

The modified Lax and two-dimensional Toda lattice equations associated with simple Lie algebras

GEORGE WILSON

Mathematical Institute, 24–29 St Giles, Oxford OX1 3LB, England

(Received 25 April 1981, postscript received 18 August 1981)

Abstract. We associate to each complex simple Lie algebra \mathfrak{g} a hierarchy of evolution equations; in the simplest case $\mathfrak{g} = \mathfrak{sl}(2)$ they are the modified KdV equations. These new equations are related to the two-dimensional Toda lattice equations associated with \mathfrak{g} in the same way that the modified KdV equations are related to the sinh-Gordon equation.

1. Introduction

Recently there has been considerable interest in the ‘two-dimensional (periodic) Toda lattice’ (2DTL) equations (see [5], [6], [10], [11], [13], [14]). These are the equations:

$$(1.1) \quad w_{i,x,t} = \exp(w_{i-1} - w_i) - \exp(w_i - w_{i+1})$$

for n unknown functions $w_1(x, t), \dots, w_n(x, t)$ with $\sum w_i = 0$ (the subscripts are read mod n , so that $w_0 \equiv w_n, w_{n+1} \equiv w_1$). If we consider solutions that depend only on the variable $\tau = x + t$, then (1.1) reduces to the usual (one-dimensional) periodic Toda lattice.

Now, as Bogoyavlensky [3] pointed out in the one-dimensional case, we can form an equation resembling (1.1) from any (complex) simple Lie algebra \mathfrak{g} : (1.1) is the case $\mathfrak{g} = \mathfrak{sl}(n)$. I refer to [11], [14] and § 5 below for the exact definition of these ‘generalized’ 2DTL equations, but the main clue is that the expressions $w_i - w_{i+1}$ occurring in the exponents in (1.1) remind us of the simple roots (and minus the maximal root) for $\mathfrak{sl}(n)$. The papers [5], [11] and [14] find ‘zero curvature representations’ (see [19])

$$(1.2) \quad [\partial - U, \partial_t - V] = 0 \quad (\partial \equiv \partial/\partial x, \partial_t \equiv \partial/\partial t)$$

for these generalized equations; moreover, [14] indicates a construction of an infinite number of local conservation laws for the equations (in the case $\mathfrak{g} = \mathfrak{sl}(n)$ that had been done before in [6] and [13]). However, the paper [14] does not determine which of the conservation laws are non-trivial, surely a serious omission, since it leaves open the possibility that only a finite number, or even none of them are.

In [10], B. A. Kupershmidt and I made a more careful study of the formal properties of the 2DTL equations, but only in the case $\mathfrak{g} = \mathfrak{sl}(n)$ (actually we also treated the equations of [14] corresponding to the root systems C_k, BC_k , though we did not realize that this is what they were until we saw a preprint of [14]). We determined which of the conservation laws for (1.1) are non-trivial, and also showed that to each conservation law there corresponds a symmetry of (1.1); the symmetries are evolution equations for the variables $u_i = \partial w_i$, and they can be interpreted as the Hamiltonian flows corresponding to the conserved quantities of (1.1). The term ‘symmetry’ means, intuitively speaking, that the flows of the evolution equations leave invariant the space of solutions of (1.1).

The purpose of the present paper is to extend to the case of a general simple Lie algebra \mathfrak{g} everything that is done in [10] in the case $\mathfrak{g} = \mathfrak{sl}(n)$. The main results are set out in (1.3) below. Recall that, if \mathfrak{g} is a simple Lie algebra of rank l , the invariant polynomials on \mathfrak{g} form a polynomial algebra with l homogeneous generators of degrees d_1, \dots, d_l , say. The largest of the d_i is called the *Coxeter number* of \mathfrak{g} , and is denoted by h . The numbers $m_i = d_i - 1$ are called the *exponents* of \mathfrak{g} . For the reader’s convenience the values of d_i for the various simple Lie algebras are tabulated in the appendix below.

(1.3) THEOREM. *The 2DTL equation associated with a simple Lie algebra \mathfrak{g} has*

- (i) *a non-trivial conserved density of each degree congruent mod h to one of the d_i ;*
- (ii) *a non-trivial symmetry of each degree congruent mod h to one of the $m_i = d_i - 1$.*

The meaning of the term ‘degree’ in these statements will become clear in the course of the paper. As examples, we see from the table in the appendix that for

$$\mathfrak{g} = \mathfrak{sl}(n) (A_{n-1})$$

we have a conserved density of each degree not congruent to 1 mod n , while for C_l we have one of every even degree. Naturally, this agrees with the results obtained by a different method in [10].

The symmetries of the 2DTL equation are what I call the ‘modified Lax equations’ associated with \mathfrak{g} ; for $\mathfrak{g} = \mathfrak{sl}(n)$ they are indeed the equations studied under that name in [9]; for $\mathfrak{g} = \mathfrak{sl}(2)$ they are the well known modified KdV equations. Although the modified Lax equations are our main object of study, they do not appear in the paper until § 4. §§ 2 and 3 study what, for want of a better name, I call the ‘general zero curvature equations’ associated with \mathfrak{g} : the modified Lax equations are ‘specializations’ or ‘reductions’ of these. Let me explain that further. The U and V in equation (1.2) can be thought of as functions of x and t with values in \mathfrak{g} which, moreover, depend on a ‘spectral’ parameter λ . In this paper U will always have the form

$$U = -u(x, t) + \lambda F,$$

where $F \in \mathfrak{g}$ is a (constant) semi-simple element; and V will depend polynomially on λ (except in § 5). In the ‘general’ equations of §§ 2 and 3, no restriction is placed on u except that it must take values in the image of $\text{ad } F$; but in the modified Lax equations, the values of u are required to lie in a certain Cartan subalgebra of \mathfrak{g} .

Many of the properties of the modified Lax equations are inherited immediately from those of the general equations, and I think the theory becomes clearer if we discuss these first (they are a little easier). The idea of specialization is discussed in considerable detail in [9], § 3.

For brevity, in the proofs of the basic results in § 3 I have used the ugly device of choosing a faithful representation of \mathfrak{g} , so that the U and V in (1.2) become matrices of functions. The advantage of this is that any equation of the form (1.2) (with $U = -u + \lambda F$) can then be rewritten as a Lax equation

$$\partial_t L = [P, L]$$

with L a first-order matrix differential operator (in the case where V is polynomial in λ , P will be a differential operator too). Thus we can read off many results about our equation (1.2) from known facts about Lax equations. In §§ 2, 4 and 5, however, we proceed ‘intrinsically’ (that is, without using a representation of \mathfrak{g}). It would probably be enlightening to formulate all the proofs in intrinsic terms, but I leave that for another occasion.

The construction of the conservation laws that we use here is superficially different from the one suggested in [14]. However, the equivalence of the two constructions is proved in [18] (in the wider context of matrix Lax equations).

The exponents of \mathfrak{g} , which play such a prominent role in (1.3), can be defined in several different ways: the definition I gave above is close to the original one, which was in terms of the cohomology of a group with Lie algebra \mathfrak{g} . For the purposes of the present paper it would be more appropriate to adopt the definition given in [4], in terms of the eigenvalues of a Coxeter transformation (the equivalence of the two definitions is proved in [4], p. 121). The definition using the invariant polynomials suggests that (1.3) probably has a nice interpretation in terms of the formalism of Adler [2] or Reiman and Semenov-Tian-Shansky [15], but I shall not try to discuss that here.

