Commun. Math. Phys. 81, 189–202 (1981)

Communications in
Mathematical
Physics

Conservation Laws and Symmetries of Generalized Sine-Gordon Equations

B. A. Kupershmidt[1] and George Wilson[2]

1 University of Michigan, Ann Arbor, MI 48109, USA
2 Mathematical Institute, 24–29 St Giles, Oxford OX1 3LB, England

Abstract. We study some systems of non-linear PDE's (Eqs. 1.1 below) which can be regarded either as generalizations of the sine-Gordon equation or as two-dimensional versions of the Toda lattice equations. We show that these systems have an infinite number of non-trivial conservation laws and an infinite number of symmetries. The second result is deduced from the first by a variant of the Hamiltonian formalism for evolution equations. We also consider some specializations of the systems.

1. Introduction

The title refers to the following system of equations for n unknown functions $R_0(x,t), \dots, R_{n-1}(x,t)$:

$$R_{i,xt} = c_{i-1} \exp(R_{i-1} - R_i) - c_i \exp(R_i - R_{i+1}). \tag{1.1}$$

The c_i are constants, and the suffixes are read mod n where necessary. It follows from (1.1) that $(\Sigma R_i)_{xt} = 0$, and the most interesting case is when $\Sigma R_i = 0$ too, so that there are really only $n-1$ independent unknown functions in (1.1), say $R_0, \dots, R_{n-2}$. However, even in this case it is more pleasant to write the equations in the symmetrical form (1.1).

These equations have been studied recently by several other authors (see [4, 9, 10]). The work [10] is in some respects more advanced than ours: we did not see either this paper or [9] until the present manuscript had been completed. We feel that since our point of view is rather different from that of [10], it is best to present our results without any alteration. However, at the end of the introduction we have inserted a few comments comparing our results with those of [10].

Let us first explain how the Eqs. (1.1) arise from our point of view. In the simplest case $n = 2$ and $R_0 + R_1 = 0$, we have just one unknown $R = R_0$, and the equation is

$$R_{xt} = c_1 \exp(-2R) - c_0 \exp(2R). \tag{1.2}$$

For suitable values of c_i we get the well known sinh-Gordon equation

$$R_{xt} = \sinh 2R.$$

The factor 2 is of course inessential, and could be removed by rescaling. Replacing

0010-3616/81/0081/0189/$02.80

R by iR, we could get the even more popular sine-Gordon equation

$$R_{xt} = \sin R.$$

That is why we call the systems (1.1) 'generalized sine-Gordon equations'. In this paper we are concerned only with purely algebraic properties of the equations, so the substitution $R \to iR$ is harmless.

Now, it is well known that the sinh-Gordon equation, or more generally Eq. 1.2, is closely related to the Korteweg–de Vries (KdV) equation. The connection is via the 'modified' KdV equation for the variable $r = R_x$: the modified KdV equation and the sinh–Gordon equation both have Lax representations

$$\partial_t L = [L, P] \tag{1.3}$$

with the same 'scattering operator' L (it is a first order operator with 2×2 matrix coefficients). On the other hand, the KdV equation itself has a Lax representation in which L is the Schrödinger operator $\xi^2 + u$ ($\xi \equiv \partial/\partial x$). The position of the system (1.1) can now be explained as follows: it is related to the Lax equations based on a scalar n^{th} order operator

$$L = \xi^n + u_{n-2}\xi^{n-2} + \ldots + u_1\xi + u_0$$

in the same way that Eq. 1.2 is related to the KdV equation. The 'modified Lax equations' needed to make this connection were introduced in [6].

To avoid confusion, we point out that although we speak of a 'Lax representation' for the system (1.1), this system is not what we call a 'Lax equation': as in, for example [6, 12], we reserve that term for equations having a representation (1.3) in which (at least) the entries in the coefficients of P are differential polynomials in those of L. A Lax equation is thus always an evolution equation for the entries in L. The operator P in the Lax representation for (1.1) is not of this kind (indeed, (1.1) is not even an evolution equation); however, we shall see that it is near enough so that we can handle the system (1.1) by only a slight extension of the usual algebraic machinery for Lax equations.

Our two main results about the system (1.1) are simply extensions of the fundamental results on Lax equations. They state that (i) all the (infinitely many) conserved densities for the modified Lax equations are also conserved densities for (1.1); (ii) all the modified Lax equations are symmetries of (1.1) in a rather strong sense that is explained precisely in Sect. 4. We should like to emphasize that both these properties are of a purely algebraic nature: in our view it would be inappropriate to formulate or prove them in terms involving functions decreasing at infinity, transmission coefficients, or other irrelevant complications. The system (1.1) can be viewed algebraically as follows: it is an example of the class of what might be called 'quasi-evolution equations' of the form

$$\partial_t r = F(R); \tag{1.4}$$

here $r_i = R_{i,x}$, r and R are vectors, and F is a function of the R_i and (possibly) their x-derivatives $R_i^{(j)}$. The ∂_t in (1.4) can now be regarded as defining a derivation of the algebra of functions of $r_i^{(j)}$, with values in the larger algebra of functions of $R_i^{(j)}$. Studying the algebraic properties of the Eq. 1.4 amounts to studying this derivation. The situation should be compared with the more familiar one of an evolution equation $\partial_t r = f(r)$: here ∂_t can be viewed as a derivation of the algebra of functions of $r_i^{(j)}$ into itself.

