Explicit multipeakon solutions of Novikov's cubically nonlinear integrable Camassa–Holm type equation

Andrew N. W. Hone* Hans Lundmark† Jacek Szmigielski‡

March 17, 2009

Abstract

Recently Vladimir Novikov found a new integrable analogue of the Camassa–Holm equation, admitting peaked soliton (*peakon*) solutions, which has nonlinear terms that are cubic, rather than quadratic. In this paper, the explicit formulas for multipeakon solutions of Novikov's cubically nonlinear equation are calculated, using the matrix Lax pair found by Hone and Wang. By a transformation of Liouville type, the associated spectral problem is related to a cubic string equation, which is dual to the cubic string that was previously found in the work of Lundmark and Szmigielski on the multipeakons of the Degasperis–Procesi equation.

1 Introduction

Integrable PDEs with nonsmooth solutions have attracted much attention in recent years, since the discovery of the Camassa–Holm shallow water wave equation and its peak-shaped soliton solutions called *peakons* [5]. Our purpose in this paper is to explicitly compute the multipeakon solutions of a new integrable PDE, equation (3.1) below, which is of the Camassa–Holm form $u_t - u_{xxt} = F(u, u_x, u_{xx}, \dots)$, but has cubically nonlinear terms instead of quadratic. This equation was found by Vladimir Novikov, and published in a recent paper by Hone and Wang [17].

We will apply inverse spectral methods. The spatial equation in the Lax pair for Novikov's equation turns out to be equivalent to what we call the *dual cubic string*, a spectral problem closely related to the *cubic string* that was used for finding the multipeakon solutions to the Degasperis–Procesi equation

*Institute of Mathematics, Statistics & Actuarial Science, University of Kent, Canterbury CT2 7NF, United Kingdom; anwh@kent.ac.uk

†Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden; halun@mai.liu.se

‡Department of Mathematics and Statistics, University of Saskatchewan, 106 Wiggins Road, Saskatoon, Saskatchewan, S7N 5E6, Canada; szmigiel@math.usask.ca

[25, 26, 21]. Once this relation is established, the Novikov peakon solution can be derived in a straightforward way using the results obtained in [21]. The constants of motion have a more complicated structure than in the Camassa–Holm and Degasperis–Procesi cases, and the study of this gives as an interesting by-product a combinatorial identity concerning the sum of all minors in a symmetric matrix, which we have dubbed the *Canada Day Theorem* (Theorem 4.1, proved in Appendix A).

The peakon problem for Novikov's equation presents in addition one important challenge. Unlike its Camassa–Holm or Degasperis–Procesi counterparts, the Lax pair for the Novikov equation is originally ill-defined in the peakon sector. The problem is caused by terms which involve multiplication of a singular measure by a discontinuous function. We prove in Appendix B that there exists a regularization of the Lax pair which preserves integrability of the peakon sector, thus allowing us to use spectral and inverse spectral methods to obtain the multipeakon solutions to the Novikov equation. This regularization problem has a subtle but nevertheless real impact on the formulas. In general, the use of Lax pairs to construct distributional solutions to nonlinear equations which are Lax integrable in the smooth sector but may not be so in the whole non-smooth sector is relatively uncharted territory, and the case of Novikov's equation may provide some relevant insight in this regard.

2 Background

The main example of a PDE admitting peaked solitons is the family

$$u_t - u_{xxt} + (b+1)uu_x = bu_x u_{xx} + uu_{xxx}, \tag{2.1}$$

often written as

$$m_t + m_x u + bmu_x = 0, \qquad m = u - u_{xx}, \tag{2.2}$$

which was introduced by Degasperis, Holm and Hone [8], and is Hamiltonian for all values of b [15]. It includes the Camassa–Holm equation as the case $b = 2$, and another integrable PDE called the Degasperis–Procesi equation [9, 8] as the case $b = 3$. These are the only values of b for which the equation is integrable, according to a variety of integrability tests [9, 28, 16, 18]. (However, we note that the case $b = 0$ is excluded from the aforementioned integrability tests; yet this case provides a regularization of the inviscid Burgers equation that is Hamiltonian and has classical solutions globally in time [4].) *Multipeakons* are weak solutions of the form

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

formed through superposition of n *peakons* (peaked solitons of the shape $e^{-|x|}$). This ansatz satisfies the PDE (2.2) if and only if the positions $(x_1, \ldots, x_n)$ and

momenta $(m_1, \ldots, m_n)$ of the peakons obey the following system of $2n$ ODEs:

$$\dot{x}_k = \sum_{i=1}^{n} m_i \, e^{-|x_k - x_i|}, \qquad \dot{m}_k = (b-1)\, m_k \sum_{i=1}^{n} m_i \, \operatorname{sgn}(x_k - x_i)\, e^{-|x_k - x_i|}. \tag{2.4}$$

Here, $\operatorname{sgn} x$ denotes the signum function, which is $+1$, -1 or 0 depending on whether x is positive, negative or zero. In shorthand notation, with $\langle f(x) \rangle$ denoting the average of the left and right limits,

$$\langle f(x) \rangle = \frac{1}{2}\big(f(x^-) + f(x^+)\big), \tag{2.5}$$

the ODEs can be written as

$$\dot{x}_k = u(x_k), \qquad \dot{m}_k = -(b-1)\, m_k \, \langle u_x(x_k) \rangle. \tag{2.6}$$

In the Camassa–Holm case $b = 2$, this is a canonical Hamiltonian system generated by $h = \frac{1}{2} \sum_{j,k=1}^{n} m_j \, m_k \, e^{-|x_j - x_k|}$. Explicit formulas for the n-peakon solution of the Camassa–Holm equation were derived by Beals, Sattinger and Szmigielski [1, 2] using inverse spectral methods, and the same thing for the Degasperis–Procesi equation was accomplished by Lundmark and Szmigielski [25, 26].

It requires some care to specify the exact sense in which the peakon solutions satisfy the PDE. The formulation (2.2) suffers from the problem that the product $m u_x$ is ill-defined in the peakon case, since the quantity $m = u - u_{xx} = 2 \sum_{i=1}^{n} m_i \, \delta_{x_i}$ is a discrete measure, and it is multiplied by a function u_x which has jump discontinuities exactly at the points x_k where the Dirac deltas in the measure m are situated. To avoid this problem, one can instead rewrite (2.1) as

$$(1 - \partial_x^2) u_t + (b + 1 - \partial_x^2)\, \partial_x \left(\tfrac{1}{2} u^2\right) + \partial_x \left(\tfrac{3-b}{2} u_x^2\right) = 0. \tag{2.7}$$

Then a function $u(x,t)$ is said to be a solution if

- $u(\cdot, t) \in W^{1,2}_{\mathrm{loc}}(\mathbf{R})$ for each fixed t, which means that $u(\cdot,t)^2$ and $u_x(\cdot,t)^2$ are locally integrable functions, and therefore define distributions of class $\mathcal{D}'(\mathbf{R})$ (i.e., continuous linear functionals acting on compactly supported C^∞ test functions on the real line $\mathbf{R}$),
- the time derivative $u_t(\cdot, t)$, defined as the limit of a difference quotient, exists as a distribution in $\mathcal{D}'(\mathbf{R})$ for all t,
- equation (2.7), with ∂_x taken to mean the usual distributional derivative, is satisfied for all t in the sense of distributions in $\mathcal{D}'(\mathbf{R})$.

It is worth mentioning that functions in the space $W^{1,2}_{\mathrm{loc}}(\mathbf{R})$ are continuous, by the Sobolev embedding theorem. However, the term u_x^2 is absent from equation (2.7) if $b = 3$, so in that particular case one requires only that $u(\cdot,t) \in L^2_{\mathrm{loc}}(\mathbf{R})$; this means that the Degasperis–Procesi can admit solutions u that are not continuous [6, 7, 24].

3 Novikov's equation

The new integrable equation found by Vladimir Novikov is

$$u_t - u_{xxt} + 4u^2 u_x = 3uu_x u_{xx} + u^2 u_{xxx}, \tag{3.1}$$

which can be written as

$$m_t + (m_x u + 3mu_x)\, u = 0, \qquad m = u - u_{xx}, \tag{3.2}$$

to highlight the similarity in form to the Degasperis–Procesi equation, or as

$$(1-\partial_x^2)u_t + (4-\partial_x^2)\,\partial_x\left(\tfrac{1}{3}\,u^3\right) + \partial_x\left(\tfrac{3}{2}\,uu_x^2\right) + \tfrac{1}{2}\,u_x^3 = 0 \tag{3.3}$$

in order to rigorously define weak solutions as above, except that here one requires that $u(\cdot,t) \in W^{1,3}_{\mathrm{loc}}(\mathbf{R})$ for all t, so that u^3 and u_x^3 are locally integrable and therefore define distributions in $\mathcal{D}'(\mathbf{R})$; it then follows from Hölder's inequality with the conjugate indices 3 and 3/2 that uu_x^2 is locally integrable as well, and (3.3) can thus be interpreted as a distributional equation. Since functions in $W^{1,p}_{\mathrm{loc}}(\mathbf{R})$ with $p \geq 1$ are automatically continuous, Novikov's equation is similar to the Camassa–Holm equation in that it only admits continuous distributional solutions (as opposed to the Degasperis–Procesi equation, which has discontinuous solutions as well).

Like the equations in the b-family (2.1), Novikov's equation admits (in the weak sense just defined) multipeakon solutions of the form (2.3), but in this case the ODEs for the positions and momenta are

$$\begin{aligned} \dot{x}_k &= u(x_k)^2 = \left(\sum_{i=1}^n m_i\, e^{-|x_k - x_i|}\right)^2, \\ \dot{m}_k &= -m_k\, u(x_k)\,\big\langle u_x(x_k)\big\rangle \\ &= m_k \left(\sum_{i=1}^n m_i\, e^{-|x_k-x_i|}\right)\left(\sum_{j=1}^n m_j\ \mathrm{sgn}(x_k - x_j)\, e^{-|x_k - x_j|}\right). \end{aligned} \tag{3.4}$$

These equations were stated in [17], where it was also shown that they constitute a Hamiltonian system $\dot{x}_k = \{x_k, h\}$, $\dot{m}_k = \{m_k, h\}$, generated by the same Hamiltonian $h = \frac{1}{2}\sum_{j,k=1}^n m_j m_k\, e^{-|x_j - x_k|}$ as the Camassa–Holm peakons, but with respect to a different, non-canonical, Poisson structure given by

$$\begin{aligned} \{x_j, x_k\} &= \mathrm{sgn}(x_j - x_k)\,\big(1 - E_{jk}^2\big), \\ \{x_j, m_k\} &= m_k E_{jk}^2, \\ \{m_j, m_k\} &= \mathrm{sgn}(x_j - x_k)\, m_j m_k E_{jk}^2, \qquad \text{where } E_{jk} = e^{-|x_j - x_k|}. \end{aligned} \tag{3.5}$$

As will be shown below, (3.4) is a Liouville integrable system (Theorem 4.7); in fact, it is even explicitly solvable in terms of elementary functions (Theorem 9.1).

4 Forward spectral problem

In order to integrate the Novikov peakon ODEs, we are going to make use of the matrix Lax pair found by Hone and Wang [17], specified by the following matrix linear system:

$$\frac{\partial}{\partial x}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix} = \begin{pmatrix}0 & zm & 1\\ 0 & 0 & zm\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix}, \tag{4.1}$$

$$\frac{\partial}{\partial t}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix} = \begin{pmatrix}-uu_x & u_x z^{-1} - u^2 mz & u_x^2\\ uz^{-1} & -z^{-2} & -u_x z^{-1} - u^2 mz\\ -u^2 & uz^{-1} & uu_x\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix}. \tag{4.2}$$

(Compared with reference [17] we have added a constant multiple of the identity to the matrix on the right hand side of (4.2), and used z in place of λ.) In the peakon case, when $u = \sum_{i=1}^n m_i\, e^{-|x-x_i|}$, the quantity $m = u - u_{xx} = 2\sum_{i=1}^n m_i\, \delta_{x_i}$ is a discrete measure. We assume that $x_1 < x_2 < \cdots < x_n$ (which at least remains true for a while if it is true for $t = 0$). These points divide the x axis into $n+1$ intervals which we number from 0 to n, so that the kth interval runs from x_k to x_{k+1}, with the convention that $x_0 = -\infty$ and $x_{n+1} = +\infty$. Since m vanishes between the point masses, equation (4.1) reduces to $\partial_x\psi_1 = \psi_3$, $\partial_x\psi_2 = 0$ and $\partial_x\psi_3 = \psi_1$ in each interval, so that in the kth interval we have

$$\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix} = \begin{pmatrix}A_k\, e^x + z^2\, C_k\, e^{-x}\\ 2z\, B_k\\ A_k\, e^x - z^2\, C_k\, e^{-x}\end{pmatrix} \qquad \text{for } x_k < x < x_{k+1}, \tag{4.3}$$

where the factors containing z have been inserted for later convenience. These piecewise solutions are then glued together at the points x_k. The proper interpretation of (4.1) at these points turns out to be that ψ_3 must be continuous, while ψ_1 and ψ_2 are allowed to have jump discontinuities; moreover, in the term $zm\psi_2$, one should take $\psi_2(x)\delta_{x_k}$ to mean $\langle\psi_2(x_k)\rangle\delta_{x_k}$. This point is fully explained in Appendix B. This leads to

$$\begin{aligned}\begin{pmatrix}A_k\\ B_k\\ C_k\end{pmatrix} &= \begin{pmatrix}1-\lambda\, m_k^2 & -2\lambda\, m_k\, e^{-x_k} & -\lambda^2\, m_k^2\, e^{-2x_k}\\ m_k\, e^{x_k} & 1 & \lambda\, m_k\, e^{-x_k}\\ m_k^2\, e^{2x_k} & 2\, m_k\, e^{x_k} & 1+\lambda m_k^2\end{pmatrix}\begin{pmatrix}A_{k-1}\\ B_{k-1}\\ C_{k-1}\end{pmatrix}\\ &=: S_k(\lambda)\begin{pmatrix}A_{k-1}\\ B_{k-1}\\ C_{k-1}\end{pmatrix}, \qquad \text{where } \lambda = -z^2.\end{aligned} \tag{4.4}$$

We impose the boundary condition $(A_0, B_0, C_0) = (1, 0, 0)$, which is consistent with the time evolution given by (4.2) for $x < x_1$. Then all (A_k, B_k, C_k) are determined by successive application of the jump matrices $S_k(\lambda)$ as in (4.4). For $x > x_n$, equation (4.2) implies that $(A, B, C) := (A_n, B_n, C_n)$ evolves as

$$\dot{A} = 0, \quad \dot{B} = \frac{B - AM_+}{\lambda}, \quad \dot{C} = \frac{2M_+\,(B - AM_+)}{\lambda}, \tag{4.5}$$

where $M_+ = \sum_{k=1}^N m_k\, e^{x_k}$. Thus A is invariant. It is the $(1,1)$ entry of the total jump matrix

$$S(\lambda) = S_n(\lambda)\dots S_2(\lambda)S_1(\lambda), \tag{4.6}$$

and therefore it is a polynomial in λ of degree n,

$$A(\lambda) = \sum_{k=0}^n H_k(-\lambda)^k = \left(1-\frac{\lambda}{\lambda_1}\right)\dots\left(1-\frac{\lambda}{\lambda_n}\right), \tag{4.7}$$

where $H_0 = 1$ (since $S(0) = I$, the identity matrix), and where the other coefficients $H_1,\dots,H_n$ are Poisson commuting constants of motion (see Theorems 4.2 and 4.7 below).

The first linear equation (4.1), together with the boundary conditions expressed by the requirements that $B_0 = C_0 = 0$ and $A_n(\lambda) = 0$, is a spectral problem which has the zeros $\lambda_1,\dots,\lambda_n$ of $A(\lambda)$ as its eigenvalues. (To be precise, one should perhaps say that it is the corresponding values of $z = \pm\sqrt{-\lambda}$ that are the eigenvalues, but we will soon show that the λ_k are positive, at least in the pure peakon case, and therefore more convenient to work with than the purely imaginary values of z; see (4.19) below.)

