An Algebraic Study of Zero Curvature Equations

by

Ian R. McIntosh

September 1988

Submitted to the University of London for the degree of Doctor of Philosophy, and for the Diploma of Imperial College.

Dept. of Mathematics
Imperial College
Queen's Gate Rd.
London, SW7 2BZ.

Abstract.

A zero curvature equation is, in this thesis, a partial differential equation which can be written in the form

$$[\ \partial_x + q(x,t) - z\Lambda\ ,\ \partial_t - V_+(x,t,z)\ ] = 0$$

where q, V_+ are matrix-valued functions and Λ is a constant semisimple (i.e. diagonalizable) matrix.

In this thesis the work of Drinfel'd, V.G. & Sokolov, V.V. (*J. Sov. Math.*, 30 , *1975-2036,* (1985)) is generalized. We investigate the construction of these equations and prove that each one constructed belongs to a hierarchy of equations all of whose flows commute.

It is shown that each hierarchy is characterized by a triple $(\mathfrak{g},\theta,\Lambda)$ where $(\mathfrak{g},\theta)$ is a periodically graded semisimple Lie algebra and Λ is a semisimple element of degree one in the grading. We investigate which gradings admit such an element and classify these cases for the simple Lie algebras of rank $\le$ 4. We also investigate the equivalence of the hierarchies in terms of the conjugacy classes of semisimple elements of $\mathfrak{g}$.

A method is presented for constructing transformations of Miura-type, based on a method due to Drinfel'd & Sokolov (*op.cit.*). We show that, in certain cases where $\mathfrak{g}$ = $\mathfrak{sl}(\ell+1,\mathbb{C})$, these transformations can be explained using differential Galois theory (after Wilson, to be published). In these cases the appropriate Galois group induces a group of transformations on the set of solutions to the zero curvature equation.

CONTENTS

Acknowledgments.

It is my pleasure to thank Dr. George Wilson for his careful supervision during the past three years, I have learnt a great deal from him.

I had the good fortune to speak with many other experienced mathematicians during this time, in particular I wish to thank Drs. J.D.Gibbon, J.Gibbons, B.A.Kupershmidt and M.W.Liebeck for taking an interest in my work and keeping me informed. I am grateful to Prof.V.G.Kac for informing me of the theorem of Kempf & Ness stated in chapter 2. Without this knowledge chapter 2 would have been considerably longer and unnecessarily complicated. I also thank Prof.I.G.Macdonald for his opinion on some matters arising in chapter 2.

Finally, I am very grateful to the Association of Commonwealth Universities for their generous scholarship and to the British Council for their efficiency in the administration of this scholarship.

To Mum, Dad, Peter and Jeannette; another book for the coffee table.

Introduction.

By a zero curvature equation we will mean a p.d.e. which can be written as the integrability condition for a pair of first order differential equations

$$(\partial_x + q(x,t))\Omega = z\Lambda\Omega \tag{1}$$

$$\partial_t\Omega = V_+(x,t,z)\Omega \tag{2}$$

for the invertible matrix $\Omega(x,t,z)$. The functions $q(x,t)$, $V_+(x,t,z)$ are matrix-valued, with Λ a constant semisimple (i.e. diagonalizable) matrix. The first equation is a spectral problem for $\Omega(x,t,z)$, with z as the spectral parameter. If we assume the spectrum is t-independent the integrability condition for this system is given by

$$[\ \partial_x + q - z\Lambda\ ,\ \partial_t - V_+\ ] = 0 \tag{3}$$

This is the zero curvature equation associated with the spectral (or scattering) problem (1) with t-evolution (2). We are interested in these equations when they produce a system of p.d.e's describing the t-evolution of the (coordinates of the) matrix $q(x,t)$.

Strictly speaking, p.d.e's of this type were first considered in the paper of Ablowitz *et al.* (1974), although the idea is a natural extension of the work of Lax (1968) and, later, Zakharov & Shabat (1972, 1974).

In general zero curvature equations are non-linear p.d.e's. They have generated a great deal of interest because there is a well-formed theory about how to construct exact solutions (and, in particular, soliton solutions). In principle, solutions to a given zero curvature equation can be constructed using the "inverse scattering method" pioneered by Gardner *et al.* (1967). We will not attempt any analysis of solutions to the equations we investigate; a rigorous treatment of inverse scattering for zero curvature equations can be found in the paper by Beals & Coifman (1984).

In this thesis we will be wholly interested in the algebraic properties of zero curvature equations, in particular, their construction, their classification and their relationship to other equations via transformations of the variables.

We remove any need for analysis by considering the variables in the equations to be indeterminates generating a differential algebra. Thus, for example, the modified Korteweg-de Vries (mKdV) equation

$$q_t = q_{xxx} - 6q^2q_x \tag{4}$$

describes the derivation ∂_t on the differential algebra $\mathbb{C}\{q\}$ of polynomials in the symbols q , q_x , q_{xx} and so forth.

In fact we treat each matrix as the representation of an element of a Lie algebra (we will only be interested in semisimple Lie algebras). In particular, a matrix whose entries are polynomials in z will be seen as the representation of an element of a loop algebra over a semisimple Lie algebra. Thus $V_+(x,t,z)$ is considered as an element of the tensor product of the appropriate differential algebra with a given loop algebra. A brief description of the relevant facts about loop algebras over semisimple Lie algebras will be presented in the second section of chapter 1. It is assumed that the reader is familiar with the basic facts about the root space decomposition of semisimple Lie algebras; the reference most often used within is Helgason (1978, A.P.).