The paper is organized as follows. Section 2 summarizes the material on modified Lax equations that we need from our previous paper [6]: we refer the reader to [6] for proofs and more details. Section 3 derives the Lax representation for Eq. 1.1, following the example of AKNS [1] for the sine-Gordon equation. The formulation and proof of the two main results are in Sect. 4. The formulations are probably of more interest than the proofs, which we omit anyway, since they are exactly the same as for Lax equations. The two results can be proved independently of each other, but it is more interesting to deduce one from the other by Hamiltonian formalism: the necessary extension of the usual Hamiltonian formalism is explained in Sect. 5. Section 6 considers the ‘specializations’ of (1.1) obtained by making the basic operator L in the Lax representation skew-adjoint; this gives equations resembling (1.1), but in about half as many variables. The main problem here is to check which of the conservation laws for (1.1) remain non-trivial for the specialization. Finally, in an appendix we have discussed the relationship of the approaches of Lax [7] and AKNS [1] to ‘integrable’ equations. This will (we hope) be well known to many readers, and is essentially explained already in Lax’s article [8]; nevertheless, we have the impression that the situation is not as widely understood as it should be, so it seemed worth while to set it out in black and white.

To end the introduction we offer a few comments on the relationship between our work and the papers [9, 10]. These papers approach the system (1.1) from a different angle: they view it as a two-dimensional version of the periodic Toda lattice. Indeed, if in (1.1) we regard R_i as functions of just one variable t, replace the left hand side by $R_{i,tt}$ and take all the constants $c_i = 1$, we have exactly the equations of the periodic Toda lattice. (The finite non-periodic Toda lattice is also included as the case when $c_0 = 0$.) Now, as Bogoyavlensky first pointed out [2], the n-periodic Toda lattice is the special case A_{n-1} of a construction that can be carried out starting off from any irreducible root system: the paper [10] considers the two-dimensional versions of these generalized Toda lattices. The equations that we obtain by specialization in Sect. 6 are among these: for $n = 2k$ we get the system corresponding to C_k, and for $n = 2k + 1$ we get the system of [10] corresponding to the non-reduced root system BC_k (the possibility of the specialization $A_{2k} \to BC_k$ is mentioned in [9, 10]). However, the remaining irreducible root systems do not seem to arise so naturally from our point of view. In [10] the authors obtain ‘zero curvature’ representations for these systems, and use them to construct conservation laws (though they do not state which of the conservation laws constructed are non-trivial). The papers [9, 10] do not contain any analogue of our second main result, concerning the symmetries; however, we note that previously Zhiber and Shabat [13] had proved the existence of infinitely many symmetries for the equation

$$R_{xt} = \alpha \exp(-2R) + \beta \exp R.$$

This is our specialized equation in the simplest case $n=3$ (corresponding to the root system BC_1).

We are grateful to A. M. Perelomov for letting us see a preprint of the interesting paper [10].

2. Modified Lax Equations

The modified Lax equations of [6] are evolution equations for unknown functions $v_1(x,t), \ldots, v_{n-1}(x,t)$. They have a Lax representation based on the following first order $n \times n$ matrix operator L: let $\omega = \exp(2\pi i/n)$, let Ω be the diagonal matrix

$$\Omega = \operatorname{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}), \tag{2.1}$$

and let circ $(a_0, \ldots, a_{n-1})$ denote the circulant matrix[1] whose first row is the vector indicated. Then

$$L = \Omega . \operatorname{circ}(\xi, v_1, \ldots, v_{n-1}) \tag{2.2}$$

(recall that $\xi \equiv \partial/\partial x$).

Let $B = \mathbb{C}[v_i^{(j)}], j \geqq 0$, be the algebra of differential polynomials in the v_i with the usual derivation $\partial v_i^{(j)} = v_i^{(j+1)}$. We denote by $M_n(B)$ the algebra of $n \times n$ matrices with entries in B, and by $M_n(B)[\xi, \xi^{-1}]$ the algebra of formal pseudo-differential operators with coefficients in $M_n(B)$. Each element X of this algebra has a unique decomposition $X = X_+ + X_-$, where X_+ and X_- are of the form

$$X_+ = \sum_0^r x_i \xi^i, \quad X_- = \sum_{-\infty}^{-1} x_i \xi^i, \quad x_i \in M_n(B). \tag{2.3}$$

Proposition 2.3. *Let L be given by* (2.2). *Then there is a unique element $X \in M_n(B)[\xi, \xi^{-1}]$ with the following properties:*

(i) *X commutes with L*

(ii) *X is homogeneous of degree* 1 *with respect to the natural grading* ($\deg \xi = 1$, $\deg v_i^{(j)} = j+1$)

(iii) $X = \mathrm{Id}.\ \xi +$ (*terms of negative order*)

(iv) *X is a circulant.*

The modified Lax equations are now defined by

$$\partial_q L = [X_+^q, L] = [L, X_-^q]. \tag{2.4}$$

That is, ∂_q is the '$\partial/\partial t$' of the q^{th} equation of the hierarchy. The equations are non-trivial except when q is a multiple of n.

