\circ y_n|\, y_{n,\xi} y_\xi \, d\xi \leq \varepsilon$ so that

$$\lim_{n\to\infty} \int_{\mathbb{R}} |g\circ y - g\circ y_n|\, y_{n,\xi} y_\xi \, d\xi = 0,$$

and, from (5.4) and (5.5), it follows that $\zeta_{n,\xi} \to \zeta_\xi$ in $L^1(\mathbb{R})$. Since $X_n \in \mathcal{F}_0$, $\zeta_{n,\xi}$ is bounded in $L^\infty(\mathbb{R})$ and we finally get that $\zeta_{n,\xi} \to \zeta_\xi$ in $L^2(\mathbb{R})$ and, by (3.21b), $H_{n,\xi} \to H_\xi$ in $L^2(\mathbb{R})$. It remains to prove that $U_n \to U$ in $H^1(\mathbb{R})$. Let $C_n = \{x \in \mathbb{R} \mid g_n(x) > 1\}$. Chebychev's inequality yields $\operatorname{meas}(C_n) \leq \|g_n\|_{L^1(\mathbb{R})}$. Let $B_n = \{\xi \in \mathbb{R} \mid y_{n,\xi}(\xi) < \frac{1}{2}\}$. Since $y_{n,\xi}(g_n\circ y_n + 1) = 1$ almost everywhere, $g_n \circ y_n > 1$ on B_n and therefore $y_n(B_n) \subset C_n$. From (5.1), we get that

$$\operatorname{meas}(y_n(B)) + \int_{y_n(B)} g_n(\xi)\, d\xi = \operatorname{meas}(B) \tag{5.7}$$

for any set B equal to a countable union of disjoint open intervals. Any Borel set B can be "approximated" by such countable union of disjoint open intervals and therefore, using the fact that y_n is Lipschitz and one-to-one, we infer that (5.7) holds for any Borel set B. After taking $B = B_n$, (5.7) yields

$$\begin{aligned}\operatorname{meas}(B_n) &\leq \operatorname{meas}(y_n(B_n)) + \|g_n\|_{L^1(\mathbb{R})}\\ &\leq \operatorname{meas}(C_n) + \|g_n\|_{L^1(\mathbb{R})}\end{aligned}$$

and therefore $\operatorname{meas}(B_n) \leq 2\|g_n\|_{L^1(\mathbb{R})}$. For any function $f_1, f_2 \in H^1(\mathbb{R})$, we have

$$\|f_1\circ y_n - f_2\circ y_n\|_{L^2(\mathbb{R})}^2 = \int_{B_n} (f_1\circ y_n - f_2\circ y_n)^2\, d\xi + \int_{B_n^c} (f_1\circ y_n - f_2\circ y_n)^2\, d\xi \tag{5.8}$$

and, as $y_{n,\xi} \geq 0$ on B_n^c,

$$\int_{B_n^c} (f_1\circ y_n - f_2\circ y_n)^2\, d\xi \leq 2\int_{B_n^c} (f_1\circ y_n - f_2\circ y_n)^2 y_{n,\xi}\, d\xi \leq 2\|f_1 - f_2\|_{L^2(\mathbb{R})}^2.$$

Hence,

$$\|f_1\circ y_n - f_2\circ y_n\|_{L^2(\mathbb{R})}^2 \leq \operatorname{meas}(B_n)\|f_1 - f_2\|_{L^\infty(\mathbb{R})}^2 + 2\|f_1 - f_2\|_{L^2(\mathbb{R})}^2$$

and, since $\operatorname{meas}(B_n) \leq 2\left\|g_n\right\|_{L^1(\mathbb{R})}$,

$$\begin{aligned}\left\|f_1 \circ y_n - f_2 \circ y_n\right\|_{L^2(\mathbb{R})}^2 &\leq 2\left\|g_n\right\|\left\|f_1 - f_2\right\|_{L^\infty(\mathbb{R})}^2 + 2\left\|f_1 - f_2\right\|_{L^2(\mathbb{R})}^2 \\ &\leq C\left\|f_1 - f_2\right\|_{H^1(\mathbb{R})}^2 \end{aligned}\tag{5.9}$$

for some constant C which is independent of n. We have

$$\left\|U_n - U\right\|_{L^2(\mathbb{R})} \leq \left\|u_n \circ y_n - u \circ y_n\right\|_{L^2(\mathbb{R})} + \left\|u \circ y_n - u \circ y\right\|_{L^2(\mathbb{R})}. \tag{5.10}$$

After using (5.9) for $f_1 = u_n$ and $f_2 = u$ and since, by assumption, $u_n \to u$ in $H^1(\mathbb{R})$, we obtain that $\lim_{n\to\infty}\left\|u_n \circ y_n - u \circ y_n\right\|_{L^2(\mathbb{R})} = 0$. We can find continuous functions with compact support h which are arbitrarily close to u in $H^1(\mathbb{R})$. Then, from (5.9), $h \circ y_n$ and $h \circ y$ are arbitrarily closed in $L^2(\mathbb{R})$ to $u \circ y_n$ and $u \circ y$, respectively, and independently of n. By the Lebesgue dominated convergence theorem, as $y_n \to y$ in $L^\infty(\mathbb{R})$, we get that $h \circ y_n \to h \circ y$ in $L^2(\mathbb{R})$. Hence,

$$\begin{aligned}\left\|u \circ y_n - u \circ y\right\|_{L^2(\mathbb{R})} &\leq \left\|u \circ y_n - h \circ y_n\right\|_{L^2(\mathbb{R})} \\ &\quad + \left\|h \circ y_n - h \circ y\right\|_{L^2(\mathbb{R})} + \left\|h \circ y - u \circ y\right\|_{L^2(\mathbb{R})}\end{aligned}$$

implies that $\lim_{n\to\infty}\left\|u \circ y_n - u \circ y\right\|_{L^2(\mathbb{R})} = 0$ and, finally, from (5.10), we conclude that $U_n \to U$ in $L^2(\mathbb{R})$. It remains to prove that $U_{n,\xi} \to U_\xi$ in $L^2(\mathbb{R})$. Since $H_{n,\xi} = 1 - y_{n,\xi}$, (2.24c) can be rewritten as

$$U_{n,\xi}^2 = H_{n,\xi} - H_{n,\xi}^2 - U_n^2 + H_{n,\xi}^2 U_n^2 \tag{5.11}$$

and there holds the corresponding identity holds for U_ξ. We have $\left\|U_n\right\|_{L^\infty(\mathbb{R})} = \left\|u_n\right\|_{L^\infty(\mathbb{R})}$ and therefore $\left\|U_n\right\|_{L^\infty(\mathbb{R})}$ is uniformly bounded in n. Hence, since $U_n \to U$ in $L^2(\mathbb{R})$, $H_n \to H$ in V and $\left\|U_n\right\|_{L^\infty(\mathbb{R})}$, $\left\|H_{n,\xi}\right\|_{L^\infty(\mathbb{R})}$ are uniformly bounded in n, we get from (5.11) that

$$\lim_{n\to\infty}\left\|U_{n,\xi}\right\|_{L^2(\mathbb{R})} = \left\|U_\xi\right\|_{L^2(\mathbb{R})}. \tag{5.12}$$

Once we have proved that $U_{n,\xi}$ converges weakly to U_ξ, then (5.11) will imply that $U_{n,\xi} \to U_\xi$ strongly in $L^2(\mathbb{R})$, see, for example, [27, section V.1]. For any continuous function ϕ with compact support, we have

$$\int_{\mathbb{R}} U_{n,\xi}\phi\, d\xi = \int_{\mathbb{R}} u_{n,x} \circ y_n y_{n,\xi}\phi\, d\xi = \int_{\mathbb{R}} u_{n,x}\, \phi \circ {y_n}^{-1}\, d\xi. \tag{5.13}$$

By assumption, we have $u_{n,x} \to u_x$ in $L^2(\mathbb{R})$. Since $y_n \to y$ in $L^\infty(\mathbb{R})$, the support of $\phi \circ {y_n}^{-1}$ is contained in some compact that can be chosen to be independent of n. Thus, after using Lebesgue's dominated convergence theorem, we obtain that $\phi \circ {y_n}^{-1} \to \phi \circ y^{-1}$ in $L^2(\mathbb{R})$ and therefore

$$\lim_{n\to\infty}\int_{\mathbb{R}} U_{n,\xi}\phi\, d\xi = \int_{\mathbb{R}} u_x\, \phi \circ y^{-1}\, d\xi = \int_{\mathbb{R}} U_\xi \phi\, d\xi. \tag{5.14}$$

From (5.12), we have that $U_{n,\xi}$ is bounded and therefore, by a density argument, (5.14) holds for any function ϕ in $L^2(\mathbb{R})$ and $U_{n,\xi} \rightharpoonup U_\xi$ weakly in $L^2(\mathbb{R})$. $\square$

Proposition 5.2. *Let (u_n, μ_n) be a sequence in $\mathcal{D}$ that converges to (u, μ) in $\mathcal{D}$. Then*

$$u_n \to u \text{ in } L^\infty(\mathbb{R}) \quad \text{and} \quad \mu_n \overset{*}{\rightharpoonup} \mu.$$

Proof. We denote by $X_n = (y_n, U_n, H_n)$ and $X = (y, U, H)$ the representative of $L(u_n, \mu_n)$ and $L(u, \mu)$ given by (3.21). For any $x \in \mathbb{R}$, there exists ξ_n and ξ, which may not be unique, such that $x = y_n(\xi_n)$ and $x = y(\xi)$. We set $x_n = y_n(\xi)$. We have

$$u_n(x) - u(x) = u_n(x) - u_n(x_n) + U_n(\xi) - U(\xi) \tag{5.15}$$

and

$$\begin{aligned}
|u_n(x) - u_n(x_n)| &= \left| \int_\xi^{\xi_n} U_{n,\xi}(\eta)\, d\eta \right| \\
&\leq \sqrt{\xi_n - \xi} \left(\int_\xi^{\xi_n} U_{n,\xi}^2\, d\eta \right)^{1/2} && \text{(Cauchy–Schwarz)} \\
&\leq \sqrt{\xi_n - \xi} \left(\int_\xi^{\xi_n} y_{n,\xi} H_{n,\xi}\, d\eta \right)^{1/2} && \text{(from (2.24c))} \\
&\leq \sqrt{\xi_n - \xi} \sqrt{|y_n(\xi_n) - y_n(\xi)|} && \text{(since } H_{n,\xi} \leq 1) \\
&= \sqrt{\xi_n - \xi} \sqrt{y(\xi) - y_n(\xi)} \\
&\leq \sqrt{\xi_n - \xi}\, \|y - y_n\|_{L^\infty(\mathbb{R})}^{1/2}.
\end{aligned} \tag{5.16}$$

From (3.22), we get

$$|\xi_n - \xi| \leq 2\mu_n(\mathbb{R}) + |y_n(\xi_n) - y_n(\xi)| = 2 \lim_{\xi \to \infty} H_n(\xi) + |y(\xi) - y_n(\xi)|$$

and, therefore, since $H_n \to H$ and $y_n \to y$ in $L^\infty(\mathbb{R})$, $|\xi_n - \xi|$ is bounded by a constant C independent of n. Then, (5.16) implies

$$|u_n(x) - u_n(x_n)| \leq C\, \|y - y_n\|_{L^\infty(\mathbb{R})}^{1/2}. \tag{5.17}$$

Since $y_n \to y$ and $U_n \to U$ in $L^\infty(\mathbb{R})$, it follows from (5.15) and (5.17) that $u_n \to u$ in $L^\infty(\mathbb{R})$. By weak-star convergence, we mean that

$$\lim_{n \to \infty} \int_{\mathbb{R}} \phi\, d\mu_n = \int_{\mathbb{R}} \phi\, d\mu \tag{5.18}$$

for all continuous functions with compact support. It follows from (3.32b) that

$$\int_{\mathbb{R}} \phi\, d\mu_n = \int_{\mathbb{R}} \phi \circ y_n H_{n,\xi}\, d\xi \quad \text{and} \quad \int_{\mathbb{R}} \phi\, d\mu = \int_{\mathbb{R}} \phi \circ y H_\xi\, d\xi \tag{5.19}$$

see [1, definition 1.70]. Since $y_n \to y$ in $L^\infty(\mathbb{R})$, the support of $\phi \circ y_n$ is contained in some compact which can be chosen independently of n and, from Lebesgue's dominated convergence theorem, we have that $\phi \circ y_n \to \phi \circ y$ in $L^2(\mathbb{R})$. Hence, since $H_{n,\xi} \to H_\xi$ in $L^2(\mathbb{R})$,

$$\lim_{n\to\infty} \int_{\mathbb{R}} \phi \circ y_n H_{n,\xi} \, d\xi = \int_{\mathbb{R}} \phi \circ y H_\xi \, d\xi,$$

and (5.18) follows from (5.19).

□

Acknowledgments. The authors gratefully acknowledge the hospitality of the Mittag-Leffler Institute, Sweden, creating a great working environment for research, during the Fall of 2005.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* The Clarendon Press, Oxford University Press, New York, 2000.

[2] V. I. Arnold and B. A. Khesin. *Topological Methods in Hydrodynamics.* Springer-Verlag, New York, 1998.

[3] R. Beals, D. Sattinger, and J. Szmigielski. Peakon-antipeakon interaction. *J. Nonlinear Math. Phys.* 8:23–27, 2001.

[4] A. Bressan and A. Constantin. Global conservative solutions of the Camassa–Holm equation. *Preprint.* Submitted, 2005.

[5] A. Bressan and M. Fonte. An optimal transportation metric for solutions of the Camassa–Holm equation. *Methods Appl. Anal.*, to appear.

[6] H. Brezis. *Analyse Fonctionnelle. Théorie et applications.* Masson, Paris, 1983.

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

[8] R. Camassa, D. D. Holm, and J. Hyman. A new integrable shallow water equation. *Adv. Appl. Mech.*, 31:1–33, 1994.

[9] G. M. Coclite, and H. Holden, and K. H, Karlsen. Well-posedness for a parabolic-elliptic system. *Discrete Cont. Dynam. Systems* 13:659–682, 2005.

[10] G. M. Coclite, and H. Holden, and K. H, Karlsen. Global weak solutions to a generalized hyperelastic-rod wave equation. *SIAM J. Math. Anal.*, 37:1044–1069, 2005.

[11] A. Constantin. Existence of permanent and breaking waves for a shallow water equation: a geometric approach. *Ann. Inst. Fourier (Grenoble)*, 50(2):321–362, 2000.

[12] A. Constantin and J. Escher. Global existence and blow-up for a shallow water equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 26(2):303–328, 1998.

[13] A. Constantin and B. Kolev. On the geometric approach to the motion of inertial mechanical systems. *J. Phys. A*, 35(32):R51–R79, 2002.

[14] A. Constantin and L. Molinet. Global weak solutions for a shallow water equation. *Comm. Math. Phys.*, 211(1):45–61, 2000.

[15] H.-H. Dai. Exact travelling-wave solutions of an integrable equation arising in hyperelastic rods. *Wave Motion*, 28(4):367–381, 1998.

[16] H.-H. Dai. Model equations for nonlinear dispersive waves in a compressible Mooney–Rivlin rod. *Acta Mech.*, 127(1-4):193–207, 1998.

[17] H.-H. Dai and Y. Huo. Solitary shock waves and other travelling waves in a general compressible hyperelastic rod. *R. Soc. Lond. Proc. Ser. A*, 456(1994):331–363, 2000.

[18] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton, Florida, USA, 1992.

[19] G. B. Folland. *Real Analysis.* John Wiley & Sons Inc., New York, second edition, 1999.

[20] M. Fonte. Conservative solution of the Camassa Holm equation on the real line. *arXiv:math.AP/0511549*,3 2005.

[21] H. Holden, K. H. Karlsen, and N. H. Risebro. Convergent difference schemes for the Hunter–Saxton equation. *Math. Comp.*, to appear.

[22] H. Holden and X. Raynaud. Convergence of a finite difference scheme for the Camassa–Holm equation. *SIAM J. Numer. Anal.* 2006, to appear. **paper I** in this thesis.

[23] H. Holden and X. Raynaud. A convergent numerical scheme for the Camassa–Holm equation based on multipeakons. *Discrete Contin. Dyn. Syst.*, 14:505–523, 2006. **paper II** in this thesis.

[24] H. Holden and X. Raynaud. Global multipeakon solutions of the Camassa–Holm equation. *J. Hyperbolic Differ. Equ.* to appear, 2006. **paper IV** in this thesis.

[25] R. S. Johnson. Camassa–Holm, Korteweg–de Vries and related models for water waves. *J. Fluid Mech.*, 455:63–82, 2002.

[26] E. Wahlén. On the peakon-antipeakon interaction. *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.*, to appear.

[27] K. Yosida. *Functional Analysis.* Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

Paper V

Global conservative multipeakon solutions of the Camassa-Holm equation.

H. Holden and X. Raynaud

To appear in *Journal of Hyperbolic Differential Equation.*

GLOBAL CONSERVATIVE MULTIPEAKON SOLUTIONS OF THE CAMASSA–HOLM EQUATION

HELGE HOLDEN AND XAVIER RAYNAUD

Abstract. We show how to construct globally defined multipeakon solutions of the Camassa–Holm equation. The construction includes in particular the case with peakon-antipeakon collisions. The solutions are conservative in the sense that the associated energy is constant for almost all times. Furthermore, we construct a new set of ordinary differential equations that determines the multipeakons globally. The system remains globally well-defined.

1. Introduction

The Cauchy problem for the Camassa–Holm equation [8, 9]

$$u_t - u_{xxt} + 2\kappa u_x + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0, \quad u|_{t=0} = u_0, \tag{1.1}$$

has received considerable attention the last decade. With κ positive it models, see [19], propagation of unidirectional gravitational waves in a shallow water approximation, with u representing the fluid velocity. The Camassa–Holm equation has a bi-Hamiltonian structure and is completely integrable. It has infinitely many conserved quantities. In particular, for smooth solutions the quantities

$$\int u\,dx, \quad \int (u^2 + u_x^2)\,dx, \quad \int (u^3 + uu_x^2)\,dx \tag{1.2}$$

are all time independent.

In this article we consider the case $\kappa = 0$ on the real line, that is,

$$u_t - u_{xxt} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0, \tag{1.3}$$

and henceforth we refer to (1.3) as the Camassa–Holm equation.

Solutions of the Camassa–Holm equation may experience wave breaking in the sense that the solution develops singularities in finite time, while keeping the H^1 norm finite. Continuation of the solution beyond the time of wave breaking is a challenging problem. It is most easily explained in the context of multipeakons, which are special solutions of the Camassa–Holm equation of the form

$$u(t,x) = \sum_{i=1}^{n} p_i(t) e^{-|x - q_i(t)|}, \tag{1.4}$$

Key words and phrases. Camassa–Holm equation, multipeakons, conservative solutions.

where the $(p_i(t), q_i(t))$ satisfy the explicit system of ordinary differential equations

$$\dot{q}_i = \sum_{j=1}^{n} p_j e^{-|q_i - q_j|}, \quad \dot{p}_i = \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i - q_j|}.$$

Observe that the solution (1.4) is not smooth even with continuous functions $(p_i(t), q_i(t))$; one possible way to interpret (1.4) as a weak solution of (1.3) is to rewrite Eq. (1.3) as

$$u_t + \big(\frac{1}{2}u^2 + (1 - \partial_x^2)^{-1}(u^2 + \frac{1}{2}u_x^2)\big)_x = 0.$$

Peakons interact in a way similar to that of solitons of the Korteweg–de Vries equation, and wave breaking may appear when at least two of the q_i's coincide. If all the $p_i(0)$ have the same sign, the peakons move in the same direction. Furthermore, in that case the solution experiences no wave breaking, and one has a unique global solution. Higher peakons move faster than the smaller ones, and when a higher peakon overtakes a smaller, there is an exchange of mass, but no wave breaking takes place. Furthermore, the $q_i(t)$ remain distinct, and thus there is no collision. However, if some of $p_i(0)$ have opposite sign, wave breaking or collision may incur, see, e.g., [4, 20]. For simplicity, consider the case with $n = 2$ and one peakon $p_1(0) > 0$ (moving to the right) and one antipeakon $p_2(0) < 0$ (moving to the left). In the symmetric case ($p_1(0) = -p_2(0)$ and $q_1(0) = -q_2(0) < 0$) the solution will vanish pointwise at the collision time t^* when $q_1(t^*) = q_2(t^*)$, that is, $u(t^*, x) = 0$ for all $x \in \mathbb{R}$, see Fig. 1. Clearly, at least two scenarios are possible; one is to let $u(t, x)$ vanish identically for $t > t^*$, and the other possibility is to let the peakon and antipeakon "pass through" each other in a way that is consistent with the Camassa–Holm equation. In the first case the energy $\int (u^2 + u_x^2)\, dx$ decreases to zero at t^*, while in the second case, the energy remains constant except at t^*. Clearly, the well-posedness of the equation is a delicate matter in this case. The first solution could be denoted a dissipative solution, while the second one could be called conservative, which is the class of solutions we study here. Other solutions are also possible. Global dissipative solutions of a more general class of equations were derived by Coclite, Holden, and Karlsen [12, 13]. In their approach the solution was obtained by first regularizing the equation by adding a small diffusion term ϵu_{xx} to the equation, and subsequently analyzing the vanishing viscosity limit $\epsilon \to 0$.

Global conservative solutions of the Camassa–Holm were recently studied by using a completely new approach, see [5, 6, 15, 18]. In this approach the Camassa–Holm equation is reformulated as a system of ordinary differential equations taking values in a Banach space, see Sect. 2. This allows for the construction of a global and stable solution. To obtain a well-posed initial-value problem it is necessary to introduce the associated energy as an additional variable.

We here study in detail this construction in the context of multipeakons, following [18] where the transformation into new variables can be interpreted as a transformation from Eulerian into Lagrangian coordinates. The explicit nature of multipeakons make them very interesting objects to study in a relation to wave breaking. In particular, the singularity corresponds to a focusing of the energy into a Dirac delta-function.

The general construction in [18] is rather complicated, making the case of multipeakons involved. We show that multipeakons are given as continuous solutions u that on intervals $[y_i(t), y_{i+1}(t)]$ satisfy

$$u - u_{xx} = 0 \text{ with boundary conditions } u(t, y_i(t)) = u_i(t),\ u(t, y_{i+1}(t)) = u_{i+1}(t).$$

The (y_i, u_i) are given by a set of ordinary differential equations, which in addition includes a third variable that measures the energy of the system. The variables y_i and u_i denote the location of the point (for fixed time) where the solution u has a discontinuous spatial derivative (the "peak"), and the value of u at that point, respectively. The system of ordinary differential equations, which is new, remains globally well-defined.

In addition to allowing for a detailed study of the property of solutions near wave breaking, multipeakons are important as building blocks for general solutions. Indeed, if the initial data u_0 is in H^1 and $m_0 := u_0 - u_0''$ is a positive Radon measure, then it can proved, see [17], that one can construct a sequence of multipeakons that converges in $L^\infty_{\mathrm{loc}}(\mathbb{R}; H^1_{\mathrm{loc}}(\mathbb{R}))$ to the unique global solution of the Camassa–Holm equation. See also [6, 15].

The method is illustrated by explicit calculations in the cases $n = 1$ and $n = 2$ (see also [2, 3, 20]) and by numerical computations in the case $n = 4$ with and without wave breaking.

Furthermore, the methods presented in this paper can be used to derive numerical methods that converge to conservative solutions rather than dissipative solutions. This contrasts finite difference methods that normally converge to dissipative solutions, see [16]. See also [17]. Results will be presented separately.

2. Global conservative solutions

The goal of this section is to introduce the results obtained in [18], namely the construction of the continuous semigroup of conservative solutions of the Camassa–Holm equation with a change of variable to Lagrangian coordinates. The equation can be rewritten as the following system

$$u_t + uu_x + P_x = 0, \tag{2.1a}$$

$$P - P_{xx} = u^2 + \frac{1}{2}u_x^2. \tag{2.1b}$$

It is not hard to check that the energy density $u^2 + u_x^2$ fulfills the following transport equation

$$(u^2 + u_x^2)_t + (u(u^2 + u_x^2))_x = (u^3 - 2Pu)_x. \tag{2.2}$$

We denote $y_t(t, \xi) = u(t, y(t, \xi))$ the characteristics and set

$$U(t, \xi) = u(t, y(t, \xi)) \text{ and } H(t, \xi) = \int_{-\infty}^{y(t,\xi)} (u^2 + u_x^2)\, dx,$$

which corresponds to the Lagrangian velocity and the Lagrangian cumulative energy distribution, respectively. We set $\zeta(\xi)=y(\xi)-\xi$. From the definition of the characteristics, it follows that

$$U_t(t,\xi)=u_t(t,y)+y_t(t,\xi)u_x(t,y)=-P_x\circ y\,(t,\xi). \tag{2.3}$$

This last term can be expressed uniquely in term of U, y, and H. From (2.1b), we obtain the following explicit expression for P,

$$P(t,x)=\frac{1}{2}\int_{\mathbb{R}}e^{-|x-z|}\big(u^2(t,z)+\frac{1}{2}u_x^2(t,z)\big)\,dz. \tag{2.4}$$

Thus we have

$$P_x\circ y\,(t,\xi)=-\frac{1}{2}\int_{\mathbb{R}}\operatorname{sgn}(y(t,\xi)-z)e^{-|y(t,\xi)-z|}\big(u^2(t,z)+\frac{1}{2}u_x^2(t,z)\big)\,dz,$$

and, after the change of variables $z=y(t,\eta)$,

$$\begin{aligned}P_x\circ y(t,\xi)=-\frac{1}{2}\int_{\mathbb{R}}\Big[&\operatorname{sgn}(y(t,\xi)-y(t,\eta))e^{-|y(t,\xi)-y(t,\eta)|}\\&\times\Big(u^2(t,y(t,\eta))+\frac{1}{2}u_x^2(t,y(t,\eta))\Big)\,y_\xi(t,\eta)\Big]\,d\eta.\end{aligned}$$

Finally, since $H_\xi=(u^2+u_x^2)\circ y\,y_\xi$,

$$P_x\circ y\,(\xi)=-\frac{1}{4}\int_{\mathbb{R}}\operatorname{sgn}(y(\xi)-y(\eta))\exp(-|y(\xi)-y(\eta)|)\big(U^2y_\xi+H_\xi\big)(\eta)\,d\eta \tag{2.5}$$

where the t variable has been dropped to simplify the notation. It turns out that $y_\xi(t,\xi)\geq 0$ for all t and almost every ξ, see Definition 2.1 and [18, Theorem 2.8]. Thus, $P_x\circ y$ is can be replaced by Q where

$$Q(t,\xi)=-\frac{1}{4}\int_{\mathbb{R}}\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi+H_\xi\big)(\eta)\,d\eta, \tag{2.6}$$

and, slightly abusing the notation, we write

$$P(t,\xi)=\frac{1}{4}\int_{\mathbb{R}}\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big)\big(U^2y_\xi+H_\xi\big)(\eta)\,d\eta. \tag{2.7}$$

Thus $P_x\circ y$ and $P\circ y$ can be replaced by equivalent expressions given by (2.6) and (2.7) which only depend on our new variables U, H, and y. From (2.2), it follows that

$$H_t=\int_{-\infty}^{y}(u^2+u_x^2)_t\,dx+y_t\circ y(u^2+u_x^2)\circ y=\int_{-\infty}^{y}(u^3-2Pu)_x\,dx=U^3-2PU. \tag{2.8}$$

Finally, from (2.3) and (2.8), we infer that the Camassa–Holm equation is formally equivalent to the following system

$$\begin{cases}\zeta_t=U,\\ U_t=-Q,\\ H_t=U^3-2PU.\end{cases} \tag{2.9}$$

We look at (2.9) as a system of ordinary differential equations in the Banach space

$$E = V \times H^1(\mathbb{R}) \times V$$

where $V = \{f \in C_b(\mathbb{R}) \mid f_\xi \in L^2(\mathbb{R})\}$. By a contraction argument we establish the short-time existence of solutions ([18, Theorem 2.3]). We have

$$Q_\xi = -\frac{1}{2}H_\xi - \left(\frac{1}{2}U^2 - P\right) y_\xi \quad \text{and} \quad P_\xi = Qy_\xi, \tag{2.10}$$

and, differentiating (2.9) yields

$$\begin{cases} \zeta_{\xi t} = U_\xi \text{ (or } y_{\xi t} = U_\xi), \\ U_{\xi t} = \frac{1}{2}H_\xi + \left(\frac{1}{2}U^2 - P\right) y_\xi, \\ H_{\xi t} = -2Q\, U y_\xi + \big(3U^2 - 2P\big)\, U_\xi. \end{cases} \tag{2.11}$$

The system (2.11) is semilinear with respect to the variables y_ξ, U_ξ and H_ξ.