Elimination of ψ_1 from (4.1) gives $\partial_x\psi_2 = zm\psi_3$ and $(\partial_x^2 - 1)\,\psi_3 = zm\psi_2$, and the boundary conditions above imply that $(\psi_2,\psi_3) \to (0,0)$ as $x \to -\infty$ and $\psi_3 \to 0$ as $x \to +\infty$. Using the Green's function $-e^{-|x|}/2$ for the operator $\partial_x^2 - 1$ with vanishing boundary conditions, we can rephrase the problem as a system of integral equations,

$$\begin{aligned}\psi_2(x) &= z\int_{-\infty}^x \psi_3(y)\,dm(y),\\ \psi_3(x) &= -z\int_{-\infty}^\infty \frac{1}{2}\,e^{-|x-y|}\psi_2(y)\,dm(y),\end{aligned} \tag{4.8}$$

with integrals taken with respect to the discrete measure $m = 2\sum_{i=1}^n m_i\,\delta_{x_i}$. Here, there is again the problem of Dirac deltas multiplying a function ψ_2 with jump discontinuities, and we take $\psi_2(x)\delta_{x_k}$ to mean the average $\langle\psi_2(x_k)\rangle\delta_{x_k}$, in full agreement with the earlier definition of the singular term appearing in the spectral problem. Then

$$\begin{aligned}\langle\psi_2(x_j)\rangle &= z\left(2\sum_{k=1}^{j-1}\psi_3(x_k)\,m_k + \psi_3(x_j)\,m_j\right),\\ \psi_3(x_j) &= -z\sum_{k=1}^n e^{-|x_j-x_k|}\langle\psi_2(x_k)\rangle\,m_k,\end{aligned} \tag{4.9}$$

which can be written in block matrix notation as

$$\begin{pmatrix}\langle\Psi_2\rangle\\ \Psi_3\end{pmatrix} = z\begin{pmatrix}0 & TP\\ -EP & 0\end{pmatrix}\begin{pmatrix}\langle\Psi_2\rangle\\ \Psi_3\end{pmatrix}, \tag{4.10}$$

where

$$\begin{aligned}
\Psi_3 &= \big(\psi_3(x_1), \dots, \psi_3(x_n)\big)^t, \\
\big\langle \Psi_2 \big\rangle &= \big(\langle \psi_2(x_1)\rangle, \dots, \langle \psi_2(x_n)\rangle\big)^t, \\
P &= \operatorname{diag}(m_1, \dots, m_n), \\
E &= (E_{jk})_{j,k=1}^n, \qquad \text{where } E_{jk} = e^{-|x_j - x_k|}, \\
T &= (T_{jk})_{j,k=1}^n, \qquad \text{where } T_{jk} = 1 + \operatorname{sgn}(j-k).
\end{aligned} \tag{4.11}$$

(In words, T is the lower triangular $n \times n$ matrix that has 1 on the main diagonal and 2 everywhere below it.) In terms of $\big\langle \Psi_2 \big\rangle$ alone, we have

$$\big\langle \Psi_2 \big\rangle = -z^2 TPEP \big\langle \Psi_2 \big\rangle, \tag{4.12}$$

so the eigenvalues are given by $0 = \det(I + z^2 TPEP) = \det(I - \lambda TPEP)$, where of course I denotes the $n \times n$ identity matrix. Since the eigenvalues are the zeros of $A(\lambda)$, and since $A(0) = 1$, it follows that

$$A(\lambda) = \det(I - \lambda TPEP). \tag{4.13}$$

This gives us a fairly concrete representation of the constants of motion H_k, which by definition are the coefficients of $A(\lambda)$ (see (4.7)), and it can be made even more explicit thanks to the curious combinatorial result in Theorem 4.1. We remind the reader that a $k \times k$ *minor* of an $n \times n$ matrix X is, by definition, the determinant of a submatrix $X_{IJ} = (X_{ij})_{i \in I,\, J \in J}$ whose rows and columns are selected among those of X by two index sets $I, J \subseteq \{1, \dots, n\}$ with k elements each, and a *principal minor* is one for which $I = J$. Compare the result of the theorem with the well-known fact that the coefficient of s^k in $\det(I + sX)$ equals the sum of all *principal* $k \times k$ minors of X, regardless of whether X is symmetric or not.

Theorem 4.1 ("The Canada Day Theorem"). *Let the matrix T be defined as in* (4.11) *above. Then, for any symmetric $n \times n$ matrix X, the coefficient of s^k in the polynomial* $\det(I + sTX)$ *equals the sum of all $k \times k$ minors (principal and non-principal) of X.*

Proof. The proof is presented in Appendix A. It relies on the Cauchy–Binet formula, Lindström's Lemma, and some rather intricate dependencies among the minors of X due to the symmetry of the matrix. □

Theorem 4.1 is named after the date when we started trying to prove it: July 1, 2008, Canada's national day. (It turned out that the proof was more difficult than we expected, so we didn't finish it until a few days later.) Summarizing the results so far, we now have the following description of the constants of motion:

Theorem 4.2. *The Novikov peakon ODEs* (3.4) *admit n constants of motion $H_1, \dots, H_n$, where H_k equals the sum of all $k \times k$ minors (principal and non-principal) of the $n \times n$ symmetric matrix $PEP = (m_j m_k E_{jk})_{j,k=1}^n$. (See* (4.11) *for notation.)*

Proof. This follows at once from (4.7), (4.13), and Theorem 4.1. □

Example 4.3. The sum of all 1×1 minors of PEP is of course just the sum of all entries,

$$H_1 = \sum_{j,k=1}^{n} m_j m_k E_{jk} = \sum_{j,k=1}^{n} m_j m_k \, e^{-|x_j - x_k|}, \tag{4.14}$$

and the Hamiltonian of the peakon ODEs (3.4) is $h = \frac{1}{2} H_1$. Higher order minors of PEP are easily computed using Lindström's Lemma, as explained in Section A.3 in the appendix. In particular, the constant of motion of highest degree in the m_k is

$$H_n = \det(PEP) = \prod_{j=1}^{n-1} (1 - E_{j,j+1}^2) \prod_{j=1}^{n} m_j^2. \tag{4.15}$$

Example 4.4. Written out in full, the constants of motion in the case $n = 3$ are

$$\begin{aligned} H_1 &= m_1^2 + m_2^2 + m_3^2 + 2m_1 m_2 E_{12} + 2m_1 m_3 E_{13} + 2m_2 m_3 E_{23}, \\ H_2 &= (1 - E_{12}^2)\, m_1^2\, m_2^2 + (1 - E_{13}^2)\, m_1^2\, m_3^2 + (1 - E_{23}^2)\, m_2^2\, m_3^2 \\ &\quad + 2(E_{23} - E_{12}E_{13})\, m_1^2\, m_2\, m_3 + 2(E_{12} - E_{13}E_{23})\, m_1\, m_2\, m_3^2, \\ H_3 &= (1 - E_{12}^2)(1 - E_{23}^2)\, m_1^2\, m_2^2\, m_3^2. \end{aligned} \tag{4.16}$$

From now on we mainly restrict ourselves to the pure peakon case when $m_k > 0$ for all k (no antipeakons). Our first reason for this is that we can then use the positivity of H_1 and H_n to show global existence of peakon solutions.

Theorem 4.5. *Let*

$$\mathcal{P} = \{x_1 < \cdots < x_n, \ \textit{all } m_k > 0\} \tag{4.17}$$

be the phase space for the Novikov peakon system (3.4) *in the pure peakon case. If the initial data are in $\mathcal{P}$, then the solution $(\mathbf{x}(t), \mathbf{m}(t))$ exists for all $t \in \mathbf{R}$, and remains in $\mathcal{P}$.*

Proof. Local existence in $\mathcal{P}$ is automatic in view of the smoothness of the ODEs there. By (4.14) and (4.15), both H_1 and H_n are strictly positive on $\mathcal{P}$. Since $m_k^2 < H_1$, all m_k remain bounded from above. The positivity of H_n ensures that the m_k are bounded away from zero, and that the positions remain ordered. The velocities $\dot{x}_k$ are all bounded by $(\sum m_k)^2$, hence $0 < \dot{x}_k \leq C$ for some constant C, and the positions $x_k(t)$ are therefore finite for all $t \in \mathbf{R}$. Since neither x_k nor m_k can blow up in finite time, the solution exists globally in time. □

Remark 4.6. The peakon ODEs (3.4) are invariant under the transformation $(m_1, \ldots, m_n) \mapsto (-m_1, \ldots, -m_n)$, so the analogous result holds also when all m_k are negative.

Theorem 4.7. *The constants of motion $H_1, \ldots, H_n$ of Theorem 4.2 are functionally independent and commute with respect to the Poisson bracket* (3.5)*, so the Novikov peakon system* (3.4) *is Liouville integrable on the phase space $\mathcal{P}$.*

Proof. To prove functional independence, one should check that $J := dH_1 \wedge dH_2 \wedge \ldots \wedge dH_n$ does not vanish on any open set in $\mathcal{P}$. Since J is rational in the variables $\{m_k, e^{x_k}\}_{k=1}^n$, it vanishes identically if it vanishes on an open set, so it is sufficient to show that J is not identically zero. To see this, note that

$$H_k = e_k(m_1^2, \ldots, m_n^2) + O(E_{pq}), \tag{4.18}$$

where e_k denotes the kth elementary symmetric function in n variables, and $O(E_{pq})$ denotes terms involving exponentials of the positions x_j. It is well known that the first n elementary symmetric functions are independent (they provide a basis for symmetric functions of n variables [27]), and therefore the leading part of J (neglecting the $O(E_{pq})$ terms) does not vanish. Since the $O(E_{pq})$ terms can be made arbitrarily small by taking the x_k far apart, we see that there is a region in $\mathcal{P}$ where J does not vanish, and we are done.

To prove that the quantities H_k Poisson commute with respect to the bracket (3.5), it is convenient to adapt some arguments of Moser that he applied to the scattering of particles in the Toda lattice and the rational Calogero–Moser system [29]. The Poisson bracket of two constants of motion is itself a constant of motion, so $\{H_j, H_k\}$ is independent of time. Consider now this bracket at a fixed point $(\mathbf{x}^0, \mathbf{m}^0) := (x_1^0, x_2^0, \ldots, x_n^0, m_1^0, m_2^0, \ldots m_n^0) \in \mathcal{P}$ which we consider as an initial condition for the peakon flow $(\mathbf{x}(t), \mathbf{m}(t))$, which exists globally in time by Theorem 4.5. Theorem 9.4, which will be proved later without using what we are proving here, shows that the peakons scatter as $t \to -\infty$; more precisely, m_k^2 tends to $1/\lambda_k$, while the x_k move apart, so that the terms $O(E_{pq})$ tend to zero. (It should also be possible to prove these scattering properties directly from the peakon ODEs, along the lines of what was done for the Degasperis–Procesi equation in [26, Theorem 2.4], but we have not done that.) Thus, from (4.18), $\{H_j, H_k\}(\mathbf{x}^0, \mathbf{m}^0) = \{H_j, H_k\}(\mathbf{x}(t), \mathbf{m}(t)) = \lim_{t\to-\infty}\{H_j, H_k\}(\mathbf{x}(t), \mathbf{m}(t)) = \lim_{t\to-\infty}\{e_j, e_k\}(\mathbf{x}(t), \mathbf{m}(t))$. Now the Poisson brackets of these symmetric functions are given by linear combinations of the Poisson brackets $\{m_j, m_k\}$ with coefficients dependent only on the amplitudes. However, from (3.5) it is clear that $\{m_j, m_k\}(\mathbf{x}(t), \mathbf{m}(t)) = O(E_{pq}) \to 0$, from which it follows that $\{e_j, e_k\}(\mathbf{x}(t), \mathbf{m}(t)) \to 0$ as $t \to -\infty$, and hence $\{H_j, H_k\}(\mathbf{x}^0, \mathbf{m}^0) = 0$ as required. □

Remark 4.8. Since the vanishing of the Poisson bracket is a purely algebraic relation, the H_k Poisson commute at each point of R^{2n}, not just in the region $\mathcal{P}$.

The λ_k, which are defined as the zeros of $A(\lambda)$, are the eigenvalues of the inverse of the matrix $TPEP$, since $A(\lambda) = \det(I - \lambda TPEP)$. Another reason why we restrict our attention to the case with all $m_k > 0$ is that the matrix $TPEP$ can then be shown to be oscillatory (see Section A.2 in the appendix), which implies that its eigenvalues are positive and simple. Consequently, the

λ_k are also positive and simple, and for definiteness we will number them such that

$$0 < \lambda_1 < \cdots < \lambda_n. \tag{4.19}$$

(For another proof that the spectrum is positive and simple, see Theorem 6.1.)

Turning now to $B = S(\lambda)_{21}$ and $C = S(\lambda)_{31}$, we find from (4.6) and (4.4) that they are polynomials in λ of degree $n-1$, with $B(0) = M_+$ and $C(0) = M_+^2$, where $M_+ = \sum_{k=1}^N m_k\, e^{x_k}$ as before. This means that the two *Weyl functions*

$$\omega(\lambda) = -\frac{B(\lambda)}{A(\lambda)} \qquad \text{and} \qquad \zeta(\lambda) = -\frac{C(\lambda)}{2A(\lambda)} \tag{4.20}$$

are rational functions of order $O(1/\lambda)$ as $\lambda \to \infty$, having poles at the eigenvalues λ_k. Let b_k and c_k denote the residues,

$$\omega(\lambda) = \sum_{k=1}^n \frac{b_k}{\lambda - \lambda_k}, \qquad \zeta(\lambda) = \sum_{k=1}^n \frac{c_k}{\lambda - \lambda_k}. \tag{4.21}$$

The time evolution of (A, B, C), given by (4.5), translates into

$$\dot\omega(\lambda) = \frac{\omega(\lambda) - \omega(0)}{\lambda}, \qquad \dot\zeta(\lambda) = -\omega(0)\,\dot\omega(\lambda). \tag{4.22}$$

Comparing residues on both sides in (4.22) gives

$$\dot b_k = \frac{b_k}{\lambda_k}, \qquad \dot c_k = -\omega(0)\,\frac{b_k}{\lambda_k} = \sum_{m=1}^n \frac{b_m b_k}{\lambda_m \lambda_k}. \tag{4.23}$$

This at once implies $b_k(t) = b_k(0)\, e^{t/\lambda_k}$, and integrating $\dot c_k(\tau)$ from $\tau = -\infty$ (assuming that c_k vanishes there) to $\tau = t$ then gives

$$c_k = \sum_{m=1}^n \frac{b_m b_k}{\lambda_m + \lambda_k}. \tag{4.24}$$

A purely algebraic proof of this relation between the Weyl functions, not relying on time dependence and integration, will be given below; see Theorem 6.1. We note the identities $\sum_1^n c_k/\lambda_k = \frac{1}{2}(\sum_1^n b_k/\lambda_k)^2$ and $\sum_1^n \lambda_k c_k = \frac{1}{2}(\sum_1^n b_k)^2$.

The multipeakon solution is obtained as follows. The initial data $x_k(0)$, $m_k(0)$ (for $k = 1, \ldots, n$) determine initial spectral data $\lambda_k(0)$, $b_k(0)$, which after time t have evolved to $\lambda_k(t) = \lambda_k(0)$, $b_k(t) = b_k(0)\, e^{t/\lambda_k}$ (since the λ_k are the zeros of the time-invariant polynomial $A(\lambda)$, and since the b_k satisfy (4.23)). Solving the inverse spectral problem for these spectral data at time t gives the solution $x_k(t)$, $m_k(t)$. The remainder of the paper is devoted to this inverse spectral problem.

5 The dual cubic string

Just like for the Camassa–Holm and Degasperis–Procesi equations, some terms in the Lax pair's spatial equation (equation (4.1) in this case, repeated as (5.1) below) can be removed by a change of both dependent and independent variables. We refer to this as a Liouville transformation, since it is reminiscent of the transformation used for bringing a second-order Sturm–Liouville operator to a simple normal form. This simplification reveals an interesting connection between the Novikov equation and the Degasperis–Procesi equation, and allows us to solve the inverse spectral problem by making use of the tools developed in the study of the latter.