The connection of these equations with scalar Lax equations is as follows. Let

$$s: (\text{circulant operators}) \to B[\xi, \xi^{-1}]$$

be the homomorphism that adds up the entries in one row (or column) of a circulant. Let $\tilde{L} = s(L^n)$, $\tilde{X} = s(X)$. Thus $\tilde{L}$ is a scalar operator of the form

$$\tilde{L} = \xi^n + u_{n-2}\xi^{n-2} + \ldots + u_0, \quad u_i \in B.$$

1 That is, the matrix whose (i,j) entry is a_{j-i} (indices run from 0 to $n-1$, and $j-i$ is read mod n)

Proposition 2.5. (i) *The equations* $\partial_q \tilde{L} = [\tilde{X}^q_+, \tilde{L}]$ *implied by (2.4) are just the usual Lax equations formed from* $\tilde{L}$. (ii) *The operator* $\tilde{L}$ *factorizes*

$$\tilde{L} = (\xi + r_{n-1}) \dots (\xi + r_1)(\xi + r_0),$$

where we have set

$$r_i = \sum_j \omega^{ij} v_j. \tag{2.6}$$

Remark 2.7. In the paper [4] the factorization of $\tilde{L}$ is taken as the starting point. From this point of view our operator L arises as follows. There is a standard way of using the factorization of $\tilde{L}$ to rewrite the n^{th} order equation $\tilde{L}\psi = \lambda^n \psi$ as a first order system; however, in the resulting system $L\psi = \lambda\psi$ the leading coefficient of L is not diagonal (it is a circulant). If we diagonalize it, we arrive at the L in (2.2).

Remark 2.8. It is sometimes convenient to consider slightly more general equations involving an extra variable v_0: we start off from the operator

$$L = \Omega.\text{circ}(\xi + v_0, v_1, \dots, v_{n-1})$$

and proceed as before. We have $\partial_q v_0 = 0$ for all q, so the case $v_0 = 0$ is the one of most interest. In the general case $v_0 \neq 0$, the roots $r_0, \dots, r_{n-1}$ of $\tilde{L}$ are independent variables ($v_0 = 0$ clearly corresponds to $\Sigma r_i = 0$): that makes the Hamiltonian form of the modified equations a little more pleasant in this case.

The Hamiltonian form is as follows. Define the 'Hamiltonians' $H_q \in B$ by

$$H_q = q^{-1} \text{ tr res } X^q. \tag{2.9}$$

(We recall that the *residue* of an operator is the coefficient of ξ^{-1}.)

Proposition 2.10. *The modified Lax equations (2.4) can be written in the form*

$$\partial_q v_i = -(1/n)\partial \frac{\delta H_q}{\delta v_{n-i}}$$

It will be useful to have the corresponding expression in terms of the variables r_i in (2.6). From [6], Sect. 6, we see that the skew matrix defining the Hamiltonian structure in terms of these variables is $D\ell D^*$, where $D = (\omega^{ij})$ is the (Fréchet) Jacobian of r with respect to v, and ℓ is the skew matrix for the variables v_i. A short calculation gives the following.

Proposition 2.11. *When written in terms of the variables* r_i, *the modified Lax equations take the form*

$$\partial_q r = S\partial \frac{\delta H_q}{\delta r},$$

where

(i) *in the case* $v_0 \neq 0$, $r = (r_0, \dots, r_{n-1})^t$ *and* $S = -\text{Id}$

(ii) *in the case* $v_0 = 0$, $r = (r_0, \dots, r_{n-2})^t$ *and* $S = n^{-1}E - \text{Id}$; *here* E *is the* $(n-1) \times (n-1)$ *matrix with* $E_{ij} = 1$ *for all* i, j.

Proposition 2.12. *The Hamiltonians* H_q *are conserved densities for all the modified*

Lax equations (2.4), that is, we have $\partial_r H_q \in \partial B$ for all r. Except when q is a multiple of n, these conserved densities are non-trivial, that is, $H_q \notin \partial B$.

It is clear that H_q is homogeneous of degree $q+1$ with respect to the grading on B ($\deg v_i^{(j)} = j+1$). Thus the modified Lax equations have a non-trivial conserved density of every degree not of the form $an+1$.

Finally, in Sect. 6 we shall use the fact that the H_q can be calculated from the scalar operator $\tilde{L}$.