The idea of using loop algebras to investigate zero curvature equations is due to Drinfel'd & Sokolov (1981,1985) (see also Wilson (1981)). Their first important result was to prove that to each principally graded loop algebra there corresponds a hierarchy of zero curvature equations. Their proof makes use of a variation of the "method of dressing"; an important concept first described by Zakharov & Shabat (1974,1979). This concept will be explained in the first section of chapter 1, along with the variation used by Drinfel'd & Sokolov.

The rest of the first chapter is a direct extension of the work of Drinfel'd & Sokolov (1981,1985) along the lines suggested by Wilson (1981). It is shown that to each semisimple element of the form $z\Lambda$ in the loop algebra there corresponds a hierarchy of zero curvature equations, indexed by the centre of the centralizer of $z\Lambda$. All these equations possess the same spectral problem (1). We show that all the derivations in a given hierarchy commute, and that for each non-trivial equation there is a non-trivial conserved density which is conserved by all the "flows" in the hierarchy (in the literature these properties are sometimes used to justify calling the equations in a hierarchy "completely integrable"). We end the first chapter with an example of an equation constructed using a loop algebra over $\mathfrak{sl}(3,\mathbb{C})$.

The second chapter makes some headway towards classifying zero curvature equations. We begin by describing an equivalence relation on the set of hierarchies. From chapter 1 it becomes clear that each hierarchy is characterized by a triple $(\mathfrak{g},\theta,\Lambda)$ where $\mathfrak{g}$ is a semisimple Lie algebra, θ is a automorphism of $\mathfrak{g}$ of finite order and Λ is a semisimple element of $\mathfrak{g}$ satisfying $\theta(\Lambda)=\omega\Lambda$ for a primitive root of unity ω with the same order as θ. It is shown that two hierarchies $(\mathfrak{g},\theta,\Lambda)$ and $(\mathfrak{g},\theta,\Lambda')$ are equivalent (any equation from one can be transformed into an equation in the other) if Λ and Λ' are G_0-conjugate, where G_0 is the adjoint group of the Lie subalgebra $\mathfrak{g}_0 \leq \mathfrak{g}$ fixed pointwise by θ.

Larger equivalence classes are found for certain specializations of the hierarchies which result from setting equal to zero certain variables which are stationary with respect to every flow in the hierarchy. It is found that two such specialized hierarchies are equivalent if Λ and Λ' have their centralizers conjugate to one another under G_0. In particular, it is demonstrated that there is, up to equivalence, only one hierarchy corresponding to a principally graded loop algebra.

The problem of determining the existence of a hierarchy

$(\mathfrak{g},\theta,\Lambda)$, that is, the existence of a non-trivial semisimple element Λ with $\theta(\Lambda)=\omega\Lambda$, proves to be very difficult to solve in general. We are reduced to a virtually case by case analysis of the periodically graded simple Lie algebras and settle for solving the problem for rank$(\mathfrak{g})\le 4$. Nevertheless some results hold for arbitrary rank, for example, it is shown that every periodic grading on $\mathfrak{sl}(\ell+1,\mathbb{C})$ corresponding to an (inner) automorphism of type $(s_0,\ldots,s_\ell;1)$, where $s_i\in\{0,1\}$, admits a suitable semisimple element. The existence proof is constructive, that is to say, in each case where at least one suitable element exists the proof shows how to construct such a semisimple element. For the cases where rank$(\mathfrak{g})\le 4$ this information is contained in the table at the end of the fourth section of chapter 2.

Chapter 3 is entirely concerned with the investigation of transformations of "Miura type". The original Miura transformation (Miura (1968)) consists of setting

$$u = q_x - q^2 \qquad (5)$$

which transforms the mKdV equation (4) into the Korteweg-de Vries (KdV) equation

$$u = u_{xxx} + 6uu_x \qquad (6)$$

This equation can be written in "scalar Lax form" as

$$L_t = [\ P_3\ ,\ L\ ]$$

$$L = \partial_x^2 + u \qquad P_3 = 4\partial_x^3 + 6u\partial_x + 3u_x$$

The Miura transformation actually transforms each equation in the mKdV hierarchy into an equation which can be written in scalar Lax form using the same operator L as above. This transformation has an elegant explanation (due to Wilson, to be published) involving the use of differential Galois theory.

The differential Galois theory predicts the existence of a differential field $\mathbb{C}\langle\psi,\phi\rangle$ (the field of rational expressions in the quantities ϕ, ψ, ϕ_x, ψ_x and so on) upon which the group SL(2,$\mathbb{C}$) acts, such that the variable q is invariant under the action of a solvable subgroup of SL(2,$\mathbb{C}$), while u is invariant under the whole SL(2,$\mathbb{C}$) action. Wilson has shown that there exists a derivation ∂_t

on this field which is SL(2,$\mathbb{C}$)-equivariant, such that we obtain both the mKdV and KdV equations when we restrict ∂_t to the fixed field of the appropriate subgroup. The connection with the scalar Lax form is that the field $\mathbb{C}\langle\phi,\psi\rangle$ is the Picard-Vessiot extension associated with the operator L, and SL(2,$\mathbb{C}$) is the Galois group of this field over $\mathbb{C}\langle u\rangle$ (the reference for this is the book by Kaplansky (1957, Hermann)).