I have tried to make the paper reasonably self-contained, except for some of the proofs: the proofs of the results on the modified Lax equations in § 3 depend on facts proved in [17] and [9], § 7; and in § 5, on the 2DTL equations, I have referred to [10] for the proofs, since the generalization from $\mathfrak{sl}(n)$ to \mathfrak{g} presents no new features here. Also, I have assumed known the most basic facts about simple Lie algebras (they can be learned, for example, from [16]).

This paper is a continuation of work begun in collaboration with Boris Kuper-shmidt. I am indebted to him for introducing me to the 2DTL equations, and for sharing with me his ideas on these matters.

Most of the results to be found below are also announced in the paper [21] of Drinfel’d & Sokolov (which appeared after the present manuscript was submitted for publication). I have added a ‘postscript’ at the end of this paper with the aim of clarifying the relationship between my work and [21].

2. The general zero curvature equations: results

Let \mathfrak{g} be a simple Lie algebra over \mathbb{C} , $F \in \mathfrak{g}$ a regular semi-simple element. ‘Regular’ means that F lies in a unique Cartan subalgebra: this assumption is not essential

(see remark (3.13) below). We denote by \mathfrak{h} the Cartan subalgebra containing F . Let $R \subset \mathfrak{h}^*$ be the set of roots of $(\mathfrak{g}, \mathfrak{h})$, and for each $\alpha \in R$ let E_α be a non-zero element of the corresponding (one-dimensional) root space: then we have the direct sum decomposition

$$(2.1) \quad \mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in R} \mathbb{C}E_\alpha.$$

Let $u_\alpha, \alpha \in R$, be differentially independent variables; we denote by B the algebra $\mathbb{C}[u_\alpha^{(j)}]$ of differential polynomials in the u_α , with the usual derivation ∂ defined by

$$\partial u_\alpha^{(j)} = u_\alpha^{(j+1)}.$$

We make B into a graded algebra by giving each u_α degree 1 and making ∂ increase degree by 1: thus

$$\text{deg } u_\alpha^{(j)} = 1 + j.$$

Set $\tilde{\mathfrak{g}} = B \otimes \mathfrak{g}$. Then $\tilde{\mathfrak{g}}$ is a Lie algebra over B , and ∂ has an obvious extension to $\tilde{\mathfrak{g}}$ with $\partial(1 \otimes \mathfrak{g}) = 0$. We also extend the grading to $\tilde{\mathfrak{g}}$ by letting the elements of \mathfrak{g} have degree zero.

The algebra $\tilde{\mathfrak{g}}$ can be thought of concretely in various ways: if $\{Y_i\}$ is any basis for \mathfrak{g} , then $\tilde{\mathfrak{g}}$ consists of formal sums $\sum f_i Y_i, f_i \in B$ (we write $f_i Y_i$ instead of $f_i \otimes Y_i$). Thus in the case $\mathfrak{g} = \mathfrak{sl}(n)$, an element of $\tilde{\mathfrak{g}}$ is simply a matrix (of trace zero) with entries in B . If we think of the u_α as maps from A to \mathbb{C} (where A could be an open set in \mathbb{C}^2 , or the circle, or other possibilities depending on the context), then the elements of $\tilde{\mathfrak{g}}$ will be maps from A to \mathfrak{g} .

Now set

$$u = \sum u_\alpha E_\alpha \in \tilde{\mathfrak{g}};$$

thus u can be thought of as a ‘general’ map into the subspace of \mathfrak{g} spanned by the E_α (which can also be described as the image of $\text{ad } F$). Let λ be a formal parameter commuting with everything else. We set

$$U = -u + \lambda F.$$

The equations we are going to construct will be evolution equations for the variables $u_\alpha (\alpha \in R)$, and will be represented in the form

$$(2.2) \quad [\partial - U, \partial_t - V] = 0,$$

where V is a polynomial in λ (with coefficients in $\tilde{\mathfrak{g}}$). Equation (2.2) is equivalent to

$$(2.3) \quad \partial_t u = -\partial V + [U, V].$$

For this to reduce to an evolution equation for the u_α , we have to choose V so that the right-hand side of (2.3) (like the left-hand side) is independent of λ and lies in the image of $\text{ad } F$. The possible choices of V are obtained using the next lemma.

(2.4) LEMMA. *For each $v_0 \in \mathfrak{h}$, there is a unique formal series \mathcal{V} of the form*

$$(2.5) \quad \mathcal{V} = \sum_0^\infty v_i \lambda^{-i}, \quad v_i \in \tilde{\mathfrak{g}},$$

such that

- (i) v_i is homogeneous of degree i ;
- (ii) $\partial\mathcal{V} = [U, \mathcal{V}]$.

Granting the lemma, which will be proved in § 3, the equations (2.3) are constructed as follows: for each $r \geq 1$, set

$$V = \sum_0^r v_i \lambda^{r-i}, \quad V_- = \sum_{r+1}^\infty v_i \lambda^{r-i}.$$

Thus

$$V + V_- = \lambda^r \mathcal{V},$$

so that we have

$$-\partial V + [U, V] = \partial V_- - [U, V_-].$$

In this equality the left-hand side clearly does not involve any negative powers of λ , and the right-hand side does not involve any positive powers. Hence the expression must be independent of λ ; picking out the terms of order zero, we obtain our equations (2.3) in two alternative forms

$$(2.6) \quad -\partial_t u = \partial v_r + [u, v_r] = [F, v_{r+1}].$$

The second form shows that the right-hand side does indeed lie in $\text{Im ad } F$, as required. Thus, for each $v_0 \in \mathfrak{h}$ and each integer $r \geq 1$, (2.6) defines a system of evolution equations for the u_α ; these are our ‘general zero curvature equations’. Algebraically, we regard the ∂_t in (2.6) as defining an evolutionary (commuting with ∂) derivation of the algebra B . We shall write $\partial_t(v_0, r)$ when it is necessary to specify a particular equation. Note that the derivation $\partial_t(v_0, r)$ increases degree by r : we shall sometimes (as in (1.3)) refer to the corresponding evolution equation as having degree r .

We set

$$l = \dim \mathfrak{h} = \text{rank } \mathfrak{g}.$$

(2.7) PROPOSITION. (i) If $v_0 \neq 0$, then $\partial_t(v_0, r) \neq 0$ for all $r \geq 1$; that is, all the equations constructed above are non-trivial.

(ii) For fixed r , as v_0 runs over a basis for \mathfrak{h} , the right-hand sides of the l equations corresponding to $\partial_t(v_0, r)$ are linearly independent.

Briefly, we have constructed l essentially different equations of each degree.

Let K denote the Killing form:

$$K(X, Y) = \text{tr}(\text{ad } X \text{ ad } Y).$$

We also denote by K the obvious extension to a bilinear map from $\tilde{\mathfrak{g}} \times \tilde{\mathfrak{g}}$ to B . For $r \geq 1$, we set

$$(2.8) \quad H_r = H_r(v_0) = r^{-1} K(v_{r+1}, F) \in B.$$

(The factor r^{-1} is inserted to improve (2.11) below.) Note that H_r is homogeneous of degree $r + 1$.

(2.9) PROPOSITION. The elements $H_r(v_0)$ ($r \geq 1$, any choice of $v_0 \in \mathfrak{h}$) are conserved densities for all the zero curvature equations (that is, $\partial_t H_r \in \partial B$ for all the ∂_t).

(2.10) PROPOSITION. *These conserved densities are all non-trivial (that is, $H_r \notin \partial B$); as v_0 varies over a basis for \mathfrak{h} , the l conserved densities of each degree r are linearly independent as elements of $B/\partial B$.*

Briefly, the equations have l essentially different common conserved densities of each degree $r \geq 2$.