Global solutions of (2.9) may not exist for all initial data in E. However they exist when the initial data $\bar{X} = (\bar{y}, \bar{U}, \bar{H})$ belongs to the set $\mathcal{G}$ ([18, Theorem 2.8]) where $\mathcal{G}$ is defined as follows:

Definition 2.1. The set $\mathcal{G}$ is composed of all $(\zeta, U, H) \in E$ such that

$$(\zeta, U, H) \in \left[W^{1,\infty}(\mathbb{R})\right]^3, \tag{2.12a}$$

$$y_\xi \geq 0, H_\xi \geq 0, y_\xi + H_\xi > 0 \text{ almost everywhere, and } \lim_{\xi \to -\infty} H(\xi) = 0, \tag{2.12b}$$

$$y_\xi H_\xi = y_\xi^2 U^2 + U_\xi^2 \text{ almost everywhere,} \tag{2.12c}$$

where we denote $y(\xi) = \zeta(\xi) + \xi$.

The proof of the global existence of the solution for initial data in $\mathcal{G}$ ([18, Theorem 2.8]) relies essentially on the fact that the set $\mathcal{G}$ is preserved by the flow, that is, if $X(0) \in \mathcal{G}$, then $X(t) \in \mathcal{G}$ for all time t, for any solution $X(t)$ of (2.9) with initial data in $\mathcal{G}$ ([18, Lemma 2.7]). We also have that, for almost every t, $y_\xi(t, \xi) > 0$ for almost every ξ, which implies that for almost every t, $\xi \mapsto y(t, \xi)$ is invertible [18].

To obtain a semigroup of solution for (2.1), we have to consider the space $\mathcal{D}$, which characterizes the solutions in *Eulerian coordinates*:

Definition 2.2. The set $\mathcal{D}$ is composed of all pairs (u, μ) such that u belongs to $H^1(\mathbb{R})$ and μ is a positive finite Radon measure whose absolute continuous part, μ_{ac}, satisfies

$$\mu_{\text{ac}} = (u^2 + u_x^2)\, dx. \tag{2.13}$$

The set $\mathcal{D}$ allows the energy density to have a singular part and a positive amount of energy can concentrate on a set of Lebesgue measure zero. In [14], the Camassa–Holm equation is derived as a geodesic equation on the group of diffeomorphism equipped with a right-invariant metric. The right-invariance of the metric can be interpreted as an invariance with respect to relabeling as noted in [1]. This is a property that we also observe in our setting. We denote by G the subgroup of the group of homeomorphisms from $\mathbb{R}$ to $\mathbb{R}$ such that

$$f - \mathrm{Id} \text{ and } f^{-1} - \mathrm{Id} \text{ both belong to } W^{1,\infty}(\mathbb{R}) \tag{2.14}$$

where Id denotes the identity function. The set G can be interpreted as the set of relabeling functions. Let $\mathcal{F}$ be the following subset of $\mathcal{G}$

$$\mathcal{F} = \{X = (y, U, H) \in \mathcal{G} \mid y + H \in G\}.$$

For the sake of simplicity, for any $X = (y, U, H) \in \mathcal{F}$ and any function $f \in G$, we denote $(y \circ f, U \circ f, H \circ f)$ by $X \circ f$. The map $(f, X) \mapsto X \circ f$ defines an action of the group G on $\mathcal{F}$ ([18, Proposition 3.4]), and we denote by $\mathcal{F}/G$ the quotient space of $\mathcal{F}$ with respect to the action of the group G. The equivalence relation on $\mathcal{F}$ is defined as follows: For any $X, X' \in \mathcal{F}$, X and X' are equivalent if there exists $f \in G$ such that $X' = X \circ f$, that is, if X and X' are equal up to a relabeling.

As proved in [18, Lemma 3.3], $\mathcal{F}$ is preserved by the flow. Let us denote by $S\colon \mathcal{F} \times \mathbb{R}_+ \to \mathcal{F}$ the continuous semigroup which to any initial data $\bar{X} \in \mathcal{F}$ associates the solution $X(t)$ of the system of differential equation (2.9) at time t. The Camassa–Holm equation is invariant with respect to relabeling, that is,

$$S_t(X \circ f) = S_t(X) \circ f \tag{2.15}$$

for any initial data $X \in \mathcal{F}$, any time t and any $f \in \mathcal{F}$. Thus the map $\tilde{S}_t$ from $\mathcal{F}/G$ to $\mathcal{F}/G$ given by $\tilde{S}_t([X]) = [S_t X]$ is well-defined and it generates a continuous semigroup. The topology on $\mathcal{F}/G$ is defined by a complete metric which is derived from the E-norm restricted to $\mathcal{F}$.

In order to transport the continuous semigroup obtained in the Lagrangian framework (solutions in $\mathcal{F}/G$) into the Eulerian framework (solutions in $\mathcal{D}$), we want to establish a bijection between $\mathcal{F}/G$ and $\mathcal{D}$. Let us denote by $L\colon \mathcal{D} \to \mathcal{F}/G$ the map transforming Eulerian coordinates into Lagrangian coordinates defined as follows: For any (u, μ) in $\mathcal{D}$, let

$$y(\xi) = \sup\{y \mid \mu((-\infty, y)) + y < \xi\}, \tag{2.16a}$$
$$H(\xi) = \xi - y(\xi), \tag{2.16b}$$
$$U(\xi) = u \circ y(\xi). \tag{2.16c}$$

We define $L(u, \mu) \in \mathcal{F}/G$ to be the equivalence class of (y, U, H). In the other direction, we obtain μ, the energy density in Eulerian coordinates, by pushing forward by y the energy density in Lagrangian coordinates, $H_\xi\, d\xi$. Recall that the push-forward of a measure ν by a measurable function f is the measure $f_\#\nu$ defined as

$$f_\#\nu(B) = \nu(f^{-1}(B))$$

for all Borel sets B. Given any element $[X]$ in $\mathcal{F}/G$, let (u, μ) be

$$u(x) = U(\xi) \text{ for any } \xi \text{ such that } x = y(\xi), \tag{2.17a}$$
$$\mu = y_\#(H_\xi\, d\xi). \tag{2.17b}$$

Then (u,μ) belongs to $\mathcal{D}$ and is independent of the representative $X=(y,U,H)\in\mathcal{F}$ we choose for $[X]$. We denote by $M\colon \mathcal{F}/G\to\mathcal{D}$ the map which to any $[X]$ in $\mathcal{F}/G$ associates (u,μ) as given by (2.17). The map M corresponds to the transformation from Lagrangian to Eulerian coordinates. In [18, Theorems 3.8, 3.11], it is proven that the maps L and M are well-defined and that $L^{-1}=M$, see [18, Theorem 3.12].

We define the metric $d_{\mathcal{D}}$ on $\mathcal{D}$ as

$$d_{\mathcal{D}}((u,\mu),(\bar u,\bar\mu))=d_{\mathcal{F}/G}(L(u,\mu),L(\bar u,\bar\mu)).$$

Since $\mathcal{F}/G$ equipped with $d_{\mathcal{F}/G}$ is a complete metric space, $\mathcal{D}$ equipped with the metric d_D is a complete metric space. For each $t\in\mathbb{R}$, we define the map T_t from $\mathcal{D}$ to $\mathcal{D}$ as

$$T_t=M\tilde S_tL.$$

We have the following commutative diagram:

$$\begin{tikzcd} \mathcal{D} & \mathcal{F}/G \arrow[l, "M"'] \\ \mathcal{D} \arrow[u, "T_t"] \arrow[r, "L"] & \mathcal{F}/G \arrow[u, "\tilde S_t"'] \end{tikzcd} \tag{2.18}$$

Finally, we have the following main result from [18].

Theorem 2.3. *$T\colon \mathcal{D}\times\mathbb{R}_+\to\mathcal{D}$ (where $\mathcal{D}$ is defined by Definition 2.2) defines a continuous semigroup of solutions of the Camassa–Holm equation, that is, given $(\bar u,\bar\mu)\in\mathcal{D}$, if we denote $t\mapsto(u(t),\mu(t))=T_t(\bar u,\bar\mu)$ the corresponding trajectory, then u is a weak solution of the Camassa–Holm equation* (2.1)*. Moreover μ is a weak solution of the following transport equation for the energy density*

$$\mu_t+(u\mu)_x=(u^3-2Pu)_x. \tag{2.19}$$

Furthermore, we have that

$$\mu(t)(\mathbb{R})=\mu(0)(\mathbb{R})\ \textit{for all } t \tag{2.20}$$

and

$$\mu(t)(\mathbb{R})=\mu_{ac}(t)(\mathbb{R})=\|u(t)\|_{H^1}^2=\mu(0)(\mathbb{R})\ \textit{for almost all } t. \tag{2.21}$$

Remark 2.4. We denote the unique solution described in the theorem as a *conservative* weak solution of the Camassa–Holm equation.

3. Characterization of multipeakon solutions

Peakons are given by

$$u(t,x)=\sum_{i=1}^n p_i(t)e^{-|x-q_i(t)|} \tag{3.1}$$

where p_i, q_i satisfy the system of ordinary differential equations

$$\dot{q}_i = \sum_{j=1}^{n} p_j e^{-|q_i - q_j|}, \tag{3.2a}$$

$$\dot{p}_i = \sum_{j=1}^{n} p_i p_j \operatorname{sgn}(q_i - q_j) e^{-|q_i - q_j|}. \tag{3.2b}$$

Note that (3.2) is a Hamiltonian system, viz.,

$$\dot{q}_i = \frac{\partial H(p,q)}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H(p,q)}{\partial q_i},$$

with Hamiltonian

$$H(p,q) = \frac{1}{2}\sum_{i,j=1}^{n} p_i p_j e^{-|q_i - q_j|}.$$

Clearly, if the q_i remain distinct, the system (3.2) allows for a global smooth solution. By inserting that solution into (3.1) we find that u is a global weak solution of the Camassa–Holm equation. See, e.g., [17] for details. In the case where $p_i(0)$ have the same sign for all $i \in \{1, \dots, n\}$, then the $q_i(t)$ remain distinct, and (3.2) admits a unique global solution, see [7, 11, 10, 17]. In this case, the peakons are traveling in the same direction. However, when two peakons have opposite signs, see, e.g., [5, 6], collisions may occur, and if so, the system (3.2) blows up, or, more precisely, some of the p_i blow up.

Our aim is to use the variables (y, U, H) to characterize multipeakons in a way that avoids the problems related to blow up. In particular, we will derive a new system of ordinary differential equations for the multipeakon solutions which is well-posed even when collisions occur.

We consider initial data $\bar{u}$ given by

$$\bar{u}(x) = \sum_{i=1}^{n} p_i e^{-|x - \xi_i|}. \tag{3.3}$$

Without loss of generality, we assume that the p_i are all nonzero, and that the ξ_i are all distinct. From Theorem 2.3 we know that there exists a unique and global weak solution with initial data (3.3), and the aim is to characterize this solution explicitly. The most natural way to define a multipeakon is to say that, given a time t, there exist p_i and ξ_i such that u can be expressed in the form given in (3.1). However, the variables p_i are not appropriate since they blow up at collisions. That is why we will prefer the following characterization of multipeakons. Given the position of the peaks x_i and the values u_i of u at the peaks, u is defined on each interval $[x_i, x_{i+1}]$ as the solution of the Dirichlet problem

$$u - u_{xx} = 0,\ u(x_i) = u_i,\ u(x_{i+1}) = u_{i+1}.$$

Clearly, the function (3.1) satisfies this for each fixed time t, but we will now show that this property persists for conservative solutions.

A multipeakon is piecewise C^∞ with discontinuous first derivative at the peaks. From (3.2a), we infer that

$$\dot{q}_i = u(q_i)$$

which means that the peaks and therefore the discontinuities follow the characteristics. In this case, the Lagrangian point of view becomes very convenient, as the location of the peaks is known a priori. Let us prove that $\bar{X} = (\bar{y}, \bar{U}, \bar{H})$ given by

$$\bar{y}(\xi) = \xi, \tag{3.4a}$$

$$\bar{U}(\xi) = \bar{u}(\xi), \tag{3.4b}$$

$$\bar{H}(\xi) = \int_{-\infty}^{\xi} (u^2 + u_x^2)\,dx, \tag{3.4c}$$

is a representative of u in Lagrangian coordinates, that is, $[\bar{X}] = L(\bar{u}, (\bar{u}^2 + \bar{u}_x^2)dx)$. First we have to check that $\bar{X} \in \mathcal{F}$. Since $\bar{u}$ is a multipeakon, from (3.3), we have that $\bar{u} \in W^{1,\infty}(\mathbb{R}) \cap H^1(\mathbb{R})$. Hence, $\bar{U}$ and $\bar{H}$ both belong to $W^{1,\infty}(\mathbb{R})$ while $\bar{y} - \mathrm{Id}$ is identically zero. Due to the exponential decay of $\bar{u}$ and $\bar{u}_x$ and since $\bar{H}_\xi \in L^\infty(\mathbb{R})$, we have $\bar{H}_\xi \in L^2(\mathbb{R})$. The properties (2.12) are straightforward to check. Furthermore, it is not hard to check that $M([\bar{X}]) = (\bar{u}, (\bar{u}^2 + \bar{u}_x^2)dx)$. Hence, since $L \circ M = \mathrm{Id}$, we get $[\bar{X}] = L(\bar{u}, (\bar{u}^2 + \bar{u}_x^2)dx)$. We set $\mathcal{A} = \mathbb{R} \setminus \{\xi_1, \ldots, \xi_n\}$. The functions $\bar{U}$ and $\bar{H}$ belong to $C^2(\mathcal{A})$ (they even belong to $C^\infty(\mathcal{A})$). This property is preserved by the equation, as the next proposition shows.

Proposition 3.1. *Given $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) \in \mathcal{F}$ such that $\bar{X} \in [C^2(\mathcal{A})]^3$, the solution $X = (y, U, H)$ of (2.9) with $\bar{X}$ as initial data belongs to $C^1(\mathbb{R}_+, [C^2(\mathcal{A})]^3)$.*

Proof. We prove this proposition by repeating the contraction argument of [18, Theorem 2.3], replacing E by

$$\bar{E} = E \cap [C^2(\mathcal{A})]^3.$$

The norm on $\bar{E}$ is given by

$$\|X\|_{\bar{E}} = \|X\|_E + \|y - \mathrm{Id}\|_{W^{2,\infty}(\mathcal{A})} + \|U\|_{W^{2,\infty}(\mathcal{A})} + \|H\|_{W^{2,\infty}(\mathcal{A})}$$

We have to prove that $\mathcal{P} : X \mapsto P$ and $\mathcal{Q} : X \mapsto Q$ are Lipschitz maps from bounded sets of $\bar{E}$ into $H^1(\mathbb{R}) \cap C^2(\mathcal{A})$. Given a bounded set $B = \{X \in \bar{E} \mid \|X\|_{\bar{E}} \leq C_B\}$ where C_B is a positive constant we have, from [18, Lemma 2.1], that

$$\left\|\mathcal{Q}(X) - \mathcal{Q}(\bar{X})\right\|_{L^\infty(\mathbb{R})} \leq C\left\|X - \bar{X}\right\|_E \leq C\left\|X - \bar{X}\right\|_{\bar{E}}$$

for a constant C which only depends on C_B. The derivative of Q is given by (2.10). When a map is Lipschitz on bounded sets, we will say that it is B-Lipschitz. It is not hard to prove that the products of two B-Lipschitz maps from $\bar{E}$ into $C(\mathcal{A})$ is also a B-Lipschitz map from $\bar{E}$ into $C(\mathcal{A})$. Hence, from (2.10), $\mathcal{Q}$ is B-Lipschitz from $\bar{E}$ into $C^1(\mathcal{A})$. In the same way, we obtain the same result for $\mathcal{P}$. We can compute the derivative of P_ξ and Q_ξ on $\mathcal{A}$, and we obtain

$$Q_{\xi\xi} = -\frac{1}{2}H_{\xi\xi} - (UU_\xi - Qy_\xi)y_\xi - (\frac{1}{2}U^2 - P)y_{\xi\xi}, \tag{3.5}$$

$$P_{\xi\xi} = Qy_{\xi\xi} - \frac{1}{2}H_\xi y_\xi - (\frac{1}{2}U^2 - P)y_\xi^2. \tag{3.6}$$

Since $Q_{\xi\xi}$ and $P_{\xi\xi}$ are given as sums and products of B-Lipschitz maps from $\bar{E}$ into $C(\mathcal{A})$, we have that $\mathcal{Q}$ and $\mathcal{P}$ are B-Lipschitz from $\bar{E}$ into $C^2(\mathcal{A})$. The system of equation (2.9) can be written in the condensed form

$$X_t = F(X)$$

where $F\colon \bar{E} \to \bar{E}$ is given by $F(X) = [U, -Q, U^3 - 2PU]$. We can see that each component of F consist of products and sums of B-Lipschitz maps from $\bar{E}$ into $C^2(\mathcal{A})$. Hence, F is B-Lipschitz from $\bar{E}$ to $\bar{E}$ and, by the standard contraction argument, we obtain the short-time existence of solutions in $\bar{E}$. As far as global existence is concerned, we know that, for initial data in $W^{1,\infty}(\mathbb{R})$, $\|X\|_{W^{1,\infty}(\mathbb{R})}$ does not blow up, see [18, Lemma 2.4]. For the second derivative, we have, for any $\xi \in \mathcal{A}$, that

$$\begin{aligned}
y_{\xi\xi t} &= U_{\xi\xi},\\
U_{\xi\xi t} &= \frac{1}{2}H_{\xi\xi} + \left[\frac{1}{2}U^2 - P\right] y_{\xi\xi} + \left[UU_\xi y_\xi - Qy_\xi^2\right],\\
H_{\xi\xi t} &= \left[-2QU\right] y_{\xi\xi} + \left[3U^2 - 2P\right] U_{\xi\xi}\\
&\quad + \left[Uy_\xi H_\xi + U^3 y_\xi - 2PUy_\xi^2 + 6UU_\xi^2 - 4QU_\xi y_\xi\right].
\end{aligned} \tag{3.7}$$

The system (3.7) is affine (it equals the sum of a linear transformation and a constant) with respect to $y_{\xi\xi}$, $U_{\xi\xi}$ and $H_{\xi\xi}$. Hence, on any time interval $[0, T)$, we have

$$\|X_{\xi\xi}(t,\,\cdot\,)\|_{L^\infty(\mathcal{A})} \le \|X_{\xi\xi}(0,\,\cdot\,)\|_{L^\infty(\mathcal{A})} + C + C\int_0^t \|X_{\xi\xi}(\tau,\,\cdot\,)\|_{L^\infty(\mathcal{A})}\,d\tau$$

where C is a constant that only depends on $\sup_{t\in[0,T)} \|X(t,\,\cdot\,)\|_{W^{1,\infty}(\mathbb{R})}$, which is bounded. Gronwall's lemma allows us to conclude that $\|X(t,\,\cdot\,)\|_{W^{2,\infty}(\mathcal{A})}$ does not blow up, and therefore the solution is globally defined in $\bar{E}$. □

Next we want to prove that the solution given by Theorem 2.3 with initial data (3.3) satisfies $u - u_{xx} = 0$ between the peaks. Assuming that $y_\xi(t,\xi) \neq 0$, we formally have

$$u_x \circ y = \frac{U_\xi}{y_\xi}$$

and

$$u_{xx} \circ y = \left(\frac{U_\xi}{y_\xi}\right)_\xi \frac{1}{y_\xi} = \frac{U_{\xi\xi}y_\xi - y_{\xi\xi}U_\xi}{y_\xi^3}.$$

Hence,

$$(u - u_{xx})\circ y = \frac{Uy_\xi^3 - U_{\xi\xi}y_\xi + y_{\xi\xi}U_\xi}{y_\xi^3}, \tag{3.8}$$

and we are naturally led to analyze the quantity

$$A = Uy_\xi^3 - U_{\xi\xi}y_\xi + y_{\xi\xi}U_\xi. \tag{3.9}$$

For a given *fixed* $\xi \in \mathcal{A}$, we differentiate (3.9) with respect to time and, after using (2.9), (2.11), and (3.7), we obtain

$$\begin{aligned}\frac{dA}{dt} &= 3UU_\xi y_\xi^2 - Qy_\xi^3 - U_\xi U_{\xi\xi} - y_\xi\big(\frac{1}{2}H_{\xi\xi} + UU_\xi y_\xi + \frac{1}{2}U^2y_{\xi\xi} - Qy_\xi^2 - Py_{\xi\xi}\big)\\ &\quad + \left(\frac{1}{2}H_\xi + (\frac{1}{2}U^2 - P)y_\xi\right)y_{\xi\xi} + U_\xi U_{\xi\xi}\\ &= 2U_\xi Uy_\xi^2 - \frac{1}{2}y_\xi H_{\xi\xi} + \frac{1}{2}H_\xi y_{\xi\xi}. \end{aligned}\tag{3.10}$$

We differentiate (2.12c) with respect to ξ and get

$$y_{\xi\xi}H_\xi + y_\xi H_{\xi\xi} = 2y_\xi y_{\xi\xi}U^2 + 2y_\xi^2UU_\xi + 2U_\xi U_{\xi\xi}. \tag{3.11}$$

After inserting the value of $y_\xi H_{\xi\xi}$ given by (3.11) into (3.10) and multiplying the equation by y_ξ, we get

$$y_\xi\frac{dA}{dt} = y_\xi^3U_\xi U + (H_\xi y_\xi y_{\xi\xi} - y_\xi^2y_{\xi\xi}U^2) - U_\xi y_\xi U_{\xi\xi}.$$

Hence, by (2.12c),

$$y_\xi\frac{dA}{dt} = U_\xi A$$

or, since $y_{\xi t} = U_\xi$,

$$y_\xi\frac{dA}{dt} = y_{\xi t}A. \tag{3.12}$$

Let us prove that $\frac{A}{y_\xi}$ is C^1 in time (we recall that we keep ξ fixed in $\mathcal{A}$). We have

$$\frac{A}{y_\xi} = Uy_\xi^2 - U_{\xi\xi} + \frac{y_{\xi\xi}U_\xi}{y_\xi} \tag{3.13}$$

$$= Uy_\xi^2 - U_{\xi\xi} + \frac{y_{\xi\xi}U_\xi}{y_\xi + H_\xi} + \frac{y_{\xi\xi}H_\xi U_\xi}{(y_\xi + H_\xi)y_\xi}. \tag{3.14}$$

After multiplying (3.11) by $\frac{U_\xi}{y_\xi}$, we obtain

$$\begin{aligned}\frac{y_{\xi\xi}H_\xi U_\xi}{y_\xi} &= -H_{\xi\xi}U_\xi + 2y_{\xi\xi}U^2U_\xi + 2y_\xi UU_\xi^2 + 2\frac{U_\xi^2}{y_\xi}U_{\xi\xi}\\ &= -H_{\xi\xi}U_\xi + 2y_{\xi\xi}U^2U_\xi + 2y_\xi UU_\xi^2 + 2(H_\xi - y_\xi U^2)U_{\xi\xi}\end{aligned}\tag{3.15}$$

because $\frac{U_\xi^2}{y_\xi} = H_\xi + y_\xi U^2$, from (2.12c). Hence, we can rewrite (3.14) as

$$\frac{A}{y_\xi} = \frac{J(X, X_\xi, X_{\xi\xi})}{y_\xi + H_\xi}$$

for some polynomial J. Since $X \in C^1(\mathbb{R}, \bar{E})$, we have X, X_ξ and $X_{\xi\xi}$ are C^1 in time. Since $X(t)$ remains in $\mathcal{G}$ for all t, from (2.12b), we have $y_\xi + H_\xi > 0$ and therefore $1/(y_\xi + H_\xi)$ is C^1 in time. Hence, A/y_ξ is C^1 in time. For any time t such that $y_\xi(t) \neq 0$, that is, for almost every t (see [18, Lemma 2.7]) we have

$$\frac{d}{dt}\left(\frac{A}{y_\xi}\right) = \frac{A_t y_\xi - y_{\xi t} A}{y_\xi^2} = 0$$

from (3.12). Hence, $\frac{A}{y_\xi}$ is constant in time, i.e.,

$$A(t,\xi) = K(\xi) y_\xi(t,\xi), \tag{3.16}$$

for some constant $K(\xi)$ independent of time. This leads to

$$y_\xi^2 (u - u_{xx}) \circ y = K(\xi)$$

which corresponds to the conservation of spatial angular momentum as defined in [1], see [14]. For the multipeakons at time $t = 0$, we have $y(0,\xi) = \xi$ and $(u - u_{xx})(0,\xi) =$ for all $\xi \in \mathcal{A}$. Hence,

$$\frac{A}{y_\xi}(t,\xi) = 0 \tag{3.17}$$

for all time t and all $\xi \in \mathcal{A}$.

Proposition 3.2. *The energy μ admits a singular part μ_s only when two peaks collide and the support of μ_s corresponds to the points of collision of the peaks. Moreover, no more than two peaks can collide at the same time.*

Proof. Let x be a singular point of μ. We claim that $y^{-1}(\{x\})$ then is a closed interval of length $\mu_{\mathrm{s}}(\{x\})$. Let us prove this. For any ξ, from the definition (2.16a) of y, there exists an increasing sequence x_i such that $\lim_{i\to\infty} x_i = y(\xi)$ and

$$\mu((-\infty, x_i)) + x_i \leq \xi. \tag{3.18}$$

Since $(-\infty, x_i)$ is an increasing sequence of sets and $(-\infty, y(\xi)) = \cup_{i\in\mathbb{N}}(-\infty, x_i)$, we have $\lim_{i\to\infty} \mu((-\infty, x_i)) = \mu((-\infty, y(\xi)))$, and it follows from (3.18) that

$$\mu((-\infty, y(\xi))) + y(\xi) \leq \xi. \tag{3.19}$$

We set $\bar{\xi} = \mu((-\infty, x)) + x$ and, using (3.19), it is not hard to prove that $\bar{\xi}$ is the smallest element of $y^{-1}(\{x\})$. Let $\xi \in y^{-1}(\{x\})$, by definition of y, there exists a decreasing sequence x_i which converges to x such that

$$\mu((-\infty, x_i)) + x_i > \xi.$$

Letting i tend to infinity, we obtain

$$\begin{aligned} \xi &\leq \mu((-\infty, x]) + x \\ &\leq \mu((-\infty, x)) + \mu_{\mathrm{s}}(\{x\}) + x \\ &\leq \bar{\xi} + \mu_{\mathrm{s}}(\{x\}). \end{aligned}$$

Hence, $\xi \in [\bar{\xi}, \bar{\xi} + \mu_{\mathrm{s}}(\{x\})]$ and $y^{-1}(\{x\}) \subset [\bar{\xi}, \bar{\xi} + \mu_{\mathrm{s}}(\{x\})]$. Conversely, let us consider $\xi \in [\bar{\xi}, \bar{\xi} + \mu_{\mathrm{s}}(\{x\})]$. Since y is increasing, $y(\xi) \geq y(\bar{\xi}) = x$. Assume that

$y(\xi) > x$. Then, it follows from the definition of y that there exists $x' > x$ such that

$$\mu((-\infty), x') + x' \leq \xi.$$

Since $x' > x$, we have

$$\begin{aligned}\mu((-\infty, x')) &\geq \mu((-\infty, x]) \\ &= \mu((-\infty, x)) + \mu_{\mathrm{s}}(\{x\}) \\ &= \bar{\xi} - x + \mu_{\mathrm{s}}(\{x\}).\end{aligned}$$

Hence, $\bar{\xi} - x + \mu_{\mathrm{s}}(\{x\}) + x' \leq \xi$ which implies $\bar{\xi} + \mu_{\mathrm{s}}(\{x\}) < \xi$. This contradicts the fact that $\xi \in [\bar{\xi}, \bar{\xi} + \mu_{\mathrm{s}}(\{x\})]$. Our claim is proved. This claim is a general result and does not depend on the multipeakon structure of the initial data. For solutions with multipeakon initial data, we have the following result.