Proposition 2.13. *We have*

$$H_q \equiv (n/q)\operatorname{res} \tilde{X}^q \mod \partial B.$$

3. The Lax representation of (1.1)

According to AKNS [1], the sinh-Gordon equation can be represented as the compatibility condition for the system

$$\left.\begin{aligned} L\psi &= \lambda\psi \\ \partial_t\psi &= A\lambda^{-1}\psi \end{aligned}\right\},$$

where A is a matrix of functions and L is our operator (2.2) for $n=2$. As Lax pointed out (see [8] and the appendix below), this means that the equation has a Lax representation

$$\partial_t L = [L, P] \tag{3.1}$$

in which P is a matrix of functions times L^{-1}. We shall now calculate the corresponding equations of this kind for any value of n, L being given by (2.2). For (3.1) to be consistent, we want P to be a circulant: we therefore seek P in the form

$$P = \Omega B L^{-1}, \qquad B = \operatorname{circ}(b_0, \dots, b_{n-1});$$

the b_i are functions whose relationship to the v_i is yet to be determined. Let us set $V = \operatorname{circ}(v_0, \dots, v_{n-1})$, so that $L = \Omega(\xi + V)$; (the reader can set $v_0 = 0$ if he wishes). A short calculation then shows that (3.1) is equivalent to the equations

$$\left.\begin{aligned} \partial B &= (V - \Omega^{-1} V \Omega) B \\ \partial_t V &= \Omega B \Omega^{-1} - B \end{aligned}\right\}. \tag{3.2}$$

(The first of these equations expresses the condition $[L,P]_- = 0$.) Explicitly, Eqs. (3.2) say that

$$\left.\begin{aligned} \partial b_i &= \sum_k (1-\omega^k) v_k b_{i-k} \\ \partial_t v_i &= (\omega^{-i} - 1) b_i \end{aligned}\right\} \tag{3.3}$$

(All indices and summations run from 0 to $n-1$, and indices are read mod n.) Introduce new variables

$$r_i = \sum_j \omega^{ij} v_j, \quad x_i = \sum_j \omega^{ij} b_j$$

(so the r_i are as in Sect. 2). Then Eqs. (3.3) become

$$\left.\begin{aligned}\partial x_i &= (r_i - r_{i+1})x_i \\ \partial_t r_i &= x_{i-1} - x_i\end{aligned}\right\}.$$

Introducing variables $R_0, \ldots, R_{n-1}$ with $\partial R_i = r_i$, we can 'solve' the first of these equations to get

$$x_i = c_i \exp(R_i - R_{i+1});$$

putting this into the second equation yields the system (1.1).

Remark. Some readers may wonder why we do not go further and consider equations of the form (3.1) with P a polynomial in L^{-1} of order 2, 3, 4, ... The reason is that the coefficients of such operators P (and the resulting equations) would involve increasing numbers of 'integrations' of elements of our original algebra of differential polynomials in the v_i; we do not like that.

4. Conservation Laws and Symmetries

In order to discuss the algebraic properties of the system (1.1) without introducing irrelevant analytic considerations, we proceed as follows. Let

$$B = \mathbb{C}[v_0^{(j)}, \ldots, v_{n-1}^{(j)}] = \mathbb{C}[r_0^{(j)}, \ldots, r_{n-1}^{(j)}]$$

be as in Sect. 2. (For definiteness let us consider the general case $v_0 \neq 0$; the case $v_0 = 0$ is exactly the same: we then take $B = \mathbb{C}[r_0^{(j)}, \ldots, r_{n-2}^{(j)}]$.) We form the larger algebra

$$\hat{B} = B[\exp R_i, \exp(-R_i)], \quad 0 \leqq i \leqq n-1.$$

Thus an element of $\hat{B}$ is a Laurent polynomial in the symbols $\exp R_i$ with coefficients in B. The derivation ∂ is extended to $\hat{B}$ by setting

$$\partial \exp R_i = r_i \exp R_i.$$

(Naturally, we have in mind that $\partial R_i = r_i$, but at this stage we do not want to introduce the R_i themselves into our algebra.) The grading of B is extended to $\hat{B}$ by giving $\exp R_i$ degree zero. The derivation ∂ still increases degree by 1, and its kernel still consists just of the constants: these properties will ensure the validity of the arguments from [12] that we refer to below in the proofs of Theorems 4.1 and 4.2. The algebra $\hat{B}$ is the smallest algebra containing all the expressions arising in our study of the system (1.1).

We now let

$$\partial_t : B \to \hat{B}$$

be the derivation defined by the properties

(i) ∂_t commutes with ∂

(i) $\partial_t r_i = c_{i-1} \exp(R_{i-1} - R_i) - c_i \exp(R_i - R_{i+1})$.

It is clear that there is a unique ∂_t with these properties. The derivation ∂_t embodies the algebraic properties of the system (1.1).