It is shown in chapter 3 that in general for SL(ℓ+1,$\mathbb{C}$) there exists a similar setup whereby an SL(ℓ+1,$\mathbb{C}$)-invariant equation exists which can be transformed into a zero curvature equation by dividing out by the action of a solvable subgroup of SL(ℓ+1,$\mathbb{C}$). We also show that the SL(ℓ+1,$\mathbb{C}$) action induces a group of transformations on the set of solutions to the zero curvature equation.

We then explain how these transformations of Miura type fit in with a different theory developed by Drinfel'd & Sokolov (1985). In particular we prove a result which generalizes the result obtained by Drinfel'd & Sokolov. The result we prove can be stated briefly as follows.

Each zero curvature equation belonging to the hierarchy $(\mathfrak{g},\theta,\Lambda)$ defines a derivation ∂_t on the differential algebra $\mathbb{C}\{q_i\}$. To each choice of a coarser grading $(\mathfrak{g},\sigma)$ (by which we mean every fixed point of θ is a fixed point of σ) we define a number of indeterminates r_m, c_m which we adjoin to $\mathbb{C}\{q_i\}$ and generate the differential extension $\mathcal{A}'=\mathbb{C}\{q_i,r_m,c_m\}$ of $\mathbb{C}\{q_i\}$. Provided Λ satisfies certain properties dependent upon the choice of $(\mathfrak{g},\sigma)$, we can prove that there exists a freely generated differential subalgebra $\mathbb{C}\{u_i\} \subset \mathbb{C}\{q_i\}$ characterized by the following property: $\mathbb{C}\{u_i\}$ is obtained by setting all r_m and c_m equal to zero in the subalgebra $\mathcal{C}'$ of $\mathcal{A}'$ fixed by a certain one-parameter group S of automorphisms of $\mathcal{A}'$ (let us stress the point that we cannot show, in general, that $\mathbb{C}\{u_i\}$ is the fixed algebra of a group of automorphisms of $\mathbb{C}\{q_i\}$).

Our main aim is to prove that ∂_t preserves $\mathbb{C}\{u_i\}$. This is done by showing that, after setting all $c_m \equiv 0$, ∂_t is identical to a derivation ∇_t which is S-equivariant on $\mathscr{A}'$, hence ∂_t preserves the algebra $\mathscr{C}$ obtained from $\mathscr{C}'$ by setting all $c_m \equiv 0$. Thus ∂_t restricts to $\mathbb{C}\{u_i\}$ after setting all $r_m \equiv 0$ in $\mathscr{C}$.

The variables u_i are obtained by a transformation we call the Miura-Drinfel'd-Sokolov transformation; this is the generalization of the transformation described by Drinfel'd & Sokolov (1985). At the end of chapter 3 we present some examples of these transformations, in particular we show that the equation constructed at the end of chapter 1 transforms into an equation used for modelling Langmuir waves.

CHAPTER 1

§1.1 Preview: the modified Korteweg-de Vries equation.

The archetypal zero curvature equation is the modified Korteweg-de Vries (mKdV) equation

$$q_t = q_{xxx} - 6q^2 q_x \qquad 1.1.1$$

It belongs to the hierarchy of equations which have the zero curvature representation

$$[\partial_x + q - z\Lambda \ , \ \partial_t - V_+] = 0 \qquad 1.1.2$$

where

$$q - z\Lambda = \begin{pmatrix} q & 0 \\ 0 & -q \end{pmatrix} - z\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and V_+ has the form

$$\begin{pmatrix} f(z^2,q) & zg(z^2,q) \\ zh(z^2,q) & -f(z^2,q) \end{pmatrix} \qquad 1.1.3$$

Here f, g and h are polynomials in z^2 whose coefficients are polynomials in q, q_x, $q_{xx}, \ldots$. For example, the mKdV equation corresponds to the choice

$$4z^3\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - 4z^2\begin{pmatrix} q & 0 \\ 0 & -q \end{pmatrix} + 2z\begin{pmatrix} 0 & q_x - q^2 \\ -q^2 - q_x & 0 \end{pmatrix} + \begin{pmatrix} 2q^3 - q_{xx} & 0 \\ 0 & q_{xx} - 2q^3 \end{pmatrix}$$

The purpose of this preview is to present some of the facts about this hierarchy of equations which are common to all the other hierarchies we will be dealing with more abstractly later. In particular, this first chapter is concerned with the construction of equations given a spectral operator $\partial_x + q - z\Lambda$.

The form of V_+ is governed by the condition

$$[V_+ \ , \ \partial_x + q - z\Lambda] = F(q, q_x, \ldots)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad 1.1.4$$

where $F(q, q_x, \ldots)$ is a polynomial in $q, q_x, \ldots$ (in particular this expression is independent of z). This is necessary to ensure that 1.1.2 is equivalent to an evolution equation for q; it follows that

$$\partial_t q = F(q, q_x, \ldots)$$

It is possible to describe all the matrices V_+ having the form 1.1.3 and satisfying 1.1.4. This is most clearly seen using a "dressing" method similar to the idea developed by Zakharov & Shabat (1974,1979). The first step is to demonstrate another characterization of V_+.