Lemma 3.3. *If $y_\xi(t,\xi)$ vanishes at some point $\bar{\xi}$ in the interval (ξ_i, ξ_{i+1}), then $y_\xi(t,\xi)$ vanishes everywhere in (ξ_i, ξ_{i+1}).*

Proof of Lemma 3.3. Let B be the set

$$B = \{\xi \in (\xi_i, \xi_{i+1}) \mid y_\xi(t,\xi) = 0\}.$$

The set B is not empty as $\bar{\xi} \in B$. Since $y_\xi(t,\,\cdot\,) \in C(\mathcal{A})$, B is closed (relatively in (ξ_i, ξ_{i+1})). Let us prove that B is also open. Take a point $\xi_0 \in B$. We have $y_\xi(t,\xi_0) = 0$ and, by (2.12b), it implies $H_\xi(t,\xi_0) > 0$. Since $H_\xi(t,\,\cdot\,) \in C(\mathcal{A})$, there exists an open interval I around ξ_0 such that $H_\xi(t,\xi) > 0$ for all $\xi \in I$. After multiplying (3.17) by U_ξ and using (2.12c), we obtain

$$UU_\xi y_\xi^2 - U_\xi U_{\xi\xi} + y_{\xi\xi}H_\xi - y_{\xi\xi}y_\xi U^2 = 0. \tag{3.20}$$

We differentiate (2.12c) with respect to ξ and obtain

$$U_{\xi\xi}U_\xi = \frac{1}{2}(y_{\xi\xi}H_\xi + H_{\xi\xi}y_\xi) - y_\xi y_{\xi\xi}U^2 - y_\xi^2 U_\xi U. \tag{3.21}$$

Inserting this into (3.20), we end up with an equation of the form

$$y_{\xi\xi}H_\xi = f(\xi)y_\xi \tag{3.22}$$

where $f(\xi)$ is a continuous function of ξ. Since $H_\xi \neq 0$ on I, we obtain

$$\begin{aligned}y_{\xi\xi}(\xi) &= \frac{f(\xi)}{H_\xi(\xi)}y_\xi(\xi), \\ y_\xi(\xi_0) &= 0.\end{aligned}$$

The unique solution of this ordinary differential equation where $y_\xi(\xi)$ plays the role of the unknown, is $y_\xi(\xi) = 0$. Hence, $y_\xi(\xi) = 0$ for all $\xi \in I$. This implies that $I \subset B$ and therefore B is open. Thus B is an open and closed set, relatively in (ξ_i, ξ_{i+1}). Since (ξ_i, ξ_{i+1}) is a connected set, it implies that $B = (\xi_i, \xi_{i+1})$, which concludes the proof of the lemma. ☐

Let us consider a time T when μ admits a singular point that we denote $\{x\}$. Then, the interval of strictly positive length $y^{-1}(\{x\})$ intersects $\mathcal{A}$ and there exists a point $\bar{\xi} \in (\xi_i, \xi_{i+1})$ for some $i \in \{1, \dots, n\}$ such that $y_\xi(T, \bar{\xi}) = 0$ (with the convention $\xi_0 = -\infty$ and $\xi_{n+1} = \infty$). From Lemma 3.3, we get that $[\xi_i, \xi_{i+1}] \subset y^{-1}(\{x\})$. In particular, $y(\xi_i) = y(\xi_{i+1}) = x$, which means that the point x where the energy concentrates, is located at the collision point between two peaks. We claim that

$$[\xi_i, \xi_{i+1}] = y^{-1}(\{x\}), \tag{3.23}$$

which in particular means that no other peak than the ones originating from ξ_i and ξ_{i+1} can be found at x. Assume that (3.23) is not true, then, due to Lemma 3.3, $y^{-1}(\{x\})$ must take the form

$$y^{-1}(\{x\}) = [\xi_j, \xi_k]$$

where $j \leq i$, $k \geq i+1$ and $k - j \geq 2$. We introduce $\bar{X} = (\bar{y}, \bar{U}, \bar{H})$ defined as $\bar{y}(\xi) = y(T, \xi)$, $\bar{U}(\xi) = U(T, \xi)$, and

$$\bar{H}(\xi) = \begin{cases} H(T, \xi_j)\dfrac{\xi_k - \xi}{\xi_k - \xi_j} + H(T, \xi_k)\dfrac{\xi - \xi_j}{\xi_k - \xi_j} & \text{when } \xi \in (\xi_j, \xi_k), \\ H(T, \xi) & \text{otherwise,} \end{cases}$$

so that $\bar{H}$ is linear in (ξ_j, ξ_k) and continuous. Since $y_\xi(T, \xi) = U_\xi(T, \xi) = 0$ and $H_\xi(T, \xi) > 0$ in (ξ_j, ξ_j), it is not hard to check that all the conditions (2.12) are fulfilled and $\bar{X} \in \mathcal{F}$. Let us look at $X(T)$ and $\bar{X}$ in Eulerian coordinates. We write $(u, \mu) = M([X(T)])$ and $(\bar{u}, \bar{\mu}) = M([\bar{X}])$. Since $\bar{y}(\xi) = y(T, \xi)$ and $\bar{U}(\xi) = U(T, \xi)$, it is clear that $\bar{u} = u$. We have, using (2.17b),

$$\mu(\{x\}) = \int_{[\xi_j, \xi_k]} H_\xi \, d\xi = H(\xi_k) - H(\xi_j) = \int_{[\xi_j, \xi_k]} \bar{H}_\xi \, d\xi = \bar{\mu}(\{x\}).$$

Hence, for any Borel set A,

$$\mu(A) = \mu(A \setminus \{x\}) + \mu(\{x\}) = \bar{\mu}(A \setminus \{x\}) + \bar{\mu}(\{x\}) = \bar{\mu}(A)$$

and $\bar{\mu} = \mu$. Since M is injective, we have $[X(T)] = [\bar{X}]$, which means that $X(T)$ and $\bar{X}$ are equivalent and there exists $f \in \mathcal{F}$ such that

$$X(T) \circ f = \bar{X}. \tag{3.24}$$

The point is that $\bar{X}$ is linear in (ξ_j, ξ_k), and therefore it possesses a priori more regularity than $X(T)$ on this interval. Introduce $\tilde{\mathcal{A}} = \mathbb{R} \setminus \{\xi_1, \dots, \xi_j, \xi_k, \dots, \xi_n\}$. We can solve (2.9) backward in time and, slightly abusing the notation, we denote $\bar{X}(t)$ the solution which satisfies $\bar{X}(T) = \bar{X}$ at time T. Proposition 3.1 gives us that $\bar{X}(t) \in [C^2(\tilde{\mathcal{A}})]^3$ for all time t. Since $X(T)$ and $\bar{X}(T)$ are equivalent and satisfy (3.24), by (2.15), we obtain that $X(t) \circ f = \bar{X}(t)$ for all time t. At time $t = 0$, it yields

$$f(\xi) = \bar{y}(0, \xi)$$

because $y(0, \xi) = \xi$. Since $\bar{y}(0, \xi) \in C^2((\xi_j, \xi_k))$, $f \in C^2((\xi_j, \xi_k))$. By definition, see (2.14), the derivative of f^{-1} is bounded. It implies that f_ξ is bounded strictly away from zero, see [18, Lemma 3.2] for a detailed proof of this result. Hence, $f_\xi > 0$ in (ξ_j, ξ_k) and, by the implicit function theorem, f^{-1} belongs to $C^2((\xi_j, \xi_k))$. Hence,

$$u(0,\xi) = U(0,\xi) = \bar{U}(0, f^{-1}(\xi))$$

also belongs to $C^2((\xi_j, \xi_k))$. This contradicts the fact that (ξ_j, ξ_k) contains either ξ_i or ξ_{i+1}, which are points where the derivative of $u(0,\xi)$ is discontinuous. □

Given $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$, there exists ξ, which may not be unique, such that $x = y(t,\xi)$. If $\xi \in \mathcal{A}^c$, then x corresponds to the position of a peak. For $\xi \in \mathcal{A}$, if $y_\xi(t,\xi) = 0$, then, by Lemma 3.3, $y_\xi(t,\xi') = 0$ for all $\xi' \in (\xi_i, \xi_{i+1})$ where i is such that $\xi \in (\xi_i, \xi_{i+1})$, and x again corresponds to a peak. If $y_\xi(t,\xi) \neq 0$ then, using again Lemma 3.3, we have $y_\xi(t,\xi') \neq 0$ for all $\xi' \in (\xi_i, \xi_{i+1})$. By the implicit function theorem, we obtain that $y(t,\,\cdot\,)$ is invertible in (ξ_i, ξ_{i+1}) and its inverse is C^2. It follows that $u(t,x) = U(t, y^{-1}(t,x'))$ is C^2 with respect to the spatial variable and the quantity $(u - u_{xx})(t,x)$ is defined in the classical sense. Moreover, by (3.17) and (3.8), we have

$$(u - u_{xx})(t,x) = \frac{A(t,\xi)}{y_\xi^3(t,\xi)} = 0. \tag{3.25}$$

We summarize our results in the following theorem.

Theorem 3.4. *Given an initial multipeakon solution $\bar{u}(x) = \sum_{i=1}^n p_i e^{-|x-\xi_i|}$, let (y,U,H) be the solution of the system* (2.9) *with initial data $(\bar{y}, \bar{U}, \bar{H})$ given by* (3.4)*. Between adjacent peaks, say $x_i = y(t,\xi_i) \neq x_{i+1} = y(t,\xi_{i+1})$, the solution $u(t,x)$ is twice differentiable with respect to the space variable, and we have*

$$(u - u_{xx})(t,x) = 0 \textit{ for } x \in (x_i, x_{i+1}).$$

We are now in position to start the derivation of a system of ordinary differential equations for multipeakons.

4. A system of ordinary differential equations for multipeakons

For each $i \in \{1, \dots, n\}$, we have, from (2.9),

$$\begin{cases} \dfrac{dy_i}{dt} = u_i, \\ \dfrac{du_i}{dt} = -Q_i, \\ \dfrac{dH_i}{dt} = u_i^3 - 2P_i u_i \end{cases} \tag{4.1}$$

where y_i, u_i, H_i, P_i and Q_i denote $y(t,\xi_i)$, $U(t,\xi_i)$, $H(t,\xi_i)$, $P(t,\xi_i)$ and $Q(t,\xi_i)$, respectively. For almost every t, the function $y(t,\,\cdot\,)$ is invertible. We can make the change of variables $x = y(t,\xi)$ so that P_i and Q_i can be rewritten as

$$P_i = \frac{1}{2}\int_{\mathbb{R}} e^{-|y_i - x|}(u^2 + \frac{1}{2}u_x^2)\,dx, \tag{4.2}$$

$$Q_i = -\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(y_i - x)e^{-|y_i - x|}(u^2 + \frac{1}{2}u_x^2)\,dx. \tag{4.3}$$

Theorem 3.4 gives us a priori the shape of u and allows us to express P_i and Q_i as a function of the variables u_i, H_i and y_i only, thereby transforming (4.1) into a well-posed $3n$-dimensional system of ordinary differential equations.

For almost every time t, $y_\xi(t,\xi) > 0$ for almost every ξ and $\xi \mapsto y(t,\xi)$ is invertible, see Sect. 2 . From now on, we will consider such time t and omit it in the notation when there is no ambiguity. For such time, by Theorem 3.4, no peak coincide. From the same theorem, we know that between two adjacent peaks located at y_i and y_{i+1}, u satisfies $u - u_{xx} = 0$ and therefore u can be written as

$$u(x) = A_i e^x + B_i e^{-x} \text{ for } x \in [y_i, y_{i+1}], \quad i = 1, \dots, n-1. \tag{4.4}$$

The constants A_i and B_i depend on u_i, u_{i+1}, y_i and y_{i+1} and read

$$A_i = \frac{e^{-\bar{y}_i}}{2}\left[\frac{\bar{u}_i}{\cosh(\delta y_i)} + \frac{\delta u_i}{\sinh(\delta y_i)}\right], \tag{4.5}$$

$$B_i = \frac{e^{\bar{y}_i}}{2}\left[\frac{\bar{u}_i}{\cosh(\delta y_i)} - \frac{\delta u_i}{\sinh(\delta y_i)}\right], \tag{4.6}$$

where we for convenience have introduced the variables

$$\begin{aligned} \bar{y}_i &= \frac{1}{2}(y_i + y_{i+1}), \quad \delta y_i = \frac{1}{2}(y_{i+1} - y_i), \\ \bar{u}_i &= \frac{1}{2}(u_i + u_{i+1}), \quad \delta u_i = \frac{1}{2}(u_{i+1} - u_i). \end{aligned} \tag{4.7}$$

The constants A_i and B_i uniquely determine u on the interval $[y_i, y_{i+1}]$. Thus, we can compute

$$\begin{aligned} \delta H_i = H_{i+1} - H_i &= \int_{y_i}^{y_{i+1}} (u^2 + u_x^2)\,dx \\ &= 2\bar{u}_i^2 \tanh(\delta y_i) + 2\delta u_i^2 \coth(\delta y_i). \end{aligned} \tag{4.8}$$

At this point, we can get some more understanding of what is happening at a time of collision. Let t^* be a time when the two peaks located at y_i and y_{i+1} collide, i.e., such that $\lim_{t\uparrow t^*} \delta y_i(t) = 0$. Since the solution u remains in H^1 for all time, the function u remains continuous so that we have $\lim_{t\uparrow t^*} \delta u_i = 0$. Still, A_i and B_i may have a finite limit when t tends to t^*. However, we know that the first derivative blows up (see [5]), and this implies $\lim_{t\uparrow t^*} B_i = -\lim_{t\uparrow t^*} A_i = \infty$. Thus δu_i tends to zero but slower than δy_i. We can now be more precise: Letting t tend to t^* in (4.8), we obtain, to first order in δy_i, that

$$\delta u_i = \sqrt{\frac{\delta H_i}{2}}\sqrt{\delta y_i} + o(\delta y_i).$$

Recall that H and y are increasing functions, and therefore δH_i and δy_i are positive (δH_i is even strictly positive in this case). Hence, we see that δu_i tends to zero at the same rate as $\sqrt{\delta y_i}$. Let us now turn to the computation of P_i as given by (4.2). This computation is quite long but not difficult. We will not give all the intermediate steps but enough so that a courageous reader will have no problems filling in the gaps. We start by writing u as

$$u(t,x) = \sum_{j=0}^{n} (A_j e^x + B_j e^{-x}) \chi_{(y_j, y_{j+1})}(x).$$

We have set $y_0 = -\infty$, $y_{n+1} = \infty$, $u_0 = u_{n+1} = 0$, and $A_0 = u_1 e^{-y_1}$, $B_0 = 0$, $A_n = 0$, $B_n = u_n e^{y_n}$. We have

$$u^2 + \frac{1}{2}u_x^2 = \sum_{j=0}^{n} \left(\frac{3}{2} A_j^2 e^{2x} + A_j B_j + \frac{3}{2} B_j^2 e^{-2x} \right) \chi_{(y_j, y_{j+1})}. \tag{4.9}$$

Introduce

$$\kappa_{ij} = \begin{cases} -1 & \text{if } j \geq i, \\ 1 & \text{otherwise.} \end{cases}$$

Inserting (4.9) into (4.2), we obtain

$$P_i = \frac{1}{2} \sum_{j=0}^{n} \int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i - x)} \left(\frac{3}{2} A_j^2 e^{2x} + A_j B_j + \frac{3}{2} B_j^2 e^{-2x} \right) dx. \tag{4.10}$$

We have

$$\begin{aligned} \int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i - x)} A_j^2 e^{2x}\, dx &= e^{-\kappa_{ij} y_i} A_j^2 \frac{e^{(2+\kappa_{ij}) y_{j+1}} - e^{(2+\kappa_{ij}) y_j}}{2 + \kappa_{ij}} \\ &= e^{-\kappa_{ij} y_i} A_j^2 \exp\left((2+\kappa_{ij})\bar{y}_j\right) \frac{2\sinh\left((2+\kappa_{ij})\delta q_j\right)}{2+\kappa_{ij}}. \end{aligned} \tag{4.11}$$

From (4.5) and (4.8), we get

$$A_j^2 = \frac{e^{-2\bar{y}_j}}{\sinh^2(2\delta y_i)} \left[\bar{u}_j^2 \sinh^2(\delta y_j) + 2\bar{u}_j \delta u_j \sinh(\delta y_j)\cosh(\delta y_j) + \delta u_j^2 \cosh^2(\delta y_j) \right]$$

and

$$A_j^2 = \frac{e^{-2\bar{y}_j}}{4\sinh(2\delta y_i)} \left[\delta H_j + 4\bar{u}_j \delta u_j \right]. \tag{4.12}$$

Similarly, we obtain

$$B_j^2 = \frac{e^{2\bar{y}_j}}{4\sinh(2\delta y_i)} \left[\delta H_j - 4\bar{u}_j \delta u_j \right], \tag{4.13}$$

and

$$A_j B_j = \frac{1}{4\sinh(2\delta y_i)} \left[4\bar{u}_j^2 \tanh(\delta y_j) - \delta H_j \right]. \tag{4.14}$$

Hence, inserting (4.12) into (4.11), we get

$$\int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i-x)}A_j^2e^{2x}\,dx \\ = \frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{2(2+\kappa_{ij})\sinh(2\delta y_j)}\sinh\left((2+\kappa_{ij})\delta y_j\right)[\delta H_j+4\bar{u}_j\delta u_j]. \quad (4.15)$$

In the same way we find

$$\int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i-x)}A_jB_j\,dx = \frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{2\sinh(2\delta y_j)}\sinh(\delta y_j)[4\bar{u}_j^2\tanh(\delta y_j)-\delta H_j)], \quad (4.16)$$

and

$$\int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i-x)}B_j^2e^{-2x}\,dx \\ = \frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{2(\kappa_{ij}-2)\sinh(2\delta y_j)}\sinh\left((\kappa_{ij}-2)\delta y_j\right)[\delta H_j-4\bar{u}_j\delta u_j]. \quad (4.17)$$

After collecting (4.15), (4.16) and (4.17), we can rewrite P_i in (4.10) as

$$\begin{aligned} P_i = \sum_{j=0}^{n} &\frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{4\sinh(2\delta y_j)} \qquad (4.18)\\ &\times\Bigg[\delta H_j\Big[\frac{3}{2}\Big(\frac{\sinh\left((2+\kappa_{ij})\delta y_j\right)}{2+\kappa_{ij}}+\frac{\sinh\left((\kappa_{ij}-2)\delta y_j\right)}{\kappa_{ij}-2}\Big)-\sinh(\delta y_j)\Big]\\ &\quad+6\bar{u}_j\delta u_j\Big[\frac{\sinh\left((2+\kappa_{ij})\delta y_j\right)}{2+\kappa_{ij}}-\frac{\sinh\left((\kappa_{ij}-2)\delta y_j\right)}{\kappa_{ij}-2}\Big]\\ &\quad+4\bar{u}_j^2\sinh(\delta y_j)\tanh(\delta y_j)\Bigg]. \end{aligned}$$

By using only trigonometric manipulations and the fact that $\kappa_{ij}^2=1$, we get the following two identities

$$\frac{3}{2}\Big(\frac{\sinh\left((2+\kappa_{ij})\delta y_j\right)}{2+\kappa_{ij}}+\frac{\sinh\left((\kappa_{ij}-2)\delta y_j\right)}{\kappa_{ij}-2}\Big)-\sinh(\delta y_j)=2\sinh(\delta y_j)\cosh^2(\delta y_j)$$

and

$$\frac{\sinh\left((2+\kappa_{ij})\delta y_j\right)}{2+\kappa_{ij}}-\frac{\sinh\left((\kappa_{ij}-2)\delta y_j\right)}{\kappa_{ij}-2}=\frac{4\kappa_{ij}}{3}\sinh^3(\delta y_j)$$

that we use to simplify (4.18). We end up with

$$P_i=\sum_{j=0}^{n}\frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{8\cosh(\delta y_j)}\left[2\delta H_j\cosh^2(\delta y_j)+8\kappa_{ij}\bar{u}_j\delta u_j\sinh^2(\delta y_j)+4\bar{u}_j^2\tanh(\delta y_j)\right],$$

or

$$P_i=\sum_{j=0}^{n}P_{ij} \quad (4.19)$$

with

$$P_{ij} = \begin{cases} e^{(y_1-y_i)}\frac{u_1^2}{4} & \text{for } j=0, \\ \frac{e^{-\kappa_{ij}y_i}e^{\kappa_{ij}\bar{y}_j}}{8\cosh(\delta y_j)}\big[2\delta H_j\cosh^2(\delta y_j) & \\ \quad +8\kappa_{ij}\bar{u}_j\delta u_j\sinh^2(\delta y_j)+4\bar{u}_j^2\tanh(\delta y_j)\big] & \text{for } j=1,\dots,n-1, \\ e^{(y_i-y_n)}\frac{u_n^2}{4} & \text{for } j=n. \end{cases} \tag{4.20}$$

The term Q_i can be computed in the same way. We have

$$\begin{aligned} Q_i &= \sum_{j=0}^{n} -\frac{1}{2}\int_{y_j}^{y_{j+1}} \operatorname{sgn}(q_i-x)e^{-\kappa_{ij}(y_i-x)}(u^2+\frac{1}{2}u_x^2)\,dx \\ &= \sum_{j=0}^{n} -\kappa_{ij}\int_{y_j}^{y_{j+1}} e^{-\kappa_{ij}(y_i-x)}\left(\frac{3}{2}A_j^2e^{2x}+A_jB_j+\frac{3}{2}B_j^2e^{-2x}\right)dx, \end{aligned}$$

so that we end up with

$$Q_i = -\sum_{j=0}^{n}\kappa_{ij}P_{ij}, \tag{4.21}$$

where P_{ij} is given by (4.20).

We summarize the result in the following theorem.

Theorem 4.1. *Given a multipeakon initial data $\bar{u}$, as given by (3.3), let $\bar{y}_i = \xi_i$, $\bar{u}_i = \bar{u}(\xi_i)$ and $\bar{H}_i = \int_{-\infty}^{\xi_i}(\bar{u}^2+\bar{u}_x^2)\,dx$ for $i=1,\dots,n$. Then, there exists a global in time solution (y_i,u_i,H_i) of (4.1), (4.19)–(4.21) with initial data $(\bar{y}_i,\bar{u}_i,\bar{H}_i)$. For each time t, we define $u(t,x)$ as the solution of the Dirichlet problem*

$$u-u_{xx}=0 \text{ with boundary conditions } u(t,y_i(t))=u_i(t),\ u(t,y_{i+1}(t))=u_{i+1}(t)$$

on each interval $[y_i(t),y_{i+1}(t)]$. Then, u is a conservative solution of the Camassa–Holm equation, and we denote it the multipeakon solution.

The simplest cases can be computed explicitly, and for completeness we include the cases $n=1,2$. In addition we present the case $n=4$ numerically (with and without collisions).

Example 4.2. (i) Let $n=1$. Here we find that $P_1=\frac{1}{2}u_1^2$ and $Q_1=0$. Thus $u_1=c$ and $y_1=ct+a$ for constants a,c, and we finally find the familiar one peakon $u(t,x)=ce^{-|x-ct-a|}$. Note that $H_1=c^2$ is constant. However, we did not use the energy to compute the solution. This is a general result; when there is no collision, the first two equations in (4.1) decouple from the last one, and the energy equation is not needed.

(ii) Let $n=2$. We will solve analytically the case of an antisymmetric pair of peakons when the two peakons collide. In this case, at the collision point the energy concentrates in a single point, see [5]. We take the origin of time equal to the time of collision. The initial conditions are

$$y_1(0)=y_2(0)=u_1(0)=u_2(0)=\delta H_0(0)=\delta H_2(0)=0, \quad \delta H_1(0)=E^2$$

where $E > 0$ corresponds to the energy of the system. The solution remains antisymmetric. Let us assume this for the moment and write

$$\begin{aligned} y &= y_2 = -y_1, \\ u &= u_2 = -u_1, \\ h &= \delta H_1, \end{aligned} \tag{4.22}$$

and, since the total energy is preserved ($H(t,\infty)$ is constant), we have $\delta H_0 = \delta H_2 = \frac{1}{2}(E^2 - h)$. We compute P_i and Q_i using (4.19) and (4.21). After some calculations, we obtain that, whenever the solution is antisymmetric, $P_1 = P_2 = P$ and $Q_1 = -Q_2 = -Q$ where

$$\begin{aligned} P &= (2u^2 + h)\frac{1 + e^{-2y}}{8}, \\ Q &= u^2\frac{1 - e^{-2y}}{4} - h\frac{1 + e^{-2y}}{8}. \end{aligned} \tag{4.23}$$

We are led to the following system of ordinary differential equations

$$\begin{aligned} y_t &= u, \\ u_t &= -Q, \\ h_t &= 2(u^3 - 2Pu), \end{aligned} \tag{4.24}$$

with initial conditions $y(0) = u(0) = 0$ and $h(0) = E^2$. This system can be solved and, after retrieving the original variables by (4.22), since the identities $P_1 = P_2$ and $Q_1 = -Q_2$ hold, $(y_1, y_2, u_1, u_2, H_1, H_2)$ is the unique solution of (4.1) and therefore it is antisymmetric. From (4.23), we get $Q = \frac{1}{2}u^2 - P$. Hence, $h_t = 4uQ = -4uu_t$ and, after integration,

$$h = -2u^2 + E^2.$$

We insert this in (4.24) which yields the following second-order differential equation

$$y_{tt} + \frac{y_t^2}{2} = \frac{E^2}{8}(1 + e^{-2y})$$

with initial data $y(0) = y_t(0) = 0$. We can get rid of the factor E^2 by rescaling the time variable, $t \mapsto Et$, and the equation we have to solve is

$$y_{tt} + \frac{y_t^2}{2} = \frac{1}{8}(1 + e^{-2y}), \tag{4.25}$$

$$y(0) = y_t(0) = 0. \tag{4.26}$$

By a phase-plane analysis, one can prove that $y_t(t) < 0$ for $t < 0$, $y_t(t) > 0$ for $t > 0$ and $y(t) > 0$ for all $t \neq 0$. We make the change of variables $z = e^{-2y}$ and (4.25) becomes

$$-4zz_{tt} + 5z_t^2 = z^2(1 + z). \tag{4.27}$$

We multiply the equation by $z^\alpha z_t$ where α is a constant to be determined and get

$$-4z^{\alpha+1}z_t z_{tt} + 5z_t^3 z^\alpha = z^{2+\alpha}(1 + z)z_t. \tag{4.28}$$

The term on the left is the derivative of $z^\alpha z_t^2$ if $\alpha = -\frac{5}{2}$. Taking this value for α, (4.28) can be integrated and we obtain, after some calculations,

$$z_t^2 = z^2(1-z). \tag{4.29}$$

Hence,

$$z_t = -\varepsilon z\sqrt{1-z} \tag{4.30}$$

where $\varepsilon = \operatorname{sgn}(t)$. We use the change of variables $v = \sqrt{1-z}$ and obtain $v_t = -\varepsilon/2(1-v^2)$, which can be integrated and gives $v(t) = \varepsilon\tanh(\frac{t}{2})$. Finally, going back to the original variables, we obtain

$$y(t) = \ln\cosh(\frac{Et}{2}),$$
$$u(t) = \frac{E}{2}\tanh(\frac{Et}{2}).$$

Note that the ordinary differential equation (4.30) does not satisfy the Lipschitz condition and therefore does not have a unique solution. However, the solution we are looking for is in fact the solution of the second-order ordinary differential equation

$$z_{tt} = z - \frac{3}{2}z^2 \tag{4.31}$$

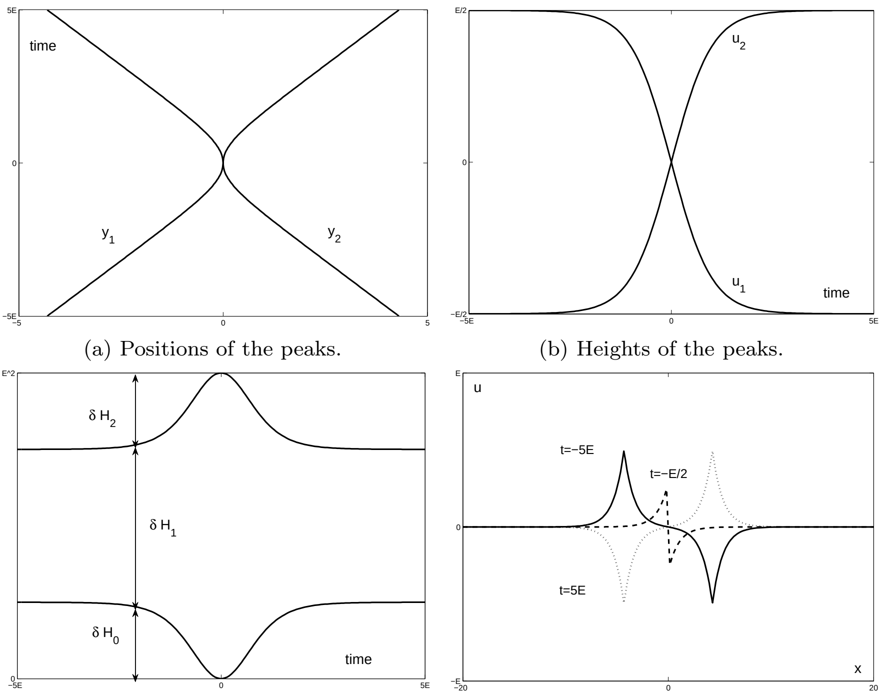
which is obtained from (4.27) by inserting (4.29), which is perfectly well-posed. It is not hard to check that the solution $z(t)$ we obtained indeed satisfies (4.31). In Fig. 1c, we plot δH_0, δH_1, and δH_2 which represent the energy contained between $-\infty$ and y_1, y_1 and y_2, y_2 and $+\infty$, respectively. We see how the energy concentrates at collision time.