The reader may wonder why we do not assert that the elements $K(v_1, F)$ are conserved densities: the reason is that $K(v_1, F) = 0$. That is easily checked using the first three equations in (3.1) below and the ad-invariance of K .

Finally, we look at the Hamiltonian form of (2.6).

(2.11) PROPOSITION. *The equation (2.6) can be written in the Hamiltonian form*

$$\partial_t(v_0, r)u_\alpha = -\langle \alpha, F \rangle \frac{\delta H_{r+1}}{\delta u_{-\alpha}}, \quad \alpha \in R,$$

where $H_{r+1} = H_{r+1}(v_0)$. (We use $\langle \cdot, \cdot \rangle$ to denote the natural pairing between \mathfrak{h} and \mathfrak{h}^* , and $\delta/\delta u_\alpha$ is the formal variational derivative

$$\frac{\delta f}{\delta u_\alpha} = \sum_{i \geq 0} (-\partial)^i \frac{\partial f}{\partial u_\alpha^{(i)}}, \quad f \in B.)$$

Thus the skew matrix defining the Hamiltonian structure (see, for example, [9], § 4) is a direct sum of multiples of the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(one for each positive root).

Combining (2.9) and (2.11), we have the following.

(2.12) PROPOSITION. *The derivations $\partial_t(r, v_0)$ for different choices of $r \geq 1$ and $v_0 \in \mathfrak{h}$ all commute with each other.*

(2.13) Remark. Of course, it is not essential to coordinatize $\text{Im ad } F$ by the root spaces (or at all); a different choice of coordinates would correspond simply to a linear change of the variables u_α .

(2.14) Remark. We are taking \mathfrak{g} to be complex only for simplicity: everything would go through equally well in the real case.

(2.15) Remark. More important, the assumption that the variables u_α are differentially independent is also not essential: we can consider ‘specializations’ of our general equations in which the u_α are differential polynomials in some smaller number of other variables. The modified Lax equations in § 4 are of this kind. The construction of the equations and conservation laws goes through just as above; but the problem is that some of the equations may turn out to be inconsistent, and some of the conserved densities trivial. § 4 is devoted to finding out which these are for the modified Lax equations.

3. The general zero curvature equations: proofs

We begin with the basic lemma (2.4). Writing out the condition (ii) in (2.4) in long-hand, we obtain the equations

$$(3.1) \quad \left\{ \begin{array}{l} [F, v_0] = 0 \\ \partial v_0 + [u, v_0] = [F, v_1] \\ \vdots \\ \partial v_i + [u, v_i] = [F, v_{i+1}] \\ \vdots \end{array} \right.$$

We have to show that these equations determine the v_i successively. Since F is semi-simple, we have the decomposition

$$(3.2) \quad \tilde{\mathfrak{g}} = \text{Ker ad } F \oplus \text{Im ad } F,$$

and since F is regular,

$$\text{Ker ad } F = \tilde{\mathfrak{h}} = B \otimes \mathfrak{h}.$$

We shall for the moment refer to the components of an element $v \in \tilde{\mathfrak{g}}$ with respect to the splitting (3.2) as the ‘diagonal’ and ‘off-diagonal’ parts of v (since that is what they are in the case $\mathfrak{g} = \mathfrak{sl}(n)$, $F = \{\text{diagonal matrix with distinct entries}\}$). It is trivial to check that

- (i) $\text{ad } F$ acts as an isomorphism on $\text{Im ad } F$;
- (ii) bearing in mind that $u \in \text{Im ad } F$, the diagonal part of $[u, v]$ depends only on the off-diagonal part of v .

Now, the first equation in (3.1) shows that $v_0 \in \tilde{\mathfrak{h}}$; taking the diagonal part of the second equation and using (ii) above gives $\partial v_0 = 0$, that is, $v_0 \in \mathfrak{h}$. We fix $v_0 \in \mathfrak{h}$. Suppose inductively that from the first i equations in (3.1) we have determined v_1, \dots, v_{i-1} and the off-diagonal part of v_i . We split the next equation $\partial v_i + \dots$ in (3.1) into its diagonal and off-diagonal parts. Using (i) and (ii) above, this determines the off-diagonal part of v_{i+1} and the diagonal part of ∂v_i . To obtain the diagonal part of v_i , we have to ‘integrate’; the homogeneity property (i) in (2.4) forces us to choose the ‘constants of integration’ to be zero, giving a unique v_i . Continuing inductively, we obtain unique solutions for all the v_i , given $v_0 \in \mathfrak{h}$.

(3.3) *Remark.* The argument shows a little more than was stated in two respects. First, we showed that any series \mathcal{V} satisfying (ii) in (2.4) must in fact have leading term v_0 lying in \mathfrak{h} . Secondly, note that if we tried to determine the polynomial V of degree r in (2.2) directly from the requirement that the right-hand side of (2.3) should have order zero in λ , that would amount to repeating the first r or so steps of the above argument. Thus we have shown that all possible equations (2.2) with V polynomial in λ arise from the construction given just after (2.4).

There is, however, still a serious gap in the argument given above: since $\partial: B \rightarrow B$ is far from being surjective, it is not clear that the ‘integrations’ required to obtain the diagonal parts of the v_i can be performed without going outside B . Thus so far

we have shown only that we can find unique v_i lying in $\bar{B} \otimes \mathfrak{g}$, where \bar{B} is (say) the algebra (containing B) with *surjective* derivation constructed in [17]. Happily, in this paper we shall not have to deal with \bar{B} , because of the next proposition.

(3.4) PROPOSITION. *The elements v_i satisfying (3.1) do indeed lie in $B \otimes \mathfrak{g}$.*

To prove this (and the other statements in § 2) we shall rewrite the equations (2.2) in the form of matrix Lax equations, and then appeal to results proved in [17] and [9]. To do that, we fix some faithful representation of \mathfrak{g} by $N \times N$ matrices, so we have an inclusion of Lie algebras $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(N)$. We shall identify \mathfrak{g} with the subalgebra $\rho(\mathfrak{g})$ of $\mathfrak{gl}(N)$ and suppress ρ from the notation. Thus F , for example, is now some semi-simple (that is, diagonalizable) $N \times N$ matrix, and the v_i are $N \times N$ matrices with entries in B . The matrix F may not be invertible, but we can make it so by adding on a multiple of the identity: we suppose that this has been done without changing the notation. We observe that adding a multiple of the identity to F clearly does not change equations (3.1).

Now, the equations (2.2) can be viewed as the integrability condition for the system

$$(3.5) \quad \begin{cases} \partial\psi = U\psi \\ \partial_t\psi = V\psi, \quad V = \sum_0^r v_i \lambda^{r-i}. \end{cases}$$

This can be rewritten in the form

$$(3.6) \quad \begin{cases} L\psi = \lambda\psi \\ \partial_t\psi = P_+\psi, \end{cases}$$

where we have set

$$(3.7) \quad L = F^{-1}(\partial + u), \quad P_+ = \sum_{i=0}^r v_i L^{r-i}.$$

The integrability condition for this is the Lax equation

$$\partial_t L = [P_+, L].$$

This suggests the next proposition.

(3.8) PROPOSITION. *The zero curvature equation (2.2) is equivalent to the matrix Lax equation*

$$\partial_t L = [P_+, L],$$

where P_+ and L are given by (3.7). More generally, the equation

$$\partial \mathcal{V} = [U, \mathcal{V}]$$

in (2.4) for the series

$$\mathcal{V} = \sum_0^\infty v_i \lambda^{-i}$$

is equivalent to the equation

$$[P, L] = 0$$

for the formal pseudo-differential operator

$$P = \sum_{i=0}^{\infty} v_i L^{r-i}.$$

I leave it to the reader to construct a rigorous proof. Of course, the value of r in the last statement is immaterial, since P commutes with L if and only if PL^q does (for any q).