Theorem 5.1. *The spectral problem*

$$\frac{\partial}{\partial x}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix} = \begin{pmatrix}0 & zm(x) & 1\\ 0 & 0 & zm(x)\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\end{pmatrix} \tag{5.1}$$

on the real line $x \in \mathbf{R}$, with boundary conditions

$$\begin{aligned} \psi_2(x) &\to 0, \quad \text{as } x \to -\infty,\\ e^x\psi_3(x) &\to 0, \quad \text{as } x \to -\infty,\\ e^{-x}\psi_3(x) &\to 0, \quad \text{as } x \to +\infty, \end{aligned} \tag{5.2}$$

is equivalent (for $z \neq 0$), under the change of variables

$$\begin{aligned} y &= \tanh x,\\ \phi_1(y) &= \psi_1(x)\cosh x - \psi_3(x)\sinh x,\\ \phi_2(y) &= z\,\psi_2(x),\\ \phi_3(y) &= z^2\,\psi_3(x)/\cosh x,\\ g(y) &= m(x)\,\cosh^3 x,\\ \lambda &= -z^2, \end{aligned} \tag{5.3}$$

to the "dual cubic string" problem

$$\frac{\partial}{\partial y}\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix} = \begin{pmatrix}0 & g(y) & 0\\ 0 & 0 & g(y)\\ -\lambda & 0 & 0\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix} \tag{5.4}$$

on the finite interval $-1 < y < 1$, with boundary conditions

$$\phi_2(-1) = \phi_3(-1) = 0 \qquad \phi_3(1) = 0. \tag{5.5}$$

In the discrete case $m = 2\sum_{k=1}^n m_k\,\delta_{x_k}$, the relation between the measures m and g should be interpreted as

$$g(y) = \sum_{k=1}^n g_k\delta_{y_k}, \qquad y_k = \tanh x_k, \qquad g_k = 2\,m_k\cosh x_k. \tag{5.6}$$

Proof. Straightforward computation using the chain rule and, for the discrete case, $\delta_{x_k} = \frac{dy}{dx}(x_k)\,\delta_{y_k}$. □

Remark 5.2. The cubic string equation, which plays a crucial role in the derivation of the Degasperis–Procesi multipeakon solution [26], is

$$\partial_y^3 \phi = -\lambda g \phi, \tag{5.7}$$

which can be written as a system by letting $\Phi = (\phi_1, \phi_2, \phi_3) = (\phi, \phi_y, \phi_{yy})$:

$$\frac{\partial}{\partial y}\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix} = \begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ -\lambda g(y) & 0 & 0\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix}. \tag{5.8}$$

The duality between (5.4) and (5.8) manifests itself in the discrete case as an interchange of the roles of masses g_k and distances $l_k = y_{k+1} - y_k$; see Section 6. When the mass distribution is given by a continuous function $g(y) > 0$, the systems are instead related via the change of variables defined by

$$\frac{d\tilde{y}}{dy} = g(y) = \frac{1}{\tilde{g}(\tilde{y})}, \tag{5.9}$$

where y and $g(y)$ refer to the primal cubic string (5.8), and $\tilde{y}$ and $\tilde{g}(\tilde{y})$ to the dual cubic string (5.4) (or the other way around; the transformation (5.9) is obviously symmetric in y and $\tilde{y}$, so that the dual of the dual is the original cubic string again).

Remark 5.3. The concept of a dual string figures prominently in the work of Krein on the ordinary string equation $\partial_y^2 \phi = -\lambda g \phi$ (as opposed to the cubic string). For a comprehensive account of Krein's theory, see [10].

Remark 5.4. As a motivation for the transformation (5.3), we note that one can eliminate ψ_1 from (5.1), which gives $\partial_x \psi_2 = zm\psi_3$, $(\partial_x^2 - 1)\,\psi_3 = zm\psi_2$. From the study of Camassa–Holm peakons [2] it is known that the transformation $y = \tanh x$, $\phi(y) = \psi(x)/\cosh x$ takes the expression $(\partial_x^2 - 1)\,\psi$ to a multiple of ϕ_{yy}, so it is not far-fetched to try something similar on ψ_3 while leaving ψ_2 essentially unchanged.

From now on we concentrate on the discrete case. The Liouville transformation maps the piecewise defined (ψ_1, ψ_2, ψ_3) given by (4.3) to

$$\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\end{pmatrix} = \begin{pmatrix} A_k(\lambda) - \lambda\, C_k(\lambda)\\ -2\lambda\, B_k(\lambda)\\ -\lambda\, A_k(\lambda)\,(1+y) - \lambda^2\, C_k(\lambda)\,(1-y)\end{pmatrix} \quad \text{for } y_k < y < y_{k+1}. \tag{5.10}$$

The initial values $(A_0, B_0, C_0) = (1, 0, 0)$ thus correspond to $\Phi(-1;\lambda) = (1,0,0)^t$, where $\Phi(y;\lambda) = \big(\phi_1, \phi_2, \phi_3\big)^t$, and at the right endpoint $y = 1$ we have

$$\Phi(1;\lambda) = \begin{pmatrix} A_n(\lambda) - \lambda\, C_n(\lambda)\\ -2\lambda\, B_n(\lambda)\\ -2\lambda\, A_n(\lambda)\end{pmatrix}. \tag{5.11}$$

In particular, the condition $A_n(\lambda) = 0$ defining the spectrum corresponds to $\phi_3(1;\lambda) = 0$, except that the latter condition gives an additional eigenvalue $\lambda_0 = 0$ which is only present on the finite interval. (This is not a contradiction, since the Liouville transformation from the line to the interval is not invertible when $z = -\lambda^2 = 0$.)

The component ϕ_3 is continuous and piecewise linear, while ϕ_1 and ϕ_2 are piecewise constant with jumps at the points y_k where the measure g is supported. More precisely, at point mass number k we have

$$\begin{aligned} \phi_1(y_k^+) - \phi_1(y_k^-) &= g_k \langle \phi_2(y_k) \rangle, \\ \phi_2(y_k^+) - \phi_2(y_k^-) &= g_k\, \phi_3(y_k), \end{aligned} \tag{5.12}$$

and in interval number k, with length $l_k = y_{k+1} - y_k$,

$$\phi_3(y_{k+1}^-) - \phi_3(y_k^+) = l_k\, \partial_y \phi_3(y_k^+) = -\lambda\, l_k\, \phi_1(y_k^+). \tag{5.13}$$

In terms of the vector Φ these relations take the form

$$\Phi(y_k^+) = \begin{pmatrix} 1 & g_k & \frac{1}{2} g_k^2 \\ 0 & 1 & g_k \\ 0 & 0 & 1 \end{pmatrix} \Phi(y_k^-), \tag{5.14}$$

and

$$\Phi(y_{k+1}^-) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda l_k & 0 & 1 \end{pmatrix} \Phi(y_k^+), \tag{5.15}$$

respectively. If we introduce the notation

$$G(x,\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda x & 0 & 1 \end{pmatrix}, \qquad L(x) = \begin{pmatrix} 1 & x & \frac{1}{2}x^2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}, \tag{5.16}$$

it follows immediately that

$$\Phi(1;\lambda) = G(l_n,\lambda)\; L(g_n)\; \ldots\; G(l_2,\lambda)\; L(g_2)\; G(l_1,\lambda)\; L(g_1)\; G(l_0,\lambda) \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right). \tag{5.17}$$

We define the Weyl functions W and Z of the dual cubic string to be

$$W(\lambda) = -\frac{\phi_2(1;\lambda)}{\phi_3(1;\lambda)}, \qquad Z(\lambda) = -\frac{\phi_1(1;\lambda)}{\phi_3(1;\lambda)}. \tag{5.18}$$

It is clear from (5.11) that they are related to the Weyl functions ω and ζ previously defined on the real line (see (4.20)) as follows:

$$\begin{aligned} W(\lambda) &= -\frac{B_n(\lambda)}{A_n(\lambda)} = \omega(\lambda) = \sum_{k=1}^{n} \frac{b_k}{\lambda - \lambda_k}, \\ Z(\lambda) &= \frac{A_n(\lambda) - \lambda C_n(\lambda)}{2\lambda A_n(\lambda)} = \frac{1}{2\lambda} + \zeta(\lambda) = \frac{1}{2\lambda} + \sum_{k=1}^{n} \frac{c_k}{\lambda - \lambda_k}. \end{aligned} \tag{5.19}$$

6 Relation to the Neumann-like cubic string

Kohlenberg, Lundmark and Szmigielski [21] studied the discrete cubic string with Neumann-like boundary conditions. We will briefly recall some results from that paper, with notation and sign conventions slightly altered to suit our needs here. The spectral problem in question is

$$\begin{aligned} \phi_{yyy}(y) &= -\lambda g(y)\phi(y) \quad \text{for } y \in \mathbf{R}, \\ \phi_y(-\infty) = \phi_{yy}(-\infty) &= 0, \qquad \phi_{yy}(\infty) = 0, \end{aligned} \tag{6.1}$$

where $g = \sum_{k=0}^{n} g_k\, \delta_{y_k}$ is a discrete measure with $n+1$ point masses $g_0, \ldots, g_n$ at positions $y_0 < y_1 < \cdots < y_n$; between these points are n finite intervals of length $l_1, \ldots, l_n$ (where $l_k = y_k - y_{k-1}$). Since $\phi_{yyy} = 0$ away from the point masses, the boundary conditions can equally well be written as

$$\phi_y(y_0^-) = \phi_{yy}(y_0^-) = 0, \qquad \phi_{yy}(y_n^+) = 0.$$

Using the normalization $\phi(-\infty) = 1$ (or $\phi(y_0^-) = 1$) and the notation $\Phi = (\phi, \phi_y, \phi_{yy})^t$, one finds

$$\Phi(y_n^+;\lambda) = G(g_n,\lambda)\; L(l_n)\; \ldots\; G(g_2,\lambda)\; L(l_2)\; G(g_1,\lambda)\; L(l_1)\; G(g_0,\lambda) \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right), \tag{6.2}$$

with matrices G and L as in (5.16) above. Under the assumption that all $g_k > 0$, the zeros of $\phi_{yy}(y_n^+;\lambda)$, which constitute the spectrum, are

$$0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n,$$

and the Weyl functions are

$$\begin{aligned} W(\lambda) &= -\frac{\phi_y(y_n^+;\lambda)}{\phi_{yy}(y_n^+;\lambda)} = \sum_{k=1}^{n} \frac{b_k}{\lambda - \lambda_k}, \\ Z(\lambda) &= -\frac{\phi(y_n^+;\lambda)}{\phi_{yy}(y_n^+;\lambda)} = \frac{1}{\gamma\lambda} + \sum_{k=1}^{n} \frac{c_k}{\lambda - \lambda_k}, \qquad \gamma = \sum_{k=0}^{n} g_k, \end{aligned} \tag{6.3}$$

with all $b_k > 0$. They satisfy the identity

$$Z(\lambda) + Z(-\lambda) + W(\lambda)W(-\lambda) = 0, \tag{6.4}$$

from which it follows, by taking the residue at $\lambda = \lambda_k$, that

$$c_k = \sum_{m=1}^{n} \frac{b_m b_k}{\lambda_m + \lambda_k}. \tag{6.5}$$

Thus $Z(\lambda)$ is uniquely determined by the function $W(\lambda)$ and the constant γ.

Now note that (6.2) is exactly the same kind of relation as (5.17), except that the roles of g_k and l_k are interchanged, and the right endpoint is called $y = y_n^+$

instead of $y = 1$. The definitions of the Weyl functions (6.3) also correspond perfectly to the Weyl functions (5.18) for the dual cubic string. Therefore, all the results above are also true in the setting of the dual cubic string. The assumption that the n distances l_k and the $n+1$ point masses g_k are all positive for the Neumann cubic string corresponds of course to the requirement that the n point masses g_k and the $n+1$ distances l_k are positive for the dual cubic string. The constant $\gamma = \sum_{k=0}^{n} g_k$ in the term $1/\gamma\lambda$ in (6.3) corresponds to the constant 2 in the term $1/2\lambda$ in (5.19), since $\sum_{k=0}^{n} l_k = 2$ is the length of the interval $-1 < y < 1$. In summary:

Theorem 6.1. *Assume that all point masses g_k are positive. Then the discrete dual cubic string of Theorem 5.1 has nonnegative and simple spectrum, with eigenvalues $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n$, and its Weyl functions* (5.18) *have positive residues and satisfy* (6.4) *and* (6.5)*. In particular, the second Weyl function $Z(\lambda)$ is uniquely determined by the first Weyl function $W(\lambda)$.*

7 Inverse spectral problem

The inverse spectral problem for the discrete dual cubic string consists in recovering the positions and masses $\{y_k, g_k\}_{k=1}^{n}$ given the spectral data consisting of eigenvalues and residues $\{\lambda_k, b_k\}_{k=1}^{n}$ (or, equivalently, given the first Weyl function $W(\lambda)$). The corresponding problem for the Neumann-like cubic string was solved in [21], and we need only translate the results, as in Section 6. See also [26] for more information about inverse problems of this kind and [3] for the underlying theory of Cauchy biorthogonal polynomials.

To begin with, we state the result in terms of the bimoment determinants $\mathcal{D}_m^{(ab)}$ and $\mathcal{D}'_m$ defined below. Things will become more explicit in the next section (Corollary 8.4), where the determinants are expressed directly in terms of the λ_k and b_k.

Definition 7.1. Suppose μ is a measure on $\mathbf{R}_+$ (the positive part of the real line) such that its moments,

$$\beta_a = \int \kappa^a \, d\mu(\kappa), \tag{7.1}$$

and its bimoments with respect to the Cauchy kernel $K(x, y) = 1/(x + y)$,

$$I_{ab} = I_{ba} = \iint \frac{\kappa^a \, \lambda^b}{\kappa + \lambda} \, d\mu(\kappa) \, d\mu(\lambda), \tag{7.2}$$

are finite. For $m \geq 1$, let $\mathcal{D}_m^{(ab)}$ denote the determinant of the $m \times m$ bimoment matrix which starts with I_{ab} in the upper left corner:

$$\mathcal{D}_m^{(ab)} = \begin{vmatrix} I_{ab} & I_{a,b+1} & \cdots & I_{a,b+m-1} \\ I_{a+1,b} & I_{a+1,b+1} & \cdots & I_{a+1,b+m-1} \\ I_{a+2,b} & I_{a+2,b+1} & \cdots & I_{a+2,b+m-1} \\ \vdots & & & \vdots \\ I_{a+m-1,b} & I_{a+m-1,b+1} & \cdots & I_{a+m-1,b+m-1} \end{vmatrix} = \mathcal{D}_m^{(ba)}. \tag{7.3}$$

Let $\mathcal{D}_0^{(ab)} = 1$, and $\mathcal{D}_m^{(ab)} = 0$ for $m < 0$.

Similarly, for $m \geq 2$, let $\mathcal{D}'_m$ denote the $m \times m$ determinant

$$\mathcal{D}'_m = \begin{vmatrix} \beta_0 & I_{10} & I_{11} & \dots & I_{1,m-2} \\ \beta_1 & I_{20} & I_{21} & \dots & I_{2,m-2} \\ \beta_2 & I_{30} & I_{31} & \dots & I_{3,m-2} \\ \vdots & & & & \vdots \\ \beta_{m-1} & I_{m0} & I_{m1} & \dots & I_{m,m-2} \end{vmatrix}, \tag{7.4}$$

and define $\mathcal{D}'_1 = \beta_0$ and $\mathcal{D}'_m = 0$ for $m < 1$.