Theorem 4.1. *The Hamiltonians H_q for the modified Lax equations (see Sect. 2) are also conserved densities for Eq. 1.1; that is, we have*

$$\partial_t H_q = \partial J_q \quad \textit{for some } J_q \in \hat{B}.$$

Proof. Using the Lax representation for (1.1) given in Sect. 3, it is easy to check that the proof of the corresponding fact for Lax equations given in [12] is still valid.

Next we formulate the result about symmetries, which is the counterpart for (1.1) of the fact that the flows of Lax equations based on the same operator L commute. As in Sect. 2, let $\partial_q : B \to B$ be the derivations corresponding to the modified Lax equations. From the Hamiltonian form of these equations (see (2.11)), the ∂_q have the form

$$\partial_q r_i = \partial y_i$$

for some $y_i \in B$ (depending of course on q). It follows that ∂_q has a natural extension to an evolutionary[2] derivation of any algebra of C^∞ functions of the variables $R_i^{(j)}$: the extension is determined by

$$\partial_q R_i = y_i.$$

In particular, we have a natural extension of ∂_q to a derivation $\hat{\partial}_q : \hat{B} \to \hat{B}$, with

$$\hat{\partial}_q \exp R_i = y_i \exp R_i.$$

Theorem 4.2. *The modified Lax equations are symmetries of (1.1) in the sense that the following diagrams commute:*

$$\begin{array}{ccc} B & \xrightarrow{\partial_t} & \hat{B} \\ {\scriptstyle \partial_q}\downarrow & & \downarrow{\scriptstyle \hat{\partial}_q} \\ B & \xrightarrow{\partial_t} & \hat{B} \end{array}$$

Proof. It is enough to show that

$$(\partial_t \partial_q - \hat{\partial}_q \partial_t) L = 0$$

where L is the operator (2.2). That can be done by the same argument as for Lax equations (see [12], Sect. 3). Alternatively, one can deduce (4.2) from (4.1) by (an extension of) the Hamiltonian formalism: we shall do that in the next section.

Remark. Our introduction of the algebra $\hat{B}$ in this section was motivated by the desire to work in the smallest algebra possible (clearly, the smaller the algebra in which the 'fluxes' J_q can be asserted to lie, the more content theorem 4.1 will have). In the next section, however, for the sake of indicating a general theory we shall work with larger algebras: we shall regard ∂_t as a derivation from the algebra of C^∞ functions of $r_i^{(j)}$ to the algebra of C^∞ functions of $R_i^{(j)}$. For the purpose of proving the commutativity of the diagrams in (4.2) it makes no difference which kind of algebra we work with, because in both situations evolutionary derivations are uniquely determined by their values on the basic variables r_i.

2 That is, commuting with ∂

5. Hamiltonian Formalism

Let $R_0, \dots, R_{N-1}$ be independent variables (in our application we shall have $N = n$ or $n - 1$, but for the moment we want to describe some general machinery). Let $A(R)$ denote the differential algebra of C^∞ functions of $R_i^{(j)}, j \geqq 0$ (with some domain of definition that we need not specify). We set $r_i = \partial R_i$, so that we have an inclusion of differential algebras

$$A(r) \subset A(R).$$

Let S be a constant symmetric $N \times N$ matrix, and suppose we are given on $A(r)$ the Hamiltonian structure defined by the skew operator $S\partial$. That means that to each function $f \in A(r)$ we assign the 'Hamiltonian vector field' (evolutionary derivation) $\partial_f : A(r) \to A(r)$ such that

$$\partial_f r = S\partial \frac{\delta f}{\delta r} \tag{5.1}$$

(vector notation: $r = (r_0, \dots, r_{N-1})^t$, etc.).

Now, (5.1) can be written

$$\partial \partial_f R = \partial S \frac{\delta f}{\delta r}.$$

Hence we have the following.

Proposition 5.2. *Every Hamiltonian vector field ∂_f, $f \in A(r)$, extends to an evolutionary derivation $\hat{\partial}_f$ of $A(R)$ defined by*

$$\hat{\partial}_f R = S \frac{\delta f}{\delta r}.$$

Now, each $f \in A(r)$ can be regarded as lying in $A(R)$, so we can form the variational derivatives $\delta f / \delta R_i$. It is easy to see that we have

$$\frac{\delta f}{\delta R} = -\partial \frac{\delta f}{\delta r}.$$

Thus (5.1) can also be written

$$\partial_f r = -S \frac{\delta f}{\delta R}.$$

In this form the formula still makes sense for $f \in A(R)$.

Proposition-definition 5.3. *Let $F \in A(R)$. Then we assign to F the (unique) evolutionary derivation*

$$\partial_F : A(r) \to A(R)$$

defined by

$$\partial_F r = -S \frac{\delta F}{\delta R}. \tag{5.4}$$

If it happens that $F \in A(r)$, then this derivation takes values in $A(r)$ and coincides with the ∂_F defined by (5.1).

In contrast to (5.2), note that if $F \notin A(r)$, then ∂_F does not necessarily have any extension to a derivation of $A(R)$ into itself.