From the point of view of the matrix Lax equations, the operator L in (3.7) is a complicated-looking specialization (since u is constrained to take values in \mathfrak{g}). However, as explained in [9], § 3, this does not affect the truth of the results from [17] that we are going to use.

Now let us prove (3.4). By (3.8), we have to show that all the entries in the coefficients of P lie in B (rather than \bar{B}). By [17], (5.10), that will be so provided the leading coefficient of P , which is $v_0 F^{-r}$, lies in the centre of the centralizer (in $\mathfrak{gl}(N)$) of the leading coefficient of L , which is F^{-1} . Of course, the centralizer of F^{-1} is the same as that of F . Let $E = \bigoplus E^\omega$ be the weight space decomposition of the space $E = \mathbb{C}^N$ of our representation of \mathfrak{g} ; that is, each $X \in \mathfrak{h}$ acts on E^ω as the scalar $\langle \omega, X \rangle$, $\omega \in \mathfrak{h}^*$. By definition, the weights ω are distinct elements of \mathfrak{h}^* ; however, it may happen that the values of two weights coincide on the element F . We suppose first that this does not happen (so we are excluding elements F lying in a finite set of hyperplanes in \mathfrak{h}). Then the E^ω are precisely the eigenspaces of F . Thus the centralizer of F in $\mathfrak{gl}(N)$ consists of the maps that preserve each E^ω ; and its centre consists of the maps that act as scalars on each E^ω . But all the elements of \mathfrak{h} , hence also the matrix $v_0 F^{-r}$, clearly have this property. That completes the proof of (3.4) for most choices of F .

The argument seems to break down if two weights coincide on F ; however, the result remains true. We can easily see that as follows. The entries in the v_i are polynomials in the $u_\alpha^{(i)}$ and (perhaps) some other elements involving ‘integrations’ introduced to form the algebra \bar{B} (see [17]). We consider how these polynomials change as F varies in \mathfrak{h} . From a glance at (3.1) it is clear that each coefficient is a rational function on \mathfrak{h} (with denominator non-vanishing if F is regular). Now, suppose some entry in a v_i were to contain a monomial that is not in B . The argument given above shows that the coefficient of this term vanishes on a dense open set in \mathfrak{h} . Hence it vanishes identically.

Next we prove the non-triviality of our equations.

(3.9) LEMMA. *The equation*

$$\partial v + [u, v] = 0$$

has no non-zero solution $v \in \tilde{\mathfrak{g}}$ which is homogeneous of positive degree.

Proof. The term ∂v involves a higher derivative of at least one u_α than the other term $[u, v]$. (The assumption of homogeneity is made just to exclude the obvious constant solutions $v \in \mathfrak{h}$.) □

Proof of proposition (2.7). (i) Suppose one of the equations (2.6) was trivial, that is,

$$\partial v_r + [u, v_r] = 0.$$

Then by (3.9), $v_r = 0$. Working our way up the equations (3.1), we find that $v_0 = 0$.

(ii) This follows at once from (i), since a linear combination of equations of the same degree comes from the corresponding linear combination of the elements v_0 . □

Now the conservation laws. From [17], we know that as P varies over the operators in (3.8) obtained from various choices of v_0 and r , the elements $\text{tr res } P \in B$ are conserved densities for all our equations. Clearly

$$\text{res } P = v_{r+1}F.$$

Now, the restriction to \mathfrak{g} of the form $(X, Y) \mapsto \text{tr } XY$ is a non-zero multiple cK of the Killing form on \mathfrak{g} : to simplify the notation we shall assume that $c = 1$ (the reader will easily check that c cancels out of our final formulae). Thus we have

$$\text{tr res } P = K(v_{r+1}, F) = rH_r.$$

Hence (2.9) follows.

Next, let $\delta: B \rightarrow \Omega^1(B)$ be the universal derivation (total variation operator): thus $\Omega^1(B)$ is the free B -module on the symbols $\delta u_\alpha^{(j)}$ ($j \geq 0$), and δ is given by

$$\delta f = \sum_{\alpha, j} \frac{\partial f}{\partial u_\alpha^{(j)}} \delta u_\alpha^{(j)}, \quad f \in B.$$

The derivation δ is extended to $\Omega^1(B)$ so as to commute with δ . For more details see [12] or [9], § 7.

(3.10) LEMMA. *We have*

$$\delta H_r \equiv K(\delta u, v_r) \pmod{\partial \Omega^1(B)}$$

(where H_r is given by (2.8)).

Proof. From [9], § 7, we know that

$$\delta H_r = \delta(r^{-1} \text{tr res } P) \equiv \text{tr res } (\delta L \cdot Q) \pmod{\partial \Omega^1(B)},$$

where $Q = PL^{-1}$. Here we have

$$\delta L = F^{-1} \delta u, \quad \text{res } Q = v_r F,$$

so (3.10) follows at once. □

Proof of proposition (2.11). We recall that the variational derivatives $\delta/\delta u_\alpha$ are characterized by the formula

$$\delta f \equiv \sum \frac{\delta f}{\delta u_\alpha} \delta u_\alpha \pmod{\partial \Omega^1(B)}.$$

The root spaces $\mathbb{C}E_\alpha$ are orthogonal (with respect to K) to \mathfrak{h} , and also we have

$$K(E_\alpha, E_\beta) = 0 \quad \text{unless } \alpha = -\beta;$$

and we may assume the E_α chosen so that we have

$$(3.11) \quad K(E_\alpha, E_{-\alpha}) = 1.$$

We have $\delta u = \sum \delta u_\alpha E_\alpha$; set

$$v_{r+1} = \sum v_{r+1, \alpha} E_\alpha + (\text{component in } \tilde{\mathfrak{h}}), \quad v_{r+1, \alpha} \in B.$$

Then from (3.10) (with $r + 1$ instead of r) we read off that

$$\frac{\delta H_{r+1}}{\delta u_{-\alpha}} = v_{r+1,\alpha}.$$

On the other hand, the second form of equation (2.6) can be written

$$-\partial_t u_\alpha = \langle \alpha, F \rangle v_{r+1,\alpha},$$

since, of course, by definition

$$[F, E_\alpha] = \langle \alpha, F \rangle E_\alpha.$$

Proposition (2.11) now follows at once. □

It remains to prove the non-triviality of the conservation laws.

Proof of proposition (2.10). This follows immediately from (2.7) and (2.11), since $H_{r+1} \in \partial B$ implies $\delta H_{r+1} / \delta u_\alpha = 0$ for all α ; thus if the conserved density H_{r+1} were trivial, the corresponding equation (2.6) would be too. □

(3.12) *Remark.* All the statements in § 2 would remain unaltered if we changed K to be some non-zero multiple of the Killing form. (If that seems surprising, note that because of (3.11), renormalizing K would entail a renormalization of the variables u_α .) This remark is useful in the case $\mathfrak{g} = \mathfrak{sl}(n)$, where we normally take the E_α to be the elementary matrices E_{ij} and $K(X, Y) = \text{tr } XY$ (strictly speaking the Killing form is $2n$ times this).

(3.13) *Remark.* All the theory would still go through if F were semi-simple but not regular: we should still have the crucial splitting (3.2). We should have to restrict u to lie in $\text{Im ad } F$, and also the elements v_0 would have to lie in the centre of $\text{Ker ad } F$, otherwise (3.4) would not be true.