The case with two peakons has been computed by Wahlén [20] (see also [2, 3, 4]). For completeness, we reproduce his results here. We have[1]

$$\begin{aligned}
y_1 &= \ln\left(\frac{c_1-c_2}{c_1e^{-c_1t}-c_2e^{-c_2t}}\right), & y_2 &= \ln\left(\frac{c_1e^{c_1t}-c_2e^{c_2t}}{c_1-c_2}\right),\\
u_1 &= \frac{c_1^2-c_2e^{(c_1-c_2)t}}{c_1-c_2e^{(c_1-c_2)t}}, & u_2 &= \frac{c_2^2-c_1e^{(c_1-c_2)t}}{c_2-c_1e^{(c_1-c_2)t}},\\
H_1 &= u_1^2, & H_2 &= 2c_1^2+2c_2^2-u_2^2,
\end{aligned} \tag{4.32}$$

where c_1, c_2 denotes the speed of the peaks y_1 and y_2, respectively, when t tends to infinity. The initial data is set so that, if there is a collision, it occurs at time $t = 0$.

[1] The expressions in (4.32) differ slightly from [20] where two different expressions are given for positive and negative time. This is due to the fact that a relabeling of the solution is implicitly made at collision time so that the two peaks interchange their role at that time. This has no consequence in the Eulerian picture and the resulting function u in Eulerian coordinates is in both cases a conservative solution of the Camassa–Holm equation.



(a) Positions of the peaks.

(b) Heights of the peaks.

(c) Redistribution of the energy between the peaks.

(d) Plot of the solution at different times

Figure 1: antisymmetric multipeakon collision

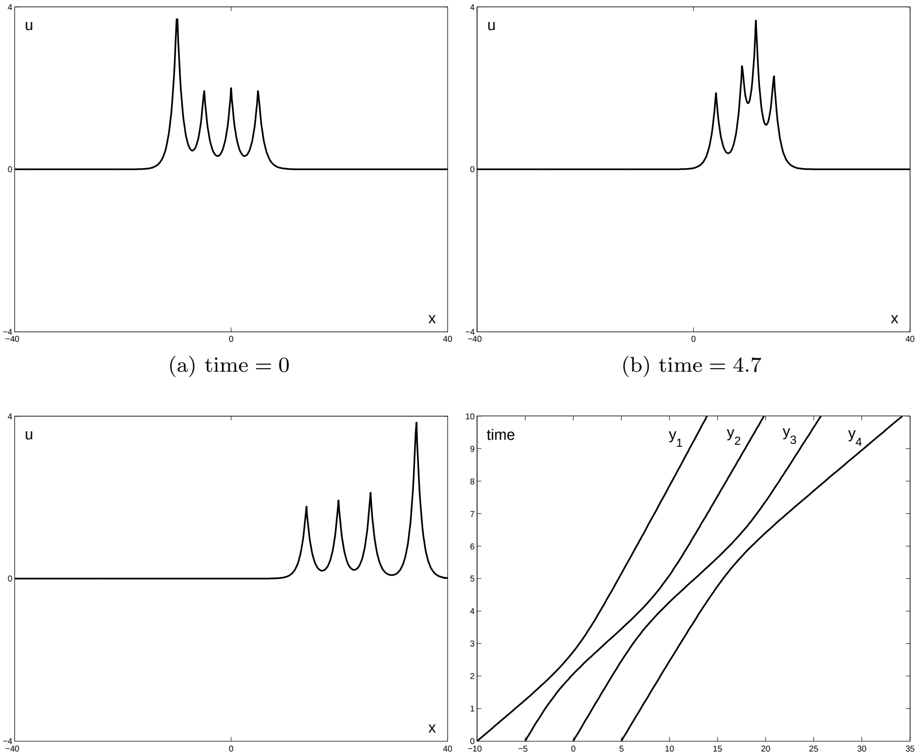
(iii) Let $n = 4$. Consider first the case where there is no wave breaking with all $p_i(0)$ positive for $i = 1, 2, 3, 4$. We take

$$y_1(0) = -10,\ y_2(0) = -5,\ y_3(0) = 0,\ y_4(0) = 5,$$
$$u_1(0) = 4,\ u_2(0) = u_3(0) = u_4(0) = 2.$$

The results are plotted in Fig. 2. Note that the characteristics do not intersect.

Consider finally the case when $p_i(0)$ is positive for $i = 1, 2, 3$, but $p_4(0)$ is negative. The system (4.1) of ordinary differential equations can be solved numerically. We use the explicit Runge–Kutta solver `ode45` for ordinary differential equations from MATLAB. In Fig. 3, we present the results obtained for the initial data

$$y_1(0) = -10,\ y_2(0) = -5,\ y_3(0) = 0,\ y_4(0) = 5,$$
$$u_1(0) = u_2(0) = u_3(0) = 2,\ u_4(0) = -2.$$



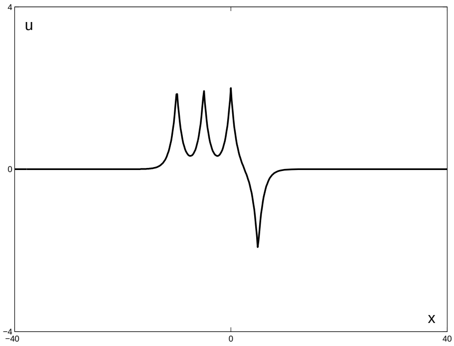
(a) time = 0

(b) time = 4.7

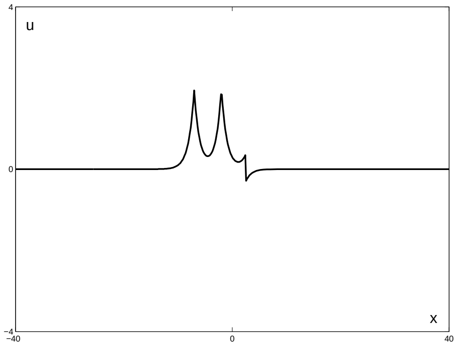
(c) time = 10

(d) positions of the peaks

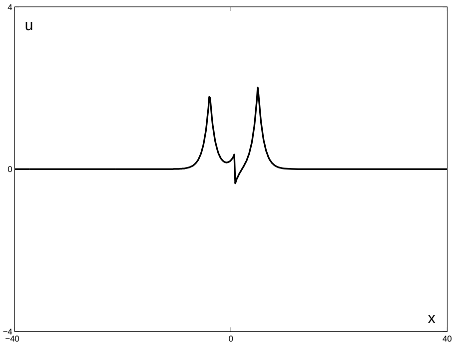
Figure 2: Example of multipeakon without collision.



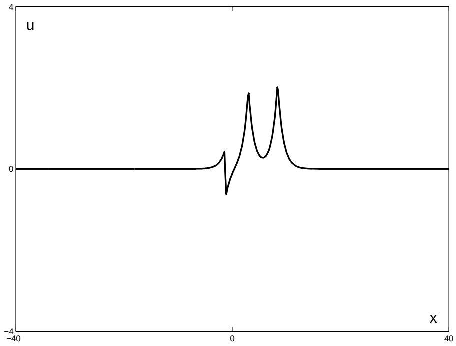
(a) time = 0



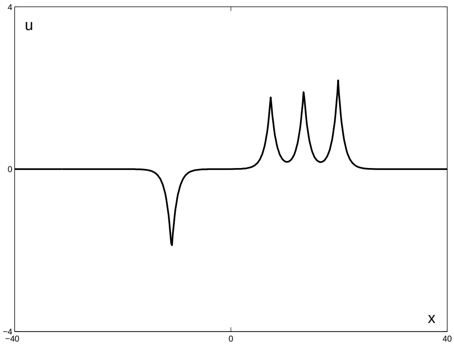
(b) time = 1.5



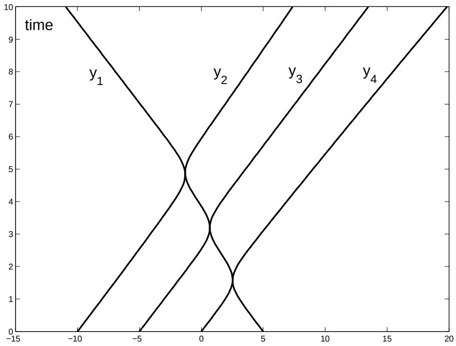
(c) time = 3.1



(d) time = 4.7



(e) time = 10



(f) positions of the peaks

Figure 3: Example of multipeakon with collision

Acknowledgments. The authors gratefully acknowledge the hospitality of the Mittag-Leffler Institute, Sweden, creating a great working environment for research, during the Fall of 2005. This work has been supported in part by the Research Council of Norway.

References

[1] V. I. Arnold and B. A. Khesin. *Topological Methods in Hydrodynamics.* Springer-Verlag, New York, 1998.

[2] R. Beals, D. Sattinger, and J. Szmigielski. Multi-peakons and a theorem of Stieltjes. *Inverse Problems* 15:L1–L4, 1999.

[3] R. Beals, D. Sattinger, and J. Szmigielski. Multipeakons and the classical moment problem. *Adv. Math.* 154:229–257, 2000.

[4] R. Beals, D. Sattinger, and J. Szmigielski. Peakon-antipeakon interaction. *J. Nonlinear Math. Phys.* 8:23–27, 2001.

[5] A. Bressan and A. Constantin. Global conservative solutions of the Camassa–Holm equation. *Preprint*, 2005, submitted.

[6] A. Bressan and M. Fonte. An optimal transportation metric for solutions of the Camassa–Holm equation. *Methods Appl. Anal.*, to appear.

[7] R. Camassa. Characteristics and the initial value problem of a completely integrable shallow water equation. *Discrete Contin. Dyn. Syst. Ser. B*, 3(1):115–139, 2003.

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

[9] R. Camassa, D. D. Holm, and J. Hyman. A new integrable shallow water equation. *Adv. Appl. Mech.*, 31:1–33, 1994.

[10] R. Camassa, J. Huang, and L. Lee. Integral and integrable algorithms for a nonlinear shallow-water wave equation. *Preprint*, 2004, submitted.

[11] R. Camassa, J. Huang, and L. Lee. On a completely integrable numerical scheme for a nonlinear shallow-water equation. *Preprint*, 2004, submitted.

[12] G. M. Coclite, and H. Holden, and K. H, Karlsen. Well-posedness for a parabolic-elliptic system. *Discrete Cont. Dynam. Systems* 13:659–682, 2005.

[13] G. M. Coclite, and H. Holden, and K. H, Karlsen. Global weak solutions to a generalized hyperelastic-rod wave equation. *SIAM J. Math. Anal.*, 37:1044–1069, 2005.

[14] A. Constantin and B. Kolev. On the geometric approach to the motion of inertial mechanical systems. *J. Phys. A*, 35(32):R51–R79, 2002.

[15] M. Fonte. Conservative solution of the Camassa Holm equation on the real line. *arXiv:math.AP/0511549*, 3, 2005.

[16] H. Holden and X. Raynaud. Convergence of a finite difference scheme for the Camassa–Holm equation. *SIAM J. Numer. Anal.* 2006, to appear. **paper I** in this thesis.

[17] H. Holden and X. Raynaud. A convergent numerical scheme for the Camassa–Holm equation based on multipeakons. *Discrete Contin. Dyn. Syst.*, 14(3):505–523, 2006. **paper II** in this thesis.

[18] H. Holden and X. Raynaud. Global conservative solutions of the Camassa–Holm equation — a Lagrangian point of view. *Preprint*, 2006, submitted. **paper IV** in this thesis.

[19] R. S. Johnson. Camassa–Holm, Korteweg–de Vries and related models for water waves. *J. Fluid Mech.*, 455:63–82, 2002.

[20] E. Wahlén. On the peakon-antipeakon interaction. *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.,* 2006, to appear.

Paper VI

Global conservative solutions of the generalized hyperelastic-rod wave equation.

H. Holden and X. Raynaud

Submitted for publication.

GLOBAL CONSERVATIVE SOLUTIONS OF THE GENERALIZED HYPERELASTIC-ROD WAVE EQUATION

HELGE HOLDEN AND XAVIER RAYNAUD

Abstract. We prove existence of global and conservative solutions of the Cauchy problem for the nonlinear partial differential equation $u_t - u_{xxt} + f(u)_x - f(u)_{xxx} + (g(u) + \frac{1}{2}f''(u)(u_x)^2)_x = 0$ where f is strictly convex or concave and g is locally uniformly Lipschitz. This includes the Camassa–Holm equation ($f(u) = u^2/2$ and $g(u) = \kappa u + u^2$) as well as the hyperelastic-rod wave equation ($f(u) = \gamma u^2/2$ and $g(u) = (3-\gamma)u^2/2$) as special cases. It is shown that the problem is well-posed for initial data in $H^1(\mathbb{R})$ if one includes a Radon measure that corresponds to the energy of the system with the initial data. The solution is energy preserving. Stability is proved both with respect to initial data and the functions f and g. The proof uses an equivalent reformulation of the equation in terms of Lagrangian coordinates.

1. Introduction

We solve the Cauchy problem on the line for the equation

$$u_t - u_{xxt} + f(u)_x - f(u)_{xxx} + (g(u) + \frac{1}{2}f''(u)(u_x)^2)_x = 0 \tag{1.1}$$

for strictly convex or concave functions f and locally uniformly Lipschitz functions g with initial data in $H^1(\mathbb{R})$. This equation includes the Camassa–Holm equation [4], the hyperelastic-rod wave equation [11] and its generalization [6, 7] as special cases.

For $f(u) = \frac{u^2}{2}$ and $g(u) = \kappa u + u^2$, we obtain the Camassa–Holm equation:

$$u_t - u_{xxt} + \kappa u_x + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0, \quad u|_{t=0} = \bar{u}, \tag{1.2}$$

which has been extensively studied the last decade [4, 5]. It was first introduced as a model describing propagation of unidirectional gravitational waves in a shallow water approximation, with u representing the fluid velocity, see [18]. The Camassa–Holm equation has a bi-Hamiltonian structure, it is completely integrable, and it has infinitely many conserved quantities.

2000 *Mathematics Subject Classification.* Primary: 65M06, 65M12; Secondary: 35B10, 35Q53.

Key words and phrases. Generalized hyperelastic-rod wave equation, Camassa–Holm equation, conservative solutions.

Supported in part by the Research Council of Norway and the Mittag-Leffler Institute, Sweden.

For $f(u) = \frac{\gamma u^2}{2}$ and $g(u) = \frac{3-\gamma}{2}u^2$, we obtain the hyperelastic-rod wave equation:

$$u_t - u_{txx} + 3uu_x - \gamma(2u_x u_{xx} + uu_{xxx}) = 0,$$

which was introduced by Dai [11, 10, 12] in 1998. It describes far-field, finite length, finite amplitude radial deformation waves in cylindrical compressible hyperelastic rods, and u represents the radial stretch relative to a pre-stressed state.

Furthermore, for $f(u) = \frac{\gamma u^2}{2}$ we find the generalized hyperelastic-rod equation

$$u_t - u_{xxt} + \frac{1}{2}g(u)_x - \gamma(2u_x u_{xx} + uu_{xxx}) = 0, \tag{1.3}$$

which was recently studied by Coclite, Holden, and Karlsen [6, 7], extending earlier results for the Camassa–Holm equation by Xin and Zhang, see [20]. They analyzed the initial value problem for this equation, using an approach based on a certain viscous regularization. By carefully studying the behavior of the limit of vanishing viscosity they derived the existence of a solution of (1.3). This solution could be called a diffusive solution, and will be distinct from the solutions studied here. We will discuss this in more detail below.

We will not try to cover the extensive body of results regarding various aspects of the Camassa–Holm equation. Suffice it here to note that in the case with $\kappa = 0$ the solution may experience wave breaking in finite time in the sense that the function remains bounded while the spatial derivative becomes unbounded with finite H^1-norm. Various mechanisms and conditions are known as to when and if wave breaking occurs. Specifically we mention that Constantin, Escher, and Molinet [8, 9] showed the following result: If the initial data $u|_{t=0} = \bar{u} \in H^1(\mathbb{R})$ and $\bar{m} := \bar{u} - \bar{u}''$ is a positive Radon measure, then equation (1.2) with $\kappa = 0$ has a unique global weak solution $u \in C([0,T], H^1(\mathbb{R}))$, for any T positive, with initial data $\bar{u}$. However, any solution with odd initial data $\bar{u}$ in $H^3(\mathbb{R})$ such that $\bar{u}_x(0) < 0$ blows up in a finite time.

The problem of continuation of the solution beyond wave breaking is intricate. It can be illustrated in the context of a peakon–antipeakon solution. The one peakon is given by $u(t,x) = c\exp(-|x-ct|)$. If c is positive, the solution is called a peakon, and with c negative it is called an antipeakon. One can construct solutions that consist of finitely many peakons and antipeakons. Peakons move to the right, antipeakons to the left. If initial data are given appropriately, one can have a peakon colliding with an antipeakon. In a particular symmetric case they exactly annihilate each other at collision time t^*, thus $u(t^*, x) = 0$. This immediately raises the question about well-posedness of the equation and allows for several distinct ways to continue the solution beyond collision time. For an extensive discussion of this case, we refer to [17] and references therein. We here consider solutions, called conservative, that preserve the energy. In the example just mentioned this corresponds to the peakon and antipeakon passing through each other, and the energy accumulating as a Dirac delta-function at the origin at the time of collision. Thus the problem cannot be well-posed by considering the solution u only. Our approach for the general equation (1.1) is based on the inclusion of the energy, in the form of a (non-negative Radon) measure, together with the function u as initial data. We have seen that singularities occur in these variables. Therefore we transform to a different set of variables, which corresponds to a Lagrangian formulation of the flow, where the singularities do not occur.

Let us comment on the approach in [6, 7]. The equation (1.3) is rewritten as

$$u_t + \gamma u u_x + P_x = 0, \qquad P - P_{xx} = \frac{1}{2}(g(u) - \gamma u^2) + \gamma (u_x)^2. \tag{1.4}$$

By adding the term ϵu_{xx} to the first equation, it is first shown that the modified system has a unique solution.[1] Subsequently, it is proved that the vanishing viscosity limit $\epsilon \to 0$ exists. The limit is shown to be weak solution of (1.3). In particular, that means that $\|u(t,\,\cdot\,)\|_{H^1} \le \|u(0,\,\cdot\,)\|_{H^1}$ and that the solution satisfies an entropy condition $u_x(t,x) \le K + 2/(\gamma t)$ for some constant K. The solution described above with a peakon and an antipeakon "passing through" each other will not satisfy this entropy condition. Thus the solution concept is different in the two approaches.

Here we take a rather different approach. Based on recent techniques developed for the Camassa–Holm equation, see [1, 2, 15, 16], we prove that (1.1) possesses a global weak and conservative solution. Furthermore, we show that the problem is well-posed. In particular we show stability with respect to both perturbations in the initial data and the functions f and g in a suitable topology.

The present approach is based on the fact that the equation can be reformulated as a system of ordinary differential equations taking values in a Banach space. It turns out to be advantageous first to rewrite the equation as

$$u_t + f(u)_x + P_x = 0, \tag{1.5a}$$

$$P - P_{xx} = g(u) + \frac{1}{2} f''(u) u_x^2 \tag{1.5b}$$

where we assume[2]

$$\begin{cases} f \in W^{3,\infty}_{\mathrm{loc}}(\mathbb{R}),\ f''(u) \neq 0,\ u \in \mathbb{R}, \\ g \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}),\ g(0) = 0. \end{cases} \tag{1.6}$$

We will use this assumption throughout the paper.

Specifically, the characteristics are given by

$$y_t(t,\xi) = f'(u(t, y(t,\xi))).$$

[1] In fact, it it is proved that a more general parabolic-elliptic system, allowing, e.g., for explicit spatial and temporal dependence in the various functions, has a solution.

[2] Without loss of generality we may and will assume that $g(0) = 0$. Otherwise, (1.5b) should be replaced by $P - P_{xx} = g(u) - g(0) + \frac{1}{2} f''(u) u_x^2$

Define subsequently

$$\begin{aligned} U(t,\xi) &= u(t,y(t,\xi)), \\ H(t,\xi) &= \int_{-\infty}^{y(t,\xi)} \left(u^2+u_x^2\right)\,dx, \end{aligned}$$

where U and H correspond to the Lagrangian velocity and the Lagrangian cumulative energy distribution, respectively. It turns out that one can derive the following system of ordinary differential equations taking values in an appropriately chosen Banach space, viz.

$$\begin{cases} y_t = U, \\ U_t = -Q, \\ H_t = G(U) - 2PU, \end{cases}$$

where the quantities G, Q, and P can be expressed in terms of the unknowns (y,U,H). Short-term existence is derived by a contraction argument. Global existence as well as stability with respect to both initial data and functions f and g, is obtained for a class of initial data that includes initial data $u|_{t=0}=\bar u$ in $H^1(\mathbb{R})$, see Theorem 2.8. The transition of this result back to Eulerian variables is complicated by several factors, one being the reduction of three Lagrangian variables to two Eulerian variables. There is a certain redundancy in the Lagragian formulation which is identified, and we rather study equivalence classes that correspond to relabeling of the same Eulerian flow. The main existence result is Theorem 2.9. It is shown that the flow is well-posed on this space of equivalence classes in the Lagrangian variables, see Theorem 3.6. A bijection is constructed between Lagrangian and Eulerian variables, see Theorems 3.8–3.11. On the set $\mathcal{D}$ of Eulerian variables we introduce a metric that turns $\mathcal{D}$ into a complete metric space, see Theorem 3.12.

The main result, Theorem 3.13, states the following: There exists a continuous semigroup on $\mathcal{D}$ which to any initial data $(\bar u,\bar\mu)\in\mathcal{D}$ associates the pair $(u(t),\mu(t))\in\mathcal{D}$ such that $u(t)$ is a weak solution of (1.5) and the measure $\mu=\mu(t)$ with $\mu(0)=\bar\mu$, evolves according to the linear transport equation $\mu_t+(u\mu)_x=(G(u)-2Pu)_x$ where the functions G and P are explicitly given. Continuity with respect to all variables, including f and g, is proved. The total energy as measured by μ is preserved, i.e., $\mu(t)(\mathbb{R})=\bar\mu(\mathbb{R})$ for all t.

The abstract construction is illustrated on the one and two peakon solutions for the Camassa–Holm equation.

The paper is organized as follows. In Section 2 the equation is reformulated in terms of Lagrangian variables, and existence is first proved in Lagrangian variables before the results are transformed back to the original Eulerian variables. Stability of the semigroup is provided in Section 3, and the construction is illustrated on concrete examples in Section 4.

2. Existence of solutions

2.1. **Transport equation for the energy density and reformulation in terms of Lagrangian variables.** In (1.5b), P can be written in explicit form:

$$P(t,x) = \frac{1}{2}\int_{\mathbb{R}} e^{-|x-z|}\big(g\circ u + \frac{1}{2}f''\circ u u_x^2\big)(t,z)\,dz. \tag{2.1}$$

We will derive a transport equation for the energy density $u^2+u_x^2$. Assuming that u is smooth, we get, after differentiating (1.5a) with respect to x and using (1.5b), that

$$u_{xt} + \frac{1}{2}f''(u)u_x^2 + f'(u)u_{xx} + P - g(u) = 0. \tag{2.2}$$

Multiply (1.5a) by u, (2.2) by u_x, add the two to find the following equation

$$(u^2+u_x^2)_t + (f'(u)(u^2+u_x^2))_x = -2(Pu)_x + (2g(u)+f''(u)u^2)u_x. \tag{2.3}$$

Define $G(v)$ as

$$G(v) = \int_0^v (2g(z)+f''(z)z^2)\,dz, \tag{2.4}$$

then (2.3) can be rewritten as

$$(u^2+u_x^2)_t + (f'(u)(u^2+u_x^2))_x = (G(u)-2Pu)_x, \tag{2.5}$$

which is transport equation for the energy density $u^2+u_x^2$.