Theorem 7.2. *Given constants $0 < \lambda_1 < \cdots < \lambda_n$ and $b_1, \dots, b_n > 0$, define the spectral measure*

$$\mu = \sum_{i=1}^{n} b_i \, \delta_{\lambda_i}, \tag{7.5}$$

and let I_{ab} be its bimoments,

$$I_{ab} = \iint \frac{\kappa^a \, \lambda^b}{\kappa + \lambda} \, d\mu(\kappa) \, d\mu(\lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\lambda_i^a \lambda_j^b}{\lambda_i + \lambda_j} \, b_i b_j. \tag{7.6}$$

Then the unique discrete dual cubic string (with positive masses g_k) having the Weyl function

$$W(\lambda) = \sum_{k=1}^{n} \frac{b_k}{\lambda - \lambda_k} = \int \frac{d\mu(\kappa)}{\lambda - \kappa}$$

is given by

$$y_{k'} = \frac{\mathcal{D}_k^{(00)} - \frac{1}{2}\mathcal{D}_{k-1}^{(11)}}{\mathcal{D}_k^{(00)} + \frac{1}{2}\mathcal{D}_{k-1}^{(11)}}, \qquad g_{k'} = 2\frac{\mathcal{D}_k^{(00)} + \frac{1}{2}\mathcal{D}_{k-1}^{(11)}}{\mathcal{D}'_k}, \tag{7.7}$$

where $k' = n + 1 - k$ for $k = 0, \dots, n+1$. The distances between the masses are given by

$$l_{k'-1} = y_{k'} - y_{k'-1} = \frac{\left(\mathcal{D}_k^{(10)}\right)^2}{\left(\mathcal{D}_k^{(00)} + \frac{1}{2}\mathcal{D}_{k-1}^{(11)}\right)\left(\mathcal{D}_{k+1}^{(00)} + \frac{1}{2}\mathcal{D}_k^{(11)}\right)}. \tag{7.8}$$

Proof. For $0 \leq k \leq n$, let $a^{(2k+1)}(\lambda)$ be the product of the first $2k+1$ factors in (5.17),

$$a^{(2k+1)}(\lambda) = G(l_n, \lambda) \; L(g_n) \; G(l_{n-1}, \lambda) \; L(g_{n-1}) \; \dots \\ \dots \; G(l_{k'}, \lambda) \; L(g_{k'}) \; G(l_{k'-1}, \lambda), \tag{7.9}$$

where $k' = n+1-k$. By Lemma 4.1 and Theorem 4.2 in [21], the entries in the first column of $a = a^{(2k+1)}(\lambda)$,

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} =: \begin{pmatrix} \widehat{P} \\ P \\ Q \end{pmatrix},$$

satisfy what in [21] was called a "Type I" approximation problem. This means that $(\widehat{P}(\lambda), P(\lambda), Q(\lambda))$ are polynomials in λ of degree k, k, $k+1$, respectively, satisfying the normalization conditions

$$\widehat{P}(0) = 1, \qquad P(0) = 0, \qquad Q(0) = 0,$$

the approximation conditions

$$Q(\lambda)W(\lambda) + P(\lambda) = O(1), \qquad Q(\lambda)Z(\lambda) + \widehat{P}(\lambda) = O(\lambda^{-1}), \qquad \text{as } \lambda \to \infty,$$

and the symmetry condition

$$Q(\lambda)Z(-\lambda) - P(\lambda)W(-\lambda) - \widehat{P}(\lambda) = O(\lambda^{-k-1}), \qquad \text{as } \lambda \to \infty.$$

According to Theorem 4.15 in [21], this determines $(\widehat{P}, P, Q)$ uniquely; in particular, the coefficients of $a_{31}^{(2k+1)}(\lambda) = Q(\lambda) = \sum_{i=1}^{k+1} q_i \lambda^i$ are given by the nonsingular linear system

$$\begin{pmatrix} I_{00} + \frac{1}{2} & I_{01} & \cdots & I_{0k} \\ I_{10} & I_{11} & \cdots & I_{1k} \\ I_{20} & I_{21} & \cdots & I_{2k} \\ \vdots & & & \vdots \\ I_{k0} & I_{k1} & \cdots & I_{kk} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_{k+1} \end{pmatrix} = - \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{7.10}$$

From (7.9) one finds that

$$\begin{aligned} a_{31}^{(2k+1)}(\lambda) &= (-\lambda)(l_n + l_{n-1} + \cdots + l_{k'-1}) + \ldots \\ &\quad + (-\lambda)^{k+1} \left(\frac{g_n^2}{2} \frac{g_{n-1}^2}{2} \ldots \frac{g_{k'}^2}{2} l_n l_{n-1} \ldots l_{k'-1} \right), \end{aligned} \tag{7.11}$$

and the lowest and highest coefficients are then extracted from (7.10) using Cramer's rule:

$$\begin{aligned} -q_1 &= \frac{\mathcal{D}_k^{(11)}}{\mathcal{D}_{k+1}^{(00)} + \frac{1}{2}\mathcal{D}_k^{(11)}} = \sum_{j=k'-1}^{n} l_j = 1 - y_{k'-1}, \\ (-1)^{k+1} q_{k+1} &= \frac{\mathcal{D}_k^{(10)}}{\mathcal{D}_{k+1}^{(00)} + \frac{1}{2}\mathcal{D}_k^{(11)}} = \left(\prod_{j=k'}^{n} \frac{g_j^2 \, l_j}{2} \right) l_{k'-1}. \end{aligned} \tag{7.12}$$

The first equation gives a formula for $y_{k'-1}$ right away, and of course also for $y_{k'}$ (with $1 \le k \le n+1$) after renumbering. This formula (7.7) for $y_{k'}$ holds also

for $k=0$, since it gives $y_{0'}=y_{n+1}=+1$ because of the way $\mathcal{D}_m^{(ab)}$ is defined for $m\le 0$. (That it indeed gives $y_{(n+1)'}=y_0=-1$ when $k=n+1$ is not as obvious; this depends on $\mathcal{D}_{n+1}^{(00)}$ being zero when the measure μ is supported on only n points. See [21, Appendix B].) Subtraction gives a formula for $l_{k'-1}$ which simplifies to (7.8) with the help of "Lewis Carroll's identity" [22, Prop. 10] applied to the determinant $\mathcal{D}_{k+1}^{(00)}$:

$$\mathcal{D}_{k+1}^{(00)}\mathcal{D}_{k-1}^{(11)}=\mathcal{D}_k^{(00)}\mathcal{D}_k^{(11)}-\mathcal{D}_k^{(10)}\mathcal{D}_k^{(01)}. \tag{7.13}$$

Finally, the second formula in (7.12), divided by the corresponding formula with k replaced by $k-1$, gives an expression for $\frac{1}{2}\,g_{k'}^2\,l_{k'-1}$ from which one obtains

$$g_{k'}=\Big(\mathcal{D}_k^{(00)}+\frac{1}{2}\mathcal{D}_{k-1}^{(11)}\Big)\sqrt{\frac{2}{\mathcal{D}_k^{(10)}\mathcal{D}_{k-1}^{(10)}}}.$$

The formula for $g_{k'}$ presented in (7.7) now follows from the identity $(\mathcal{D}_k')^2=2\mathcal{D}_k^{(10)}\mathcal{D}_{k-1}^{(10)}$ and the positivity of $\mathcal{D}_k'$, which are immediate consequences of (8.6) below. (The determinant identity can also be proved directly by expanding $\mathcal{D}_k'$ along the first column, squaring, and using $\beta_i\beta_j=I_{i+1,j}+I_{i,j+1}$.) □

Remark 7.3. We take this opportunity to correct a couple of mistakes in [21]: the formula in Corollary 4.17 should read $[Q_{3k+2}]=(-1)^{k+1}\mathcal{D}_k/\mathcal{A}_{k+1}$, and consequently it should be $m_{n-k}=\dfrac{\mathcal{D}_k^2}{2\mathcal{A}_{k+1}\mathcal{A}_k}$ in (4.54).

8 Evaluation of bimoment determinants

The aim of this section is just to state some formulas for the bimoment determinants $\mathcal{D}_m^{(ab)}$ and $\mathcal{D}_m'$, taken from [26, Lemma 4.10] and [21, Appendix B]. Quite a lot of notation is needed.

Definition 8.1. For $k\ge 1$, let

$$\begin{aligned}
t_k&=\frac{1}{k!}\int_{\mathbf{R}^k}\frac{\Delta(x)^2}{\Gamma(x)}\,\frac{d\mu^k(x)}{x_1x_2\dots x_k},\\
u_k&=\frac{1}{k!}\int_{\mathbf{R}^k}\frac{\Delta(x)^2}{\Gamma(x)}d\mu^k(x),\\
v_k&=\frac{1}{k!}\int_{\mathbf{R}^k}\frac{\Delta(x)^2}{\Gamma(x)}\,x_1x_2\dots x_k\,d\mu^k(x),
\end{aligned} \tag{8.1}$$

where

$$\begin{aligned}
\Delta(x)&=\Delta(x_1,\dots,x_k)=\prod_{i<j}(x_i-x_j),\\
\Gamma(x)&=\Gamma(x_1,\dots,x_k)=\prod_{i<j}(x_i+x_j).
\end{aligned} \tag{8.2}$$

(When $k = 0$ or 1, let $\Delta(x) = \Gamma(x) = 1$.) Also let $t_0 = u_0 = v_0 = 1$, and $t_k = u_k = v_k = 0$ for $k < 0$.

When $\mu = \sum_{k=1}^{n} b_k \, \delta_{\lambda_k}$, the integrals t_k, u_k, v_k reduce to the sums T_k, U_k, V_k below.

Definition 8.2. For $k \geq 0$, let $\binom{[1,n]}{k}$ denote the set of k-element subsets $I = \{i_1 < \cdots < i_k\}$ of the integer interval $[1, n] = \{1, \ldots, n\}$. For $I \in \binom{[1,n]}{k}$, let

$$\Delta_I = \Delta(\lambda_{i_1}, \ldots, \lambda_{i_k}), \qquad \Gamma_I = \Gamma(\lambda_{i_1}, \ldots, \lambda_{i_k}), \tag{8.3}$$

with the special cases $\Delta_\emptyset = \Gamma_\emptyset = \Delta_{\{i\}} = \Gamma_{\{i\}} = 1$. Furthermore, let

$$\lambda_I = \prod_{i \in I} \lambda_i, \qquad b_I = \prod_{i \in I} b_i,$$

with $\lambda_\emptyset = b_\emptyset = 1$. Using the abbreviation $\Psi_I = \dfrac{\Delta_I^2}{\Gamma_I}$, let

$$T_k = \sum_{I \in \binom{[1,n]}{k}} \frac{\Psi_I b_I}{\lambda_I}, \qquad U_k = \sum_{I \in \binom{[1,n]}{k}} \Psi_I b_I, \qquad V_k = \sum_{I \in \binom{[1,n]}{k}} \Psi_I \lambda_I b_I, \tag{8.4}$$

and

$$\begin{aligned} W_k &= \begin{vmatrix} U_k & V_{k-1} \\ U_{k+1} & V_k \end{vmatrix} = U_k V_k - U_{k+1} V_{k-1}, \\ Z_k &= \begin{vmatrix} T_k & U_{k-1} \\ T_{k+1} & U_k \end{vmatrix} = T_k U_k - T_{k+1} U_{k-1}. \end{aligned} \tag{8.5}$$

(To be explicit, $U_0 = V_0 = T_0 = 1$, and $U_k = V_k = T_k = 0$ for $k < 0$ or $k > n$.)

We can now finally state the promised formulas for the bimoment determinants.

Lemma 8.3. *For all m,*

$$\begin{aligned} \mathcal{D}_m^{(00)} &= \frac{\begin{vmatrix} t_m & u_{m-1} \\ t_{m+1} & u_m \end{vmatrix}}{2^m}, \quad \mathcal{D}_m^{(11)} = \frac{\begin{vmatrix} u_m & v_{m-1} \\ u_{m+1} & v_m \end{vmatrix}}{2^m}, \\ \mathcal{D}_m^{(10)} &= \frac{(u_m)^2}{2^m}, \qquad\qquad \mathcal{D}'_m = \frac{u_m u_{m-1}}{2^{m-1}}. \end{aligned} \tag{8.6}$$

In the discrete case when $\mu = \displaystyle\sum_{k=1}^{n} b_k \, \delta_{\lambda_k}$, this reduces to

$$\mathcal{D}_m^{(00)} = \frac{Z_m}{2^m}, \quad \mathcal{D}_m^{(11)} = \frac{W_m}{2^m}, \quad \mathcal{D}_m^{(10)} = \frac{(U_m)^2}{2^m}, \quad \mathcal{D}'_m = \frac{U_m U_{m-1}}{2^{m-1}}. \tag{8.7}$$

Corollary 8.4. *The solution to the inverse spectral problem for the discrete dual cubic string (Theorem 7.2) can be expressed as*

$$y_{k'} = \frac{Z_k - W_{k-1}}{Z_k + W_{k-1}}, \qquad g_{k'} = \frac{Z_k + W_{k-1}}{U_k U_{k-1}}, \tag{8.8}$$

$$l_{k'-1} = y_{k'} - y_{k'-1} = \frac{2\,(U_k)^4}{(Z_k + W_{k-1})(Z_{k+1} + W_k)}. \tag{8.9}$$

The expression W_k can be evaluated explicitly in terms of λ_k and b_k, although the formula is somewhat involved [26, Lemma 2.20]:

$$\begin{aligned} W_k = &\sum_{I \in \binom{[1,n]}{k}} \frac{\Delta_I^4}{\Gamma_I^2} \lambda_I b_I^2 \\ &+ \sum_{m=1}^{k} \sum_{\substack{I \in \binom{[1,n]}{k-m} \\ J \in \binom{[1,n]}{2m} \\ I \cap J = \emptyset}} b_I^2 b_J \left\{ 2^{m+1} \frac{\Delta_I^4 \Delta_{I,J}^2 \lambda_{I \cup J}}{\Gamma_I\, \Gamma_{I \cup J}} \left(\sum_{\substack{C \cup D = J \\ |C| = |D| = m \\ \min(C) < \min(D)}} \Delta_C^2 \Delta_D^2 \Gamma_C \Gamma_D \right) \right\}, \end{aligned} \tag{8.10}$$

where $\Delta_{I,J}^2 = \prod_{i \in I, j \in J} (\lambda_i - \lambda_j)^2$. The corresponding formula for Z_k is obtained by replacing b_i with b_i/λ_i everywhere.

9 The multipeakon solution

In order to obtain the solution to the inverse spectral problem on the real line, which provides the multipeakon solution, we merely have to map the formulas for the interval (Corollary 8.4) back to the line via the Liouville transformation (5.6).

We remind the reader that in this paper we primarily study the pure peakon case where it is assumed that all $m_k > 0$ and also that $x_1 < \cdots < x_n$. This assumption guarantees that the solutions are globally defined in time (Theorem 4.5) and, regarding the spectral data, that all $b_k > 0$ and $0 < \lambda_1 < \cdots < \lambda_n$ (Theorem 6.1). Details regarding mixed peakon-antipeakon solutions are left for future research, but we point out that since the velocity $\dot{x}_k = u(x_k)^2$ is always nonnegative, Novikov antipeakons move to the *right* just like peakons (unlike the b-family (2.1), where pure peakons move to the right and antipeakons to the left, if they are sufficiently far apart). Nevertheless, peakons and antipeakons may collide after finite time also for the Novikov equation, causing division by zero in the solution formula for m_k in (9.1) below, and this breakdown leads to the usual subtle questions regarding continuation of the solution beyond the collision.