4. The modified Lax equations

The modified Lax equations will have the form (2.3), but with a ‘specialized’ u ; if we think of u as a map into $\text{Im ad } F$, the specialization will consist in requiring u to take values in a certain Cartan subalgebra \mathfrak{h}_1 (naturally, not the one containing F , which we shall continue to denote by \mathfrak{h}). We begin by summarizing some facts from the theory of simple Lie algebras that are perhaps not as well known as those we have used so far.

So let $\mathfrak{h}_1 \subset \mathfrak{g}$ be a Cartan subalgebra. From now on α, E_α will denote roots and root vectors with respect to \mathfrak{h}_1 (not \mathfrak{h}). We fix a set of simple roots $\alpha_1, \dots, \alpha_l$ for $(\mathfrak{g}, \mathfrak{h}_1)$. This gives rise to a grading on \mathfrak{g} : to avoid confusion with the grading on $\tilde{\mathfrak{g}}$ introduced in § 2, we shall use the term height, rather than degree, to refer to this new grading. It is defined as follows: each root α has a unique expression in the form

$$\alpha = \sum_1^l n_i \alpha_i,$$

where the n_i are integers (all of the same sign); the root vector E_α is defined to be homogeneous of height $\sum n_i$. The elements of \mathfrak{h}_1 are defined to be homogeneous of height zero. This makes \mathfrak{g} into a graded Lie algebra. There is a unique root ψ

of maximal height; its height is $h - 1$, where h is the Coxeter number of \mathfrak{g} (see [4], p. 169, or [7], theorem 8.4, p. 1022).

We now define the element $F \in \mathfrak{g}$ by

$$(4.1) \quad F = \sum_1^l E_{\alpha_i} + E_{-\psi}.$$

That is, F is the sum of the root vectors corresponding to the simple roots and to minus the maximal root. These are precisely the roots whose heights are congruent to 1 mod h ; we shall want F to be homogeneous of height 1, so from now on, although we really have a \mathbb{Z} -grading on \mathfrak{g} , we shall disregard that and consider it as a mod h grading. Kostant [7] calls the F in (4.1) a cyclic element of \mathfrak{g} . The reason is that for $\mathfrak{g} = \mathfrak{sl}(n)$, $\mathfrak{h}_1 = \{\text{diagonal matrices}\}$ and the usual choice of simple roots, we have

$$(4.2) \quad F = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

(4.3) PROPOSITION (Kostant). *The cyclic element F given by (4.1) is regular semi-simple (for all \mathfrak{g}).*

For the proof, see [7], lemma 6.3, p. 1007 and corollary 6.4, p. 1009.

We denote by \mathfrak{h} the unique Cartan subalgebra containing F . (For $\mathfrak{g} = \mathfrak{sl}(n)$, with F given by (4.2), \mathfrak{h} consists of the circulant matrices. Note incidentally that the relation between \mathfrak{h} and \mathfrak{h}_1 is symmetrical: if (as is done in [9]) we choose $\mathfrak{h}_1 = \{\text{circulants}\}$, then we have

$$F = \text{diag}(1, \omega, \dots, \omega^{n-1}),$$

where $\omega = \exp(2\pi i/n)$, and \mathfrak{h} will be the diagonal matrices.)

(4.4) THEOREM (Kostant). (i) *The Cartan subalgebra \mathfrak{h} has a basis consisting of elements that are homogeneous in the mod h grading.*

(ii) *The heights m_1, \dots, m_l , say, of the elements of this basis are the exponents of \mathfrak{g} ; thus we have $m_i = d_i - 1$, where the d_i are given by the table in the appendix.*

I refer to [7] for the proof (see especially theorem 6.7, p. 1014 and corollary 8.6, p. 1026). I shall just make a few comments. Let $Q \in \mathfrak{h}_1$ be the element defined by $\langle \alpha_i, Q \rangle = 1$ for all simple roots α_i ; let $P = \exp(2\pi i/h \cdot Q)$: thus P lies in (say) the adjoint group of \mathfrak{g} . It is easy to see that an element $Y \in \mathfrak{g}$ is homogeneous of height $q \pmod{h}$ precisely when

$$\text{Ad } P(Y) = \omega^q Y,$$

where $\omega = \exp(2\pi i/h)$. In particular,

$$\text{Ad } P(F) = \omega F,$$

from which it follows at once that $\text{Ad } P$ preserves \mathfrak{h} , the centralizer of F . Part (i) of (4.4) is now clear. In corollary 8.6 of [7], Kostant shows that the restriction of

Ad P to \mathfrak{h} is a Coxeter transformation of \mathfrak{h} ; part (ii) of (4.4) thus amounts to the (now) standard fact that the eigenvalues of a Coxeter transformation are ω^{m_i} , where m_i runs over the exponents of \mathfrak{g} (see [4], p. 121).

(4.5) LEMMA. We have $\mathfrak{h}_1 \subset \text{Im ad } F$.

Proof. We recall that $[E_\alpha, E_\beta] = 0$ unless $\alpha + \beta$ is a root or zero. It follows that for each simple root α_i , we have

$$[F, E_{-\alpha_i}] = [E_{\alpha_i}, E_{-\alpha_i}],$$

and these elements form a basis for \mathfrak{h}_1 . □

Now we are ready to study the modified Lax equations. Let u_1, \dots, u_l be differentially independent variables; we set $B = \mathbb{C}[u_i^{(j)}]$, $\tilde{\mathfrak{g}} = B \otimes \mathfrak{g}$. Thus on $\tilde{\mathfrak{g}}$ we have a $\mathbb{Z} \times \mathbb{Z}/h$ bigrading: in the \mathbb{Z} -grading (by ‘degree’), $\text{deg } u_i^{(j)} = 1 + j$ and $\text{deg } Y = 0$ for $Y \in \mathfrak{g}$; the mod h grading (by ‘height’) is extended to $\tilde{\mathfrak{g}}$ by setting $\text{ht } f = 0$ for $f \in B$. We fix any basis X_1, \dots, X_l for \mathfrak{h}_1 , and we set

$$u = \sum_1^l u_i X_i \in \tilde{\mathfrak{h}}_1, \quad U = -u + \lambda F.$$

(4.6) LEMMA. Let $\mathcal{V} = \sum_0^\infty v_i \lambda^{-i}$, $v_i \in \tilde{\mathfrak{g}}$, be a series satisfying

$$(4.7) \quad \partial \mathcal{V} = [U, \mathcal{V}]$$

with v_i \mathbb{Z} -homogeneous (of degree i). If $v_0 \in \mathfrak{h}$ is \mathbb{Z}/h -homogeneous of height q , then v_i is \mathbb{Z}/h -homogeneous of height $q - i$.

Proof. Extend the bigrading to formal series in λ by giving λ bidegree $(1, -1)$, so that U is bihomogeneous (of bidegree $(1, 0)$). Then it is clear that, if \mathcal{V} satisfies (4.7), each bihomogeneous component of \mathcal{V} does too. But we know from § 3 that a \mathbb{Z} -homogeneous solution of (4.7) is uniquely determined by its leading term $v_0 \in \mathfrak{h}$. It follows that if v_0 is \mathbb{Z}/h -homogeneous of height q , so is \mathcal{V} . □

It is now easy to tackle the problem of consistency for our modified Lax equations

$$(4.8) \quad -\partial_r u = \partial v_r + [u, v_r] = [F, v_{r+1}].$$

The left-hand side of (4.8) lies in $\tilde{\mathfrak{h}}_1$, so for consistency we need the right-hand side to do the same: that is, to be \mathbb{Z}/h -homogeneous of height zero. In view of (4.6) and the non-triviality assertion (2.7) (the proof of which was given in § 3 and is still valid), that will happen precisely when we choose v_0 to be homogeneous of height $r \pmod{h}$, so that v_r is homogeneous of height zero. Combining this with (4.4), we obtain the following.