If now $F \in A(R)$ and $g \in A(r)$, we can define their 'Poisson bracket' by

$$\{F, g\} = \partial_F g = -\hat{\partial}_g F \in A(R)/\mathrm{Im}\,\partial.$$

Then the main fact expressing the 'quasi-Hamiltonian' character of this bracket is as follows.

Proposition 5.5. *For any $F \in A(R)$, $g \in A(r)$, we have*

$$\partial_{\{F,g\}} = [\partial_F, \partial_g].$$

(The last bracket means $\partial_F \partial_g - \hat{\partial}_g \partial_F$, so both sides of the equation are derivations from $A(r)$ to $A(R)$.)

We omit the proof of (5.5), which is just like the proof that the operator $S\partial$ is Hamiltonian in the usual sense (see, for example [5]).

It follows in particular from (5.5) that if $g \in A(r)$ is a conserved density for Eq. 5.4, that is, if $\partial_F g \in \mathrm{Im}\,\partial$, then $[\partial_F, \partial_g] = 0$; that is, the equation $\partial_t r = S\partial\delta g/\delta r$ is a symmetry of Eq. 5.4 in the sense discussed in Sect. 4. Since the modified Lax equations have this form (see (2.11)), in order to deduce (4.2) from (4.1) we have only to check that our Eq. (1.1) can be written in the form (5.4), with S as in (2.11).

Proposition 5.6. *The system (1.1) can be written in the form*

$$\partial_t r = -S\frac{\delta H}{\delta R},$$

where S is as in (2.11), and

$$H = -\operatorname{tr}\operatorname{res} P = -\Sigma c_i \exp(R_i - R_{i+1}).$$

That is most easily checked by direct calculation.

The 'quasi-Hamiltonian formalism' that we have been using is a special case of a fairly general set-up: we end this section by sketching the general theory. Readers who find the following discussion too brief could consult [6], Sect. 5 and 6. Suppose we have two sets of variables (u_i), (v_i), and an inclusion

$$A(u) \subset A(v).$$

(More generally, we could consider a homomorphism of differential algebras $\varphi: A(u) \to A(v)$, but to simplify the notation we suppose φ injective and suppress it.) Let D be the Fréchet Jacobian of u with respect to v. Let ℓ be a skew matrix defining a Hamiltonian structure on $A(u)$; and suppose that ℓ has the form

$$\ell = CD^* = -DC^* \tag{5.7}$$

for some matrix C of differential operators (with coefficients in $A(v)$). Then for each $f \in A(u)$, the Hamiltonian vector field ∂_f on $A(u)$ determined by ℓ has an extension to an evolutionary derivation $\hat{\partial}_f$ of $A(v)$, defined by

$$\hat{\partial}_f v = -C^*\frac{\delta f}{\delta u}, \quad f \in A(u).$$

(That follows at once from the formula

$$\partial_t u = D\partial_t v,$$

valid for any evolutionary derivation ∂_t of $A(v)$.)

Now, for $F \in A(v)$, we can define an evolutionary derivation

$$\partial_F : A(u) \to A(v)$$

by the requirement

$$\partial_F u = C\frac{\delta F}{\delta v}.$$

The formula

$$\frac{\delta f}{\delta v} = D^* \frac{\delta f}{\delta u}, \quad f \in A(u),$$

shows that if $F \in A(u)$, then this ∂_F agrees with our original one. We can now define Poisson brackets $\{F, g\}$ as before, and we say the triple (C, D, ℓ) satisfying (5.7) is *quasi-Hamiltonian* if the bracket-preserving condition in (5.5) holds. Earlier, we were dealing with the special case $\ell = S\partial, D = \mathrm{Id}.\partial, C = -S$. Slightly more generally, arguments like those in [5] would show that any triple of constant coefficient operators (C, D, ℓ) is quasi-Hamiltonian (also in the generalization to the case of more than one space variable).

This set-up should be compared with the one discussed in [6], Sect. 6. There ℓ had the form $\ell = D\ell_1 D^*$, where ℓ_1 was a skew matrix defining a Hamiltonian structure on $A(v)$. That is of course a special case of what we had above, where C has the form $C = D\ell_1$. The effect of this special form for C is that all the derivations $\partial_F : A(u) \to A(v)$ defined above in fact extend to derivations of $A(v)$ into itself (the extension is just the Hamiltonian vector field determined by ℓ_1). In the case where C is not of the form $D\ell_1$, however, ∂_F has in general no such extension, and we have, so to speak, only one and a half Hamiltonian structures rather than two.

6. Specializations

In this section we consider the specializations (sometimes called 'reductions') of the modified Lax equations, and of the system (1.1), obtained by requiring the basic operator L of (2.2) to be skew-adjoint.[3] That has the effect of cutting down the number of independent variables v_i or r_i to $[n/2]$, that is, to $n/2$ if n is even and to $(n-1)/2$ if n is odd. Of course we take $n > 2$, since for $n = 2$, L is already skew.