Suppose $V(q,z)$ is a formal Laurent series in z (possibly with an infinite principal part V_-) so that we may write

$$V(q,z) = V_+ + V_-$$
$$= (v_n z^n + \dots + v_0) + (v_{-1}z^{-1} + \dots) \qquad 1.1.5$$

The coefficients v_i are traceless 2×2 matrices whose entries depend upon q. If $V(q,z)$ is constructed so that it commutes with ∂_x+q−zΛ it follows that

$$[V_+ \, , \, \partial_x + q - z\Lambda] = [\partial_x + q - z\Lambda \, , \, V_-] \qquad 1.1.6$$

Notice that the left hand side contains no terms with negative powers of z, whereas the right hand side contains no terms with positive powers of z. Therefore 1.1.6 is independent of z. Moreover, if $V(q,z)$ has the form 1.1.3, with f, g and h now Laurent series in z^2, then 1.1.6 will be a diagonal matrix, as we require. We see then that each matrix V_+ is given by a series $V(q,z)$ commuting with $\partial_x+q-z\Lambda$. Therefore by describing the centralizer of $\partial_x+q-z\Lambda$ (in the algebra of formal Laurent series in z which have matrix coefficients depending upon q, q_x, etc.) we will be able to find the polynomials V_+ with the desired properties. This centralizer can be described in the following way.

The basic idea is to transform $\partial_x+q-z\Lambda$ into an operator whose centralizer is easier to determine, namely $\partial_x-z\Lambda$; this is done by constructing a formal series

$$\kappa = \kappa_{-1}z^{-1} + \dots . \qquad 1.1.7$$

whose coefficients $\kappa_{-k}(q)$ take values in $\mathfrak{sl}(2,\mathbb{C})$ so that

$$e^{\kappa}(\partial_x + q - z\Lambda)e^{-\kappa} = \partial_x - z\Lambda \qquad 1.1.8$$

The centralizer of $\partial_x-z\Lambda$ consists of all formal Laurent series $v(z)$, independent of x, which commute with $z\Lambda$. It follows that the series

$$e^{-\kappa}\, v(z)\, e^{\kappa}$$

will commute with $\partial_x + q - z\Lambda$ since

$$[\partial_x + q - z\Lambda \ , \ e^{-\kappa} \ v(z) \ e^{\kappa}]$$
$$= e^{-\kappa} \ [\partial_x - z\Lambda \ , \ v(z)] \ e^{\kappa} = 0$$

Any such series v(z) will be a sum of the matrices

$$z^{2n+1}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad n\in\mathbb{Z} \qquad 1.1.9$$

For simplicity we take v(z) to be one of these matrices. In particular if we take $n\in\mathbb{N}$ the series

$$V(q,z) = e^{-\kappa} \ v(z) \ e^{\kappa} \qquad 1.1.10$$

has a non-trivial positive part V_+.

Unfortunately it is not clear from this construction that the function $V_+(q,z)$ necessarily produces a polynomial $F(q,q_x,\ldots)$ as required in 1.1.4. For, if we look more carefully at 1.1.8 we observe that κ is defined by a collection of differential equations. These equations do not in general have solutions which are local expressions in q (i.e. they involve some integrals of q). Nevertheless it is true that $V_+(q,z)$ only contains local expressions in q. To see this we must adopt the modified dressing method suggested by Drinfel'd & Sokolov (1981,1985).

Drinfel'd & Sokolov point out that we need only conjugate $\partial_x + q - z\Lambda$ into

$$e^{\chi}(\partial_x + q - z\Lambda)e^{-\chi} = \partial_x - Z(q,z) \qquad 1.1.11$$

where $[Z(q,z),\Lambda] = 0$. Here

$$\chi = \chi_{-1}z^{-1} + \ldots$$

where $\chi_{-k}(q)$ takes values in $sl(2,\mathbb{C})$. It is readily seen that as a formal series in z we can write

$$Z(q,z) = z\Lambda - s(q,z) \qquad 1.1.12$$

where $s(q,z)$ is the principal part of $Z(q,z)$. The centralizer of Λ is commutative, so we still have

$$[\partial_x - Z(q,z) \ , \ v(z)] = 0$$

for any v(z) from 1.1.9. Therefore the series $e^{-\chi}v(z)e^{\chi}$ commutes with $\partial_x + q - z\Lambda$. The advantage of this construction is that it is possible to show that χ depends upon $q, q_x, \ldots$ but not upon integrals of q. Moreover, a uniqueness argument can be used to show that $V(q,z) = e^{-\chi}v(z)e^{\chi}$. Therefore V_+ will contain coefficients which are local expressions in q. These two important arguments will be presented later as part of the general program.

The series $Z(q,z)$ takes values in the centralizer of Λ, which is the one dimensional subalgebra generated by the matrix Λ. In fact the coefficients of the series expansion in z are conserved densities for the mKdV equation; the integrals of the coefficients are time-independent if $q(x,t)$ is a solution to the mKdV equation. We expect this from the following heuristic argument.

A (formal) solution to the spectral equation

$$(\partial_x + q - z\Lambda)\Omega(x,z) = 0 \qquad 1.1.13$$

is given by solving the equation

$$(\partial_x - Z(q,z))(e^{\chi}.\Omega) = 0 \qquad 1.1.14$$

This equation may be solved by using the integrating factor

$$\exp\int Z(q,z)\ dx$$

since the equation 1.1.14 is an equation on an abelian (one dimensional) subalgebra of $\mathfrak{sl}(2,\mathbb{C})$. From 1.1.12 it follows that

$$\Omega(x,z) = e^{-\chi}.\exp[-\int s(q,z)\ dx].e^{xz\Lambda} \qquad 1.1.15$$

Let us suppose for a moment that the potential $q(x)$ is a smooth, asymptotically rapidly vanishing function on $\mathbb{R}$, with values in $\mathbb{C}$. The limit