(4.9) THEOREM. The modified Lax equation (4.8) is non-trivial and consistent if and only if we choose $v_0 \in \mathfrak{h}$ to be homogeneous of height r . Thus we have one modified Lax equation of each degree congruent to an exponent of \mathfrak{g} modulo the Coxeter number h .

From the table in the appendix, we see that in the case of D_{2k} there is a repeated exponent $2k - 1$. In that case the last statement in (4.9) naturally means that we have two linearly independent equations of each degree congruent to $2k - 1 \pmod{h}$.

Next we consider the conservation laws. The proof in § 3 that the elements

$$H_r = r^{-1}K(v_{r+1}, F)$$

are conserved densities for all the equations is still valid: the problem is to determine which of them are non-trivial. Using again the facts that

$$K(E_\alpha, E_\beta) = 0 \quad \text{unless } \alpha = -\beta$$

and that

$$K(E_\alpha, \mathfrak{h}_1) = 0,$$

we see that only the component of height -1 in v_{r+1} contributes anything to H_r : this comes from the component of height r in v_0 . Thus to obtain non-trivial conservation laws, we may restrict ourselves to choosing v_0 to be homogeneous of height r .

(4.10) THEOREM. *If $v_0 \in \mathfrak{h}$ is (non-zero and) homogeneous of height r , then the conserved density*

$$H_r = r^{-1}K(v_{r+1}, F)$$

is non-trivial. Thus the modified Lax equations have a (common) non-trivial conserved density of each degree congruent to one of the $d_i = m_i + 1 \pmod h$.

(In the case D_{2k} , this means that the equations have two conserved densities of each degree congruent to $2k \pmod h$, and these are linearly independent as elements of $B/\partial B$.)

Theorem (4.10) will follow from the Hamiltonian form of the equations (see (4.12) below). Lemma (3.8) is still valid in our present specialized situation, so we can use it to read off the variational derivatives of the H_r . We choose $v_0 \in \mathfrak{h}$ homogeneous of height r , so that $v_r \in \tilde{\mathfrak{h}}_1$; let

$$v_r = \sum_1^l v_{r,i} X_i, \quad v_{r,i} \in B.$$

Then since

$$\delta u = \sum \delta u_i X_i,$$

(3.8) gives

$$\frac{\delta H_r}{\delta u_i} = \sum_{j=1}^l K(X_i, X_j) v_{r,j}.$$

Since the restriction of K to \mathfrak{h}_1 is non-degenerate, the matrix $K(X_i, X_j)$ is non-singular: denote it by S^{-1} , so that we have

$$(4.11) \quad (S^{-1})_{ij} = K(X_i, X_j).$$

Then if \mathbf{v}_r (resp. $\delta H_r/\delta u$) denotes the column vector with entries $v_{r,i}$ (resp. $\delta H_r/\delta u_i$), we have

$$\mathbf{v}_r = S \frac{\delta H_r}{\delta u}.$$

On the other hand, since $v_r \in \tilde{\mathfrak{h}}_1$, we have

$$[u, v_r] = 0,$$

so that the modified Lax equation (4.8) simply takes the form

$$\partial_t u = -\partial v_r.$$

Thus we have proved the following.

(4.12) THEOREM. *The modified Lax equation (4.8) can be written in the Hamiltonian form*

$$\partial_t \mathbf{u} = -S \partial \frac{\delta H_r}{\delta \mathbf{u}},$$

where the constant matrix S is given by (4.11) (and \mathbf{u} is the column vector with entries u_1, \dots, u_l).

Of course, by choosing $\{X_i\}$ to be a K -orthonormal basis for \mathfrak{h}_1 , we can make S the identity. As another example, if $\mathfrak{g} = \mathfrak{sl}(n)$, $\mathfrak{h}_1 = \{\text{diagonal matrices}\}$, $X_i = E_{ii} - E_{nn}$ ($1 \leq i \leq n - 1$) and $K(X, Y) = \text{tr } XY$ (see remark 3.12), then we have

$$\text{tr } X_i^2 = 2, \quad \text{tr } X_i X_j = 1 \quad (i \neq j),$$

so that

$$S^{-1} = \text{Id} + E,$$

where E is the matrix with every entry equal to 1. Thus

$$S = \text{Id} - n^{-1}E,$$

confirming the result stated in [10], (2.11), where it was obtained by a more roundabout method.

Proof of theorem 4.10. By (4.12), if the conserved density H_r were trivial, so would be the corresponding equation; but we already know that it is not. □

5. The two-dimensional Toda lattice equations

We retain the notation of § 4. By definition, the 2DTL equation associated with \mathfrak{g} is the zero curvature equation

$$(5.1) \quad \partial_t u = -\partial V + [U, V],$$

where now V is supposed to be a multiple of λ^{-1} : $V = v\lambda^{-1}$ (the idea of considering such equations is suggested by the AKNS representation [1] for the sinh-Gordon equation, which is our equation (1.1) in the case $n = 2$).

As we shall see in a moment, that forces v to lie, not in our algebra $B \otimes \mathfrak{g}$, but in a larger one $\hat{B} \otimes \mathfrak{g}$, where \hat{B} is an algebra obtained from B by adjoining exponentials of integrals of the u_i . We shall write A for the set of roots occurring in the sum (4.1); that is, A consists of the simple roots and minus the maximal root. Since U is homogeneous of height zero (recall that λ has height -1), for consistency of (5.1) we should take V to have height zero too, so that v must have height -1 ; thus v takes the form

$$v = \sum_{\alpha \in A} v_\alpha E_{-\alpha}, \quad v_\alpha \in \hat{B}$$

(the algebra \hat{B} is yet to be determined). A short calculation shows that (5.1) then takes the form

$$\partial_t u = \sum_{\alpha \in A} [\langle \alpha, u \rangle v_\alpha - \partial v_\alpha] E_{-\alpha} \lambda^{-1} + [F, v].$$

For consistency, we require the terms involving λ^{-1} to vanish; let us introduce ‘potential’ variables w_i with $\partial w_i = u_i$, and set $w = \sum w_i X_i \in \tilde{\mathfrak{h}}_1$, so that $\partial w = u$. Then the vanishing of the terms involving λ^{-1} above gives

$$v_\alpha = c_\alpha \exp \langle \alpha, w \rangle, \quad \alpha \in A, c_\alpha \text{ constants.}$$

If $\alpha, \beta \in A$, then $\alpha - \beta$ is never a root, from which we find

$$[F, v] = \sum v_\alpha [E_\alpha, E_{-\alpha}].$$

Thus we obtain our 2DTL equation in the form

$$(5.2) \quad \partial_t u = \partial_t \partial w = \sum_{\alpha \in A} c_\alpha \exp \langle \alpha, w \rangle [E_\alpha, E_{-\alpha}].$$

The Hamiltonian form of the equation is perhaps more easily understood. Set

$$(5.3) \quad H = -K(F, v) = - \sum_{\alpha \in A} c_\alpha \exp \langle \alpha, w \rangle$$

(we assume as usual that $K(E_\alpha, E_{-\alpha}) = 1$).