In terms of the variables v_i or r_i, the skew-adjointness condition is

$$v_i = -\omega^i v_{-i} \quad \text{or} \quad r_i = -r_{-i-1} \tag{6.1}$$

(suffixes mod n as usual). As our basic independent variables we can take $r_0, r_1, \ldots, r_{[n/2]-1}$.

3 We always take adjoints in the 'real' sense; that is, the adjoint of a complex number is itself, not its complex conjugate

Our first task is to determine which of the modified Lax equations and conserved densities H_q survive this specialization of L (see [6], Sect. 3). That is very easy. First, the extra consistency condition for the modified Lax equation $\partial_q L = [L, X^q_-]$ is just that the right hand side should be skew-adjoint, that is, essentially, that X^q should be skew-adjoint, which happens only when q is odd (the skew-adjointness of L is equivalent to that of X). Thus only the modified Lax equations with q odd survive the specialization (remain consistent). The same happens for the conserved densities.

Proposition 6.2. *The conserved density* $H_q = q^{-1}\mathrm{tr}\ \mathrm{res}\ X^q$ *remains non-trivial when we make L skew-adjoint if and only if q is odd (and not a multiple of n).*

Proof. First, if q is even, then X^q is self-adjoint, so that $H_q = 0$. So now suppose q is odd; by (2.13), it is enough if we show that $\mathrm{res}\ \tilde{X}^q = \mathrm{res}\ \tilde{L}^{q/n}$ is not in $\mathrm{Im}\ \partial$. Now, we have

$$\tilde{L} = \xi^n + u_{n-2}\xi^{n-2} + \dots$$

where

$$u_{n-2} = -\sum_{0}^{[n/2]-1} r_i^2 + (\text{linear combination of } \partial r_i).$$

We calculate what remains of $\tilde{L}^{q/n}$ when we put all derivatives of r_i, and also all coefficients of $\tilde{L}$ except u_{n-2}, equal to zero. Since killing the derivatives makes everything commutative, that can be done by the binomial theorem: we get

$$(\xi^n + u_{n-2}\xi^{n-2})^{q/n} = \xi^q(1 + u_{n-2}\xi^{-2})^{q/n}.$$

Since a 'fractional binomial coefficient' is never zero, we see that (if q is not a multiple of n) $\mathrm{res}\ \tilde{X}^q$ contains a term that is a non-zero multiple of Σr_i^{q+1}; hence obviously $\tilde{X}^q \notin \mathrm{Im}\ \partial$.

Remark. This argument can also be used to prove the non-triviality of the conserved densities for (unspecialized) scalar Lax equations: indeed, it is the kind of argument that was originally used for the KdV equation (see [11]). However, it does not work for matrix Lax equations: that is why we preferred to give a different argument in [6], Sect. 7.

Proposition 6.3. *The generalized sine-Gordon equation* (1.1) *remains consistent under the specialization* (6.1) *if and only if the constants* c_i *satisfy the condition* $c_i = c_{-i-2}$.

That is trivial to check. Naturally, the condition on the c_i is equivalent to the operator P in the Lax representation of (1.1) being skew-adjoint. The effect of (6.3) is that the specialized systems are obtained simply by writing down the first $[n/2]$ equations in (1.1) and substituting for the extraneous variables from the conditions $R_i = -R_{-i-1}$. It follows automatically from (4.1) and (4.2) that these systems will have infinitely many conserved densities (the surviving H_q) and symmetries (the surviving modified Lax equations).

Let us write out the simplest examples, for $n = 3, 4, 5$. For $n = 3$, we have

$r_0 + r_2 = r_1 = 0$; setting $R = R_0$ we get the equation

$$R_{xt} = c_2 \exp(-2R) - c_0 \exp R. \tag{6.4}$$

For $n = 4$, we have $r_0 + r_3 = r_1 + r_2 = 0$, giving

$$\left.\begin{aligned} R_{0,xt} &= c_3 \exp(-2R_0) - c_0 \exp(R_0 - R_1) \\ R_{1,xt} &= c_0 \exp(R_0 - R_1) - c_1 \exp(2R_1) \end{aligned}\right\}. \tag{6.5}$$

For $n = 5$, we have $r_0 + r_4 = r_1 + r_3 = r_2 = 0$, giving

$$\left.\begin{aligned} R_{0,xt} &= c_4 \exp(-2R_0) - c_0 \exp(R_0 - R_1) \\ R_{1,xt} &= c_0 \exp(R_0 - R_1) - c_1 \exp R_1 \end{aligned}\right\}. \tag{6.6}$$

From (6.2) we get the following.

Proposition 6.7. (i) *The equation (6.4) has a non-trivial conserved density of every even degree not of the form* $3a + 1$.

(ii) *The system (6.5) has a non-trivial conserved density of every even degree.*

(iii) *The system (6.6) has a non-trivial conserved density of every even degree not of the form* $5a + 1$.