Theorem 9.1. *In the notation of Section 8, the n-peakon solution of Novikov's equation is given by*

$$x_{k'} = \frac{1}{2}\ln\frac{Z_k}{W_{k-1}}, \qquad m_{k'} = \frac{\sqrt{Z_k W_{k-1}}}{U_k U_{k-1}}, \tag{9.1}$$

where $k' = n+1-k$ for $k = 1,\dots,n$, and where the time evolution is given by

$$b_k(t) = b_k(0)\, e^{t/\lambda_k}. \tag{9.2}$$

Proof. The inverse of the coordinate transformation (5.6) is

$$x_k = \frac{1}{2}\ln\frac{1+y_k}{1-y_k}, \qquad m_k = \frac{g_k\sqrt{1-y_k^2}}{2},$$

which upon inserting (8.8) gives (9.1) at once. The evolution of b_k comes from equation (4.23). □

Example 9.2. The two-peakon solution is

$$\begin{aligned}
x_1 &= \frac{1}{2}\ln\frac{Z_2}{W_1} = \frac{1}{2}\ln\frac{\dfrac{(\lambda_1-\lambda_2)^4}{(\lambda_1+\lambda_2)^2\lambda_1\lambda_2}\, b_1^2 b_2^2}{\lambda_1\, b_1^2 + \lambda_2\, b_2^2 + \dfrac{4\,\lambda_1\lambda_2}{\lambda_1+\lambda_2}\, b_1 b_2}, \\
x_2 &= \frac{1}{2}\ln\frac{Z_1}{W_0} = \frac{1}{2}\ln\left(\frac{b_1^2}{\lambda_1} + \frac{b_2^2}{\lambda_2} + \frac{4}{\lambda_1+\lambda_2}\, b_1 b_2\right), \\
m_1 &= \frac{\sqrt{Z_2 W_1}}{U_2 U_1} = \frac{\left[\dfrac{(\lambda_1-\lambda_2)^4\, b_1^2 b_2^2}{(\lambda_1+\lambda_2)^2\lambda_1\lambda_2}\left(\lambda_1\, b_1^2 + \lambda_2\, b_2^2 + \dfrac{4\,\lambda_1\lambda_2}{\lambda_1+\lambda_2}\, b_1 b_2\right)\right]^{1/2}}{\dfrac{(\lambda_1-\lambda_2)^2\, b_1 b_2}{\lambda_1+\lambda_2}\,(b_1+b_2)} \\
&= \frac{\left(\lambda_1\, b_1^2 + \lambda_2\, b_2^2 + \dfrac{4\,\lambda_1\lambda_2}{\lambda_1+\lambda_2}\, b_1 b_2\right)^{1/2}}{\sqrt{\lambda_1\lambda_2}\,(b_1+b_2)}, \\
m_2 &= \frac{\sqrt{Z_1 W_0}}{U_1 U_0} = \frac{\left(\dfrac{b_1^2}{\lambda_1} + \dfrac{b_2^2}{\lambda_2} + \dfrac{4}{\lambda_1+\lambda_2}\, b_1 b_2\right)^{1/2}}{b_1+b_2},
\end{aligned} \tag{9.3}$$

where the simpler of the two expressions for m_1 is obtained under the assumption that all spectral data are positive, and therefore only can be trusted in the pure peakon case. This way of writing the solution is simpler and more explicit than that found in [17]. In order to translate (9.3) to the notation used there, write (q_k, p_k) instead of (x_k, m_k), c_k instead of $1/\lambda_k$, and t_0 instead of $(\lambda_1^{-1} - \lambda_2^{-1})^{-1}\ln\frac{b_2(0)}{b_2(0)}$; then $\tanh T = (b_1-b_2)/(b_1+b_2)$ and $\cosh^{-2} T = 4b_1b_2/(b_1+b_2)^2$, where $T = \frac{1}{2}(c_1-c_2)(t-t_0)$.

Example 9.3. The three-peakon solution is

$$\begin{aligned} x_1 &= \frac{1}{2}\ln\frac{Z_3}{W_2}, \quad x_2 = \frac{1}{2}\ln\frac{Z_2}{W_1}, \quad x_3 = \frac{1}{2}\ln\frac{Z_1}{W_0}, \\ m_1 &= \frac{\sqrt{Z_3W_2}}{U_3U_2}, \quad m_2 = \frac{\sqrt{Z_2W_1}}{U_2U_1}, \quad m_3 = \frac{\sqrt{Z_1W_0}}{U_1U_0}, \end{aligned} \tag{9.4}$$

where $U_0 = W_0 = 1$,

$$\begin{aligned} U_1 &= b_1 + b_2 + b_3, \\ U_2 &= \Psi_{12}\, b_1b_2 + \Psi_{13}\, b_1b_3 + \Psi_{23}\, b_2b_3, \\ U_3 &= \Psi_{123}\, b_1b_2b_3, \end{aligned} \tag{9.5}$$

$$\begin{aligned} W_1 &= \lambda_1\, b_1^2 + \lambda_2\, b_2^2 + \lambda_3\, b_3^2 \\ &\quad + \frac{4\,\lambda_1\lambda_2}{\lambda_1+\lambda_2}\, b_1b_2 + \frac{4\,\lambda_1\lambda_3}{\lambda_1+\lambda_3}\, b_1b_3 + \frac{4\,\lambda_2\lambda_3}{\lambda_2+\lambda_3}\, b_2b_3, \\ W_2 &= \Psi_{12}^2\,\lambda_1\lambda_2\, b_1^2b_2^2 + \Psi_{13}^2\,\lambda_1\lambda_3\, b_1^2b_3^2 + \Psi_{23}^2\,\lambda_2\lambda_3\, b_2^2b_3^2 \\ &\quad + \frac{4\,\Psi_{13}\Psi_{23}\,\lambda_1\lambda_2\lambda_3}{\lambda_1+\lambda_2}\, b_1b_2b_3^2 + \frac{4\,\Psi_{12}\Psi_{23}\,\lambda_1\lambda_2\lambda_3}{\lambda_1+\lambda_3}\, b_1b_2^2b_3 \\ &\quad + \frac{4\,\Psi_{12}\Psi_{13}\,\lambda_1\lambda_2\lambda_3}{\lambda_2+\lambda_3}\, b_1^2b_2b_3, \end{aligned} \tag{9.6}$$

$$\begin{aligned} Z_1 &= \frac{b_1^2}{\lambda_1} + \frac{b_2^2}{\lambda_2} + \frac{b_3^2}{\lambda_3} + \frac{4}{\lambda_1+\lambda_2}\, b_1b_2 + \frac{4}{\lambda_1+\lambda_3}\, b_1b_3 + \frac{4}{\lambda_2+\lambda_3}\, b_2b_3, \\ Z_2 &= \frac{\Psi_{12}^2}{\lambda_1\lambda_2}\, b_1^2b_2^2 + \frac{\Psi_{13}^2}{\lambda_1\lambda_3}\, b_1^2b_3^2 + \frac{\Psi_{23}^2}{\lambda_2\lambda_3}\, b_2^2b_3^2 \\ &\quad + \frac{4\,\Psi_{13}\Psi_{23}}{(\lambda_1+\lambda_2)\lambda_3}\, b_1b_2b_3^2 + \frac{4\,\Psi_{12}\Psi_{23}}{(\lambda_1+\lambda_3)\lambda_2}\, b_1b_2^2b_3 + \frac{4\,\Psi_{12}\Psi_{13}}{(\lambda_2+\lambda_3)\lambda_1}\, b_1^2b_2b_3, \\ Z_3 &= \frac{\Psi_{123}^2}{\lambda_1\lambda_2\lambda_3}\, b_1^2b_2^2b_3^2, \end{aligned} \tag{9.7}$$

and

$$\begin{aligned} \Psi_{12} &= \frac{(\lambda_1-\lambda_2)^2}{\lambda_1+\lambda_2}, \quad \Psi_{13} = \frac{(\lambda_1-\lambda_3)^2}{\lambda_1+\lambda_3}, \quad \Psi_{23} = \frac{(\lambda_2-\lambda_3)^2}{\lambda_2+\lambda_3}, \\ \Psi_{123} &= \frac{(\lambda_1-\lambda_2)^2(\lambda_1-\lambda_3)^2(\lambda_2-\lambda_3)^2}{(\lambda_1+\lambda_2)(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)}. \end{aligned} \tag{9.8}$$

Theorem 9.4 (Asymptotics)**.** *Let the eigenvalues be numbered so that* $0 < \lambda_1 < \cdots < \lambda_n$*. Then*

$$\begin{aligned} x_k(t) &\sim \frac{t}{\lambda_k} + \log b_k(0) - \frac{1}{2}\ln\lambda_k + \sum_{i=k+1}^{n} \ln\frac{(\lambda_i-\lambda_k)^2}{(\lambda_i+\lambda_k)\lambda_i}, \quad \textit{as } t \to -\infty, \\ x_{k'}(t) &\sim \frac{t}{\lambda_k} + \log b_k(0) - \frac{1}{2}\ln\lambda_k + \sum_{i=1}^{k-1} \ln\frac{(\lambda_i-\lambda_k)^2}{(\lambda_i+\lambda_k)\lambda_i}, \quad \textit{as } t \to +\infty, \end{aligned} \tag{9.9}$$

where $k' = n+1-k$. Moreover,

$$\lim_{t\to-\infty} m_k(t) = \frac{1}{\sqrt{\lambda_k}} = \lim_{t\to+\infty} m_{k'}(t). \tag{9.10}$$

In words: asymptotically as $t \to \pm\infty$, the kth fastest peakon has velocity $1/\lambda_k$ and amplitude $1/\sqrt{\lambda_k}$.

Proof. This is just a matter of identifying the dominant terms; $b_1(t) = b_1(0)\, e^{t/\lambda_1}$ grows much faster as $t \to +\infty$ than $b_2(t)$, which in turn grows much faster than $b_3(t)$, etc., and as $t \to -\infty$ it is the other way around. Thus, for example, $U_k \sim \Psi_{12\ldots k}\, b_1 b_2 \ldots b_k$ as $t \to +\infty$. A similar analysis of W_k and Z_k leads quickly to the stated formulas. □

The only difference compared to the x_k asymptotics for Degasperis–Procesi peakons [26, Theorem 2.25] is that (9.9) contains an additional term $-\frac{1}{2}\ln\lambda_k$. Since this term cancels in the subtraction, the phase shifts for Novikov peakons are exactly the same as for Degasperis–Procesi peakons [26, Theorem 2.26]:

$$\lim_{t\to\infty}\left(x_{k'}(t) - \frac{t}{\lambda_k}\right) - \lim_{t\to-\infty}\left(x_k(t) - \frac{t}{\lambda_k}\right) = \\ = \sum_{i=1}^{k-1} \log\frac{(\lambda_i-\lambda_k)^2}{(\lambda_i+\lambda_k)\lambda_i} - \sum_{i=k+1}^{n} \log\frac{(\lambda_i-\lambda_k)^2}{(\lambda_i+\lambda_k)\lambda_i}. \tag{9.11}$$

A Combinatorial results

This appendix contains some material related to the combinatorial structure of the constants of motion $H_1, \ldots, H_n$ of the Novikov peakon ODEs; see Section 4, and in particular Theorem 4.2. Recall that

$$A(\lambda) = 1 - \lambda H_1 + \cdots + (-\lambda)^n H_n = \det(I - \lambda TPEP),$$

where I is the $n\times n$ identity matrix, and T, E, P are $n\times n$ matrices defined by $T_{jk} = 1 + \operatorname{sgn}(j-k)$, $E_{jk} = e^{-|x_j - x_k|}$, and $P = \operatorname{diag}(m_1, \ldots, m_n)$. The first thing to prove is that the matrix $TPEP$ is oscillatory if all $m_k > 0$, which shows that the zeros of $A(\lambda)$ are positive and simple. Then we show how to easily compute the minors of PEP, and finally we prove the "Canada Day Theorem" (Theorem 4.1) which implies that H_k equals the sum of all $k\times k$ minors of PEP.

A.1 Preliminaries

In this section we have collected some facts about total positivity [19, 13, 11] that will be used below.

Definition A.1. If X is a matrix and I and J are index sets, the submatrix $(X_{ij})_{i\in I, j\in J}$ will be denoted by X_{IJ} (or sometimes $X_{I,J}$). The set of k-element subsets of the integer interval $[1,n] = \{1,2,\ldots,n\}$ will be denoted $\binom{[1,n]}{k}$, and

elements of such a subset I will always be assumed to be numbered in ascending order $i_1 < \cdots < i_k$.

Definition A.2. A square matrix is said to be *totally positive* if all its minors of all orders are positive. It is called *totally nonnegative* if all its minors are nonnegative. A matrix is *oscillatory* if it is totally nonnegative and some power of it is totally positive.

Theorem A.3. *All eigenvalues of a totally positive matrix are positive and of algebraic multiplicity one, and likewise for oscillatory matrices. All eigenvalues of a totally nonnegative matrix are nonnegative, but in general of arbitrary multiplicity.*

Theorem A.4. *The product of an oscillatory matrix and a nonsingular totally nonnegative matrix is oscillatory.*

Definition A.5. A *planar network* (Γ, ω) of order n is an acyclic planar directed graph Γ with arrows going from left to right, with n sources (vertices with outgoing arrows only) on the left side, and with n sinks (vertices with incoming arrows only) on the right side. The sources and sinks are numbered 1 to n, from bottom to top, say. All other vertices have at least one arrow coming in and at least one arrow going out. Each edge e of the graph Γ is assigned a scalar weight $\omega(e)$. The *weight* of a directed path in Γ is the product of all the weights of the edges of that path. The *weighted path matrix* $\Omega(\Gamma, \omega)$ is the $n \times n$ matrix whose (i, j) entry Ω_{ij} is the sum of the weights of the possible paths from source i to sink j.

The following theorem was discovered by Lindström [23] and made famous by Gessel and Viennot [14]. A similar theorem also appeared earlier in the work of Karlin and McGregor on birth and death processes [20].

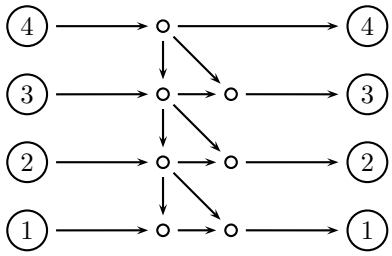
Theorem A.6 (Lindström's Lemma)**.** *Let I and J be subsets of $\{1, \ldots, n\}$ with the same cardinality. The minor $\det \Omega_{IJ}$ of the weighted path matrix $\Omega(\Gamma, \omega)$ of a planar network is equal to the sum of the weights of all possible families of nonintersecting paths (i.e., paths having no vertices in common) connecting the sources labelled by I to the sinks labelled by J. (The weight of a family of paths is defined as the product of the weights of the individual paths.)*

Corollary A.7. *If all weights $\omega(e)$ are nonnegative, then the weighted path matrix is totally nonnegative.*

Remark A.8. Beware that having *positive* weights does *not* in general imply total positivity of the path matrix Ω, since some minors $\det \Omega_{IJ}$ may be zero due to absence of nonintersecting path families from I to J, in which case Ω is only totally nonnegative.

A.2 Proof that $TPEP$ is oscillatory

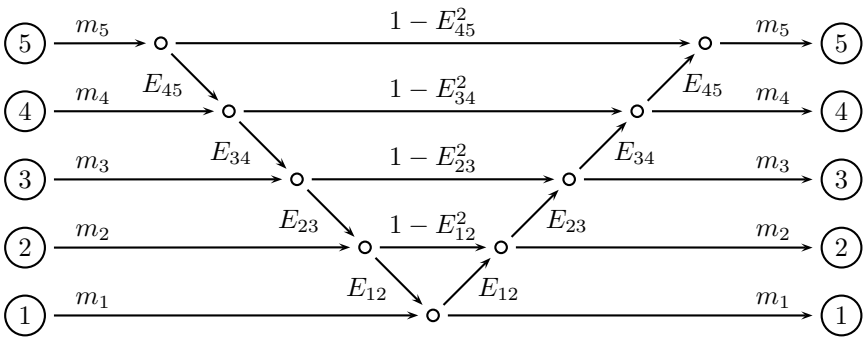
The matrix T is the path matrix of the planar network whose structure is illustrated below for the case $n = 4$ (with all edges, and therefore all paths and families of paths, having unit weight):



Indeed, there is clearly one path from source i to sink j if $i = j$, two paths if $i > j$, and none if $i < j$, and this agrees with

$$T_{ij} = 1 + \operatorname{sgn}(i - j) = \begin{cases} 1, & i = j, \\ 2, & i > j, \\ 0, & i < j. \end{cases}$$

Similarly one can check that the matrix PEP is the weighted path matrix of the planar network illustrated below for the case $n = 5$ (we are assuming that $x_1 < \cdots < x_n$, so that $E_{12}E_{23} = e^{x_1 - x_2} e^{x_2 - x_3} = E_{13}$, etc.):



By Corollary A.7, both T and PEP are totally nonnegative (if all $m_k > 0$). Furthermore, $(PEP)^N$ is the weighted path matrix of the planar network obtain by connecting N copies of the network for PEP in series, and if N is large enough, there is clearly enough wiggle room in this network to find a nonintersecting path family from any source set I to any sink set J with $|I| = |J|$. Thus $(PEP)^N$ is totally positive for sufficiently large N; in other words, PEP is oscillatory. (Another way to see this is to use a criterion [13, Chapter II, Theorem 10] which says that a totally nonnegative matrix X is oscillatory if and only if it is nonsingular and $X_{ij} > 0$ for $|i - j| = 1$.) Since T is nonsingular, Theorem A.4 implies that $TPEP$ is oscillatory, which was the first thing we wanted to prove.

A.3 Minors of PEP

Having a planar network for PEP makes it easy to compute its minors using Lindström's Lemma.