$$\lim_{x\to\infty} \Omega(x,z)e^{-xz\Lambda} = \lim_{x\to\infty} e^{-\chi}.\exp[-\int_{-\infty}^{x} s(q(x),z)\ dx] \qquad 1.1.16$$

is called the (formal) scattering matrix by Drinfel'd & Sokolov (1985). We treat this limit as a formal series in z. The so-called direct scattering problem involves evaluating the singular behaviour of this limit as a function of $z\in\mathbb{C}$ (for a good explanation of direct and inverse scattering on the line, see Beals & Coifman (1984)). Notice that, since the series χ is comprised of expressions in q, $q_x, \ldots,$ which vanish asymptotically, the limit 1.1.16 will be the formal series

$$\exp[-\int_{-\infty}^{\infty} s(q(x),z)\ dx] \qquad 1.1.17$$

If we let $q(x,t)$ be a solution to the mKdV equation with initial value $q(x,0) = q(x)$ we find that the limit in 1.1.16

evolves under the simple equation

$$\partial_t(\lim) - z^3\Lambda(\lim) = 0 \qquad 1.1.18$$

Therefore an invariant of t is given by

$$\lim_{x\to\infty} \Omega(x,t,z)\exp(-z\Lambda - tz^3\Lambda) = \exp[-\int_{-\infty}^{\infty} s(q(x,t),z)\ dx]$$

We conclude that the coefficients of the series in z

$$\int_{-\infty}^{\infty} (Z(q,z) - z\Lambda)\ dx = -\int_{-\infty}^{\infty} s(q(x,t),z)\ dx \qquad 1.1.19$$

are conserved integrals; their integrands will be conserved densities.

In fact these integrals are invariants of every one of the flows in the mKdV hierarchy. Later we will prove algebraically that a similar fact holds for any hierarchy of zero curvature equations. In general there are as many conserved densities as there are equations in a hierarchy (in fact there are infinitely many, all to be found amongst the coefficients of the appropriate series $Z(q,z)$)

Throughout this section we have been using a Lie algebra of Laurent polynomials in z with values in $\mathfrak{sl}(2,\mathbb{C})$ together with formal series in these polynomials. This Lie algebra is isomorphic to an algebra of maps from the circle S^1 into $\mathfrak{sl}(2,\mathbb{C})$ (i.e. loops in $\mathfrak{sl}(2,\mathbb{C})$) . To describe zero curvature equations in general we will replace $\mathfrak{sl}(2,\mathbb{C})$ by an arbitrary semisimple Lie algebra. In this matter we follow the lead of Drinfel'd & Sokolov (1981, 1985) and use the language of loop algebras. The next section presents a brief summary of the basic facts that will be used frequently; these facts are fully explained in, for example, the book by Helgason (1978, A.P.)

§1.2 Loop algebras.

Let $\mathfrak{g}$ be a semisimple Lie algebra of finite dimension over $\mathbb{C}$, on which a $\mathbb{Z}_m$-grading has been fixed (here m is a positive integer and $\mathbb{Z}_m$ is the cyclic group of integers modulo m). In other words

$$\mathfrak{g} = \bigoplus_{j\in\mathbb{Z}_m} \mathfrak{g}_j \quad \text{and} \quad [\mathfrak{g}_j , \mathfrak{g}_k] \subseteq \mathfrak{g}_{j+k} \tag{1.2.1}$$

for subspaces $\mathfrak{g}_j$ of $\mathfrak{g}$. We will say an element of $\mathfrak{g}_j$ has degree j in the grading.

Each $\mathbb{Z}_m$-grading corresponds to an m^{th}-order automorphism θ of $\mathfrak{g}$. Necessarily, θ has eigenvalues $\{ \omega^j | \ \omega$ a primitive m^{th}-root of unity; $j=0,\dots,m-1\}$; in this correspondence the subspaces $\mathfrak{g}_j$ are defined to be the eigenspaces of eigenvalue ω^j. It follows that each $\mathbb{Z}_m$-graded Lie algebra can be denoted by $(\mathfrak{g},\theta)$.

Consider the space $L(\mathfrak{g},e)$ of loops in $\mathfrak{g}$ which may be represented as Laurent polynomials i.e. maps of the form

$$u : S^1 \longrightarrow \mathfrak{g} \qquad \text{where} \qquad u(z) = \sum_{j=-k}^{n} u_j z^j \ , \ u_j \in \mathfrak{g}$$

This inherits a Lie algebra structure from $\mathfrak{g}$, indeed it is a covering algebra with covering homomorphism given by the evaluation map

$$\pi : L(\mathfrak{g},e) \longrightarrow \mathfrak{g}$$
$$u(z) \mapsto u(1)$$

For any graded Lie algebra $(\mathfrak{g},\theta)$ we can define the loop algebra $L(\mathfrak{g},\theta)$ to be the Lie subalgebra of $L(\mathfrak{g},e)$ containing all loops equivariant with respect to the action of ω on S^1 (by multiplication) and θ on $\mathfrak{g}$, i.e.

$$u \in L(\mathfrak{g},\theta) \qquad \text{if} \qquad \theta(u(z)) = u(\omega z).$$

It follows that $u \in L(\mathfrak{g},\theta)$ if and only if

$$u(z) = \sum_{j=-k}^{n} u_j z^j \ , \qquad u_j \in \mathfrak{g}_{j \bmod m} \tag{1.2.2}$$

The notation $L(\mathfrak{g},e)$ now makes sense if we let e denote the identity automorphism on $\mathfrak{g}$.