Here it is understood that the conserved densities are polynomials in the $r_i^{(j)}$, and we recall that $r_i^{(j)}$ has degree $j + 1$. The equation (6.4) is first mentioned (as far as we know) in the paper [3] of Dodd and Bullough, who already noticed that it had many conserved densities. (They thought it had only a finite number, but as we have seen, that was not right.)

Appendix: Lax and AKNS Representations

In the AKNS approach to 'integrable' equations [1], the equation of interest is represented as the compatibility condition for a system

$$\left.\begin{aligned} L\psi &= \lambda\psi \\ \partial_t\psi &= A(\lambda)\psi \end{aligned}\right\}. \tag{A1}$$

Here L is a first order matrix ordinary differential operator with leading coefficient an invertible constant diagonal matrix, and A is a matrix of functions (not operators) depending rationally (often polynomially) on the 'spectral parameter' λ.

On the other hand, the Lax equation $\partial_t L = [P, L]$ can be viewed as the compatibility condition for the system

$$\left.\begin{aligned} L\psi &= \lambda\psi \\ \partial_t\psi &= P\psi \end{aligned}\right\}. \tag{A2}$$

Here P is an operator (differential or formal pseudo-differential) and does not depend on λ.

The connection between the points of view (A1) and (A2) is very simple. Since L is of order 1 with invertible leading coefficient, it is clear that every (formal pseudo-differential) operator has a unique expansion in the form $\sum_{-\infty}^{r} m_i L^i$, the m_i being matrices of functions. In particular, the operator P in (A2) can be written in the form

$$P = \Sigma p_i L^i. \tag{A3}$$

(If P is a differential operator, only non-negative powers of L will occur.) The corresponding $A(\lambda)$ in (A1) is then just

$$A(\lambda) = \Sigma p_i \lambda^i. \tag{A4}$$

Conversely, given $A(\lambda)$ in the form (A4), we can write down the corresponding operator P in the form (A3). Thus (in the case when L has order 1) the only difference between the Lax and AKNS approaches is that in the latter one chooses always to represent the operator P in the form (A3) (and then suppresses it by writing (A4) instead).

References

1. Ablowitz, M. J., Kaup, D. J., Newell, A. C., Segur, H.: Stud. Appl. Math. **53**, 249–315 (1974)
2. Bogoyavlensky, O. I.: Commun. Math. Phys. **51**, 201–209 (1976)
3. Dodd, R. K., Bullough, R. K.: Proc. R. Soc. Lond. A**352**, 481–503 (1977)
4. Fordy, A. P., Gibbons, J.: Commun. Math. Phys. **77**, 21–30 (1980)
5. Gel'fand, I. M., Manin, Yu. I., Shubin, M. A.: Funct. Anal. Appl. **10:4**, 30–34 (Russian); 274–278 (English) (1976)
6. Kupershmidt, B. A., Wilson, G.: Invent. Math. **62**, 403–436 (1981)
7. Lax, P. D.: Commun. Pure Appl. Math. **21**, 467–490 (1968)
8. Lax, P. D.: Adv. Math. **16**, 368–379 (1975)
9. Mikhailov, A. V.: Pis'ma Zh. Eksp. Teor. Fiz. **30**, 443–448 (1979); JETP Lett. **30**, 414–418 (1979)
10. Mikhailov, A. V., Olshanetsky, M. A., Perelomov, A. M.: Commun. Math. Phys. **79**, 473–488 (1981)
11. Miura, R. M., Gardner, C. S., Kruskal, M. D.: J. Math. Phys. **9**, 1204–1209 (1968)
12. Wilson, G.: Math. Proc. Cambridge Philos. Soc. **86**, 131–143 (1979)
13. Zhiber, A. V., Shabat, A. B.: Dokl. Akad. Nauk SSSR **247**, 1103–1107 (1979); Sov. Phys. Dokl. **24**, 607–609 (1979)

Communicated by J. Moser
Received November 3, 1980

Notes added in proof. (i) Everything in this paper can be generalized to the equations associated with simple Lie algebras studied in [10]: these equations all have infinitely many (non-trivial) conservation laws and corresponding symmetries. There is a symmetry of each degree congruent to an exponent of the relevant Lie algebra modulo the Coxeter number, and the degrees of the conserved densities are one more than these. Details will be given in a forthcoming paper by the second author (submitted to 'Ergodic Theory and Dynamical Systems').

(ii) Recent work of Drinfel'd and Sokolov (Dokl. Akad. Nauk SSSR **258**:1, 11-16 (1981)) also contains these results, and shows clearly that the affine (Kac-Moody, Euclidean) Lie algebras provide the correct setting in which to discuss these questions; for example, the equations described in [10] as associated with the root system BC_k are best understood as coming from the 'twisted' affine algebra $A_{2k}^{(2)}$. Drinfel'd and Sokolov also have a far-reaching generalization of the results of our paper [6] concerning the Miura transformation. We are most grateful to Yu. I. Manin for informing us about this work.