Example A.9. Consider the constant of motion H_3 in the case $n = 6$.

For sources $I = \{1, 2, 3\}$ and sinks $J = \{1, 2, 3\}$ there is only one family of nonintersecting paths, namely the paths going straight across. The weights of these paths are $m_1 m_1$, $m_2(1 - E_{12}^2)m_2$ and $m_3(1 - E_{23}^2)m_3$, and the total weight of that family is therefore $(1 - E_{12}^2)(1 - E_{23}^2)\, m_1^2 m_2^2 m_3^2$, which will be the first term in H_3.

A similar term results whenever $I = J$. For instance, when $I = J = \{1, 2, 4\}$ the paths starting at sources 1 and 2 must go straight across, while the path from source 4 to to sink 4 can go straight across, or down to line 3 and up again. The contributions from these two possible nonintersecting path families add up to

$$m_1 m_1 \cdot m_2(1 - E_{12}^2)m_2 \cdot \Big(m_4(1 - E_{34}^2)m_4 + m_4 E_{34}(1 - E_{23}^2)E_{34} m_4\Big)$$
$$= (1 - E_{12}^2)(1 - E_{24}^2)\, m_1^2 m_2^2 m_4^2.$$

From $I = \{1, 2, 3\}$ to $J = \{1, 2, 4\}$ there is one nonintersecting path family, and there is another one with the same weight from $I = \{1, 2, 4\}$ to $J = \{1, 2, 3\}$; the two add up to the term $2(1 - E_{12}^2)(1 - E_{23}^2)E_{24}\, m_1^2 m_2^2 m_3 m_4$.

Continuing like this, one finds that the types of terms that appear in H_3 are

$$\begin{aligned} H_3 = {} & (1 - E_{12}^2)(1 - E_{23}^2)\, m_1^2 m_2^2 m_3^2 + \dots \\ & + 2(1 - E_{12}^2)(1 - E_{23}^2)E_{34}\, m_1^2 m_2^2 m_3 m_4 + \dots \\ & + 4(1 - E_{12}^2)(1 - E_{34}^2)E_{23} E_{45}\, m_1^2 m_2 m_3 m_4 m_5 + \dots \\ & + 8\,(1 - E_{23}^2)(1 - E_{45}^2)E_{12} E_{34} E_{56}\, m_1 m_2 m_3 m_4 m_5 m_6. \end{aligned} \tag{A.1}$$

The last term comes from the 8 possible nonintersecting path families from $I = \{i_1, i_2, i_3\}$ to $J = \{j_1, j_2, j_3\}$ where $(i_1, j_1) = (1, 2)$ or $(2, 1)$, $(i_2, j_2) = (3, 4)$ or $(4, 3)$, and $(i_3, j_3) = (5, 6)$ or $(6, 5)$.

Remark A.10. Alternatively, the m_k can be factored out from any minor of PEP, leaving the corresponding minor of E, which can be computed using a result from Gantmacher and Krein [13, Section II.3.5], since the matrix E is what they call a *single-pair matrix*. This means a symmetric $n \times n$ matrix X with entries

$$X_{ij} = \begin{cases} \psi_i \chi_j, & i \le j, \\ \psi_j \chi_i, & i \ge j. \end{cases} \tag{A.2}$$

The $k \times k$ minors of a single-pair matrix are given by the following rule: $\det X_{IJ} = 0$, unless $I, J \in \binom{[1,n]}{k}$ satisfy the condition

$$(i_1, j_1) < (i_2, j_2) < \dots < (i_k, j_k), \tag{A.3}$$

where the notation means that both numbers in one pair must be less than both numbers in the following pair; in this case,

$$\det X_{IJ} = \psi_{\alpha_1} \begin{vmatrix} \chi_{\beta_1} & \chi_{\alpha_2} \\ \psi_{\beta_1} & \psi_{\alpha_2} \end{vmatrix} \begin{vmatrix} \chi_{\beta_2} & \chi_{\alpha_3} \\ \psi_{\beta_2} & \psi_{\alpha_3} \end{vmatrix} \dots \begin{vmatrix} \chi_{\beta_{k-1}} & \chi_{\alpha_k} \\ \psi_{\beta_{k-1}} & \psi_{\alpha_k} \end{vmatrix} \chi_{\beta_k}, \tag{A.4}$$

where

$$(\alpha_m, \beta_m) = \big(\min(i_m, j_m), \max(i_m, j_m)\big). \tag{A.5}$$

In the case of E we have $\psi_i = e^{x_i}$ and $\chi_i = e^{-x_i}$ (assuming as usual that $x_1 < \dots < x_n$), and (A.4) becomes

$$\det E_{IJ} = (1 - E^2_{\beta_1\alpha_2})(1 - E^2_{\beta_2\alpha_3}) \dots (1 - E^2_{\beta_{k-1}\alpha_k}) E_{\alpha_1\beta_1} E_{\alpha_2\beta_2} \dots E_{\alpha_k\beta_k}. \tag{A.6}$$

A.4 Proof of the "Canada Day Theorem"

The result to be proved (Theorem 4.1) is that for any symmetric $n \times n$ matrix X, the coefficient of s^k in the polynomial $\det(I + s\,TX)$ equals the sum of all $k \times k$ minors of X:

$$\det(I + s\,TX) = 1 + \sum_{k=1}^{n} \left(\sum_{I \in \binom{[1,n]}{k}} \sum_{J \in \binom{[1,n]}{k}} \det X_{IJ} \right) s^k. \tag{A.7}$$

We start from the elementary fact that for any matrix Y, the coefficients in its characteristic polynomial are given by the sums of the *principal* minors,

$$\det(I + s\,Y) = 1 + \sum_{k=1}^{n} \left(\sum_{J \in \binom{[1,n]}{k}} \det Y_{JJ} \right) s^k.$$

Applying this to $Y = TX$ and computing the minors of TX using the Cauchy–Binet formula [12, Ch. I, § 2]

$$\det(TX)_{AB} = \sum_{I \in \binom{[1,n]}{k}} \det T_{AI} \det X_{IB}, \qquad \text{for } A, B \in \tbinom{[1,n]}{k}, \tag{A.8}$$

we find that

$$\det(I + s\,TX) = 1 + \sum_{k=1}^{n} \left(\sum_{I \in \binom{[1,n]}{k}} \sum_{J \in \binom{[1,n]}{k}} \det T_{JI} \det X_{IJ} \right) s^k.$$

Comparing this to (A.7), it is clear that what we need to show is that, for any k,

$$\sum_{I \in \binom{[1,n]}{k}} \sum_{J \in \binom{[1,n]}{k}} \det T_{JI} \det X_{IJ} = \sum_{I \in \binom{[1,n]}{k}} \sum_{J \in \binom{[1,n]}{k}} \det X_{IJ}. \tag{A.9}$$

The first thing to do is calculate the minors $\det T_{JI}$.

Definition A.11. Given $I, J \in \binom{[1,n]}{k}$, the set I is said to *interlace* with the set J, denoted $I \leq J$, if

$$i_1 \leq j_1 \leq i_2 \leq j_2 \leq \ldots \leq i_k \leq j_k. \tag{A.10}$$

If all the inequalities are strict, then I is said to *strictly interlace* with J, in which case we write $I < J$. If $I \leq J$, then I' and J' denote the strictly interlacing subsets (possibly empty)

$$I' = I \setminus (I \cap J), \qquad J' = J \setminus (I \cap J), \tag{A.11}$$

whose cardinality (possibly zero) will be denoted by

$$p(I, J) = |I'| = |J'|. \tag{A.12}$$

Lemma A.12. *For $I, J \in \binom{[1,n]}{k}$, the corresponding $k \times k$ minor of T is*

$$\det T_{JI} = \begin{cases} 2^{p(I,J)}, & \text{if } I \leq J, \\ 0, & \text{otherwise.} \end{cases} \tag{A.13}$$

Proof. We will use Lindström's Lemma (Theorem A.6) on the planar network for T given in Section A.2 above; the minor $\det T_{JI}$ equals the total number of families of nonintersecting paths connecting the source nodes (on the left) indexed by J to the sink nodes (on the right) indexed by I.

The proof proceeds by induction on the size n of T. The claim is trivially true for $n = 1$. Consider an arbitrary $n > 1$, and suppose the claim is true for size $n - 1$. If neither I nor J contain n, the claim follows immediately from the induction hypothesis, and likewise if I and J both contain n, because there is only one path connecting source n to sink n. If I contains n but J does not, then $\det T_{JI} = 0$ because there are no paths going upward; this agrees with the claim, since in this case I does not interlace with J.

The only remaining case is therefore $J = J_1 \cup \{n\}$, $I = I_1 \cup \{i_k\}$, with $i_k < n$. But then

$$\det T_{JI} = \det T_{J_1 I_1} \times \begin{cases} 2, & \text{if } j_{k-1} < i_k, \\ 1, & \text{if } j_{k-1} = i_k, \\ 0, & \text{if } j_{k-1} > i_k, \end{cases}$$

depending on whether the path connecting source n with sink i_k has to cross the j_{k-1} level; if it does not, there are two available paths, if it does, there is only one available path provided $j_{k-1} = i_n$, otherwise the path intersects the path coming from source j_{k-1}. In the last instance, I does not interlace with J, while in the other two $I \leq J$ if and only if $I_1 \leq J_1$, thus proving the claim. □

According to this lemma, the structure of (A.9) (which is what we want to prove) is

$$\sum_{\substack{I,J \in \binom{[1,n]}{k} \\ I \leq J}} 2^{p(I,J)} \det X_{IJ} = \sum_{A,B \in \binom{[1,n]}{k}} \det X_{AB}, \tag{A.14}$$

and we must show that those terms $\det X_{IJ}$ that occur more than once on the left-hand side exactly compensate for those that are absent. This will follow from another technical lemma:

Lemma A.13 (Relations between $k\times k$ minors of a symmetric matrix)**.** *Suppose $I, J \in \binom{[1,n]}{k}$ and $I \le J$. Then, for any symmetric $n \times n$ matrix X,*

$$\sum_{\substack{A,B\in\binom{I\cup J}{k}\\ A\cap B=I\cap J}} \det X_{AB} = 2^{p(I,J)} \det X_{IJ}. \tag{A.15}$$

Before proving Lemma A.13, we will use it to finish the proof of the main theorem. The two lemmas above show that the sum on the left-hand side of (A.14) equals

$$\sum_{\substack{I,J\in\binom{[1,n]}{k}\\ I\le J}} 2^{p(I,J)} \det X_{IJ} \;=\; \sum_{\substack{I,J\in\binom{[1,n]}{k}\\ I\le J}} \left(\sum_{\substack{A,B\in\binom{I\cup J}{k}\\ A\cap B=I\cap J}} \det X_{AB} \right), \tag{A.16}$$

which in turn equals the sum on the right-hand side of (A.14),

$$\sum_{A,B\in\binom{[1,n]}{k}} \det X_{AB}. \tag{A.17}$$

Thus (A.14) holds, and the theorem is proved. The final step from (A.16) to (A.17) is justified by the observation that any given pair (A, B) of the type summed over in (A.17) appears exactly once in the right-hand side of (A.16), namely for the sets I and J defined as follows. Let $M = A \cap B$, $A' = A \setminus M$, $B' = B \setminus M$, and let $p \ge 0$ be the cardinality of the disjoint sets A' and B' (they are empty sets if $p = 0$). Then define I' and J' by enumerating the $2p$ elements of $A' \cup B'$ in the strictly interlacing order $I' < J'$, and let $I = M \cup I'$ and $J = M \cup J'$. Conversely, no other terms than these appear in the right hand side of (A.16), and it is therefore indeed equal to (A.17).

Proof of Lemma A.13. The sets $I \le J$ and $I' < J'$ (as in Definition A.11), with

$$|I| = |J| = k, \qquad |I'| = |J'| = p(I, J) = p,$$

will be fixed throughout the proof, and for convenience we also introduce $M = I \cap J$ and $U = I \cup J$, with $|M| = k - p$ and $|U| = k + p$. We can assume that $p > 0$, since the case $p = 0$ is trivial; it occurs when $I = J$, and then both sides of (A.15) simply equal $\det X_{II}$.

The set $U \setminus M$ consists of the $2p$ numbers which belong alternatingly to I' and to J'. The sum (A.15) runs over all pairs of sets (A, B) obtained by splitting these $2p$ numbers into two disjoint p-sets A' and B' in an arbitrary way and letting $A = M \cup A'$ and $B = M \cup B'$. Write $\mathcal{Q}$ for this set; that is, $\mathcal{Q}$ denotes

the set of pairs $(A,B) \in \binom{[1,n]}{k} \times \binom{[1,n]}{k}$ such that $A \cup B = U$ and $A \cap B = M$. After expanding $\det X_{AB}$, we can then write the left-hand side of (A.15) as

$$\sum_{((A,B),\sigma)\in\mathcal{Q}\times\mathcal{S}_k} (-1)^\sigma X_{a_1 b_{\sigma(1)}} X_{a_2 b_{\sigma(2)}} \dots X_{a_k b_{\sigma(k)}}, \tag{A.18}$$

where $\mathcal{S}_k$ is the group of permutations of $\{1,2,\dots,k\}$, and $(-1)^\sigma$ denotes the sign of the permutation σ.

For each $((A,B),\sigma) \in \mathcal{Q}\times\mathcal{S}_k$, we let $A' = A\backslash M$ and $B' = B\backslash M$, and set up a (σ-dependent) bijection between A' and B' as follows: $a' \in A'$ is paired up with $b' \in B'$ if and only if the product $X_{a_1 b_{\sigma(1)}} X_{a_2 b_{\sigma(2)}} \dots X_{a_k b_{\sigma(k)}}$ contains either the factor $X_{a'b'}$ or a sequence of factors $X_{a'r}$, X_{rs}, ..., $X_{tb'}$ where $r,s,\dots,t \in M$. Let us say that a' and b' are *linked* if they are paired up in this manner. A linked pair $(a',b') \in A' \times B'$ will be called *hostile* if (a',b') belongs to $I' \times I'$ or $J' \times J'$, and *friendly* if (a',b') belongs to $I' \times J'$ or $J' \times I'$. To each term in the sum (A.18) there will thus correspond p such linked pairs, and what we will show is that the terms containing at least one hostile pair will cancel out, and that the remaining terms (with all friendly pairs) will add up to the right-hand side of (A.15).