Let

$$L(\mathfrak{g},\theta)_j = \{u_j z^j \in L(\mathfrak{g},\theta)\} \qquad \text{for any } j \in \mathbb{Z}.$$

Clearly $L(\mathfrak{g},\theta)$ is a $\mathbb{Z}$-graded algebra with homogeneous subspaces

$$L(\mathfrak{g},\theta)_j \cong \mathfrak{g}_{j \bmod m}$$

It can be shown (see, for example, Helgason (1978, A.P.)) that, in particular, $L(\mathfrak{g},\theta)_0$ is a reductive Lie subalgebra i.e. it is the direct sum of a semisimple subalgebra and its centre. We may therefore choose a maximal abelian subalgebra $\mathfrak{h}_0$ of semisimple elements of $\mathfrak{g}_0$. The adjoint representation of $\mathfrak{h}_0$ on $L(\mathfrak{g},\theta)_j$ provides a weight space decomposition for $L(\mathfrak{g},\theta)$; these weight spaces are used as the "root spaces" for the *affine roots*.

1.2.3 Definition. *An element* $\alpha = (a,j) \in \mathfrak{h}_0^* \times \mathbb{Z}$ *is an affine root for* $L(\mathfrak{g},\theta)$ *if the subspace*

$$L(\mathfrak{g},\theta)^\alpha = \{ x \in L(\mathfrak{g},\theta)_j \mid [h,x]=a(h)x \ \forall h \in \mathfrak{h}_0 \}$$

is non-zero.

1.2.4. Example. Suppose θ is a finite order inner automorphism of $\mathfrak{g}$, and therefore θ fixes pointwise a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. Let R denote the root system of $\mathfrak{g}$. We will show that the affine root system Δ of $L(\mathfrak{g},\theta)$ can be identified with a subset of $R \cup \{0\} \times \mathbb{Z}$.

Since θ fixes $\mathfrak{h}$ pointwise we know that $\mathfrak{h} \subset \mathfrak{g}_0$. It follows that each maximal abelian subalgebra of semisimple elements in $L(\mathfrak{g},\theta)_0$ is isomorphic to $\mathfrak{h}$; we choose $\mathfrak{h}_0$ such that $\pi(\mathfrak{h}_0) = \mathfrak{h}$. For any $x \in L(\mathfrak{g},\theta)^\alpha \subset L(\mathfrak{g},\theta)_j$, where $\alpha=(a,j)$, the definition 1.2.3 implies

$$[\pi(h) \ , \ \pi(x)] = a(h)\pi(x) \qquad \forall h \in \mathfrak{h}_0$$

so

$$[h' \ , \ \pi(x)] = (a \circ \rho)(h')\pi(x) \qquad \forall h' \in \mathfrak{h}$$

where the homomorphism $\rho: \mathfrak{h} \mapsto \mathfrak{h}_0$ satisfies $\rho \circ \pi$ = identity. Therefore $a \circ \rho \in \mathfrak{h}^*$ is a root for $\mathfrak{g}$, so we can identify α with an element of $R \cup \{0\} \times \mathbb{Z}$. We conclude that, in this case, the affine root system Δ can be seen as a lattice covering $R \cup \{0\}$.

Just as for finite dimensional semisimple Lie algebras, the structure of each affine root system can be distilled into a matrix; a generalized Cartan matrix A. Each affine root system Δ has a basis $\{\alpha_i=(a_i,j_i) \mid i=0,\dots,\ell\}$ of simple affine roots, that is, Δ belongs to the lattice generated by this set of affine roots and $\alpha_i-\alpha_j \notin \Delta$ for all $i \neq j$. The columns of A provide a representation for the vectors $a_i \in$

$\mathfrak{h}_0^*$. It is a fact that $\ell = \dim \mathfrak{h}_0$ and that the matrix A is the direct sum of indecomposable generalized Cartan matrices of corank 1. The matrix A may in turn be represented by a diagram $\Gamma(A)$ analogous to the Dynkin diagram of a Cartan matrix. An indecomposable matrix has a connected affine diagram with $\ell+1$ vertices (representing the simple affine roots $\alpha_0,\dots,\alpha_\ell$).

The possible generalized Cartan matrices of affine type have been classified by Kac (1969), and the diagrams corresponding to the indecomposable matrices are shown at the end of this section (taken from Helgason (1978, A.P.)). The vertices $\{\alpha_i\}$ have corresponding integer labels $\{n_i\}$ which are the normalized coefficients of linear dependence of the columns of A. The fundamental theorem of Kac (1969) states that two loop algebras are isomorphic as Lie algebras if and only if they have the same affine diagram.

To characterize a loop algebra we must provide a (generalized) Cartan matrix A of affine type and a specific $\mathbb{Z}$-grading on a particular Lie algebra $L\mathfrak{g}(A)$ constructed from A. If we let $A=(a_{ij})$ we construct the Lie algebra $L\mathfrak{g}(A)$ by fixing a set of generators $\{e_i, h_i, f_i \mid i=0,\dots,\ell\}$ and demanding that they satisfy the relations

$$[e_i\ ,\ f_j] = \delta_{ij}h_i \qquad [h_i\ ,\ h_j] = 0$$

$$[h_i\ ,\ e_j] = a_{ji}e_j \qquad [h_i\ ,\ f_j] = -a_{ji}f_j \qquad 1.2.5$$

$$\underbrace{[..[e_i,e_j],\dots,e_j]}_{1-a_{ji}\ \text{times}} = 0 \qquad \underbrace{[..[f_i,f_j],\dots,f_j]}_{1-a_{ji}\ \text{times}} = 0 \qquad i\neq j$$

$$\sum_{i=0}^{\ell} n_i h_i = 0$$

(this last relation says that $L\mathfrak{g}(A)$ is the quotient of an affine Lie algebra by its one-dimensional centre).