(5.4) PROPOSITION. *The 2DTL equation (5.2) can be written in the form*

$$\partial_t \partial w = -S \frac{\delta H}{\delta w},$$

where S is given by (4.11), H by (5.3), and $w, \delta H / \delta w$ are column vectors.

Proof. We have

$$K(X_i, [E_\alpha, E_{-\alpha}]) = K([X_i, E_\alpha], E_{-\alpha}) = K(\langle \alpha, X_i \rangle E_\alpha, E_{-\alpha}) = \langle \alpha, X_i \rangle,$$

so that from (5.2) we obtain

$$K(X_i, \partial_t u) = \sum_j (S^{-1})_{ij} \partial_t u_j = \sum_{\alpha \in A} c_\alpha \exp \langle \alpha, w \rangle \langle \alpha, X_i \rangle.$$

But this last expression is clearly equal to $-\delta H / \delta w_i$, so (5.4) follows. □

It is usual (as we did in the introduction) to set all the constants c_α equal to 1. Let us note, however, that if instead we set $c_{-\psi} = 0$ for ψ the maximal root and $c_\alpha = 1$ for α simple, then we obtain the two-dimensional versions of the ‘tied’ (non-periodic) finite Toda lattice studied, for example, in [8].

Now we can state the main facts about the 2DTL equations.

(5.5) THEOREM. (i) *The modified Lax equations of § 4 are all symmetries of the 2DTL equation in the sense explained in [10].*

(ii) *The conserved densities for the modified Lax equations are also conserved densities for the 2DTL equation.*

Combining this with what was proved in § 4, we obtain theorem (1.3) in the introduction.

I shall not comment on the proof of (5.5), since it is exactly the same as in the case $\mathfrak{g} = \mathfrak{sl}(n)$, which is explained in detail in [10].

Appendix

The table below (which I extracted from [4]) gives the values of d_1, \dots, d_l for the various simple Lie algebras. The exponents are given by $m_i = d_i - 1$ and the largest d_i is the Coxeter number h .

Lie algebra	d_1, \dots, d_l
$A_{l-1} (l \geq 2)$	$2, 3, 4, \dots, l$
$B_l, C_l (l \geq 2)$	$2, 4, 6, \dots, 2l$
$D_l (l \geq 3)$	$2, 4, 6, \dots, 2l - 2; l$
E_6	$2, 5, 6, 8, 9, 12$
E_7	$2, 6, 8, 10, 12, 14, 18$
E_8	$2, 8, 12, 14, 18, 20, 24, 30$
F_4	$2, 6, 8, 12$
G_2	$2, 6$

Postscript

In [21] Drinfel'd & Sokolov associate a hierarchy of 'modified KdV' equations to each of the class of Lie algebras variously called in the literature affine, Euclidean, Kac-Moody, contragredient or loop algebras. The list of these can be found (for example) on p. 503 of Helgason's book [22]. Among them are the loop algebras $\mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$ (\mathfrak{g} simple, finite dimensional) that we have been working with in this paper: the others can be described as the 'twisted loop algebras', $L(\mathfrak{g}, \nu)$ (see (P2) below), where ν is an (outer) automorphism coming from a symmetry of the Dynkin diagram of \mathfrak{g} . I want to describe the construction of [21], and explain why in the untwisted case it is the same as the one given above.

So, let \mathcal{G} be one of these loop algebras; recall that it is generated by elements $\{e_i, f_i, h_i\}$, $0 \leq i \leq l$, satisfying relations similar to those for finite dimensional simple Lie algebras (see [16], appendix to ch. 6). We equip \mathcal{G} with its *principal grading*, in which the e_i, f_i and h_i have heights 1, -1 and 0, respectively. Let $\mathfrak{h}_1 \subset \mathcal{G}$ be the subspace of elements of height zero: it is spanned by the h_i , and is l -dimensional. (The notation \mathfrak{h}_1 is chosen to conform with that of § 4 above.) Let $u(x, t)$ be a map with values in \mathfrak{h}_1 , and let $\Lambda = \sum_0^l e_i$. Then the modified KdV equations of [21] associated with \mathcal{G} have the zero curvature representations

$$[\partial - U, \partial_t - V] = 0,$$

where $U = -u + \Lambda$, and V is a sum of elements of non-negative height.

Now specialize to the 'untwisted' case

$$\mathcal{G} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}].$$

Then I claim that these equations are the same as the modified Lax equations associated with \mathfrak{g} constructed in the present paper. The zero curvature representation for them in [21] is not quite the same as ours, but it is related to it by a very simple 'gauge transformation' (inner automorphism of \mathcal{G}). To see that, recall (see [22], p. 504) that if

$$\mathcal{G} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$$

then \mathfrak{h}_1 and e_1, \dots, e_l have the same meaning[†] as in § 4, and

$$e_0 = \lambda e_{-\psi}$$

(ψ the maximal root). In the principal grading, λ therefore has height h (the Coxeter number). Let μ be a parameter with $\mu^h = \lambda$, and let $\bar{\mathcal{G}} \subset \mathfrak{g} \otimes \mathbb{C}[\mu, \mu^{-1}]$ be the subalgebra consisting of sums $\sum X_i \mu^i$ such that $X_i \in \mathfrak{g}$ has height i in the mod h grading of § 4. Let ϕ be the automorphism of $\mathfrak{g} \otimes \mathbb{C}[\mu, \mu^{-1}]$ defined as follows: if $X \in \mathfrak{g}$ has height i (in the \mathbb{Z} -grading), then

$$\phi(X\mu^i) = X\mu^{r-i}.$$

The following proposition is trivial.

(P1) PROPOSITION. *We have an isomorphism of algebras*

$$\phi: \bar{\mathcal{G}} \rightarrow \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}].$$

On the other hand, if F is as in § 4, we clearly have $F\mu \in \bar{\mathcal{G}}$ and

$$\phi(F\mu) = \sum_1^l e_i + \mu^h e_{-\psi} = \Lambda;$$

and ϕ leaves the elements of \mathfrak{h}_1 fixed. It follows easily that ϕ throws our construction (with our λ replaced by μ) onto that of [21]. Of course, the construction of [21] has the advantage that it works equally well in the twisted case.

To conclude, I should like to make a few short remarks.

(1) The proper setting for the arguments of this paper seems to be the following. Let σ be an automorphism of finite order m of a simple Lie algebra \mathfrak{g} , and let

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_{m-1}$$

be the corresponding mod m grading of \mathfrak{g} (that is, \mathfrak{g}_i is the subspace on which σ acts as ω^i , ω a primitive m 'th root of unity). Let

$$(P2) \quad L(\mathfrak{g}, \sigma) = \left\{ \sum_{i \in \mathbb{Z}} X_i \lambda^i \mid X_i \in \mathfrak{g}_{i \bmod m} \right\}$$

be the corresponding twisted loop algebra. Let $F \in \mathfrak{g}_1$ be a semi-simple element, $u(x, t)$ a map with values in $\mathfrak{g}_0 \cap \text{Im ad } F$, and set $U = -u + \lambda F$. Then we can form a hierarchy of evolution equations

$$[\partial - U, \partial_t - V] = 0$$

just as in § 2: we shall have one equation for each element of the centre of the centralizer of F in $L(\mathfrak{g}, \sigma)$. The equations of § 2 are the case when σ is the identity; the modified Lax equations are the case when σ is the Coxeter transformation (conjugation by the element denoted by P in § 4). The twisted loop algebra in that case is the algebra $\bar{\mathcal{G}}$ above. Note that Kac (see [23]) has shown that the classification of automorphisms of finite order reduces to classifying the gradings on the 'standard' loop algebras \mathcal{G} , which is trivial: however, I do not know how to classify the automorphisms such that \mathfrak{g}_1 contains a (non-zero) semi-simple element.