Next we define what we mean by *flipping* a linked pair (a',b'). This means that, in the product $X_{a_1 b_{\sigma(1)}} X_{a_2 b_{\sigma(2)}} \dots X_{a_k b_{\sigma(k)}}$, those factors $X_{a'r} X_{rs} \dots X_{tb'}$ that link a' to b' are replaced by $X_{b't} \dots X_{sr} X_{ra'}$, with all the indices in reversed order. (When the linking involves just a single factor $X_{a'b'}$, flipping means replacing it by $X_{b'a'}$.) Since the matrix X is symmetric, this does not change the value of the product, but it changes the way it is indexed. The number a' which used to be in the first slot (in $X_{a'r}$) is now in the second slot (in $X_{ra'}$), and vice versa for b'. The connecting indices $r,s,\dots,t \in M$ do not contribute to any change in the indexing sets, since, for example, the r in $X_{a'r}$ is moved from the second slot to the first, while the other r in X_{rs} is moved from the first to the second. The new product (the result of the flipping) is therefore indexed by the sets

$$\Big(A \setminus \{a'\}\Big) \cup \{b'\} =: \widetilde{A} = \{\widetilde{a}_1 < \dots < \widetilde{a}_k\}$$

and

$$\Big(B \setminus \{b'\}\Big) \cup \{a'\} =: \widetilde{B} = \{\widetilde{b}_1 < \dots < \widetilde{b}_k\}$$

respectively, and after reordering the factors so that the first indices come in ascending order, it can be written

$$X_{\widetilde{a}_1 \widetilde{b}_{\widetilde{\sigma}(1)}} X_{\widetilde{a}_2 \widetilde{b}_{\widetilde{\sigma}(2)}} \dots X_{\widetilde{a}_k \widetilde{b}_{\widetilde{\sigma}(k)}}$$

for some uniquely determined permutation $\widetilde{\sigma} \in \mathcal{S}_k$. Flipping a given pair thus takes $((A,B),\sigma)$ to $((\widetilde{A},\widetilde{B}),\widetilde{\sigma})$. This operation is invertible, with inverse given by simply flipping the same pair again, now viewed as a pair $(b',a') \in ((\widetilde{A})',(\widetilde{B})')$ linked via the indices $t,\dots,s,r$. Because of the symmetry of the matrix X, the term in (A.18) corresponding to $((\widetilde{A},\widetilde{B}),\widetilde{\sigma})$ is equal to the term corresponding

to $((A,B),\sigma)$, except possibly for a difference in sign, depending on whether the signs of σ and $\widetilde{\sigma}$ come out equal or not:

$$(-1)^{\widetilde{\sigma}} X_{\widetilde{a}_1 \widetilde{b}_{\widetilde{\sigma}(1)}} X_{\widetilde{a}_2 \widetilde{b}_{\widetilde{\sigma}(2)}} \dots X_{\widetilde{a}_k \widetilde{b}_{\widetilde{\sigma}(k)}} = \pm(-1)^{\sigma} X_{a_1 b_{\sigma(1)}} X_{a_2 b_{\sigma(2)}} \dots X_{a_k b_{\sigma(k)}}.$$

We will show below that *the permutation $\widetilde{\sigma}$ has the same sign as σ when a friendly pair is flipped, and the opposite sign when a hostile pair is flipped.* Taking this for granted for the moment, divide the set $\mathcal{Q} \times \mathcal{S}_k$ into the two sets $(\mathcal{Q} \times \mathcal{S}_k)_{\text{hostile}}$, consisting of those $((A,B),\sigma)$ for which at least one linked pair is hostile, and $(\mathcal{Q} \times \mathcal{S}_k)_{\text{friendly}}$, consisting of those $((A,B),\sigma)$ for which all p linked pairs are friendly. The mapping "flip that out of all hostile pairs (a',b') for which $\min(a',b')$ is smallest" is an involution on $(\mathcal{Q} \times \mathcal{S}_k)_{\text{hostile}}$ that pairs up each term with a partner term that is equal except for having the opposite sign (since it is a hostile pair that is flipped). Consequently these terms cancel out, and the contribution from $(\mathcal{Q} \times \mathcal{S}_k)_{\text{hostile}}$ to (A.18) is zero. The sum therefore reduces to

$$\sum_{((A,B),\sigma) \in (\mathcal{Q} \times \mathcal{S}_k)_{\text{friendly}}} (-1)^{\sigma} X_{a_1 b_{\sigma(1)}} X_{a_2 b_{\sigma(2)}} \dots X_{a_k b_{\sigma(k)}}. \tag{A.19}$$

Now equip the set $(\mathcal{Q} \times \mathcal{S}_k)_{\text{friendly}}$ with an equivalence relation; $((\widetilde{A},\widetilde{B}),\widetilde{\sigma})$ and $((A,B),\sigma)$ are equivalent if one can go from one to another by flipping friendly pairs. Each equivalence class contains 2^p elements, since each of the p friendly pairs can belong to either $I' \times J'$ or $J' \times I'$. Moreover, the terms corresponding to the elements in one equivalence class are all equal (including the sign, since only friendly pairs are flipped), and each class has a "canonical" representative with all linked pairs belonging to $I' \times J'$,

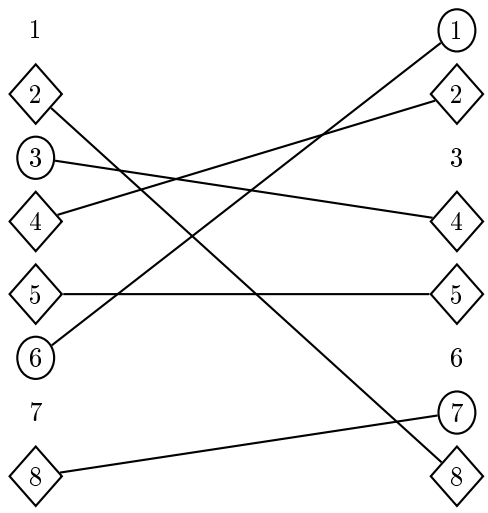
$$(-1)^{\sigma} X_{i_1 j_{\sigma(1)}} X_{i_2 j_{\sigma(2)}} \dots X_{i_k j_{\sigma(k)}},$$

where the permutation σ is uniquely determined by the equivalence class (and vice versa). Thus (A.19) becomes

$$2^p \sum_{\sigma \in \mathcal{S}_k} (-1)^{\sigma} X_{i_1 j_{\sigma(1)}} X_{i_2 j_{\sigma(2)}} \dots X_{i_k j_{\sigma(k)}} = 2^p \det X_{IJ}, \tag{A.20}$$

which is what we wanted to prove.

To finish the proof, it now remains to demonstrate the rule that $\widetilde{\sigma}$ has the same (opposite) sign as σ when a friendly (hostile) pair is flipped. To this end, we will represent $((A,B),\sigma)$ with a bipartite graph, with the numbers in $U = A \cup B$ (in increasing order) as nodes both on the left and on the right, and the left nodes $a_i \in A$ connected by edges to the corresponding right nodes $b_{\sigma(i)} \in B$. The sign of σ will then be equal to $(-1)^c$, where c is the crossing number of the graph. As an aid in explaining the ideas we will use the following example with $U = [1,8]$, where the nodes in $M = A \cap B$ are marked with diamonds, and the nodes in A' and B' are marked with circles:



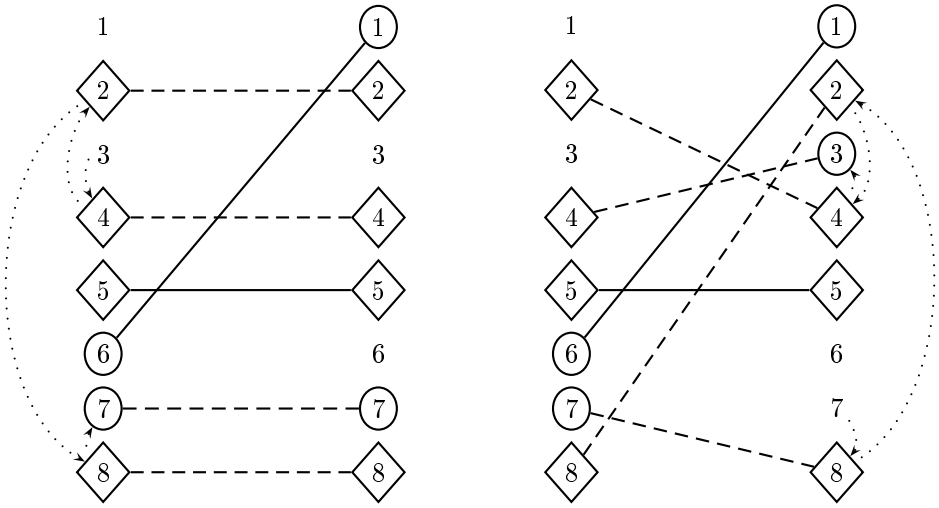
$$\begin{aligned} A &= \{2,3,4,5,6,8\} \\ &= \{2,4,5,8\} \cup \{3,6\} \\ &= M \cup A' \end{aligned} \qquad \begin{aligned} B &= \{1,2,4,5,7,8\} \\ &= \{2,4,5,8\} \cup \{1,7\} \\ &= M \cup B' \end{aligned}$$

Clearly, $A' \cup B' = \{3,6\} \cup \{1,7\} = \{1,3,6,7\} = \{i'_1 < j'_1 < i'_2 < j'_2\}$, so that $I' = \{i'_1, i'_2\} = \{1,6\}$ and $J' = \{j'_1, j'_2\} = \{3,7\}$. Consequently, $I = M \cup I' = \{1,2,4,5,6,8\}$ and $J = M \cup J' = \{2,3,4,5,6,7\}$. The chosen permutation is $\sigma(123456) = 632415$, where the notation means that $\sigma(1) = 6$, $\sigma(2) = 3$, etc.; for example, the latter equality comes from the second smallest number a_2 in A being connected to the third smallest number b_3 in B. There are 9 crossings, so σ is an odd permutation, and this graph therefore represents the term $-X_{28}X_{34}X_{42}X_{55}X_{61}X_{87}$, appearing with a minus sign in the sum (A.18). The linked pairs $(a', b') \in A' \times B'$ are $(6,1)$ (directly linked) and $(3,7)$ (linked via $4,2,8 \in M$). Both pairs are hostile, since $(6,1) \in I' \times I'$ and $(3,7) \in J' \times J'$.

We will illustrate in detail what happens when the pair $(3,7)$ is flipped. The flip is effected by replacing the factors $X_{34}X_{42}X_{28}X_{87}$ by $X_{78}X_{82}X_{24}X_{43}$ and sorting the resulting product so that the first indices come in ascending order; this gives $X_{24}X_{43}X_{55}X_{61}X_{78}X_{82}$. Thus $\widetilde{A} = \{2,4,5,6,7,8\}$, $\widetilde{B} = \{1,2,3,4,5,8\}$, and $\widetilde{\sigma}(123456) = 435162$ (an even permutation). In terms of the graph, the nodes that are involved in the flip are, on both sides, $\{2,3,4,7,8\}$ (the two nodes in the pair being flipped, plus the nodes linking them), and the edges involved are $\{34,42,28,87\}$, which get changed into $\{43,24,82,78\}$. In other words, the flip corresponds to this *active subgraph* being mirror reflected across the central vertical line. To understand how the process of reflection affects the crossing number, it can be broken down into two steps, as follows.

On the left, node 7 is unoccupied to begin with, so we can change the edge 87 to 77. This frees node 8 on the left, so that we can change the edge 28 to 88, which frees node 2 on the left. (Think of this edge as a rubber band connected at one end to node 8 on the right; we're disconnecting its other end from node

2 on the left and sliding it past all the other nodes down to node 8 on the left. Obviously the crossing number increases or decreases by one every time we slide past a node that has an edge attached to it.) Continuing like this, we get the result illustrated in Step 1 below; the edges changed are $87 \to 77$, $28 \to 88$, $42 \to 22$, $34 \to 44$.



Intermediate stage (after Step 1) Result of the flip (after Step 2)

In Step 2, we work similarly on the right-hand side: node 3 is unoccupied to begin with, so we can change edge 44 to 43, and so on. The list of edge moves is $44 \to 43$, $22 \to 24$, $88 \to 82$, $77 \to 78$. (In the graph on the right we see that the crossing number after the flip is 8, verifying the claim that $\widetilde{\sigma}$ is an even permutation.)

We need to keep track of the changes in the crossing number caused by sliding active edges past nodes that have edges attached to them. This is most easily done by following the dotted lines in the figures, and counting whether the nodes that are marked (with circles and diamonds) are passed an even or an odd number of times. However, since the active subgraph simply gets reflected, the crossings among its edges will be the same before and after the flip, so we need in fact only count how many times we pass a *passive* marked node. (The passive nodes in the example are $\{1, 5, 6\}$.)

If a passive node belonging to M is passed in Step 1, then it is passed the same number of times in Step 2 as well, since the nodes in M are marked both on the left and on the right. Therefore they do not affect the parity of the crossing number either, and we can ignore the nodes marked with diamonds, and only look at the *passive circled* nodes (all the nodes in A' and B' except for the two active nodes that are being flipped).

Passive nodes belonging to A' are counted only in Step 1 and passive nodes in B' only in Step 2; they get counted an *odd* number of times if they lie *between* the two flipped nodes (like node 6 in the example, counted once), and an *even*

number of times otherwise (like node 1, never counted). Consequently, what determines whether the parity of the crossing number changes is the number of nodes *between* the flipped ones that belong to $A' \cup B' = I' \cup J'$. And for a friendly pair, this number is even, while for a hostile pair, it is odd.

This shows that the crossing number keeps its parity (so that $(-1)^\sigma = (-1)^{\tilde\sigma}$) when a friendly pair is flipped, and the opposite when a hostile pair is flipped. The proof is finally complete. □

B Verification of the Lax pair for peakons

The purpose of this appendix is to carefully verify that the Lax pair formulation (4.1)–(4.2) of the Novikov equation really is valid for the class of distributional solutions that we are considering. This is not at all obvious, as should be clear from the computations below.

B.1 Preliminaries

We will need to be more precise regarding the notation here than in the main text. A word of warning right away: our notation for derivatives here will differ from that used in the rest of the paper (where subscripts should be interpreted as distributional derivatives).

To begin with, given n smooth functions $x = x_k(t)$ such that $x_1(t) < \cdots < x_n(t)$, let $x_0(t) = -\infty$ and $x_{n+1}(t) = +\infty$, and let Ω_k (for $k = 0, \ldots, n$) denote the region $x_k(t) < x < x_{k+1}(t)$ in the (x,t) plane.

Our computations will deal with a class that we denote PC^∞, consisting of piecewise smooth functions $f(x,t)$ such that the restriction of f to each region Ω_k is (the restriction to Ω_k of) a smooth function $f^{(k)}(x,t)$ defined on an open neighbourhood of $\overline{\Omega}_k$ (so that $f^{(k)}$ and its partial derivatives make sense on the curves $x = x_k(t)$). For each fixed t, the function $f(\cdot,t)$ defines a regular distribution T_f in the class $\mathcal{D}'(\mathbf{R})$, depending parametrically on t (and written $T_f(t)$ where needed). After having made clear exactly what is meant, we will mostly be less strict, and write f instead of T_f for simplicity.

The values of f on the curves $x = x_k(t)$ need not be defined; the function defines the same distribution T_f no matter what values are assigned to $f(x_k(t),t)$. But our assumptions imply that the left and right limits of f exist, and (suppressing the time dependence) they will be denoted by $f(x_k^-) := f^{(k-1)}(x_k)$ and $f(x_k^+) := f^{(k)}(x_k)$, respectively. The jump and the average of f at x_k will be denoted by

$$\big[f(x_k)\big] := f(x_k^+) - f(x_k^-) \qquad \text{and} \qquad \big\langle f(x_k)\big\rangle := \frac{f(x_k^+) + f(x_k^-)}{2}, \tag{B.1}$$

respectively. They satisfy the product rules

$$[fg] = \langle f\rangle[g] + [f]\langle g\rangle, \qquad \langle fg\rangle = \langle f\rangle\langle g\rangle + \tfrac{1}{4}[f]\,[g]. \tag{B.2}$$

We will use subscripts to denote partial derivatives in the classical sense, so that (for example) f_x denotes the piecewise smooth function whose restriction to Ω_k is given by $\partial f^{(k)}/\partial x$ (and whose values at $x = x_k(t)$ are in general undefined). On the other hand, D_x will denote the distributional derivative, which in addition picks up Dirac delta contributions from jump discontinuities of f at the curves $x = x_k(t)$. That is, $D_x T_f = T_{f_x} + \sum_{k=1}^n \big[f(x_k)\big]\delta_{x_k}$, or, in less strict notation,

$$D_x f = f_x + \sum_{k=1}^n \big[f(x_k)\big]\delta_{x_k}. \tag{B.3}$$

The time derivative D_t is defined as a limit in $\mathcal{D}'(\mathbf{R})$,

$$D_t T_f(t) = \lim_{h\to 0} \frac{T_f(t+h) - T_f(t)}{h}, \tag{B.4}$$

and it commutes with D_x by the continuity of D_x on $\mathcal{D}'(\mathbf{R})$. For our class PC^∞ of piecewise smooth functions, we have $D_t T_f = T_{f_t} - \sum_{k=1}^n \dot{x}_k \big[f(x_k)\big]\delta_{x_k}$, or simply

$$D_t f = f_t - \sum_{k=1}^n \dot{x}_k \big[f(x_k)\big]\delta_{x_k}, \tag{B.5}$$

where $\dot{x}_k = dx_k/dt$. We also note that $\frac{d}{dt} f(x_k^\pm(t), t) = f_x(x_k^\pm(t), t)\,\dot{x}_k(t) + f_t(x_k^\pm(t), t)$, which gives

$$\begin{aligned} \tfrac{d}{dt}\big[f(x_k)\big] &= \big[f_x(x_k)\big]\,\dot{x}_k + \big[f_t(x_k)\big], \\ \tfrac{d}{dt}\big\langle f(x_k)\big\rangle &= \big\langle f_x(x_k)\big\rangle\,\dot{x}_k + \big\langle f_t(x_k)\big\rangle. \end{aligned} \tag{B.6}$$

B.2 The problem of multiplication

If the function f is continuous at $x = x_k$, then the Dirac delta at x_k can be multiplied by the corresponding distribution T_f according to the well-known formula

$$T_f\,\delta_x = f(x_k)\,\delta_{x_k}. \tag{B.7}$$

But below we will have to consider this product for functions in the class PC^∞ described above, where the value $f(x_k)$ is not defined. It will turn out that in the present context, the right thing to do is to use the average value of f at the jump, and thus define $T_f\,\delta_x := \big\langle f(x_k)\big\rangle\,\delta_{x_k}$. However, since we want this to be a consequence of the analysis, rather than an a priori assumption, we will, to begin with, just assign a hypothetical value $f(x_k)$ and use that value in (B.7). This assignment is justified in the present context, as we will see below. However, we are not claiming that this addresses any of the deeper issues; for example, this assignment does not respect the product structure of piecewise continuous functions. See [30, Ch. 5] for more information about the structural problems associated with any attempt to define a product of distributions in $\mathcal{D}'(\mathbf{R})$.