The root spaces are determined with respect to the abelian subalgebra

$$\mathfrak{h}_0 = \langle\langle h_i \mid i=0,\dots,\ell\rangle\rangle$$

where this denotes the $\mathbb{C}$-subspace generated by the h_i. From 1.2.5 it can be shown that each one dimensional space $\langle\langle e_i\rangle\rangle$

is a root space; in fact

$$\langle\langle e_i \rangle\rangle = L(\mathfrak{g},\theta)^{\alpha_i} \qquad \text{for each } i=0,\ldots,\ell$$

A $\mathbb{Z}$-grading is then fixed on the Lie algebra $L\mathfrak{g}(A)$ in the following way. We choose $\ell+1$ non-negative integers $(s_0,\ldots,s_\ell)$, not all zero, and assign to e_i the degree s_i, to f_i the degree $-s_i$ and to h_i the degree 0. This is called a grading of type $(s_0,\ldots,s_\ell)$. We can deduce from this that the grading of type $(s_0,..,s_\ell)$ defines a map from Δ to $\mathbb{Z}$ which gives to each root the degree of its root space in the grading:

$$(s_0,\ldots,s_\ell) : \Delta \longrightarrow \mathbb{Z}$$
$$\alpha = \sum_0^\ell m_i\alpha_i \longmapsto \sum_0^\ell m_i s_i = \text{degree}(\alpha)$$

The homogeneous space of degree i is therefore

$$L(\mathfrak{g},\theta)_i = \sum (L(\mathfrak{g},\theta)^\alpha | \text{degree}(\alpha)=i)$$

Remark. An automorphism θ of $\mathfrak{g}$ corresponding to the grading of type $(s_0,\ldots,s_\ell)$ on $L\mathfrak{g}(A)$ is called an automorphism of type $(s_0,\ldots,s_\ell;k)$ if the diagram $\Gamma(A)$ is listed in table k, k = 1, 2 or 3. The integer k is the index of the automorphism θ, that is, θ induces a symmetry of order k on the Dynkin diagram of $\mathfrak{g}$. We note that the grading of type $(1,1,\ldots,1)$ on $L\mathfrak{g}(A)$ is called the k-principal grading by Kac (1985, C.U.P) when $\Gamma(A)$ belongs to table k. Unless there is a chance for confusion we shall refer to this grading simply as the principal grading. In particular, the (inner) automorphism of type $(1,1,\ldots,1;1)$ is called the Coxeter transformation, which we will denote by γ, so that the 1-principally graded loop algebra over $\mathfrak{g}$ may be denoted by $L(\mathfrak{g},\gamma)$.

Clearly any two sequences $(s_0,\ldots,s_\ell)$ which are equivalent under a symmetry of the affine diagram will induce $\mathbb{Z}$-gradings which are equivalent under an automorphism of $L\mathfrak{g}(A)$ induced by the symmetry.

We will find later that it is only the gradings of type $(s_0,\ldots,s_\ell)$ where this is a sequence of 0's and 1's (not all 0) which interest us.

1.2.6 Example. A realization of the loop algebras $L(\mathfrak{a}_1, e)$ and $L(\mathfrak{a}_1, \gamma)$.

Let us look at the diagram $\mathfrak{a}_1^{(1)}$ and the two loop algebras given by the gradings of type (1,0) and (1,1), which correspond to the trivial grading and the principal grading on $\mathfrak{a}_1$ respectively. We will represent $\mathfrak{a}_1$ as the matrix Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ and obtain a realization of each loop algebra.

A realization of the Lie algebra $L\mathfrak{g}(A)$, where $\Gamma(A)$ is the diagram $\mathfrak{a}_1^{(1)}$, is given by the assignment

$$e_0 = \begin{bmatrix} 0 & 0 \\ z & 0 \end{bmatrix} \qquad h_0 = -\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad f_0 = \begin{bmatrix} 0 & z^{-1} \\ 0 & 0 \end{bmatrix}$$

$$e_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad h_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad f_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

This provides the *standard realization* of this Lie algebra. It is not difficult to see that if we consider these as loops (functions of $z \in S^1$) in $\mathfrak{sl}(2,\mathbb{C})$ then this is a realization of the loop algebra $L(\mathfrak{a}_1, e)$ which has the grading of type (1,0). This realization is comprised of all matrices of the form

$$\begin{bmatrix} f(z) & g(z) \\ h(z) & -f(z) \end{bmatrix}$$

where $f(z), g(z)$ and $h(z)$ are Laurent polynomials in z.

The *principal realization* is given by the assignment

$$e_1 = \begin{bmatrix} 0 & z \\ 0 & 0 \end{bmatrix} \qquad h_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad f_1 = \begin{bmatrix} 0 & 0 \\ z^{-1} & 0 \end{bmatrix}$$

with e_0, h_0, f_0 as before. Once again the parameter z can be used to define the $\mathbb{Z}$-grading; in this case we obtain the grading of type (1,1) corresponding to $L(\mathfrak{g}, \gamma)$. This realization only contains those matrices of the form

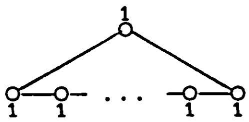
$$\begin{bmatrix} f(z^2) & zg(z^2) \\ zh(z^2) & -f(z^2) \end{bmatrix}$$

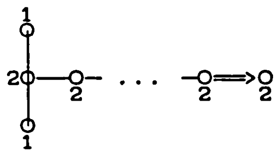
(c.f. 1.1.3)

Tables of affine diagrams.