(2) The question of the degrees of the MKdV equations of [21] and their conservation laws takes the form: for what positive integers r does the centralizer

[†] More precisely, the e_i corresponds to the E_i of § 4.

(which is commutative) of Λ in \mathcal{G} contain non-zero elements of height r (in the principal grading)? In the case

$$\mathcal{G} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$$

we have seen that these are the numbers congruent to an exponent of \mathfrak{g} modulo the Coxeter number. In the twisted cases the numbers are listed in [21]; and Kac has pointed out to me that they are also calculated (for a different reason) in his paper ([23], p. 124). They are as follows: for the algebras $A_{2l-1}^{(2)}$ and $D_{l+1}^{(2)}$, all odd numbers; for the ‘exceptional’ algebras $E_6^{(2)}$ and $D_4^{(3)}$, all numbers congruent to $\pm 1 \pmod{6}$; and for $A_{2l}^{(2)}$, all odd numbers not divisible by $2l+1$. It is amusing that the numbers for $A_{2l}^{(2)}$ were obtained in [10], by regarding the MKdV equations for $A_{2l}^{(2)}$ as specializations of those for A_{2l} ; incidentally, the algebras $A_{2l}^{(2)}$ are the only twisted algebras that can be treated in that way.

(3) The terminology ‘modified Lax equations’ will doubtless have led the reader to suspect that there should also be for each simple Lie algebra \mathfrak{g} a hierarchy of ‘Lax equations’ related to the modified equations by a ‘Miura transformation’. That is indeed the case, and the present work was intended as a preliminary to the study of such equations. However, the paper [21] shows that the situation is richer than this, at least if one does not insist on the Lax equations having *two* Hamiltonian structures: namely, Drinfel’d & Sokolov show how to construct a hierarchy of ‘KdV’ equations for each affine algebra \mathcal{G} and each vertex c of its Dynkin diagram (see [22], p. 503). The MKdV and KdV equations are related by a Miura transformation of type corresponding to the semi-simple Lie algebra \mathfrak{g} whose Dynkin diagram is obtained by removing c from that of \mathcal{G} . In fact, the situation is richer still: one can form equations related to the MKdV equations by ‘degenerate’ Miura transformations corresponding to the different parabolic subalgebras of \mathfrak{g} (the Miura transformations of [21] correspond to the Borel subalgebras). Note that in the case

$$\mathcal{G} = \mathfrak{sl}(n) \otimes \mathbb{C}[\lambda, \lambda^{-1}]$$

(which was studied in [9]), because of the exceptional symmetry of the Dynkin diagram one obtains the same equations for each vertex c : it is this that is responsible for the ‘Bäcklund transformations’ of these equations (see [9], p. 407).

(4) The recent Japanese preprints [20] indicate that the affine algebras are central, not only to the formal properties of integrable equations, but also to the study of their solutions. It should be very interesting to combine the viewpoints of [20] and [21]. For example, a cursory comparison suggests that the solutions of the KdV equations of [21] corresponding to an affine algebra \mathcal{G} and vertex c of its Dynkin diagram should be obtained from a study of the irreducible representation of \mathcal{G} with highest weight dual to c . In any case, the main message of the papers [20] and [21] seems to me very clear: the role of the affine algebras in the theory of integrable evolution equations can scarcely be exaggerated.

REFERENCES

- [1] M. J. Ablowitz, D. J. Kaup, A. C. Newell & H. Segur. The inverse scattering transform: Fourier analysis for non-linear problems. *Studies in Appl. Math.* **53** (1974), 249–315.

- [2] M. Adler. On a trace functional for formal pseudo-differential operators and the symplectic structure of the Korteweg-de Vries equations. *Inventiones Math.* **50** (1979), 219–248.
- [3] O. I. Bogoyavlensky. On perturbations of the periodic Toda lattice. *Commun. Math. Phys.* **51** (1976), 201–209.
- [4] N. Bourbaki. *Groupes et algèbres de Lie*, ch. 4, 5, 6. Hermann: Paris, 1968.
- [5] S. A. Bulgadaev. Two dimensional integrable field theories connected with simple Lie algebras. *Phys. Lett.* **96B** (1980), 151–153.
- [6] A. P. Fordy & J. Gibbons. Integrable non-linear Klein–Gordon equations and Toda lattices. *Commun. Math. Phys.* **77** (1980), 21–30.
- [7] B. Kostant. The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group. *Amer. J. Math.* **131** (1959), 973–1032.
- [8] B. Kostant. The solution to a generalized Toda lattice and representation theory. *Advances in Math.* **34** (1979), 195–338.
- [9] B. A. Kupershmidt & G. Wilson. Modifying Lax equations and the second Hamiltonian structure. *Inventiones Math.* **62** (1981), 403–436.
- [10] B. A. Kupershmidt & G. Wilson. Conservation laws and symmetries of generalized sine-Gordon equations. *Commun. Math. Phys.* **81** (1981), 189–202.
- [11] A. N. Leznov & M. V. Saveliev. Representation of zero curvature for the system of non-linear partial differential equations $x_{\alpha, z\bar{z}} = \exp(kx)_{\alpha}$ and its integrability. *Letters in Math. Phys.* **3** (1979), 498–494.
- [12] Yu. I. Manin. Algebraic aspects of non-linear differential equations. *Itogi Nauki i Tekhniki, ser. Sovremennyye Problemy Matematiki* **11** (1978), 5–152; *J. Sov. Math.* **11** (1979), 1–122.
- [13] A. V. Mikhailov. Integrability of a two-dimensional generalization of the Toda chain. *Pis'ma Zh. Eksp. Teor. Fiz.* **30** (1979), 443–448; *JETP Letters* **30** (1979), 414–418.
- [14] A. V. Mikhailov, M. A. Olshanetsky & A. M. Perelomov. Two-dimensional generalized Toda lattice. *Commun. Math. Phys.* **79** (1981), 473–488.
- [15] A. G. Reiman & M. A. Semenov-Tian-Shansky. Current algebras and non-linear partial differential equations. *Dokl. Akad. Nauk SSSR* **251** (1980), 1310–1314; *Soviet Math. Doklady* **21** (1980), 630–634.
- [16] J.-P. Serre. *Algèbres de Lie semi-simples complexes*. Benjamin: New York, 1966.
- [17] G. Wilson. Commuting flows and conservation laws for Lax equations. *Math. Proc. Camb. Phil. Soc.* **86** (1979), 131–143.
- [18] G. Wilson. On two constructions of conservation laws for Lax equations. *Quart. J. Math. Oxford* (to appear).
- [19] V. E. Zakharov & A. B. Shabat. A scheme for integrating the non-linear equations of mathematical physics by the inverse scattering method II. *Funct. Anal. Appl.* **13** (3) (1979), 13–22 (Russian), 166–174 (English).

ADDITIONAL REFERENCES

- [20] E. Date, M. Jimbo, M. Kashiwara & T. Miwa. RIMS preprints 356–362. Kyoto University 1981.
- [21] V. G. Drinfel'd & V. V. Sokolov. Equations of KdV type and simple Lie algebras. *Dokl. Akad. Nauk SSSR* **258** (1981), 11–16.
- [22] S. Helgason. *Differential Geometry, Lie Groups and Symmetric Spaces*. Academic Press: New York, 1978.
- [23] V. G. Kac. Infinite dimensional algebras, Dedekind's η -function and the very strange formula. *Advances in Math.* **30** (1978), 85–136.