Let us introduce the characteristics $y(t,\xi)$ defined as the solutions of

$$y_t(t,\xi) = f'(u(t,y(t,\xi)) \tag{2.6}$$

with $y(0,\xi)$ given. Equation (2.5) gives us information about the evolution of the amount of energy contained between two characteristics. Indeed, given ξ_1,ξ_2 in $\mathbb{R}$, let

$$H(t) = \int_{y(t,\xi_1)}^{y(t,\xi_2)} \left(u^2+u_x^2\right)\,dx$$

be the energy contained between the two characteristic curves $y(t,\xi_1)$ and $y(t,\xi_2)$. Then, we have

$$\frac{dH}{dt} = \big[y_t(t,\xi)(u^2+u_x^2)\circ y(t,\xi)\big]_{\xi_1}^{\xi_2} + \int_{y(t,\xi_1)}^{y(t,\xi_2)} (u^2+u_x^2)_t\,dx. \tag{2.7}$$

We use (2.5) and integrate by parts. The first term on the right-hand side of (2.7) cancels because of (2.6) and we end up with

$$\frac{dH}{dt} = \left[(G(u)-2Pu)\circ y\right]_{\xi_1}^{\xi_2}. \tag{2.8}$$

We now derive a system equivalent to (1.5). The calculations here are formal and will be justified later. Let y still denote the characteristics. We introduce two other variables, the Lagrangian velocity U and cumulative energy distribution H defined by

$$U(t,\xi) = u(t,y(t,\xi)), \tag{2.9}$$

$$H(t,\xi) = \int_{-\infty}^{y(t,\xi)} \left(u^2 + u_x^2\right)\,dx, \tag{2.10}$$

respectively. From the definition of the characteristics, it follows from (1.5a) that

$$\begin{aligned} U_t(t,\xi) &= u_t(t,y) + y_t(t,\xi)u_x(t,y) \\ &= (u_t + f'(u)u_x)\circ y(t,\xi) \\ &= -P_x\circ y\,(t,\xi). \end{aligned} \tag{2.11}$$

This last term can be expressed uniquely in term of U, y, and H. Namely, we have

$$P_x\circ y\,(t,\xi) = -\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(y(t,\xi)-z)e^{-|y(t,\xi)-z|}(g\circ u + \frac{1}{2}f''\circ u\,u_x^2)(t,z)\,dz$$

and, after the change of variables $z = y(t,\eta)$,

$$\begin{aligned} P_x\circ y(t,\xi) = -\frac{1}{2}\int_{\mathbb{R}} \Big[&\operatorname{sgn}(y(t,\xi)-y(t,\eta))e^{-|y(t,\xi)-y(t,\eta)|} \\ &\times \Big(g\circ u + \frac{1}{2}f''\circ u u_x^2\Big)(t,y(t,\eta))y_\xi(t,\eta)\Big]\,d\eta. \end{aligned}$$

Finally, since $H_\xi = (u^2+u_x^2)\circ y\,y_\xi$,

$$\begin{aligned} P_x\circ y\,(\xi) = -\frac{1}{2}\int_{\mathbb{R}} &\operatorname{sgn}(y(\xi)-y(\eta))e^{-|y(\xi)-y(\eta)|} \\ &\times \Big(\big(g(U) - \frac{1}{2}f''(U)U^2\big)y_\xi + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta \end{aligned} \tag{2.12}$$

where the t variable has been dropped to simplify the notation. Later we will prove that y is an increasing function for any fixed time t. If, for the moment, we take this for granted, then $P_x\circ y$ is equivalent to Q where

$$\begin{aligned} Q(t,\xi) = -\frac{1}{2}\int_{\mathbb{R}} &\operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big) \\ &\times \Big(\big(g(U) - \frac{1}{2}f''(U)U^2\big)y_\xi + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta, \end{aligned} \tag{2.13}$$

and, slightly abusing the notation, we write

$$\begin{aligned} P(t,\xi) = \frac{1}{2}\int_{\mathbb{R}} &\exp\big(-\operatorname{sgn}(\xi-\eta)(y(\xi)-y(\eta))\big) \\ &\times \Big(\big(g(U) - \frac{1}{2}f''(U)U^2\big)y_\xi + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta. \end{aligned} \tag{2.14}$$

Thus $P_x \circ y$ and $P \circ y$ can be replaced by equivalent expressions given by (2.13) and (2.14) which only depend on our new variables U, H, and y. We introduce yet another variable, $\zeta(t,\xi)$, simply defined as

$$\zeta(t,\xi) = y(t,\xi) - \xi.$$

It will turn out that $\zeta \in L^\infty(\mathbb{R})$. We have now derived a new system of equations, formally equivalent (1.5). Equations (2.11), (2.8) and (2.6) give us

$$\begin{cases} \zeta_t = U, \\ U_t = -Q, \\ H_t = G(U) - 2PU. \end{cases} \tag{2.15}$$

Detailed analysis will reveal that the system (2.15) of ordinary differential equations for $(\zeta, U, H)\colon [0,T] \to E$ is well-posed, where E is a Banach space to be defined in the next section. We have

$$Q_\xi = -\frac{1}{2} f''(U) H_\xi + \big(P + \frac{1}{2} f''(U) U^2 - g(U)\big) y_\xi,$$

and $P_\xi = Q y_\xi$. Hence, differentiating (2.15) yields

$$\begin{cases} \zeta_{\xi t} = f''(U) U_\xi \text{ (or } y_{\xi t} = f''(U) U_\xi), \\ U_{\xi t} = \frac{1}{2} f''(U) H_\xi - \big(P + \frac{1}{2} f''(U) U^2 - g(U)\big) y_\xi, \\ H_{\xi t} = -2Q\, U y_\xi + (2g(U) - f''(U) U^2 - 2P) U_\xi. \end{cases} \tag{2.16}$$

The system (2.16) is semilinear with respect to the variables y_ξ, U_ξ and H_ξ.

2.2. **Existence and uniqueness of solutions in Lagrangian variables.** In this section, we focus our attention on the system of equations (2.15) and prove, by a contraction argument, that it admits a unique solution. Let V be the Banach space defined by

$$V = \{ f \in C_b(\mathbb{R}) \mid f_\xi \in L^2(\mathbb{R}) \}$$

where $C_b(\mathbb{R}) = C(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and the norm of V is given by $\|f\|_V = \|f\|_{L^\infty(\mathbb{R})} + \|f_\xi\|_{L^2(\mathbb{R})}$. Of course $H^1(\mathbb{R}) \subset V$ but the converse is not true as V contains functions that do not vanish at infinity. We will employ the Banach space E defined by

$$E = V \times H^1(\mathbb{R}) \times V$$

to carry out the contraction map argument. For any $X = (\zeta, U, H) \in E$, the norm on E is given by

$$\|X\|_E = \|\zeta\|_V + \|U\|_{H^1(\mathbb{R})} + \|H\|_V .$$

The following lemma gives the Lipschitz bounds we need on Q and P.

Lemma 2.1. *For any $X = (\zeta, U, H)$ in E, we define the maps $\mathcal{Q}$ and $\mathcal{P}$ as $\mathcal{Q}(X) = Q$ and $\mathcal{P}(X) = P$ where Q and P are given by* (2.13) *and* (2.14)*,* *respectively. Then, $\mathcal{P}$ and $\mathcal{Q}$ are Lipschitz maps on bounded sets from E to $H^1(\mathbb{R})$. Moreover, we have*

$$Q_\xi = -\frac{1}{2}f''(U)H_\xi + \big(P + \frac{1}{2}f''(U)U^2 - g(U)\big)(1+\zeta_\xi), \tag{2.17}$$

$$P_\xi = Q(1+\zeta_\xi). \tag{2.18}$$

Proof. We rewrite $\mathcal{Q}$ as

$$\begin{aligned}\mathcal{Q}(X)(\xi) = &-\frac{e^{-\zeta(\xi)}}{2}\int_{\mathbb{R}} \chi_{\{\eta<\xi\}}(\eta)e^{-(\xi-\eta)}e^{\zeta(\eta)}\\ &\qquad\times\Big(\big(g(U) - \frac{1}{2}f''(U)U^2\big)(1+\zeta_\xi) + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta\\ &+\frac{e^{\zeta(\xi)}}{2}\int_{\mathbb{R}} \chi_{\{\eta>\xi\}}(\eta)e^{(\xi-\eta)}e^{-\zeta(\eta)}\\ &\qquad\times\Big(\big(g(U) - \frac{1}{2}f''(U)U^2\big)(1+\zeta_\xi) + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta,\end{aligned} \tag{2.19}$$

where χ_B denotes the indicator function of a given set B. We decompose $\mathcal{Q}$ into the sum $\mathcal{Q}_1 + \mathcal{Q}_2$ where $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are the operators corresponding to the two terms on the right-hand side of (2.19). Let $h(\xi) = \chi_{\{\xi>0\}}(\xi)e^{-\xi}$ and A be the map defined by $A\colon v \mapsto h \star v$. Then, Q_1 can be rewritten as

$$\mathcal{Q}_1(X)(\xi) = -\frac{e^{-\zeta(\xi)}}{2}A\circ R(\zeta, U, H)(\xi) \tag{2.20}$$

where R is the operator from E to $L^2(\mathbb{R})$ given by

$$R(\zeta, U, H)(\xi) = e^{\zeta(\xi)}\big((g(U) - \frac{1}{2}f'(U)U^2)(1+\zeta_\xi) + \frac{1}{2}f''(U)H_\xi\big)(\xi). \tag{2.21}$$

We claim that A is continuous from $L^2(\mathbb{R})$ into $H^1(\mathbb{R})$. The Fourier transform of h can easily be computed, and we obtain

$$\hat{h}(\eta) = \int_{\mathbb{R}} h(\xi)e^{-2i\pi\eta\xi}\,d\xi = \frac{1}{1+2i\pi\eta}. \tag{2.22}$$

The $H^1(\mathbb{R})$ norm can be expressed in term of the Fourier transform as follows, see, e.g., [14],

$$\|h\star v\|_{H^1(\mathbb{R})} = \Big\|(1+\eta^2)^{\frac{1}{2}}\widehat{h\star v}\Big\|_{L^2(\mathbb{R})}.$$

Since $\widehat{h\star v} = \hat{h}\hat{v}$, we have

$$\begin{aligned}\|h\star v\|_{H^1(\mathbb{R})} &= \Big\|(1+\eta^2)^{\frac{1}{2}}\hat{h}\hat{v}\Big\|_{L^2(\mathbb{R})}\\ &\le C\,\|\hat{v}\|_{L^2(\mathbb{R})} \quad \text{by (2.22)}\\ &= C\,\|v\|_{L^2(\mathbb{R})} \quad \text{by Plancherel equality}\end{aligned}$$

for some constant C. Hence, $A\colon L^2(\mathbb{R}) \to H^1(\mathbb{R})$ is continuous. We prove that $R(\zeta, U, H)$ belongs to $L^2(\mathbb{R})$ by using the assumption that $g(0) = 0$. Then, $A\circ R(\zeta, U, H)$ belongs to H^1. Let us prove that $R\colon E \to L^2(\mathbb{R})$ is locally Lipschitz. For that purpose we will use the following short lemma.

Lemma 2.2. *Let $B_M = \{X \in E \mid \|X\|_E \leq M\}$ be a bounded set of E.*
(i) If g_1 is Lipschitz from B_M to $L^\infty(\mathbb{R})$ and g_2 Lipschitz from B_M to $L^2(\mathbb{R})$, then the product $g_1 g_2$ is Lipschitz from B_M to $L^2(\mathbb{R})$.
(ii) If g_1, g_2 are two Lipschitz maps from B_M to $L^\infty(\mathbb{R})$, then the product $g_1 g_2$ is Lipschitz from B_M to $L^\infty(\mathbb{R})$.

Proof of Lemma 2.2. Let X and $\bar{X}$ be in B_M, and assume that g_1 and g_2 satisfy the assumptions of (i). We denote by L_1 and L_2, the Lipschitz constants of g_1 and g_2, respectively. We have

$$\begin{aligned}
&\left\|g_1(X)g_2(X) - g_1(\bar{X})g_2(\bar{X})\right\|_{L^2(\mathbb{R})} \\
&\quad \leq \left\|g_1(X) - g_1(\bar{X})\right\|_{L^\infty(\mathbb{R})} \|g_2(X)\|_{L^2(\mathbb{R})} + \left\|g_1(\bar{X})\right\|_{L^\infty(\mathbb{R})} \left\|g_2(X) - g_2(\bar{X})\right\|_{L^2(\mathbb{R})} \\
&\quad \leq \left[2L_1L_2M + L_1 \|g_2(0)\|_{L^2(\mathbb{R})} + L_2 \|g_1(0)\|_{L^\infty(\mathbb{R})}\right] \left\|X - \bar{X}\right\|_E
\end{aligned}$$

and (i) is proved. One proves (ii) the same way. □

Let us consider a bounded set $B_M = \{X \in E \mid \|X\|_E \leq M\}$ of E. For $X = (\zeta, U, H)$ and $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in B_M, we have $\|U\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}} \|U\|_{H^1(\mathbb{R})} \leq \frac{1}{\sqrt{2}} M$, because $\frac{1}{\sqrt{2}}$ is the constant of the Sobolev embedding from $H^1(\mathbb{R})$ into $L^\infty(\mathbb{R})$, and, similarly, $\left\|\bar{U}\right\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}} M$. Let $I_M = [-\frac{1}{\sqrt{2}}M, \frac{1}{\sqrt{2}}M]$ and

$$L_M = \|f\|_{W^{3,\infty}(I_M)} + \|g\|_{W^{1,\infty}(I_M)} < \infty. \tag{2.23}$$

Then

$$\left\|g(\bar{U}) - g(U)\right\|_{L^2(\mathbb{R})} \leq \|g\|_{W^{1,\infty}(I_M)} \left\|\bar{U} - U\right\|_{L^2(\mathbb{R})} \leq L_M \left\|\bar{U} - U\right\|_{L^2(\mathbb{R})}.$$

Hence, $g_1 \colon X \to g(U)$ is Lipschitz from B_M to $L^2(\mathbb{R})$. For $X, \bar{X}$ in B_M, we have $\|\zeta\|_{L^\infty(\mathbb{R})} \leq M$ and $\left\|\bar{\zeta}\right\|_{L^\infty(\mathbb{R})} \leq M$. The function $x \mapsto e^x$ is Lipschitz on $\{x \in \mathbb{R} \mid |x| \leq M\}$. Hence, $g_2 \colon X \mapsto e^\zeta$ is Lipschitz from B_M to $L^\infty(\mathbb{R})$. Thus, the first term in (2.21), $g_1(X)g_2(X) = e^\zeta g(U)$, is, by Lemma 2.2, Lipschitz from B_M to $L^2(\mathbb{R})$. We look at the second term, that is, $e^\zeta g(U)\zeta_\xi$. For $X, \bar{X}$ in B_M, we have

$$\left\|g(U) - g(\bar{U})\right\|_{L^\infty(\mathbb{R})} \leq \|g\|_{W^{1,\infty}(I_M)} \left\|U - \bar{U}\right\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}} L_M \left\|U - \bar{U}\right\|_{H^1(\mathbb{R})}.$$

Hence, $X \mapsto g(U)$ is Lipschitz from B_M to $L^\infty(\mathbb{R})$ and, by Lemma 2.2, as $X \mapsto e^\zeta$ is also Lipschitz from B_M to $L^\infty(\mathbb{R})$, we have that the product $X \mapsto e^\zeta g(U)$ is Lipschitz from B_M to $L^\infty(\mathbb{R})$. After using again Lemma 2.2, since $X \mapsto \zeta_\xi$, being linear, is obviously Lipschitz from B_M to $L^2(\mathbb{R})$, we obtain, as claimed, that the second term in (2.21), $e^\zeta g(U)\zeta_\xi$, is Lipschitz from B_M to $L^2(\mathbb{R})$. We can handle the other terms in (2.21) similarly and prove that R is Lipschitz from B_M to $L^2(\mathbb{R})$. Since $A \colon L^2(\mathbb{R}) \to H^1(\mathbb{R})$ is linear and continuous, $A \circ R$ is Lipschitz from B_M to $H^1(\mathbb{R})$. Then, we use the following lemma whose proof is basically the same as the proof of Lemma 2.2.

Lemma 2.3. *Let $\mathcal{R}_1\colon B_M \to V$, $\mathcal{R}_2\colon B_M \to H^1(\mathbb{R})$, and $\mathcal{R}_3\colon B_M \to V$ be three Lipschitz maps. Then, the products $X \mapsto R_1(X)R_2(X)$ and $X \mapsto R_1(X)R_3(X)$ are also Lipschitz maps from B_M to $H^1(\mathbb{R})$ and B_M to V, respectively.*

Since the map $X \mapsto e^{-\zeta}$ is Lipschitz from B_M to V, Q_1 is the product of two Lipschitz maps, one from B_M to $H^1(\mathbb{R})$ and the other from B_M to V, and therefore it is Lipschitz map from B_M to $H^1(\mathbb{R})$. Similarly, one proves that $\mathcal{Q}_2$ and therefore $\mathcal{Q}$ are Lipschitz on B_M. Furthermore, $\mathcal{P}$ is Lipschitz on B_M. The formulas (2.17) and (2.18) are obtained by direct computation using the product rule, see [13, p. 129]. □

In the next theorem, by using a contraction argument, we prove the short-time existence of solutions to (2.15).

Theorem 2.4. *Given $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in E, there exists a time T depending only on $\|\bar{X}\|_E$ such that the system* (2.15) *admits a unique solution in $C^1([0,T], E)$ with initial data $\bar{X}$.*

Proof. Solutions of (2.15) can be rewritten as

$$X(t) = \bar{X} + \int_0^t F(X(\tau))\,d\tau \tag{2.24}$$

where $F\colon E \to E$ is given by $F(X) = (f'(U), -\mathcal{Q}(X), G(U) - 2\mathcal{P}(X)U)$ where $X = (\zeta, U, H)$. The integrals are defined as Riemann integrals of continuous functions on the Banach space E. Let B_M and L_M be defined as in the proof of Lemma 2.1, see (2.23). We claim that $X = (\zeta, U, H) \mapsto f'(U)$ and $X = (\zeta, U, H) \mapsto G(U)$ are Lipschitz from B_M to V. Then, using Lemma 2.1, we can check that each component of $F(X)$ is a product of functions that satisfy one of the assumptions of Lemma 2.3 and using this same lemma, we obtain that $F(X)$ is Lipchitz on B_M. Thus, F is Lipschitz on any bounded set of E. Since E is a Banach space, we use the standard contraction argument to show the existence of short-time solutions and the theorem is proved. For any $X = (\zeta, U, H)$ and $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in B_M, we have $\|U\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}}M$ and $\|\bar{U}\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}}M$. Then,

$$\|f'(U) - f'(\bar{U})\|_{L^\infty(\mathbb{R})} \leq \|f'\|_{W^{1,\infty}(I_M)} \|U - \bar{U}\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2}} L_M \|X - \bar{X}\|_E$$

and $X \mapsto f'(U)$ is Lipschitz from B_M into $L^\infty(\mathbb{R})$. Since f' is C^1 and $U \in H^1(\mathbb{R})$, using [19, Appendix A.1], we obtain that $f'(U)_\xi \in L^2(\mathbb{R})$ and

$$f'(U)_\xi = f''(U)U_\xi.$$

As before, it is not hard to prove that $X \mapsto f''(U)$ is Lipschitz from B_M into $L^\infty(\mathbb{R})$. It is clear that $X \mapsto U_\xi$ is Lipschitz from B_M into $L^2(\mathbb{R})$. Hence, it follows from Lemma 2.2 that $X \mapsto f''(U)U_\xi$ is Lipschitz from B_M into $L^2(\mathbb{R})$. Therefore, $X \mapsto f'(U)$ is Lipschitz from B_M into V. Similarly, one proves that $X \mapsto G(U)$ is Lipschitz from B_M into V and our previous claim is proved. □

We now turn to the proof of existence of global solutions of (2.15). We are interested in a particular class of initial data that we are going to make precise later, see Definition 2.5. In particular, we will only consider initial data that belong to $E \cap \left[W^{1,\infty}(\mathbb{R})\right]^3$ where

$$W^{1,\infty}(\mathbb{R}) = \{f \in C_b(\mathbb{R}) \mid f_\xi \in L^\infty(\mathbb{R})\}.$$

Given $(\bar\zeta, \bar U, \bar H) \in E \cap [W^{1,\infty}(\mathbb{R})]^3$, we consider the short-time solution $(\zeta, U, H) \in C([0,T], E)$ of (2.15) given by Theorem 2.4. Using the fact that $\mathcal{Q}$ and $\mathcal{P}$ are Lipschitz on bounded sets (Lemma 2.1) and, since $X \in C([0,T], E)$, we can prove that P and Q belongs to $C([0,T], H^1(\mathbb{R}))$. We now consider U, P and Q as given function in $C([0,T], H^1(\mathbb{R}))$. Then, for any *fixed* $\xi \in \mathbb{R}$, we can solve the system of ordinary differential equations

$$\begin{cases} \dfrac{d}{dt}\alpha(t,\xi) = f''(U)\beta(t,\xi), \\ \dfrac{d}{dt}\beta(t,\xi) = \dfrac{1}{2}f''(U)\gamma(t,\xi) + \big(-\dfrac{1}{2}f''(U)U^2 + g(U) - P\big)(1+\alpha)(t,\xi), \\ \dfrac{d}{dt}\gamma(t,\xi) = -2(Q\,U)(1+\alpha)(t,\xi) + (2g(U) + f''(U)U^2 - 2P)\beta(t,\xi), \end{cases} \tag{2.25}$$

which is obtained by substituting ζ_ξ, U_ξ and H_ξ in (2.16) by the unknowns α, β, and γ, respectively. Concerning the initial data, we set $(\alpha(0,\xi), \beta(0,\xi), \gamma(0,\xi)) = (\bar\zeta_\xi, \bar U_\xi, \bar H_\xi)$ if $\left|\bar\zeta_\xi(\xi)\right| + \left|\bar U_\xi(\xi)\right| + \left|\bar H_\xi(\xi)\right| < \infty$ and $(\alpha(0,\xi), \beta(0,\xi), \gamma(0,\xi)) = (0,0,0)$ otherwise. In the same way as in [16, Lemma 2.4], see also Lemma 2.6 below, we can prove that solutions of (2.25) exist in $[0,T]$ and that, for all time $t \in [0,T]$,

$$(\alpha(t,\xi), \beta(t,\xi), \gamma(t,\xi)) = (\zeta_\xi(t,\xi), U_\xi(t,\xi), H_\xi(t,\xi))$$

for almost every $\xi \in \mathbb{R}$. Thus, we can select a special representative for $(\zeta_\xi, U_\xi, H_\xi)$ given by (α, β, γ), which is defined for all $\xi \in \mathbb{R}$ and which, for any given ξ, satisfies the ordinary differential equation (2.25) in $\mathbb{R}^3$. From now on we will of course identify the two and set $(\zeta_\xi, U_\xi, H_\xi)$ equal to (α, β, γ).

Our goal is to find solutions of (1.5) with initial data $\bar u$ in H^1 because H^1 is the natural space for the equation. However, Theorem 2.4 gives us the existence of solutions to (2.15) for initial data in E. Therefore we have to find initial conditions that match the initial data $\bar u$ and belong to E. A natural choice would be to use $\bar y(\xi) = y(0,\xi) = \xi$ and $\bar U(\xi) = u(\xi)$. Then $y(t,\xi)$ gives the position of the particle which is at ξ at time $t = 0$. But, if we make this choice, then $\bar H_\xi = \bar u^2 + \bar u_x^2$ and $\bar H_\xi$ does not belong to $L^2(\mathbb{R})$ in general. We consider instead $\bar y$ implicitly given by

$$\xi = \int_{-\infty}^{\bar y(\xi)} (\bar u^2 + \bar u_x^2)\,dx + \bar y(\xi) \tag{2.26a}$$

and

$$\bar{U} = \bar{u}\circ\bar{y}\,, \tag{2.26b}$$

$$\bar{H} = \int_{-\infty}^{\bar{y}} \left(\bar{u}^2 + \bar{u}_x^2\right)\,dx. \tag{2.26c}$$

In the next lemma we prove that $(\bar{y}, \bar{U}, \bar{H})$ belongs to the set $\mathcal{G}$ where $\mathcal{G}$ is defined as follows.

Definition 2.5. The set $\mathcal{G}$ is consists of all $(\zeta, U, H) \in E$ such that

$$(\zeta, U, H) \in \left[W^{1,\infty}(\mathbb{R})\right]^3, \tag{2.27a}$$

$$y_\xi \geq 0, H_\xi \geq 0, y_\xi + H_\xi > 0 \text{ almost everywhere, and } \lim_{\xi\to-\infty} H(\xi) = 0, \tag{2.27b}$$

$$y_\xi H_\xi = y_\xi^2 U^2 + U_\xi^2 \text{ almost everywhere,} \tag{2.27c}$$

where we denote $y(\xi) = \zeta(\xi) + \xi$.

Lemma 2.6. *Given* $\bar{u} \in H^1(\mathbb{R})$*, then* $(\bar{y}, \bar{U}, \bar{H})$ *as defined in* (2.26) *belongs to* $\mathcal{G}$*.*

Proof. The function $k\colon z \mapsto \int_0^z (\bar{u}^2 + \bar{u}_x^2)(x)\,dx + z$ is a strictly increasing continuous function with $\lim_{z\to\pm\infty} k(z) = \pm\infty$. Hence, k is invertible and $\bar{y}(\xi) = k^{-1}(\xi)$ is well-defined. We have to check that $(\bar{\zeta}, \bar{U}, \bar{H})$ belongs to E. It follows directly from the definition that $\bar{y}$ is a strictly increasing function. We have

$$\bar{\zeta}(\xi) = -\int_{-\infty}^{\bar{y}(\xi)} \left(\bar{u}^2 + \bar{u}_x^2\right)\,dx, \tag{2.28}$$

and therefore, since $\bar{u} \in H^1$, $\bar{\zeta}$ is bounded. For any $(\xi, \xi') \in \mathbb{R}^2$, we have

$$\begin{aligned} |\xi - \xi'| &= \left| \int_{\bar{y}(\xi')}^{\bar{y}(\xi)} (\bar{u}^2 + \bar{u}_x^2)\,dx + \bar{y}(\xi) - \bar{y}(\xi') \right| \\ &= \left| \int_{\bar{y}(\xi')}^{\bar{y}(\xi)} (\bar{u}^2 + \bar{u}_x^2)\,dx \right| + |\bar{y}(\xi) - \bar{y}(\xi')| \end{aligned} \tag{2.29}$$

because the two quantities inside the absolute values have the same sign. It follows from (2.29) that $\bar{y}$ is Lipschitz (with Lipschitz constant at most 1) and therefore almost everywhere differentiable. From (2.28), we get that, for almost every $\xi \in \mathbb{R}$,

$$\bar{\zeta}_\xi = -(\bar{u}^2 + \bar{u}_x^2)\circ\bar{y}\,\bar{y}_\xi$$

Since $\bar{y}_\xi = 1 + \bar{\zeta}_\xi$, it implies

$$\bar{\zeta}_\xi(\xi) = -\frac{\bar{u}^2 + \bar{u}_x^2}{1 + \bar{u}^2 + \bar{u}_x^2}\circ\bar{y}(\xi)\,. \tag{2.30}$$

Therefore $\bar{\zeta}_\xi$ is bounded almost everywhere and $\bar{\zeta}$ belongs to $W^{1,\infty}(\mathbb{R})$. We also have

$$\bar{y}_\xi = \frac{1}{1 + \bar{u}^2 + \bar{u}_x^2}\circ\bar{y} \tag{2.31}$$

which implies that $\bar{y}_\xi > 0$ almost everywhere. From (2.28), we see that $\bar{H} = -\bar{\zeta}$ and therefore $\bar{H}$ belongs to $W^{1,\infty}(\mathbb{R})$. Since $H^1(\mathbb{R}) \subset L^\infty(\mathbb{R})$, $\bar{U} = \bar{u} \circ \bar{y}$ is bounded. We have, for almost every $\xi \in \mathbb{R}$,

$$\bar{H}_\xi = (\bar{u}^2 + \bar{u}_x^2) \circ \bar{y}\, \bar{y}_\xi \tag{2.32}$$

which, since $\bar{U}_\xi = \bar{u}_x \circ \bar{y}\, y_\xi$ almost everywhere, gives us

$$\bar{y}_\xi \bar{H}_\xi = \bar{y}_\xi^2 \bar{U}^2 + \bar{U}_\xi^2. \tag{2.33}$$

Hence, $\bar{U}_\xi^2 \leq \bar{y}_\xi \bar{H}_\xi$ and $\bar{U}_\xi$ is bounded and $\bar{U}$ belongs to $W^{1,\infty}(\mathbb{R})$. We have $(\bar{\zeta}, \bar{U}, \bar{H}) \in [W^{1,\infty}(\mathbb{R})]^3$. It remains to prove that $\bar{U}$, $\bar{\zeta}_\xi$, $\bar{U}_\xi$ and $\bar{H}_\xi$ belong to $L^2(\mathbb{R})$. By making the change of variable $x = \bar{y}(\xi)$ and using (2.31), we obtain

$$\begin{aligned}\left\|\bar{U}\right\|_{L^2(\mathbb{R})}^2 &= \int_{\mathbb{R}} \bar{u}^2(x)(1 + \bar{u}^2 + \bar{u}_x^2)(x)\, dx \\ &\leq \left\|\bar{u}\right\|_{L^2(\mathbb{R})}^2 + \left\|\bar{u}\right\|_{L^\infty(\mathbb{R})}^2 \left\|\bar{u}\right\|_{H^1(\mathbb{R})}^2 .\end{aligned}$$

Hence, $\bar{U} \in L^2(\mathbb{R})$. Since $0 \leq \bar{H}_\xi \leq 1$, $\bar{H}$ is monotone and

$$\left\|\bar{H}_\xi\right\|_{L^2(\mathbb{R})}^2 \leq \left\|\bar{H}_\xi\right\|_{L^\infty(\mathbb{R})} \left\|\bar{H}_\xi\right\|_{L^1(\mathbb{R})} \leq \lim_{\xi \to \infty} \bar{H}(\xi) = \left\|\bar{u}\right\|_{H^1(\mathbb{R})}^2 .$$

Hence, $\bar{H}_\xi$, and therefore $\bar{\zeta}_\xi$, belong to $L^2(\mathbb{R})$. From (2.33) we get

$$\left\|\bar{U}_\xi\right\|_{L^2(\mathbb{R})}^2 \leq \left\|\bar{y}_\xi \bar{H}_\xi\right\|_{L^1(\mathbb{R})} \leq (1 + \left\|\bar{\zeta}_\xi\right\|_{L^\infty(\mathbb{R})}) \left\|\bar{H}\right\|_{L^\infty(\mathbb{R})}$$

and $\bar{U}_\xi \in L^2(\mathbb{R})$. □

For initial data in $\mathcal{G}$, the solution of (2.15) exists globally. To prove that we will use the following lemma.