B.3 Distributional Lax pair

Peakon solutions

$$u(x,t) = \sum_{k=1}^{n} m_k(t)\, e^{-|x-x_k(t)|} \tag{B.8}$$

belong to the piecewise smooth class PC^∞. They are continuous and satisfy

$$D_x u = u_x = \sum_{k=1}^{n} m_k \ \mathrm{sgn}(x_k - x)\, e^{-|x-x_k|},$$

$$D_x^2 u = D_x(u_x) = u_{xx} + \sum_{k=1}^{n} \big[u_x(x_k)\big]\, \delta_{x_k} = u + \sum_{k=1}^{n} (-2m_k)\, \delta_{x_k},$$

which implies

$$m := u - D_x^2 u = 2\sum_{k=1}^{n} m_k\, \delta_{x_k}. \tag{B.9}$$

The Lax pair (4.1)–(4.2) will involve the functions u and $D_x u$, as well as the purely singular distribution m. We will take ψ_1, ψ_2, ψ_3 to be functions in PC^∞, and separate the regular (function) part from the singular (Dirac delta) part. The formulation obtained in this way reads

$$D_x \Psi = \widehat{L}\Psi, \qquad D_t \Psi = \widehat{A}\Psi, \tag{B.10}$$

where $\Psi = (\psi_1, \psi_2, \psi_3)^t$,

$$\widehat{L} = L + 2z\left(\sum_{k=1}^{n} m_k\, \delta_{x_k}\right) N, \quad L = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{B.11}$$

and

$$\widehat{A} = A - 2z\left(\sum_{k=1}^{n} m_k\, u(x_k)^2 \delta_{x_k}\right) N, \quad A = \begin{pmatrix} -uu_x & u_x/z & u_x^2 \\ u/z & -1/z^2 & -u_x/z \\ -u^2 & u/z & uu_x \end{pmatrix}. \tag{B.12}$$

Note that (B.10) involves multiplying $N\Psi = (\psi_2, \psi_3, 0)$ by δ_{x_k}, and some value $\psi_2(x_k)$ must be assigned in order for this to be well-defined (we will soon see that ψ_3 must be continuous and therefore it is only ψ_2 that presents any problems).

Theorem B.1. *Provided that the product $m\psi_2$ is defined using the average value $\psi_2(x_k) := \big\langle \psi_2(x_k) \big\rangle$ at the jumps,*

$$m\psi_2 := 2\sum_{k=1}^{n} m_k \big\langle \psi_2(x_k) \big\rangle\, \delta_{x_k}, \tag{B.13}$$

the following statement holds. With u and m given by (B.8)–(B.9)*, and with $\Psi \in PC^\infty$, the Lax pair* (B.10)–(B.12) *satisfies the compatibility condition $D_t D_x \Psi = D_x D_t \Psi$ if and only if the peakon ODEs* (3.4) *are satisfied: $\dot{x}_k = u(x_k)^2$ and $\dot{m}_k = -m_k\, u(x_k) \big\langle u_x(x_k) \big\rangle$.*

Proof. For simplicity, we will write just $\sum$ instead of $\sum_{k=1}^{n}$. Identifying coefficients of δ_{x_k} in the two Lax equations (B.10) immediately gives $\big[\Psi(x_k)\big] = 2z\, m_k N\Psi(x_k)$ and $-\dot{x}_k\big[\Psi(x_k)\big] = -2z\, m_k u(x_k)^2 N\Psi(x_k)$, respectively. Thus, $[\psi_3(x_k)] = 0$ (in other words, ψ_3 is continuous) and $\dot{x}_k = u(x_k)^2$. Next we compute the derivatives of (B.10):

$$\begin{aligned}
D_t(D_x\Psi) &= D_t(L\Psi + 2z\left(\sum m_k\,\delta_{x_k}\right)N\Psi) \\
&= L(\widehat{A}\Psi) + 2zN\sum \tfrac{d}{dt}\big(m_k\Psi(x_k)\big)\,\delta_{x_k} - 2zN\sum m_k\Psi(x_k)\dot{x}_k\delta'_{x_k}, \\
D_x(D_t\Psi) &= D_x(A\Psi - 2z\left(\sum m_k\, u(x_k)^2\delta_{x_k}\right)N\Psi) \\
&= (A\Psi)_x + \sum\big[A\Psi(x_k)\big]\delta_{x_k} - 2zN\sum m_k\Psi(x_k)u(x_k)^2\delta'_{x_k}.
\end{aligned}$$

The regular part of (B.10) gives $\Psi_x = L\Psi$, so that $(A\Psi)_x = A_x\Psi + AL\Psi$, and it is easily verified that $LA = A_x + AL$ holds identically (since $u_{xx} = u$). This implies that the regular parts of the two expressions above are equal, and the terms involving δ'_{x_k} are also equal since $\dot{x}_k = u(x_k)^2$. Therefore the compatibility condition $D_t(D_x\Psi) = D_x(D_t\Psi)$ reduces to an equality between the coefficients of δ_{x_k},

$$-2z\, m_k u(x_k)^2 LN\Psi(x_k) + 2zN\tfrac{d}{dt}\big(m_k\Psi(x_k)\big) = \big[A\Psi(x_k)\big]. \tag{B.14}$$

Using the product rule (B.2), the expression for $\big[\Psi(x_k)\big]$ above, and $\big[u_x(x_k)\big] = -2m_k$, we find that the right-hand side of (B.14) equals

$$\begin{aligned}
&\big\langle A(x_k)\big\rangle\, 2z\, m_k N\Psi(x_k) + \big[A(x_k)\big]\big\langle\Psi(x_k)\big\rangle = \\
&2z\, m_k \begin{pmatrix} 0 & -u\big\langle u_x\big\rangle & \big\langle u_x\big\rangle/z \\ 0 & u/z & -1/z^2 \\ 0 & -u^2 & u/z \end{pmatrix}_{x_k} \Psi(x_k) + 2m_k \begin{pmatrix} u & -1/z & -2\big\langle u_x\big\rangle \\ 0 & 0 & 1/z \\ 0 & 0 & -u \end{pmatrix}_{x_k} \big\langle\Psi(x_k)\big\rangle.
\end{aligned} \tag{B.15}$$

The (3,2) entry $-u^2$ in the matrix in the first term will cancel against the whole first term on the left-hand side of (B.14), since the only nonzero entry of LN is $(LN)_{32} = 1$. Thus (B.14) is equivalent to

$$\begin{aligned}
&\dot{m}_k\, N\Psi(x_k) + m_k\, N\tfrac{d}{dt}\Psi(x_k) = \\
&m_k \begin{pmatrix} 0 & -u\big\langle u_x\big\rangle & \big\langle u_x\big\rangle/z \\ 0 & u/z & -1/z^2 \\ 0 & 0 & u/z \end{pmatrix}_{x_k} \Psi(x_k) + m_k \begin{pmatrix} u/z & -1/z^2 & -2\big\langle u_x\big\rangle/z \\ 0 & 0 & 1/z^2 \\ 0 & 0 & -u/z \end{pmatrix}_{x_k} \big\langle\Psi(x_k)\big\rangle.
\end{aligned} \tag{B.16}$$

To make it clear how the assumption (B.13) enters the proof, we want to avoid assigning a value to $\psi_2(x_k)$ for as long as possible. Therefore we can't compute $\frac{d}{dt}\Psi(x_k)$ quite yet. But $\big\langle\Psi(x_k)\big\rangle$ is well-defined, and its time derivative can be

computed using $\Psi_x = L\Psi$ and $\Psi_t = A\Psi$ in (B.6):

$$\begin{aligned} N\tfrac{d}{dt}\big\langle\Psi(x_k)\big\rangle &= N\big\langle L\Psi(x_k)\big\rangle\,\dot{x}_k + N\big\langle A\Psi(x_k)\big\rangle \\ &= N\Big(Lu(x_k)^2 + \big\langle A(x_k)\big\rangle\Big)\big\langle\Psi(x_k)\big\rangle + N\tfrac{1}{4}\big[A(x_k)\big]\big[\Psi(x_k)\big] \\ &= \begin{pmatrix} u/z & -1/z^2 & -\big\langle u_x\big\rangle/z \\ 0 & u/z & u\big\langle u_x\big\rangle \\ 0 & 0 & 0 \end{pmatrix}_{x_k} \big\langle\Psi(x_k)\big\rangle + \tfrac{1}{4}\underbrace{N\big[A(x_k)\big]N}_{=0}\,2z\,m_k\Psi(x_k). \end{aligned}$$

A bit of manipulation using this result, as well as $\big\langle\psi_3\big\rangle(x_k) = \psi_3(x_k)$, shows that the compatibility condition (B.16) can be written as

$$\begin{aligned} m_k N\tfrac{d}{dt}\Big(\Psi(x_k) - \big\langle\Psi(x_k)\big\rangle\Big) + \Big(\dot{m}_k + m_k u(x_k)\big\langle u_x(x_k)\big\rangle\Big)N\Psi(x) \\ = m_k \begin{pmatrix} 0 & 0 & 0 \\ 0 & u/z & 0 \\ 0 & 0 & 0 \end{pmatrix}_{x_k} \Big(\Psi(x_k) - \big\langle\Psi(x_k)\big\rangle\Big) \end{aligned} \tag{B.17}$$

The third row is zero, and the first two rows say that

$$\begin{aligned} \Big(\dot{m}_k + m_k u(x_k)\big\langle u_x(x_k)\big\rangle\Big)\psi_2(x_k) &= -m_k\tfrac{d}{dt}\Big(\psi_2(x_k) - \big\langle\psi_2(x_k)\big\rangle\Big), \\ \Big(\dot{m}_k + m_k u(x_k)\big\langle u_x(x_k)\big\rangle\Big)\psi_3(x_k) &= \tfrac{1}{z}m_k\,u(x_k)\Big(\psi_2(x_k) - \big\langle\psi_2(x_k)\big\rangle\Big). \end{aligned}$$

At this point we choose to assign $\psi_2(x_k) := \big\langle\psi_2\big\rangle(x_k)$, and then it is clear that (B.17) is satisfied if and only if

$$\dot{m}_k = -m_k u(x_k)\big\langle u_x(x_k)\big\rangle.$$

□

Acknowledgements

HL is supported by the Swedish Research Council (Vetenskapsrådet), and JS by the National Sciences and Engineering Research Council of Canada (NSERC).

HL and JS would like to acknowledge the hospitality of the Mathematical Research and Conference Center, Będlewo, Poland, and the Department of Mathematics and Statistics, University of Saskatchewan.

References

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

[2] Richard Beals, David H. Sattinger, and Jacek Szmigielski. Multipeakons and the classical moment problem. *Advances in Mathematics*, 154:229–257, 2000.

[3] Marco Bertola, Michael Gekhtman, and Jacek Szmigielski. Peakons and Cauchy biorthogonal polynomials. *arXiv:0711.4082 [nlin.SI]*, 2007.

[4] Harish S. Bhat and Razvan C. Fetecau. A Hamiltonian Regularization of the Burgers Equation. *Journal of Nonlinear Science*, 16(6):615–638, December 2006.

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

[6] Giuseppe M. Coclite and Kenneth H. Karlsen. On the well-posedness of the Degasperis–Procesi equation. *J. Funct. Anal.*, 233(1):60–91, 2006.

[7] Giuseppe M. Coclite and Kenneth H. Karlsen. On the uniqueness of discontinuous solutions to the Degasperis–Procesi equation. *J. Differential Equations*, 234(1):142–160, 2007.

[8] Antonio Degasperis, Darryl D. Holm, and Andrew N. W. Hone. A new integrable equation with peakon solutions. *Theoretical and Mathematical Physics*, 133:1463–1474, 2002.

[9] Antonio Degasperis and Michela Procesi. Asymptotic integrability. In A. Degasperis and G. Gaeta, editors, *Symmetry and perturbation theory (Rome, 1998)*, pages 23–37. World Scientific Publishing, River Edge, NJ, 1999.

[10] Harry Dym and Henry P. McKean. *Gaussian processes, function theory, and the inverse spectral problem.* Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. Probability and Mathematical Statistics, Vol. 31.

[11] Sergey Fomin and Andrei Zelevinsky. Total positivity: tests and parametrizations. *Math. Intelligencer*, 22(1):23–33, 2000.

[12] Felix R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998. Translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[13] Felix R. Gantmacher and Mark G. Krein. *Oscillation matrices and kernels and small vibrations of mechanical systems.* AMS Chelsea Publishing, Providence, RI, revised edition, 2002. Translation based on the 1941 Russian original, edited and with a preface by Alex Eremenko.

[14] Ira Gessel and Gérard Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3):300–321, 1985.

[15] Darryl D. Holm and Andrew N. W. Hone. A class of equations with peakon and pulson solutions. *Journal of Nonlinear Mathematical Physics*, 12(2):46–62, 2005.

[16] Andrew N. W. Hone and Jing Ping Wang. Prolongation algebras and Hamiltonian operators for peakon equations. *Inverse Problems*, 19(1):129–145, 2003.

[17] Andrew N. W. Hone and Jing Ping Wang. Integrable peakon equations with cubic nonlinearity. *Journal of Physics A: Mathematical and Theoretical*, 41(37):372002, 2008.

[18] Rossen Ivanov. On the integrability of a class of nonlinear dispersive wave equations. *J. Nonlinear Math. Phys.*, 12(4):462–468, 2005.

[19] Samuel Karlin. *Total positivity. Vol. I.* Stanford University Press, Stanford, Calif, 1968.

[20] Samuel Karlin and James McGregor. Coincidence probabilities. *Pacific J. Math.*, 9:1141–1164, 1959.

[21] Jennifer Kohlenberg, Hans Lundmark, and Jacek Szmigielski. The inverse spectral problem for the discrete cubic string. *Inverse Problems*, 23:99–121, 2007.

[22] Christian Krattenthaler. Advanced determinant calculus. *Sém. Lothar. Combin.*, 42:Art. B42q, 67 pp. (electronic), 1999. The Andrews Festschrift (Maratea, 1998).

[23] Bernt Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90, 1973.

[24] Hans Lundmark. Formation and dynamics of shock waves in the Degasperis–Procesi equation. *J. Nonlinear Sci.*, 17(3):169–198, 2007.

[25] Hans Lundmark and Jacek Szmigielski. Multi-peakon solutions of the Degasperis–Procesi equation. *Inverse Problems*, 19:1241–1245, December 2003.

[26] Hans Lundmark and Jacek Szmigielski. Degasperis–Procesi peakons and the discrete cubic string. *IMRP Int. Math. Res. Pap.*, 2005(2):53–116, 2005.

[27] Ian G. Macdonald. *Symmetric functions and Hall polynomials.* Oxford University Press, New York, second edition, 1995.

[28] Alexander V. Mikhailov and Vladimir S. Novikov. Perturbative symmetry approach. *J. Phys. A*, 35(22):4775–4790, 2002.

[29] Jürgen Moser. Three integrable systems connected with isospectral deformations. *Advances in Mathematics*, 16:197–220, 1975.

[30] Laurent Schwartz. *Théorie des distributions. Tome I.* Actualités Sci. Ind., no. 1091 = Publ. Inst. Math. Univ. Strasbourg 9. Hermann & Cie., Paris, 1950.