Lemma 2.7. *Given initial data $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$ in $\mathcal{G}$, let $X(t) = (\zeta(t), U(t), H(t))$ be the short-time solution of* (2.15) *in $C([0,T], E)$ for some $T > 0$ with initial data $\bar{X} = (\bar{\zeta}, \bar{U}, \bar{H})$. Then,*

(i) $X(t)$ belongs to $\mathcal{G}$ for all $t \in [0,T]$, that is, $\mathcal{G}$ is preserved by the flow.
(ii) for almost every $t \in [0,T]$, $y_\xi(t,\xi) > 0$ for almost every $\xi \in \mathbb{R}$,
(iii) For all $t \in [0,T]$, $\lim_{\xi \to \pm\infty} H(t,\xi)$ exists and is independent of time.

We denote by $\mathcal{A}$ the set of all $\xi \in \mathbb{R}$ for which $\left|\bar{\zeta}_\xi(\xi)\right| + \left|\bar{U}_\xi(\xi)\right| + \left|\bar{H}_\xi(\xi)\right| < \infty$ and the relations in (2.27b) and (2.27c) are fulfilled for $\bar{y}_\xi$, $\bar{U}_\xi$ and $\bar{H}_\xi$. Since by assumption $\bar{X} \in \mathcal{G}$, we have $\operatorname{meas}(\mathcal{A}^c) = 0$, and we set $(\bar{U}_\xi, \bar{H}_\xi, \bar{\zeta}_\xi)$ equal to zero on $\mathcal{A}^c$. Then, as we explained earlier, we choose a special representative for $(\zeta(t,\xi), U(t,\xi), H(t,\xi))$ which satisfies (2.16) as an ordinary differential equation in $\mathbb{R}^3$ for every $\xi \in \mathbb{R}$.

Proof. (*i*) The fact that $W^{1,\infty}(\mathbb{R})$ is preserved by the equation can be proved in the same way as in [16, Lemma 2.4] and we now give only a sketch of this proof. We look at (2.15) as a system of ordinay differential equations in $E \cap W^{1,\infty}(\mathbb{R})$. We have already established the short-time existence of solutions in E, and, since (2.16) is semilinear with respect to y_ξ, U_ξ and H_ξ (and affine with respect to ζ_ξ, U_ξ and H_ξ), it is not hard to establish, by a contraction argument, the short-time existence of solutions in $E\cap W^{1,\infty}(\mathbb{R})$. Let $C_1=\sup_{t\in[0,T]}\big(\|U(t,\,\cdot)\|_{L^\infty(\mathbb{R})}+\|P(t,\,\cdot)\|_{L^\infty(\mathbb{R})}+\|Q(t,\,\cdot)\|_{L^\infty(\mathbb{R})}\big)$ and $Z(t)=\|\zeta_\xi(t,\,\cdot)\|_{L^\infty(\mathbb{R})}+\|U_\xi(t,\,\cdot)\|_{L^\infty(\mathbb{R})}+\|H_\xi(t,\,\cdot)\|_{L^\infty(\mathbb{R})}$. We have $C_1\leq\sup_{t\in[0,T]}\|X(t,\,\cdot)\|_E<\infty$. Using again the semi-linearity of (2.16), we get that

$$Z(t)\leq Z(0)+CT+C\int_0^t Z(\tau)\,d\tau$$

for a constant C that only depends on C_1. Hence, it follows from Gronwall's lemma that $\sup_{t\in[0,T)}Z(t)<\infty$, which proves that the space $W^{1,\infty}(\mathbb{R})$ is preserved by the flow of (2.15). Let us prove that (2.27c) and the inequalities in (2.27b) hold for any $\xi\in\mathcal{A}$ and therefore almost everywhere. We consider a fixed ξ in $\mathcal{A}$ and drop it in the notations when there is no ambiguity. From (2.16), we have, on the one hand,

$$\begin{aligned}(y_\xi H_\xi)_t&=y_{\xi t}H_\xi+H_{\xi t}y_\xi\\&=f''(U)U_\xi H_\xi+\big(G'(U)U_\xi-2QUy_\xi-2PU_\xi\big)y_\xi\\&=f''(U)U_\xi H_\xi+\big(2g(U)+f''(U)U^2\big)U_\xi y_\xi-2Qy_\xi^2U-2PU_\xi y_\xi\end{aligned}$$

and, on the other hand,

$$\begin{aligned}(y_\xi^2U^2+U_\xi^2)_t&=2y_{\xi t}y_\xi U^2+2y_\xi^2U_tU+2U_{\xi t}U_\xi\\&=2f''(U)U_\xi y_\xi U^2-2y_\xi^2QU\\&\quad+2U_\xi\big(\frac12 f''(U)H_\xi-\frac12 f''(U)U^2y_\xi+g(U)y_\xi-Py_\xi\big).\end{aligned}$$

Thus, $(y_\xi H_\xi-y_\xi^2U^2-U_\xi^2)_t=0$, and since $y_\xi H_\xi(0)=(y_\xi^2U^2+U_\xi^2)(0)$, we have $y_\xi H_\xi(t)=(y_\xi^2U^2+U_\xi^2)(t)$ for all $t\in[0,T]$. We have proved (2.27c). Let us introduce t^* given by

$$t^*=\sup\{t\in[0,T]\mid y_\xi(t')\geq0\text{ for all }t'\in[0,t]\}.$$

Recall that we consider a fixed $\xi\in\mathcal{A}$, and drop it in the notation. Assume that $t^*<T$. Since $y_\xi(t)$ is continuous with respect to time, we have

$$y_\xi(t^*)=0. \tag{2.34}$$

Hence, from (2.27c) that we just proved, $U_\xi(t^*)=0$ and, by (2.16),

$$y_{\xi t}(t^*)=f''(U)U_\xi(t^*)=0. \tag{2.35}$$

From (2.16), since $y_\xi(t^*)=U_\xi(t^*)=0$, we get

$$y_{\xi tt}(t^*)=f''(U)U_{\xi t}(t^*)=\frac12 f''(U)^2H_\xi(t^*). \tag{2.36}$$

If $H_\xi(t^*)=0$, then $(y_\xi,U_\xi,H_\xi)(t^*)=(0,0,0)$ and, by the uniqueness of the solution of (2.16), seen as a system of ordinary differential equations, we must have $(y_\xi,U_\xi,H_\xi)(t)=0$ for all $t\in[0,T]$. This contradicts the fact that $y_\xi(0)$ and $H_\xi(0)$ cannot vanish at the same time ($\bar y_\xi+\bar H_\xi>0$ for all $\xi\in\mathcal{A}$). If $H_\xi(t^*)<0$, then $y_{\xi tt}(t^*)<0$ because f does not vanish and, because of (2.34) and (2.35), there exists a neighborhood $\mathcal{U}$ of t^* such that $y(t) < 0$ for all $t \in \mathcal{U} \setminus \{t^*\}$. This contradicts the definition of t^*. Hence,

$$H_\xi(t^*) > 0, \tag{2.37}$$

and, since we now have $y_\xi(t^*) = y_{\xi t}(t^*) = 0$ and $y_{\xi tt}(t^*) > 0$, there exists a neighborhood of t^* that we again denote by $\mathcal{U}$ such that $y_\xi(t) > 0$ for all $t \in \mathcal{U} \setminus \{t^*\}$. This contradicts the fact that $t^* < T$, and we have proved the first inequality in (2.27b), namely that $y_\xi(t) \geq 0$ for all $t \in [0, T]$. Let us prove that $H_\xi(t) \geq 0$ for all $t \in [0, T]$. This follows from (2.27c) when $y_\xi(t) > 0$. Now, if $y_\xi(t) = 0$, then $U_\xi(t) = 0$ from (2.27c), and we have seen that $H_\xi(t) < 0$ would imply that $y_\xi(t') < 0$ for some t' in a punctured neighborhood of t, which is impossible. Hence, $H_\xi(t) \geq 0$, and we have proved the second inequality in (2.27b). Assume that the third inequality in (2.27c) does not hold. Then, by continuity, there exists a time $t \in [0, T]$ such that $(y_\xi + H_\xi)(t) = 0$. Since y_ξ and H_ξ are positive, we must have $y_\xi(t) = H_\xi(t) = 0$ and, by (2.27c), $U_\xi(t) = 0$. Since zero is a solution of (2.16), this implies that $y_\xi(0) = U_\xi(0) = H_\xi(0)$, which contradicts $(y_\xi + H_\xi)(0) > 0$. The fact that $\lim_{\xi \to -\infty} H(t, \xi) = 0$ will be proved below in (*iii*).

(*ii*) We define the set

$$\mathcal{N} = \{(t, \xi) \in [0, T] \times \mathbb{R} \mid y_\xi(t, \xi) = 0\}.$$

Fubini's theorem gives us

$$\operatorname{meas}(\mathcal{N}) = \int_{\mathbb{R}} \operatorname{meas}(\mathcal{N}_\xi)\, d\xi = \int_{[0,T]} \operatorname{meas}(\mathcal{N}_t)\, dt \tag{2.38}$$

where $\mathcal{N}_\xi$ and $\mathcal{N}_t$ are the ξ-section and t-section of $\mathcal{N}$, respectively, that is,

$$\mathcal{N}_\xi = \{t \in [0, T] \mid y_\xi(t, \xi) = 0\}$$

and

$$\mathcal{N}_t = \{\xi \in \mathbb{R} \mid y_\xi(t, \xi) = 0\}.$$

Let us prove that, for all $\xi \in \mathcal{A}$, $\operatorname{meas}(\mathcal{N}_\xi) = 0$. If we consider the sets $\mathcal{N}_\xi^n$ defined as

$$\mathcal{N}_\xi^n = \{t \in [0, T] \mid y_\xi(t, \xi) = 0 \text{ and } y_\xi(t', \xi) > 0 \text{ for all } t' \in [t - 1/n, t + 1/n] \setminus \{t\}\},$$

then

$$\mathcal{N}_\xi = \bigcup_{n \in \mathbb{N}} \mathcal{N}_\xi^n. \tag{2.39}$$

Indeed, for all $t \in \mathcal{N}_\xi$, we have $y_\xi(t, \xi) = 0$, $y_{\xi t}(t, \xi) = 0$ from (2.27c) and (2.16) and $y_{\xi tt}(t, \xi) = \frac{1}{2} f''(U)^2 H_\xi(t, \xi) > 0$ from (2.16) and (2.27b) (y_ξ and H_ξ cannot vanish at the same time for $\xi \in \mathcal{A}$). Since f'' does not vanish, this implies that, on a small punctured neighborhood of t, y_ξ is strictly positive. Hence, t belongs to some $\mathcal{N}_\xi^n$ for n large enough. This proves (2.39). The set $\mathcal{N}_\xi^n$ consists of isolated points that are countable since, by definition, they are separated by a distance larger than $1/n$ from one another. This means that $\operatorname{meas}(\mathcal{N}_\xi^n) = 0$ and, by the subadditivity of the measure, $\operatorname{meas}(\mathcal{N}_\xi) = 0$. It follows from (2.38) and since $\operatorname{meas}(\mathcal{A}^c) = 0$ that

$$\operatorname{meas}(\mathcal{N}_t) = 0 \text{ for almost every } t \in [0,T]. \tag{2.40}$$

We denote by $\mathcal{K}$ the set of times such that $\operatorname{meas}(\mathcal{N}_t) > 0$, i.e.,

$$\mathcal{K} = \{t \in \mathbb{R}_+ \mid \operatorname{meas}(\mathcal{N}_t) > 0\}. \tag{2.41}$$

By (2.40), $\operatorname{meas}(\mathcal{K}) = 0$. For all $t \in \mathcal{K}^c$, $y_\xi > 0$ almost everywhere and, therefore, $y(t,\xi)$ is strictly increasing and invertible (with respect to ξ).

(*iii*) For any given $t \in [0,T]$, since $H_\xi(t,\xi) \geq 0$, $H(t,\xi)$ is an increasing function with respect to ξ and therefore, as $H(t,\,\cdot\,) \in L^\infty(\mathbb{R})$, $H(t,\xi)$ has a limit when $\xi \to \pm\infty$. We denote those limits $H(t,\pm\infty)$. Since $U(t,\,\cdot\,) \in H^1(\mathbb{R})$, we have $\lim_{\xi\to\pm\infty} U(t,\xi) = 0$ for all $t \in [0,T]$. We have

$$H(t,\xi) = H(0,\xi) + \int_0^t \left[G(U) - 2PU\right](\tau,\xi)\,d\tau \tag{2.42}$$

and $\lim_{\xi\to\pm\infty} G(U(t,\xi)) = 0$ because $\lim_{\xi\to\pm\infty} U(t,\xi) = 0$, $G(0) = 0$ and G is continuous. As U, $G(U)$ and P are bounded in $L^\infty([0,T]\times\mathbb{R})$, we can let ξ tend to $\pm\infty$ and apply the Lebesgue dominated convergence theorem. We get $H(t,\pm\infty) = H(0,\pm\infty)$ for all $t \in [0,T]$. Since $\bar{X} \in \mathcal{G}$, $H(0,-\infty) = 0$ and therefore $H(t,-\infty) = 0$ for all $t \in [0,T]$. □

In the next theorem, we prove global existence of solutions to (2.15). We also state that the solutions are continuous with respect to the functions $(f,g) \in \mathcal{E}$ (cf. (1.6)) that appear in (1.5). Therefore we need to specify the topology we use on $\mathcal{E}$. The space $L^\infty_{\mathrm{loc}}(\mathbb{R})$ is a locally convex linear topological space. Let K_j be a given increasing sequence of compact sets such that $\mathbb{R} = \cup_{j\in\mathbb{N}} K_j$, then the topology of $L^\infty_{\mathrm{loc}}(\mathbb{R})$ is defined by the sequence of semi-norms $h \mapsto \|h\|_{L^\infty(K_n)}$. The space $L^\infty_{\mathrm{loc}}(\mathbb{R})$ is metrizable, see [14, Proposition 5.16]. A subset B of $L^\infty_{\mathrm{loc}}(\mathbb{R})$ is bounded if, for all $n \geq 1$, there exists $C_n > 0$ such that $\|f\|_{L^\infty(K_n)} \leq C_n$ for all $f \in B$, see [21, I.7] for the general definition of bounded sets in a linear topological space. The topologies of $W^{k,\infty}_{\mathrm{loc}}(\mathbb{R})$ follows naturally from the topology of $L^\infty_{\mathrm{loc}}(\mathbb{R})$ applied to the k first derivatives. We equip $\mathcal{E}$ with the topology induced $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times L^\infty_{\mathrm{loc}}(\mathbb{R})$. We will also consider bounded subsets of $\mathcal{E}$ in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$. A subset $\mathcal{E}'$ of $\mathcal{E}$ is bounded in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$ if for all $n \geq 1$, there exists C_n such that $\|f\|_{W^{2,\infty}(K_n)} + \|g\|_{W^{1,\infty}(K_n)} \leq C_n$ for all $(f,g) \in \mathcal{E}'$. In the remaining, by bounded sets of $\mathcal{E}$ we will always implicitly mean bounded sets of $\mathcal{E}$ in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$.

Theorem 2.8. *Assume* (1.6). *For any $\bar{X} = (\bar{y},\bar{U},\bar{H}) \in \mathcal{G}$, the system* (2.15) *admits a unique global solution $X(t) = (y(t),U(t),H(t))$ in $C^1(\mathbb{R}_+,E)$ with initial data $\bar{X} = (\bar{y},\bar{U},\bar{H})$. We have $X(t) \in \mathcal{G}$ for all times. If we equip $\mathcal{G}$ with the topology induced by the E-norm, then the map $S\colon \mathcal{G}\times\mathcal{E}\times\mathbb{R}_+ \to \mathcal{G}$ defined as*

$$S_t(\bar{X},f,g) = X(t)$$

is a semigroup which is continuous with respect to all variables, on any bounded set of $\mathcal{E}$.

Proof. The solution has a finite time of existence T only if $\|X(t,\cdot)\|_E$ blows up when t tends to T because, otherwise, by Theorem 2.4, the solution can be extended by a small time interval beyond T. Thus, We want to prove that

$$\sup_{t\in[0,T)} \|X(t,\cdot)\|_E < \infty.$$

Since $X(t)\in\mathcal{G}$, $H_\xi \geq 0$, from (2.27b), and $H(t,\xi)$ is an increasing function in ξ for all t and, from Lemma 2.7, we have $\lim_{\xi\to\infty} H(t,\xi) = \lim_{\xi\to\infty} H(0,\xi)$. Hence, $\sup_{t\in[0,T)} \|H(t,\cdot)\|_{L^\infty(\mathbb{R})} = \|\bar{H}\|_{L^\infty(\mathbb{R})}$ and therefore $\sup_{t\in[0,T)} \|H(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is finite. To simplify the notation we suppress the dependence in t for the moment and denote $h = \|\bar{H}\|_{L^\infty(\mathbb{R})}$. We have

$$U^2(\xi) = 2\int_{-\infty}^{\xi} U(\eta)U_\xi(\eta)\,d\eta = 2\int_{\{\eta\leq\xi \mid y_\xi(\eta)>0\}} U(\eta)U_\xi(\eta)\,d\eta \tag{2.43}$$

since, from (2.27c), $U_\xi(\eta) = 0$ when $y_\xi(\eta) = 0$. For almost every ξ such that $y_\xi(\xi) > 0$, we have

$$|U(\xi)U_\xi(\xi)| = \left|\sqrt{y_\xi}U(\xi)\frac{U_\xi(\xi)}{\sqrt{y_\xi(\xi)}}\right| \leq \frac{1}{2}\left(U(\xi)^2 y_\xi(\xi) + \frac{U_\xi^2(\xi)}{y_\xi(\xi)}\right) = \frac{1}{2}H_\xi(\xi),$$

from (2.27c). Inserting this inequality in (2.43), we obtain $U^2(\xi) \leq H(\xi)$ and we have

$$U(t,\xi) \in I := [-\sqrt{h}, \sqrt{h}] \tag{2.44}$$

for all $t\in[0,T)$ and $\xi\in\mathbb{R}$. Hence, $\sup_{t\in[0,T)} \|U(t,\cdot)\|_{L^\infty(\mathbb{R})} < \infty$. The property (2.44) is important as it says that the $L^\infty(\mathbb{R})$-norm of U is bounded by a constant which does not depend on time. We set

$$\kappa = \|f\|_{W^{2,\infty}(I)} + \|g\|_{W^{1,\infty}(I)}.$$

By using (2.44), we obtain

$$\|f'(U)\|_{L^\infty(\mathbb{R})} \leq \|f'\|_{L^\infty(I)} \leq \kappa. \tag{2.45}$$

Hence, from the governing equation (2.15), it follows that

$$|\zeta(t,\xi)| \leq |\zeta(0,\xi)| + \kappa T,$$

and $\sup_{t\in[0,T)} \|\zeta(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is finite. Next we prove that $\sup_{t\in[0,T)} \|Q(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is finite. We decompose Q into a sum of two integrals that we denote Q_a and Q_b, respectively,

$$\begin{aligned} Q(t,\xi) = &-\frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta)e^{-|y(\xi)-y(\eta)|}y_\xi(\eta)\big(g(U) - \frac{1}{2}f''(U)U^2\big)\,d\eta \\ &-\frac{1}{4}\int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta)e^{-|y(\xi)-y(\eta)|}f''(U)H_\xi\,d\eta \\ :=\ & Q_a + Q_b. \end{aligned}$$

Since $y_\xi \geq 0$, we have

$$|Q_a(t,\xi)| \leq C_1 \int_{\mathbb{R}} e^{-|y(\xi)-y(\eta)|} y_\xi(\eta)\,d\eta = C_1 \int_{\mathbb{R}} e^{-|y(\xi)-x|}\,dx = 2C_1$$

where the constant C_1 depends only on κ and h. Since $H_\xi \geq 0$, we have

$$\begin{aligned}
|Q_b(t,\xi)| &\leq \frac{\kappa}{4} \int_{\mathbb{R}} e^{-|y(\xi)-y(\eta)|} H_\xi(\eta)\,d\eta \\
&= \frac{\kappa}{4} \int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta) e^{-|y(\xi)-y(\eta)|} H(\eta) y_\xi(\eta)\,d\eta \quad \text{(after integrating by parts)} \\
&\leq \frac{\kappa h}{4} \int_{\mathbb{R}} e^{-|y(\xi)-y(\eta)|} y_\xi(\eta)\,d\eta \quad \text{(as } y_\xi \geq 0\text{)} \\
&= \frac{\kappa h}{2} \quad \text{(after changing variables).}
\end{aligned}$$

Hence, Q_b and therefore Q are bounded by a constant that depends only on κ and h. Similarly, one proves that $\sup_{t\in[0,T)} \|P(t,\cdot)\|_{L^\infty(\mathbb{R})}$ is bounded by such constant. We denote

$$\begin{aligned}
C_2 = \sup_{t\in[0,T)} \{&\|U(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|H(t,\cdot)\|_{L^\infty(\mathbb{R})} \\
&+ \|\zeta(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|P(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|Q(t,\cdot)\|_{L^\infty(\mathbb{R})}\}.
\end{aligned} \tag{2.46}$$

We have just proved that C_2 is finite and only depends on $\|\bar{X}\|_E$, T and κ. Let $t \in [0,T)$. We have, as $g(0) = 0$,

$$\|g(U(t,\cdot))\|_{L^2(\mathbb{R})} \leq \|g\|_{W^{1,\infty}(I)} \|U\|_{L^2(\mathbb{R})} \leq \kappa \|U\|_{L^2(\mathbb{R})}. \tag{2.47}$$

We use (2.47) and, from (2.21), we obtain that

$$\|R(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})})$$

for some constant C depending only on C_2, h and κ. From now on, we denote generically by C such constants that are increasing functions of $\|\bar{X}\|_E$, T and κ. Since A is a continuous linear map from $L^2(\mathbb{R})$ to $H^1(\mathbb{R})$, it is a fortiori continuous from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$, and we get

$$\|A \circ R(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})}).$$

From (2.20), as $\|e^{-\zeta(t,\cdot)}\|_{L^\infty(\mathbb{R})} \leq C$, we obtain that

$$\|Q_1(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})}).$$

The same bound holds for Q_2 and therefore

$$\|Q(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})}). \tag{2.48}$$

Similarly, one proves

$$\|P(t,\cdot)\|_{L^2(\mathbb{R})} \leq C(\|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})}). \tag{2.49}$$

Let $Z(t) = \|U(t,\cdot)\|_{L^2(\mathbb{R})} + \|\zeta_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|U_\xi(t,\cdot)\|_{L^2(\mathbb{R})} + \|H_\xi(t,\cdot)\|_{L^2(\mathbb{R})}$, then the theorem will be proved once we have established that $\sup_{t\in[0,T)} Z(t) < \infty$.

From the integrated version of (2.15) and (2.16), after taking the $L^2(\mathbb{R})$-norms on both sides and adding the relevant terms, we use (2.49), (2.47) and obtain

$$Z(t) \leq Z(0) + C\int_0^t Z(\tau)\,d\tau.$$

Hence, Gronwall's lemma gives us that $\sup_{t\in[0,T)} Z(t) < \infty$. Thus the solution exists globally in time. Moreover we have that

$$\sup_{t\in[0,T]} \|X(t,\,\cdot)\|_E \leq C(\|\bar{X}\|_E, T, \kappa) \tag{2.50}$$

where C is an increasing function and $\kappa = \|f\|_{W^{3,\infty}(I)} + \|g\|_{W^{1,\infty}(I)}$ with $I = [-\|\bar{H}\|^{\frac{1}{2}}_{L^\infty(\mathbb{R})}, \|\bar{H}\|^{\frac{1}{2}}_{L^\infty(\mathbb{R})}]$. Note that in order to obtain (2.47), we needed a bound on g in $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$, which explains why we are working on a bounded subset of $\mathcal{E}$ in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$.

We now turn to the proof of the continuity of the semigroup. Let $\mathcal{E}'$ be a bounded set of $\mathcal{E}$ in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$. Since $\mathcal{G}$ and $\mathcal{E}$ are metrizable, it is enough to prove sequential continuity. Let $\bar{X}_n = (\bar{y}_n, \bar{U}_n, \bar{H}_n) \in \mathcal{G}$ and $(f_n, g_n) \in \mathcal{E}'$ be sequences that converge to $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) \in \mathcal{G}$ and $(f,g) \in \mathcal{E}'$. We denote $X_n(t) = S_t(\bar{X}_n, f_n, g_n)$ and $X(t) = S_t(\bar{X}, f, g)$. Let $M = \sup_{n\geq 1}\|\bar{X}_n\|_E$, we have $\|\bar{H}_n\|_{L^\infty(\mathbb{R})} \leq M$ for all $n \geq 1$. Hence, from (2.44), it follows that

$$U_n(t,\xi) \in I := [-\sqrt{M}, \sqrt{M}] \tag{2.51}$$

for all $n \geq 1$ and $(t,\xi) \in \mathbb{R}_+ \times \mathbb{R}$. Since $\mathcal{E}'$ is a bounded set of $\mathcal{E}$ in $W^{2,\infty}_{\mathrm{loc}}(\mathbb{R}) \times W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$ and $(f_n, g_n) \in \mathcal{E}'$, there exists a constant $\kappa > 0$ such that

$$\|f_n\|_{W^{2,\infty}(I)} + \|g_n\|_{W^{1,\infty}(I)} \leq \kappa$$

for all $n \geq 1$. Hence, as $I_n := [-\|\bar{H}_n\|^{\frac{1}{2}}_{L^\infty(\mathbb{R})}, \|\bar{H}_n\|^{\frac{1}{2}}_{L^\infty(\mathbb{R})}] \subset I$,

$$\kappa_n = \|f_n\|_{W^{2,\infty}(I_n)} + \|g_n\|_{W^{1,\infty}(I_n)} \leq \kappa \tag{2.52}$$

for all $n \geq 1$. Given $T > 0$, it follows from (2.50) and (2.52) that, for all $n \geq 1$,

$$\sup_{t\in[0,T]} \|X_n(t,\,\cdot)\|_E \leq C(\|\bar{X}_n\|_E, T, \kappa_n) \leq C(M, T, \kappa) = C' \tag{2.53}$$

and $\sup_{t\in[0,T]} \|X_n(t,\,\cdot)\|_E$ is bounded uniformly with respect to n. We have

$$\|X_n(t) - X(t)\|_E \leq \|\bar{X}_n - \bar{X}\|_E + \int_0^t \|F(X_n, f_n, g_n) - F(X, f, g)\|_E(s)\,ds \tag{2.54}$$

where F denotes the right-hand side of (2.15). We consider a fixed time $t \in [0,T]$ and drop the time dependence in the notation for the moment. We have

$$\begin{aligned}\|F(X_n, f_n, g_n) - F(X, f, g)\|_E \leq{}& \|F(X_n, f_n, g_n) - F(X_n, f, g)\|_E \\ &+ \|F(X_n, f, g) - F(X, f, g)\|_E. \end{aligned}\tag{2.55}$$

The map $X \mapsto F(X, f, g)$ is Lipschitz on any bounded set of E, see the proof of Theorem 2.4. Hence, after denoting L the Lipschitz function of this map on the ball $\{X \in E \mid \|X\|_E \leq C'\}$, we get from (2.53) that

$$\|F(X_n, f, g) - F(X, f, g)\|_E \leq L \|X_n - X\|_E \,. \tag{2.56}$$

Denote by Q_n and $\tilde{Q}_n$ the expressions given by the definition (2.13) of Q where we replace X, f, g by X_n, f_n, g_n and X_n, f, g, respectively. We use the same notations and define P_n and $\tilde{P}_n$ from (2.14), and G_n and $\tilde{G}_n$ from (2.4). For example, we have

$$G_n = \int_0^{U_n} (2g_n(z) + f_n''(z)z^2)\, dz \text{ and } \tilde{G}_n = \int_0^{U_n} (2g(z) + f''(z)z^2)\, dz.$$

Still using the same notations, we have, from (2.21),

$$\begin{aligned}(R_n - \tilde{R}_n)(\xi) = e^{\zeta_n}\big(g_n(U_n) - g(U_n) - \frac{1}{2}(f_n'(U_n) - f'(U_n))U_n^2)\big)(1 + \zeta_{n,\xi}) \\ + \frac{1}{2}e^{\zeta_n}(f_n''(U_n) - f''(U_n))H_{n,\xi}.\end{aligned} \tag{2.57}$$

Let $\delta_n = \|f_n - f\|_{W^{2,\infty}(I)} + \|g_n - g\|_{L^\infty(I)}$. Since $(f_n, g_n) \to (f, g)$ in $\mathcal{E}$, $\delta_n \to 0$ as $n \to \infty$. Hence, from (2.57) and (2.51), we get

$$\begin{aligned}\left\|R_n - \tilde{R}_n\right\|_{L^2(\mathbb{R})} \leq e^{C'}\big(\, &\|g_n(U_n) - g(U_n)\|_{L^2(\mathbb{R})} + C' \|g_n(U_n) - g(U_n)\|_{L^\infty(\mathbb{R})} \\ &+ \frac{1}{2}C' \|f_n'(U_n) - f'(U_n)\|_{L^\infty(\mathbb{R})} (M + \sqrt{M}) \\ &+ \frac{1}{2}C' \|f_n''(U_n) - f''(U_n)\|_{L^\infty(\mathbb{R})} \big).\end{aligned} \tag{2.58}$$

Let $\delta_n' = \|g_n(U) - g(U)\|_{L^2(\mathbb{R})}$, we then have

$$\begin{aligned}\|g(U_n) - g_n(U_n)\|_{L^2(\mathbb{R})} &\leq \|g_n(U_n) - g_n(U)\|_{L^2(\mathbb{R})} + \delta_n' + \|g(U_n) - g(U)\|_{L^2(\mathbb{R})} \\ &\leq 2\kappa \|U_n - U\|_{L^2(\mathbb{R})} + \delta_n'.\end{aligned} \tag{2.59}$$

Since $g_n \to g$ in $L^\infty(I)$, $g_n(U) \to g_n(U)$ in $L^\infty(\mathbb{R})$. As $|g_n(U) - g(U)| \leq 2\kappa |U|$ (because $g(0) = 0$ and $\|g\|_{W^{1,\infty}(I)} \leq \kappa$), we can apply the Lebesgue dominated convergence theorem and obtain that $\lim_{n\to\infty} \delta_n' = 0$. From (2.58) and (2.59), we obtain that

$$\left\|R_n - \tilde{R}_n\right\|_{L^2(\mathbb{R})} \leq C\big(\delta_n + \delta_n' + \|U_n - U\|_{L^2(\mathbb{R})}\big)$$

for some constant C which depends on M, T and κ. Again, we denote generically by C such constants that are increasing functions of M, T and κ, and are independent on n. Since A in (2.20) is continuous from $L^2(\mathbb{R})$ to $H^1(\mathbb{R})$, it follows that $\left\|Q_n - \tilde{Q}_n\right\|_{H^1(\mathbb{R})} \leq C\big(\delta_n + \delta_n' + \|U_n - U\|_{L^2(\mathbb{R})}\big)$. Similarly, one proves that

$\left\|P_n - \tilde{P}_n\right\|_{H^1(\mathbb{R})} \leq C\big(\delta_n + \delta_n' + \|U_n - U\|_{L^2(\mathbb{R})}\big)$. We have

$$\begin{aligned}\left\|G_n - \tilde{G}_n\right\|_V &= \left\|G_n - \tilde{G}_n\right\|_{L^\infty(\mathbb{R})} \\ &\quad + \left\|\big(2(g_n(U_n) - g(U_n)) + (f_n''(U_n) - f''(U_n))U_n^2\big)U_{n,\xi}\right\|_{L^2(\mathbb{R})} \\ &\leq \sqrt{M}\big(2\left\|g_n(U_n) - g(U_n)\right\|_{L^\infty(\mathbb{R})} + M\left\|f_n''(U_n) - f''(U_n)\right\|_{L^\infty(\mathbb{R})}\big) \\ &\quad + 2C'\left\|g_n(U_n) - g(U_n)\right\|_{L^\infty(\mathbb{R})} + C'M\left\|f_n''(U_n) - f''(U_n)\right\|_{L^\infty(\mathbb{R})} \\ &\leq C\delta_n\end{aligned}$$

by (2.59). Finally, we have

$$\left\|F(X_n, f_n, g_n) - F(X_n, f, g)\right\|_E \leq C\big(\delta_n + \delta_n' + \|U_n - U\|_{L^2(\mathbb{R})}\big). \tag{2.60}$$

Gathering (2.54), (2.55), (2.56) and (2.60), we end up with

$$\|X_n(t) - X(t)\|_E \leq \left\|\bar{X}_n - \bar{X}\right\|_E + CT(\delta_n + \delta_n') + (L + C)\int_0^t \|X_n - X\|_E(s)\,ds$$

and Gronwall's lemma yields

$$\|X_n(t) - X(t)\|_E \leq \big(\left\|\bar{X}_n - \bar{X}\right\|_E + CT(\delta_n + \delta_n')\big)\,e^{(L+C)T}.$$

Hence, $X_n \to X$ in E uniformly in $[0, T]$. ☐

The solutions are well-defined in our new sets of coordinates. Now we want to go back to the original variable u. We define $u(t, x)$ as

$$u(x, t) = U(\xi) \text{ for any } \xi \text{ such that } x = y(\xi). \tag{2.61}$$

Let us prove that this definition is well-posed. Given $x \in \mathbb{R}$, since y is increasing, continuous and $\lim_{\xi\to\pm\infty} y = \pm\infty$, y is surjective and there exists ξ such that $x = y(\xi)$. Suppose we have $\xi_1 < \xi_2$ with $x = y(\xi_1) = y(\xi_2)$. Then, since y is monotone, $y(\xi) = y(\xi_1) = y(\xi_2)$ for all $\xi \in (\xi_1, \xi_2)$ and $y_\xi = 0$ in this interval. From (2.27c), it follows that $U_\xi = 0$ on (ξ_1, ξ_2) and therefore $U(\xi_1) = U(\xi_2)$.

Theorem 2.9 (Existence of weak solutions)**.** *For initial data $\bar{u} \in H^1(\mathbb{R})$, let $(\bar{y}, \bar{U}, \bar{H})$ be as given by* (2.26) *and (y, U, H) be the solution of* (2.15) *with initial data $(\bar{y}, \bar{U}, \bar{H})$. Then u as defined in* (2.61) *belongs to $C(\mathbb{R}_+, L^\infty(\mathbb{R})) \cap L^\infty(\mathbb{R}_+, H^1(\mathbb{R}))$ and is a weak solution of* (1.1)*.*

Proof. Let us prove that $u \in L^\infty(\mathbb{R}, H^1(\mathbb{R}))$. We consider a fix time t and drop it in the notation when there is no ambiguity. For any smooth function ϕ, after using the change of variable $x = y(\xi)$, we obtain

$$\int_{\mathbb{R}} u\phi\,dx = \int_{\mathbb{R}} U(\phi\circ y)y_\xi\,d\xi = \int_{\mathbb{R}} U\sqrt{y_\xi}(\phi\circ y)\sqrt{y_\xi}\,d\xi.$$

Hence, by Cauchy–Schwarz,

$$\left|\int_{\mathbb{R}} u\phi\,dx\right| \leq \|\phi\|_{L^2(\mathbb{R})}\sqrt{\int_{\mathbb{R}} U^2 y_\xi\,d\xi} \leq \sqrt{H(\infty)}\,\|\phi\|_{L^2(\mathbb{R})}$$

as $U^2 y_\xi \leq H_\xi$ from (2.27c). Therefore, $u \in L^2(\mathbb{R})$ and

$$\|u(t,\cdot)\|_{L^2(\mathbb{R})} \leq \sqrt{H(t,\infty)} = \sqrt{H(0,\infty)} = \|\bar{u}\|_{H^1(\mathbb{R})}.$$

For any smooth function ϕ, we have, after using the change of variable $x = y(\xi)$,

$$\int_{\mathbb{R}} u(x)\phi_x(x)\,dx = \int_{\mathbb{R}} U(\xi)\phi_x(y(\xi))y_\xi(\xi)\,d\xi = -\int_{\mathbb{R}} U_\xi(\xi)(\phi\circ y)(\xi)\,d\xi. \tag{2.62}$$

Let $B = \{\xi \in \mathbb{R} \mid y_\xi(\xi) > 0\}$. Because of (2.27c), and since $y_\xi \geq 0$ almost everywhere, we have $U_\xi = 0$ almost everywhere on B^c. Hence, we can restrict the integration domain in (2.62) to B. We divide and multiply by $\sqrt{y_\xi}$ the integrand in (2.62) and obtain, after using the Cauchy–Schwarz inequality,

$$\left|\int_{\mathbb{R}} u\phi_x\,dx\right| = \left|\int_B \frac{U_\xi}{\sqrt{y_\xi}}(\phi\circ y)\sqrt{y_\xi}\,d\xi\right| \leq \sqrt{\int_B \frac{U_\xi^2}{y_\xi}\,d\xi}\sqrt{\int_B (\phi\circ y)^2 y_\xi\,d\xi}.$$

By (2.27c), we have $\frac{U_\xi^2}{y_\xi} \leq H_\xi$. Hence, after another change of variables, we get

$$\left|\int_{\mathbb{R}} u\phi_x\,dx\right| \leq \sqrt{H(\infty)}\,\|\phi\|_{L^2(\mathbb{R})},$$

which implies that $u_x \in L^2(\mathbb{R})$ and $\|u_x(t,\cdot)\|_{L^2(\mathbb{R})} \leq \|\bar{u}\|_{H^1(\mathbb{R})}$. Hence, $u \in L^\infty(\mathbb{R}, H^1(\mathbb{R}))$ and

$$\|u\|_{L^\infty(\mathbb{R},H^1(\mathbb{R}))} \leq 2\,\|\bar{u}\|_{H^1(\mathbb{R})}. \tag{2.63}$$

Let us prove sequential convergence in $C(\mathbb{R}_+, L^\infty(\mathbb{R}))$. Given $t \in \mathbb{R}_+$ and a sequence $t_n \in \mathbb{R}_+$ with $t_n \to t$, converges to t, we set $y_n(\xi) = y(t_n,\xi)$, $U_n(\xi) = U(t_n,\xi)$ and $H_n(\xi)) = H(t_n,\xi)$ and, slightly abusing notation, $(y(\xi), U(\xi), H(\xi)) = (y(t,\xi), U(t,\xi), H(t,\xi))$. For any $x \in \mathbb{R}$, there exist ξ_n and ξ, which may not be unique, such that $x = y_n(\xi_n)$ and $x = y(\xi)$. We set $x_n = y_n(\xi)$. We have

$$u(t_n,x) - u(t,x) = u(t_n,x) - u(t_n,x_n) + U_n(\xi) - U(\xi) \tag{2.64}$$

and

$$\begin{aligned}
|u(t_n,x) - u(t_n,x_n)| &= \left|\int_{x_n}^{x} u_x(t_n,x')\,dx'\right| \\
&\leq \sqrt{|x_n - x|}\left(\int_{x_n}^{x} u_x(t_n,x')^2\,dx'\right)^{1/2} \qquad \text{(Cauchy–Schwarz)} \\
&\leq \sqrt{|y_n(\xi) - y(\xi)|}\,\|u\|_{L^\infty(\mathbb{R},H^1(\mathbb{R}))} \\
&\leq 2\,\|\bar{u}\|_{H^1(\mathbb{R})}\,\|y - y_n\|_{L^\infty(\mathbb{R})}^{1/2},
\end{aligned} \tag{2.65}$$

by (2.63). Since $y_n \to y$ and $U_n \to U$ in $L^\infty(\mathbb{R})$, it follows from (2.64) and (2.65) that $u_n \to u$ in $L^\infty(\mathbb{R})$.

Since $u \in L^\infty(\mathbb{R}_+, H^1(\mathbb{R}))$, $g(u) + \frac{1}{2}f''(u)u_x^2 \in L^\infty(\mathbb{R}_+, L^1(\mathbb{R}))$ and, since $v \mapsto (1-\partial_{xx})^{-1}v$ is continuous from $H^{-1}(\mathbb{R})$ to $H^1(\mathbb{R})$, $P \in L^\infty(\mathbb{R}_+, H^1(\mathbb{R}))$.

We say that u is a weak solution of (1.5) if

$$\int_{\mathbb{R}_+\times\mathbb{R}} \left(-u\phi_t + f'(u)u_x\phi + P_x\phi\right)(t,x)\,dtdx = 0 \tag{2.66}$$

for all $\phi \in C^\infty(\mathbb{R}_+\times\mathbb{R})$ with compact support. For $t \in \mathcal{K}^c$, that is for almost every t (see (2.41) in Lemma 2.7), $y_\xi(t,\xi) > 0$ for almost every $\xi \in \mathbb{R}$ and $y(t,\cdot)$ is invertible, we have $U_\xi = u_x \circ y y_\xi$ and, after using the change of variables $x = y(t,\xi)$, we get

$$\begin{aligned}\int_{\mathbb{R}_+\times\mathbb{R}} &\left[-u(t,x)\phi_t(t,x) + f'(u(t,x))u_x(t,x)\phi(t,x)\right]\,dxdt\\ &= \int_{\mathbb{R}_+\times\mathbb{R}} \left[-U(t,\xi)y_\xi(t,\xi)\phi_t(t,y(t,\xi)) + f'(U(t,\xi))U_\xi(t,\xi)\phi(t,y(t,\xi))\right]\,d\xi dt.\end{aligned} \tag{2.67}$$

Using the fact that $y_t = f'(U)$ and $y_{\xi t} = f''(U)U_\xi$, one easily check that

$$(Uy_\xi\phi\circ y)_t - (f'(U)U\phi\circ y)_\xi = Uy_\xi\phi_t\circ y - f'(U)U_\xi\phi\circ y + U_ty_\xi\phi\circ y. \tag{2.68}$$

After integrating (2.68) over $\mathbb{R}_+\times\mathbb{R}$, the left-hand side of (2.68) vanishes and we obtain

$$\begin{aligned}\int_{\mathbb{R}_+\times\mathbb{R}} &\Big[-Uy_\xi\,\phi_t\circ y + f'(U)U_\xi\,\phi\circ y\Big]\,d\xi dt\\ &= \frac{1}{2}\int_{\mathbb{R}_+\times\mathbb{R}^2} \Big[\operatorname{sgn}(\xi-\eta)e^{-|y(\xi)-y(\eta)|}\\ &\qquad \times \big((g(U) - \frac{1}{2}f''(U)U^2))y_\xi + \frac{1}{2}f''(U)H_\xi\big)(\eta)y_\xi(\xi)\phi\circ y(\xi)\Big]\,d\eta d\xi dt\end{aligned} \tag{2.69}$$

by (2.15). Again, to simplify the notation, we deliberately omitted the t variable. On the other hand, by using the change of variables $x = y(t,\xi)$ and $z = y(t,\eta)$ when $t \in \mathcal{K}^c$, we have

$$\begin{aligned}-\int_{\mathbb{R}_+\times\mathbb{R}} P_x(t,x)\phi(t,x)\,dxdt &= \frac{1}{2}\int_{\mathbb{R}_+\times\mathbb{R}^2} \Big[\operatorname{sgn}(y(\xi)-y(\eta))e^{-|y(\xi)-y(\eta)|}\\ &\qquad \times \big(g\circ u + \frac{1}{2}f''\circ uu_x^2\big)(t,y(\eta))\phi(t,y(\xi))y_\xi(\eta)y_\xi(\xi)\Big]\,d\eta d\xi dt.\end{aligned}$$

For $t \in \mathcal{K}^c$, that is, for almost every t, $y_\xi(t,\xi)$ is strictly positive for almost every ξ, and we can replace $u_x(t,y(t,\eta))$ by $U_\xi(t,\eta)/y_\xi(t,\eta)$ in the equation above.

Using (2.27c), we obtain

$$
\begin{aligned}
-\int_{\mathbb{R}_+\times\mathbb{R}} & P_x(t,x)\phi(t,x)\,dxdt \qquad (2.70)\\
&= \frac{1}{2}\int_{\mathbb{R}_+\times\mathbb{R}^2}\Big[\operatorname{sgn}(\xi-\eta)e^{-|y(\xi)-y(\eta)|}\\
&\qquad\times\big((g(U)-\frac{1}{2}f''(U)U^2))y_\xi+\frac{1}{2}f''(U)H_\xi\big)(\eta)y_\xi(\xi)\phi\circ y(\xi)\Big]\,d\eta d\xi dt \qquad (2.71)
\end{aligned}
$$

Thus, comparing (2.69) and (2.71), we get

$$
\int_{\mathbb{R}_+\times\mathbb{R}}\left[-Uy_\xi\,\phi_t(t,y)+f'(U)U_\xi\,\phi\right]\,d\xi dt=-\int_{\mathbb{R}_+\times\mathbb{R}}P_x(t,x)\phi(t,x)\,dxdt
$$

and (2.66) follows from (2.67). □

3. Continuous semi-group of solutions

We denote by G the subgroup of the group of homeomorphisms from $\mathbb{R}$ to $\mathbb{R}$ such that

$$
f-\mathrm{Id}\ \text{ and }\ f^{-1}-\mathrm{Id}\ \text{ both belong to } W^{1,\infty}(\mathbb{R}) \tag{3.1}
$$

where Id denotes the identity function. The set G can be interpreted as the set of relabeling functions. For any $\alpha>1$, we introduce the subsets G_α of G defined by

$$
G_\alpha=\{f\in G\mid \|f-\mathrm{Id}\|_{W^{1,\infty}(\mathbb{R})}+\left\|f^{-1}-\mathrm{Id}\right\|_{W^{1,\infty}(\mathbb{R})}\le\alpha\}.
$$

The subsets G_α do not possess the group structure of G. We have the following characterization of G_α:

Lemma 3.1. [16, Lemma 3.2] *Let $\alpha\ge 0$. If f belongs to G_α, then $1/(1+\alpha)\le f_\xi\le 1+\alpha$ almost everywhere. Conversely, if f is absolutely continuous, $f-\mathrm{Id}\in L^\infty(\mathbb{R})$ and there exists $c\ge 1$ such that $1/c\le f_\xi\le c$ almost everywhere, then $f\in G_\alpha$ for some α depending only on c and $\|f-\mathrm{Id}\|_{L^\infty(\mathbb{R})}$.*

We define the subsets $\mathcal{F}_\alpha$ and $\mathcal{F}$ of $\mathcal{G}$ as follows

$$
\mathcal{F}_\alpha=\{X=(y,U,H)\in\mathcal{G}\mid y+H\in G_\alpha\},
$$

and

$$
\mathcal{F}=\{X=(y,U,H)\in\mathcal{G}\mid y+H\in G\}.
$$

For $\alpha=0$, we have $G_0=\{\mathrm{Id}\}$. As we will see, the space $\mathcal{F}_0$ will play a special role. These sets are relevant only because they are in some sense preserved by the governing equation (2.15) as the next lemma shows.

Lemma 3.2. *The space $\mathcal{F}$ is preserved by the governing equation* (2.15). *More precisely, given $\alpha,T\ge 0$, a bounded set $B_M=\{X\in E\mid \|X\|_E\le M\}$ of E and a bounded set $\mathcal{E}'$ of $\mathcal{E}$, we have, for any $t\in[0,T]$, $\bar X\in\mathcal{F}_\alpha\cap B_M$ and $(f,g)\in\mathcal{E}'$,*

$$
S_t(\bar X,f,g)\in\mathcal{F}_{\alpha'}
$$

where α' only depends on T, α, M and $\mathcal{E}'$.

Proof. Let $\bar{X} = (\bar{y}, \bar{U}, \bar{H}) \in \mathcal{F}_\alpha$, we denote $X(t) = (y(t), U(t), H(t))$ the solution of (2.15) with initial data $\bar{X}$. By definition, we have $\bar{y} + \bar{H} \in G_\alpha$ and, from Lemma 3.1, $1/c \leq \bar{y}_\xi + \bar{H}_\xi \leq c$ almost everywhere, for some constant $c > 1$ depending only α. Let $h = \bar{H}(\infty) = H(t, \infty)$. We have $h \leq M$ and, from (2.44), $\|U\|_{L^\infty(\mathbb{R}_+ \times \mathbb{R})} \leq \sqrt{h} \leq \sqrt{M}$. Let $I = [-\sqrt{M}, \sqrt{M}]$. Since $\mathcal{E}'$ is bounded, there exists $\kappa > 0$ such that $\|f\|_{W^{2,\infty}(I)} + \|g\|_{W^{1,\infty}(I)} \leq \kappa$ for all $(f, g) \in \mathcal{E}'$. We consider a fixed ξ and drop it in the notation. Applying Gronwall's inequality to (2.16) to the function $X(t - \tau)$, we obtain

$$|y_\xi(0)| + |H_\xi(0)| + |U_\xi(0)| \leq e^{CT} \left(|y_\xi(t)| + |H_\xi(t)| + |U_\xi(t)|\right) \tag{3.2}$$

for some constant C which depends on $\|f''(U)\|_{L^\infty(\mathbb{R})}$, $\|P\|_{L^\infty(\mathbb{R})}$, $\|g(U)\|_{L^\infty(\mathbb{R})}$, $\|Q\|_{L^\infty(\mathbb{R})}$, $\|U\|_{L^\infty(\mathbb{R})}$ and $\|G'(U)\|_{L^\infty(\mathbb{R})}$. In (2.46), we proved that $\|P\|_{L^\infty(\mathbb{R})}$ and $\|Q\|_{L^\infty(\mathbb{R})}$ only depend on M, κ, T. Hence, the constant C in (3.2) also only depends on M, T and κ. From (2.27c), we have

$$|U_\xi(t)| \leq \sqrt{y_\xi(t) H_\xi(t)} \leq \frac{1}{2}(y_\xi(t) + H_\xi(t)).$$

Hence, since y_ξ and H_ξ are positive, (3.2) gives us

$$\frac{1}{c} \leq \bar{y}_\xi + \bar{H}_\xi \leq \frac{3}{2} e^{CT} (y_\xi(t) + H_\xi(t)),$$

and $y_\xi(t) + H_\xi(t) \geq \frac{2}{3c} e^{-CT}$. Similarly, by applying Gronwall's lemma, we obtain $y_\xi(t) + H_\xi(t) \leq \frac{3}{2} c e^{CT}$. We have $\|(y + H)(t) - \xi\|_{L^\infty(\mathbb{R})} \leq \|X(t)\|_{C([0,T],E)} \leq C(M, T, \kappa)$, see (2.50). Hence, applying Lemma 3.1, we obtain that $y(t, \cdot\,) + H(t, \cdot\,) \in G_{\alpha'}$ and therefore $X(t) \in \mathcal{F}_{\alpha'}$ for some α' depending only on α, T, M and $\mathcal{E}'$. $\square$

For the sake of simplicity, for any $X = (y, U, H) \in \mathcal{F}$ and any function $r \in G$, we denote $(y \circ r, U \circ r, H \circ r)$ by $X \circ r$.

Proposition 3.3. [16, Proposition 3.4] *The map from $G \times \mathcal{F}$ to $\mathcal{F}$ given by $(r, X) \mapsto X \circ r$ defines an action of the group G on $\mathcal{F}$.*

Since G is acting on $\mathcal{F}$, we can consider the quotient space $\mathcal{F}/G$ of $\mathcal{F}$ with respect to the action of the group G. The equivalence relation on $\mathcal{F}$ is defined as follows: For any $X, X' \in \mathcal{F}$, X and X' are equivalent if there exists $r \in G$ such that $X' = X \circ r$. Heuristically it means that X' and X are equivalent up to a relabeling function. We denote by $\Pi(X) = [X]$ the projection of $\mathcal{F}$ into the quotient space $\mathcal{F}/G$. We introduce the map $\Gamma\colon \mathcal{F} \to \mathcal{F}_0$ given by

$$\Gamma(X) = X \circ (y + H)^{-1}$$

for any $X = (y, U, H) \in \mathcal{F}$. We have $\Gamma(X) = X$ when $X \in \mathcal{F}_0$. It is not hard to prove that Γ is invariant under the G action, that is, $\Gamma(X \circ r) = \Gamma(X)$ for any $X \in \mathcal{F}$ and $r \in G$. Hence, there corresponds to Γ a map $\tilde{\Gamma}$ from the quotient space $\mathcal{F}/G$ to $\mathcal{F}_0$ given by $\tilde{\Gamma}([X]) = \Gamma(X)$ where $[X] \in \mathcal{F}/G$ denotes the equivalence class of $X \in \mathcal{F}$. For any $X \in \mathcal{F}_0$, we have $\tilde{\Gamma} \circ \Pi(X) = \Gamma(X) = X$. Hence, $\tilde{\Gamma} \circ \Pi|_{\mathcal{F}_0} = \operatorname{Id}|_{\mathcal{F}_0}$. Any topology defined on $\mathcal{F}_0$ is naturally transported into $\mathcal{F}/G$ by this isomorphism. We equip $\mathcal{F}_0$ with the metric induced by the E-norm, i.e., $d_{\mathcal{F}_0}(X,X') = \|X - X'\|_E$ for all $X, X' \in \mathcal{F}_0$. Since $\mathcal{F}_0$ is closed in E, this metric is complete. We define the metric on $\mathcal{F}/G$ as

$$d_{\mathcal{F}/G}([X],[X']) = \|\Gamma(X) - \Gamma(X')\|_E\,,$$

for any $[X],[X'] \in \mathcal{F}/G$. Then, $\mathcal{F}/G$ is isometrically isomorphic with $\mathcal{F}_0$ and the metric $d_{\mathcal{F}/G}$ is complete.

Lemma 3.4. [16, Lemma 3.5] *Given $\alpha \geq 0$. The restriction of Γ to $\mathcal{F}_\alpha$ is a continuous map from $\mathcal{F}_\alpha$ to $\mathcal{F}_0$.*

Remark 3.5. The map Γ is not continuous from $\mathcal{F}$ to $\mathcal{F}_0$. The spaces $\mathcal{F}_\alpha$ were precisely introduced in order to make the map Γ continuous.

We denote by $S\colon \mathcal{F}\times\mathcal{E}\times\mathbb{R}_+ \to \mathcal{F}$ the continuous semigroup which to any initial data $\bar{X} \in \mathcal{F}$ associates the solution $X(t)$ of the system of differential equation (2.15) at time t as defined in Theorem 2.9. As we indicated earlier, the generalized hyperelastic-rod wave equation is invariant with respect to relabeling, more precisely, using our terminology, we have the following result.

Theorem 3.6. *For any $t > 0$, the map $S_t\colon \mathcal{F} \to \mathcal{F}$ is G-equivariant (for f and g given), that is,*

$$S_t(X\circ r) = S_t(X)\circ r \tag{3.3}$$

for any $X \in \mathcal{F}$ and $r \in G$. Hence, the map $\tilde{S} : \mathcal{F}/G\times\mathcal{E}\times\mathbb{R}_+ \to \mathcal{F}/G$ given by

$$\tilde{S}_t([X],f,g) = [S_t(X,f,g)]$$

is well-defined. It generates a continuous semigroup with respect to all variables, on any bounded set of $\mathcal{E}$.

Proof. For any $X_0 = (y_0, U_0, H_0) \in \mathcal{F}$ and $r \in G$, we denote $\bar{X}_0 = (\bar{y}_0, \bar{U}_0, \bar{H}_0) = X_0 \circ r$, $X(t) = S_t(X_0)$ and $\bar{X}(t) = S_t(\bar{X}_0)$. We claim that $X(t)\circ r$ satisfies (2.15) and therefore, since $X(t)\circ r$ and $\bar{X}(t)$ satisfy the same system of differential equation with the same initial data, they are equal. We denote $\hat{X}(t) = (\hat{y}(t), \hat{U}(t), \hat{H}(t)) = X(t)\circ r$. We have

$$\hat{U}_t = \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(\hat{y}(\xi)-y(\eta))\big) \\ \times\Big((g(U) - \frac{1}{2}f''(U)U^2)y_\xi + \frac{1}{2}f''(U)H_\xi\Big)(\eta)\,d\eta. \tag{3.4}$$

We have $\hat{y}_\xi(\xi) = y_\xi(r(\xi))r_\xi(\xi)$ and $\hat{H}_\xi(\xi) = H_\xi(r(\xi))r_\xi(\xi)$ for almost every $\xi \in \mathbb{R}$. Hence, after the change of variables $\eta = r(\eta')$, we get from (3.4) that

$$\hat{U}_t = \frac{1}{4}\int_{\mathbb{R}} \operatorname{sgn}(\xi-\eta)\exp\big(-\operatorname{sgn}(\xi-\eta)(\hat{y}(\xi)-\hat{y}(\eta))\big) \\ \times\Big((g(\hat{U}) - \frac{1}{2}f''(\hat{U})\hat{U}^2)\hat{y}_\xi + \frac{1}{2}f''(\hat{U})\hat{H}_\xi\Big)(\eta)\,d\eta.$$

We treat the other terms in (2.15) similarly, and it follows that $(\hat{y}, \hat{U}, \hat{H})$ is a solution of (2.15). Since $(\hat{y}, \hat{U}, \hat{H})$ and $(\bar{y}, \bar{U}, \bar{H})$ satisfy the same system of ordinary differential equations with the same initial data, they are equal, i.e.,

$$\bar{X}(t) = X(t)\circ r\,,$$

and (3.3) is proved. Let $\mathcal{E}'$ be a bounded set of $\mathcal{E}$ and $T > 0$. For $t \in [0, T]$, we have the following diagram

$$\begin{tikzcd}
\mathcal{F}_0 \arrow[r, "\Pi"] & \mathcal{F}/G \\
\mathcal{F}_\alpha \arrow[u, "\Gamma"] & \\
\mathcal{F}_0 \times \mathcal{E}' \arrow[u, "S_t"] \arrow[r, "\Pi"] & \mathcal{F}/G \times \mathcal{E}' \arrow[uu, "\tilde{S}_t"']
\end{tikzcd} \tag{3.5}$$

on a bounded domain of $\mathcal{F}_0$ whose diameter together with T and $\mathcal{E}'$ determines the constant α, see Lemma 3.2. By the definition of the metric on $\mathcal{F}/G$, the map $\tilde{\Gamma}$ is an isometry from $\mathcal{F}/G$ to $\mathcal{F}_0$. Hence, from the diagram (3.5), we see that $\tilde{S}_t \colon \mathcal{F}/G \times \mathcal{E}' \to \mathcal{F}/G$ is continuous if and only if $\Gamma \circ S_t \colon \mathcal{F}_0 \times \mathcal{E}' \to \mathcal{F}_0$ is continuous. Let us prove that $\Gamma \circ S_t \colon \mathcal{F}_0 \times \mathcal{E}' \to \mathcal{F}_0$ is sequentially continuous. We consider a sequence $X_n \in \mathcal{F}_0$ that converges to $X \in \mathcal{F}_0$ in $\mathcal{F}_0$, that is, $\lim_{n\to\infty} \|X_n - X\|_E = 0$ and a sequence $(f_n, g_n) \in \mathcal{E}'$ that converges to $(f, g) \in \mathcal{E}'$ in $\mathcal{E}$. From Theorem 2.8, we get that $\lim_{n\to\infty} \|S_t(X_n, f_n, g_n) - S_t(X, f, g)\|_E = 0$. Since $X_n \to X$ in E, there exists a constant $C \geq 0$ such that $\|X_n\| \leq C$ for all n. Lemma 3.2 gives us that for $t \in [0, T]$, $S_t(X_n, f_n, g_n) \in \mathcal{F}_\alpha$ for some α which depends on C, T and $\mathcal{E}'$ but is independent of n. Hence, $S_t(X_n, f_n, g_n) \to S_t(X, f, g)$ in $\mathcal{F}_\alpha$. Then, by Lemma 3.4, we obtain that $\Gamma \circ S_t(X_n, f_n, g_n) \to \Gamma \circ S_t(X, f, g)$ in $\mathcal{F}_0$ and uniformly in $[0, T]$. □

3.1. **From Eulerian to Lagrangian coordinates and vice versa.** As noted in [1] in the case of the Camassa-Holm equation, even if $H^1(\mathbb{R})$ is a natural space for the equation, there is no hope to obtain a semigroup of solutions by only considering $H^1(\mathbb{R})$. Thus, we introduce the following space $\mathcal{D}$, which characterizes the solutions in *Eulerian coordinates*:

Definition 3.7. The set $\mathcal{D}$ is composed of all pairs (u, μ) such that u belongs to $H^1(\mathbb{R})$ and μ is a positive finite Radon measure whose absolute continuous part, μ_{ac}, satisfies

$$\mu_{\mathrm{ac}} = (u^2 + u_x^2)\, dx. \tag{3.6}$$

There exists a bijection between Eulerian coordinates (functions in $\mathcal{D}$) and Lagrangian coordinates (functions in $\mathcal{F}/G$). Earlier we considered initial data in $\mathcal{D}$ with a special structure: The energy density μ was given by $(u^2 + u_x^2)\, dx$ and therefore μ did not have any singular part. The set $\mathcal{D}$ however allows the energy density to have a singular part and a positive amount of energy can concentrate on a set of Lebesgue measure zero. We constructed corresponding initial data in $\mathcal{F}_0$ by the means of (2.26a), (2.26b), and (2.26c). This construction can be generalized in the following way. Let us denote by $L\colon \mathcal{D} \to \mathcal{F}/G$ the map transforming Eulerian coordinates into Lagrangian coordinates whose definition is contained in the following theorem.

Theorem 3.8. [16, Theorem 3.8] *For any (u,μ) in $\mathcal{D}$, let*

$$y(\xi) = \sup\left\{y \mid \mu((-\infty,y)) + y < \xi\right\}, \tag{3.7a}$$

$$H(\xi) = \xi - y(\xi), \tag{3.7b}$$

$$U(\xi) = u{\circ}y(\xi)\,. \tag{3.7c}$$

Then $(y,U,H) \in \mathcal{F}_0$. We define $L(u,\mu) \in \mathcal{F}/G$ to be the equivalence class of (y,U,H).

Remark 3.9. If μ is absolutely continuous, then $\mu = (u^2 + u_x^2)dx$ and the function $y \mapsto \mu((-\infty,y))$ is continuous. From the definition (3.7a), we know that there exist an increasing sequence x_i and a decreasing sequence x_i' which both converge to $y(\xi)$ and such that

$$\mu((-\infty,x_i)) + x_i < \xi \text{ and } \mu((-\infty,x_i')) + x_i' \geq \xi.$$

Since $y \mapsto \mu((-\infty,y))$ is continuous, it implies, after letting i go to infinity, that $\mu((-\infty,y(\xi))) + y(\xi) = \xi$. Hence,

$$\int_{-\infty}^{y(\xi)} (u^2 + u_x^2)\,dx + y(\xi) = \xi$$

for all $\xi \in \mathbb{R}$ and we recover definition (2.26a).

At the very beginning, $H(t,\xi)$ was introduced as the energy contained in a strip between $-\infty$ and $y(t,\xi)$, see (2.10). This interpretation still holds. We obtain μ, the energy density in Eulerian coordinates, by pushing forward by y the energy density in Lagrangian coordinates, $H_\xi\,d\xi$. Recall that the push-forward of a measure ν by a measurable function f is the measure $f_\#\nu$ defined as

$$f_\#\nu(B) = \nu(f^{-1}(B))$$

for all Borel sets B. We are led to the map M which transforms Lagrangian coordinates into Eulerian coordinates and whose definition is contained in the following theorem.

Theorem 3.10. [16, Theorem 3.11] *Given any element $[X]$ in $\mathcal{F}/G$. Then, (u,μ) defined as follows*

$$u(x) = U(\xi) \text{ for any } \xi \text{ such that } x = y(\xi), \tag{3.8a}$$

$$\mu = y_\#(H_\xi\,d\xi) \tag{3.8b}$$

belongs to $\mathcal{D}$ and is independent of the representative $X = (y,U,H) \in \mathcal{F}$ we choose for $[X]$. We denote by $M\colon \mathcal{F}/G \to \mathcal{D}$ the map which to any $[X]$ in $\mathcal{F}/G$ associates (u,μ) as given by (3.8).

Of course, the definition of u coincides with the one given previously in (2.61). The transformation from Eulerian to Lagrangian coordinates is a bijection, as stated in the next theorem.

Theorem 3.11. [16, Theorem 3.12] *The map M and L are invertible. We have*

$$L \circ M = \mathrm{Id}_{\mathcal{F}/G} \text{ and } M \circ L = \mathrm{Id}_{\mathcal{D}}.$$

3.2. **Continuous semigroup of solutions on $\mathcal{D}$.** On $\mathcal{D}$ we define the distance $d_{\mathcal{D}}$ which makes the bijection L between $\mathcal{D}$ and $\mathcal{F}/G$ into an isometry:

$$d_{\mathcal{D}}((u,\mu),(\bar{u},\bar{\mu})) = d_{\mathcal{F}/G}(L(u,\mu),L(\bar{u},\bar{\mu})).$$

Since $\mathcal{F}/G$ equipped with $d_{\mathcal{F}/G}$ is a complete metric space, we have the following theorem.

Theorem 3.12. *$\mathcal{D}$ equipped with the metric d_D is a complete metric space.*

For each $t \in \mathbb{R}$, we define the map T_t from $\mathcal{D} \times \mathcal{E}$ to $\mathcal{D}$ as

$$T_t(\,\cdot\,,f,g) = M\tilde{S}_t(\,\cdot\,,f,g)L,$$

for any $(f,g) \in \mathcal{E}$. For a given pair $(f,g) \in \mathcal{E}$, we have the following commutative diagram:

$$\begin{tikzcd} \mathcal{D} & \mathcal{F}/G \arrow[l, "M"'] \\ \mathcal{D} \arrow[u, "T_t"] \arrow[r, "L"] & \mathcal{F}/G \arrow[u, "\tilde{S}_t"'] \end{tikzcd} \tag{3.9}$$

Our main theorem reads as follows.

Theorem 3.13. *Assume* (1.6)*. $T\colon \mathcal{D} \times \mathcal{E} \times \mathbb{R}_+ \to \mathcal{D}$ (where $\mathcal{D}$ is defined by Definition* 3.7*) defines a continuous semigroup of solutions of* (1.5)*, that is, given $(\bar{u},\bar{\mu}) \in \mathcal{D}$, if we denote $t \mapsto (u(t),\mu(t)) = T_t(\bar{u},\bar{\mu})$ the corresponding trajectory, then u is a weak solution of* (1.5)*. Moreover μ is a weak solution of the following transport equation for the energy density*

$$\mu_t + (u\mu)_x = (G(u) - 2Pu)_x. \tag{3.10}$$

The map T is continuous with respect to all the variables, on any bounded set of $\mathcal{E}$. Furthermore, we have that

$$\mu(t)(\mathbb{R}) = \mu(0)(\mathbb{R}) \text{ for all } t \tag{3.11}$$

and

$$\mu(t)(\mathbb{R}) = \mu_{\mathrm{ac}}(t)(\mathbb{R}) = \|u(t)\|_{H^1}^2 = \mu(0)(\mathbb{R}) \text{ for almost all } t. \tag{3.12}$$

Remark 3.14. We denote the unique solution described in the theorem as a *conservative* weak solution of (1.5).

Proof. From (3.12), it follows that $u \in L^\infty(\mathbb{R}_+, H^1(\mathbb{R}))$. The function u is a weak solution of (1.5) if it satisfies (2.66) and μ is a weak solution of (3.10) if

$$\int_{\mathbb{R}_+\times\mathbb{R}} (\phi_t + u\phi_x)(t,x)\,\mu(t,dx)dt = \int_{\mathbb{R}_+\times\mathbb{R}} ((G(u) - 2Pu)\phi_x)(t,x)\,dtdx \quad (3.13)$$

for all $\phi \in C^\infty(\mathbb{R}_+ \times \mathbb{R})$ with compact support. We already proved in Theorem 2.9 that $u(t)$ satisfies (2.66). We proceed the same way to prove that μ satisfies (3.13). We recall the proof of (3.11) and (3.12), which is the same as in [16]. From (3.8a), we obtain

$$\mu(t)(\mathbb{R}) = \int_{\mathbb{R}} H_\xi\,d\xi = H(t,\infty)$$

which is constant in time, see Lemma 2.7 (*iii*). Hence, (3.11) is proved. We know from Lemma 2.7 (*ii*) that, for $t \in \mathcal{K}^c$, $y_\xi(t,\xi) > 0$ for almost every $\xi \in \mathbb{R}$ (see (2.41) for the definition of $\mathcal{K}$, in particular, we have $\operatorname{meas}(\mathcal{K}) = 0$). Given $t \in \mathcal{K}^c$ (the time variable is suppressed in the notation when there is no ambiguity), we have, for any Borel set B,

$$\mu(t)(B) = \int_{y^{-1}(B)} H_\xi\,d\xi = \int_{y^{-1}(B)} \left(U^2 + \frac{U_\xi^2}{y_\xi^2}\right) y_\xi\,d\xi \quad (3.14)$$

from (2.27c). Since y is one-to-one when $t \in \mathcal{K}^c$ and $u_x \circ y y_\xi = U_\xi$ almost everywhere, we obtain from (3.14) that

$$\mu(t)(B) = \int_B (u^2 + u_x^2)(t,x)\,dx,$$

which, as $\operatorname{meas}(\mathcal{K}) = 0$, proves (3.12). □

3.3. **The topology on $\mathcal{D}$.** The metric $d_{\mathcal{D}}$ gives to $\mathcal{D}$ the structure of a complete metric space while it makes continuous the semigroup T_t of conservative solutions for the Camassa–Holm equation as defined in Theorem 3.13. In that respect, it is a suitable metric for the equation. However, as the definition of $d_{\mathcal{D}}$ is not straightforward, this metric is not so easy to manipulate. That is why we recall the results obtained in [16] where we compare the topology induced by $d_{\mathcal{D}}$ with more standard topologies. We have that convergence in $H^1(\mathbb{R})$ implies convergence in $(\mathcal{D}, d_{\mathcal{D}})$, which itself implies convergence in $L^\infty(\mathbb{R})$. More precisely, we have the following result.

Proposition 3.15. [16, Proposition 5.1] *The map*

$$u \mapsto (u, (u^2 + u_x^2)dx)$$

is continuous from $H^1(\mathbb{R})$ into $\mathcal{D}$. In other words, given a sequence $u_n \in H^1(\mathbb{R})$ converging to u in $H^1(\mathbb{R})$, then $(u_n, (u_n^2 + u_{nx}^2)dx)$ converges to $(u, (u^2 + u_x^2)dx)$ in $\mathcal{D}$.

Proposition 3.16. [16, Proposition 5.2] *Let (u_n, μ_n) be a sequence in $\mathcal{D}$ that converges to (u, μ) in $\mathcal{D}$. Then*

$$u_n \to u \text{ in } L^\infty(\mathbb{R}) \quad \text{and} \quad \mu_n \overset{*}{\rightharpoonup} \mu.$$

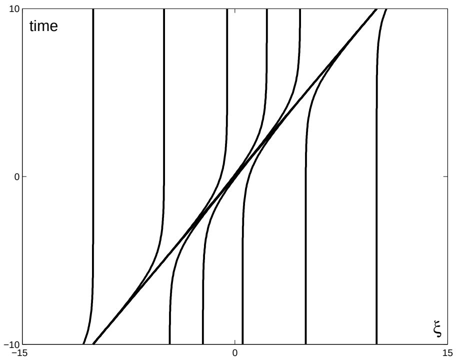


FIGURE 1. Characteristics in the single peakon case.

4. EXAMPLES

We include two examples for the Camassa–Holm equation where $f(u) = \frac{1}{2}u^2$ and $g(u) = u^2$.
(i) For initial data $\bar{u}(x) = ce^{-|x|}$, we have

$$u(t,x) = ce^{-|x-ct|}, \tag{4.1}$$

which is the familiar one peakon solution of the Camassa–Holm equation. The characteristics are the solutions of

$$y_t(t,\xi) = u(t, y(t,\xi)), \tag{4.2}$$

which can be integrated and, for initial data $\bar{y}(\xi) = \xi$, yields

$$y(t,\xi) = \operatorname{sgn}(\xi)\ln\big(e^{(\operatorname{sgn}(\xi)ct)} + e^{|\xi|} - 1\big).$$

Some characteristics are plotted in Figure 1. We have $U(t,\xi) = u(t, y(t,\xi)) = ce^{-|y(t,\xi)-ct|}$. It is easily checked that $y_\xi > 0$ almost everywhere. In this case y is invertible, there is no concentration of energy on a singular set, and we have

$$H(t,\xi) = \int_{y^{-1}((-\infty, y(t,\xi)))} H_\xi(\eta)\, d\eta = \int_{-\infty}^{y(t,\xi)} (u^2 + u_x^2)\, dx,$$

from (3.14).

(ii) The case with a peakon–antipeakon collision for the Camassa–Holm equation is considerably more complicated. In [17], we prove that the structure of the multipeakons is preserved, even through collisions. In particular, for an n-peakon u, it means that for almost all time the solution $u(t,x)$ can be written as

$$u(t,x) = \sum_{i=1}^{n} p_i(t)e^{-|x-q_i(t)|} \tag{4.3}$$

for some functions p_i and q_i that satisfy a system of ordinary differential equation that however experiences singularities at collisions. In [17] we also present a

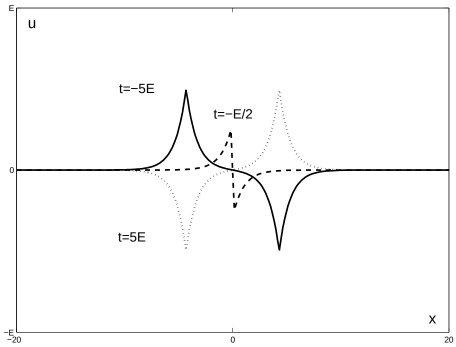


FIGURE 2. The colliding peakons case. Plot of the solution at different times.

system of ordinary differential equation satisfied by $y(t,\xi_i)$, $U(t,\xi_i)$ and $H(t,\xi_i)$ with $i=1,\dots,n$ where $y(t,\xi_i)$ and $U(t,\xi_i)$ correspond to the position and the height of the ith peak, respectively, while $H(t,\xi_i)$ represents the energy contained between $-\infty$ and the ith peak. In the antisymmetric case, this system can be solved explicitly, see [17], and we obtain

$$\begin{aligned} y(t,\xi_2) &= -y(t,\xi_1) = \ln(\cosh(\frac{Et}{2})),\\ U(t,\xi_2) &= -U(t,\xi_1) = \frac{E}{2}\tanh(\frac{Et}{2}),\\ H(t,\xi_2)-H(t,\xi_1) &= -\frac{E^2}{2}\tanh^2(\frac{Et}{2})+E^2. \end{aligned} \tag{4.4}$$

The initial conditions were chosen so that the two peaks collide at time $t=0$. From (4.3) and (4.4), we obtain

$$u(t,x) = \begin{cases} -\frac{E}{2}\sinh(\frac{Et}{2})e^{x}, & \text{for } x < -\ln\big(\cosh(\frac{Et}{2})\big),\\ \frac{E\sinh(x)}{\sinh(\frac{Et}{2})}, & \text{for } |x| < \ln\big(\cosh(\frac{Et}{2})\big),\\ \frac{E}{2}\sinh(\frac{Et}{2})e^{-x}, & \text{for } x > \ln\big(\cosh(\frac{Et}{2})\big). \end{cases} \tag{4.5}$$

See Figure 2. The formula holds for all $x\in\mathbb{R}$ and t nonzero. For $t=0$ we find formally $u(0,x)=0$. Here E denotes the total energy of the system, i.e.,

$$H(t,\infty) = \int_{\mathbb{R}}(u^2+u_x^2)\,dx = E^2, \quad t\neq 0. \tag{4.6}$$

For all $t\neq 0$ we find

$$\mu((-\infty,y)) = \mu_{\mathrm{ac}}((-\infty,y)) = \int_{-\infty}^{y}(u^2+u_x^2)\,dx. \tag{4.7}$$

For $t=0$, $H(0,\xi_2)-H(0,\xi_1)=E^2$, all the energy accumulates at the origin, and we find

$$\mu(x) = E^2\delta(x)\,dx, \quad \mu_{\mathrm{ac}}(x) = 0. \tag{4.8}$$

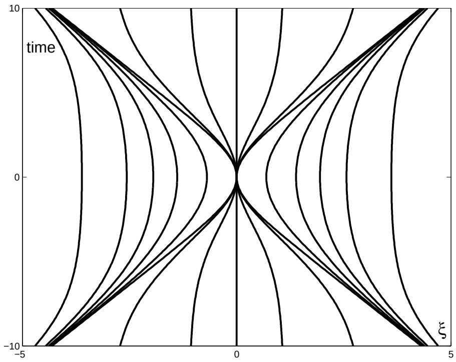


FIGURE 3. Characteristics in the colliding peakons case.

The function $u(t, x)$ is no longer Lipschitz in x, and (4.2) does not necessarily admit a unique solution. Indeed, given $T > 0$ and x_0 such that $|x_0| < \ln\big(\cosh(\frac{-TE}{2})\big)$, the characteristic arising from $(x_0, -T)$ can be continued past the origin by any characteristic that goes through (x, T) where x satisfies $|x| < \ln\big(\cosh(\frac{TE}{2})\big)$, and still be a solution of (4.2). However by taking into account the energy, the system (2.15) selects one characteristic, and in that sense the characteristics are uniquely defined. We can compute them analytically and obtain

$$y(t) = 2\tanh^{-1}\big(C\tanh^2(\frac{Et}{4})\big)$$

with $|C| < 1$, for the characteristics that collide, and

$$y(t) = \varepsilon\ln\big(C + \cosh(\frac{Et}{2})\big)$$

with $\varepsilon = \pm 1$, $C \geq 1$, for the others. See Figure 3.

Acknowledgments. The authors gratefully acknowledge the hospitality of the Mittag-Leffler Institute, Sweden, creating a great working environment for research, during the Fall of 2005.

References

[1] A. Bressan and A. Constantin. Global conservative solutions of the Camassa–Holm equation. *Preprint.* Submitted, 2005.
[2] A. Bressan and M. Fonte. An optimal transportation metric for solutions of the Camassa–Holm equation. *Methods Appl. Anal.*, to appear.
[3] H. Brezis. *Analyse Fonctionnelle.* Masson, Paris, 1983.
[4] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.
[5] R. Camassa, D. D. Holm, and J. Hyman. A new integrable shallow water equation. *Adv. Appl. Mech.*, 31:1–33, 1994.
[6] G. M. Coclite, and H. Holden, and K. H, Karlsen. Well-posedness for a parabolic-elliptic system. *Discrete Cont. Dynam. Systems* 13:659–682, 2005.

[7] G. M. Coclite, and H. Holden, and K. H, Karlsen. Global weak solutions to a generalized hyperelastic-rod wave equation. *SIAM J. Math. Anal.*, 37:1044–1069, 2005.
[8] A. Constantin and J. Escher. Global existence and blow-up for a shallow water equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 26(2):303–328, 1998.
[9] A. Constantin and L. Molinet. Global weak solutions for a shallow water equation. *Comm. Math. Phys.*, 211(1):45–61, 2000.
[10] H.-H. Dai. Exact travelling-wave solutions of an integrable equation arising in hyperelastic rods. *Wave Motion*, 28(4):367–381, 1998.
[11] H.-H. Dai. Model equations for nonlinear dispersive waves in a compressible Mooney–Rivlin rod. *Acta Mech.*, 127(1-4):193–207, 1998.
[12] H.-H. Dai and Y. Huo. Solitary shock waves and other travelling waves in a general compressible hyperelastic rod. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 456(1994):331–363, 2000.
[13] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
[14] G. B. Folland, Real Analysis, Wiley, New York, second edition, 1999.
[15] M. Fonte. Conservative solution of the Camassa Holm equation on the real line. *arXiv:math.AP/0511549*, 3, 2005.
[16] H. Holden and X. Raynaud. Global conservative solutions of the Camassa–Holm equation—a Lagrangian point of view. *Preprint*, 2006, submitted. **paper IV** in this thesis.
[17] H. Holden and X. Raynaud. Global conservative multipeakon solutions of the Camassa–Holm equation. *J. Hyperbolic Differ. Equ.* 2006, to appear. **paper V** in this thesis.
[18] R. S. Johnson. Camassa–Holm, Korteweg–de Vries and related models for water waves. *J. Fluid Mech.*, 455:63–82, 2002.
[19] J. Málek, J. Nečas, M. Rokyta, and M. Růžička. *Weak and Measure-Valued Solutions to Evolutionary PDEs*, Chapman & Hall, London, 1996.
[20] Z. Xin and P. Zhang. On the uniqueness and large time behavior of the weak solutions to a shallow water equation. *Comm. Partial Differential Equations*, 27:1815–1844, 2002.
[21] K. Yosida. *Functional Analysis.* Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.