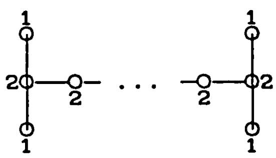
Table 1

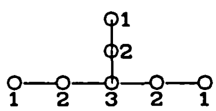
$\mathfrak{a}_1^{(1)}$ 

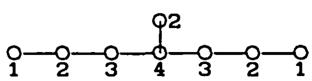
$\mathfrak{a}_n^{(1)}$ 

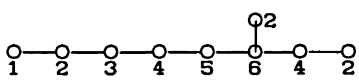
$\mathfrak{b}_n^{(1)}$ 

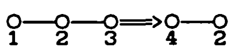
$\mathfrak{c}_n^{(1)}$ 

$\mathfrak{d}_n^{(1)}$ 

$\mathfrak{e}_6^{(1)}$ 

$\mathfrak{e}_7^{(1)}$ 

$\mathfrak{e}_8^{(1)}$ 

$\mathfrak{f}_4^{(1)}$ 

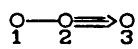
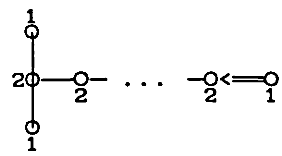
$\mathfrak{g}_2^{(1)}$ 

Table 2

$\mathfrak{a}_2^{(2)}$ 

$\mathfrak{a}_{2n}^{(2)}$ 

$\mathfrak{d}_{n+1}^{(2)}$ 

$\mathfrak{a}_{2n-1}^{(2)}$ 

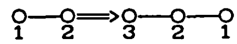
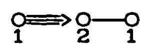
$\mathfrak{e}_6^{(2)}$ 

Table 3

$\mathfrak{d}_4^{(3)}$ 

§1.3 Formal construction of zero curvature equations.

In this section we present the method of Drinfel'd & Sokolov (1985) for the construction of zero curvature equations associated with a fixed loop algebra $L(\mathfrak{g},\theta)$. Each zero curvature equation is related to a matrix spectral problem

$$(\partial_x + q - z\Lambda)\Omega = 0 \qquad 1.3.1$$

where q is a function of x taking values in the homogeneous subspace $L(\mathfrak{g},\theta)_0$ of $L(\mathfrak{g},\theta)$, $z\Lambda \in L(\mathfrak{g},\theta)_1$ is a constant diagonalizable matrix and z is thought of as the spectral parameter.

In their paper, Drinfel'd & Sokolov chose the function $q(x)$ to be smooth and periodic. However, the dressing method introduces infinite series in z which have functions of q as coefficients. In order to avoid questions about the convergence of such expressions we will treat all such series as formal series. In any case there is no need to know the analytic properties of q. Consequently we will regard the object q as a sum $\sum q_i\varepsilon_i$, where the set $\{\varepsilon_i |$ $i=1,\ldots,n\}$ is a basis for $L(\mathfrak{g},\theta)_0$ and the q_i are indeterminates. We use these to define a differential algebra whose elements will play the part of functions of x.

Let $\mathcal{B} = \mathbb{C}[q_i^{(m)}]$ be the free algebra of polynomials in the infinitely many generators $q_i^{(m)}$, $i=1,\ldots,n$; $m \in \mathbb{N}\cup\{0\}$, where

$$q_i^{(m)} \equiv q_{i\overbrace{xx\ldots x}^{}} \text{ m-times}$$

A derivation ∂_x over $\mathbb{C}$ is defined on $\mathcal{B}$ by

$$\partial_x(q_i^{(m)}) = q_i^{(m+1)} \qquad ; \quad \partial_x(f.g) = \partial_x f.g + f.\partial_x g$$

for any $f,g \in \mathcal{B}$. The algebra with derivation $(\mathcal{B},\partial_x)$ is the differential algebra of differential polynomials in the indeterminates q_i. In future the differential algebra constructed from a set of indeterminates $\{a_i\}$ in this manner will be denoted by $\mathbb{C}\{a_i\}$.

Now define the vector spaces

$$L_+ = \bigoplus_{j=0}^{\infty} L(\mathfrak{g},\theta)_j \quad . \quad L_- = \prod_{j=1}^{\infty} L(\mathfrak{g},\theta)_{-j}$$

and let $L = L_+ \oplus L_-$. This vector space inherits a natural Lie algebra structure from $L(\mathfrak{g},\theta)$; in fact the only difference is that L_- incorporates formal infinite sums of elements of the loop algebra. Finally, for any finite dimensional vector space W define $\tilde{W} = W \otimes \mathcal{B}$. Then let

$$\tilde{L} = \tilde{L}_+ \oplus \tilde{L}_- \text{ , where } \tilde{L}_- = (\prod_{j=1}^{\infty} \tilde{L}(\mathfrak{g},\theta)_{-j})$$

The operator in 1.3.1 will be treated as an algebraic object

$$\partial_x + q - z\Lambda \in \partial_x + \tilde{L} \qquad 1.3.2$$

The idea behind the dressing method is to construct the centralizer in $\tilde{L}$ of this operator. For then, given any $V \in \tilde{L}$ such that

$$[\partial_x + q - z\Lambda \ , \ V] = 0 \qquad 1.3.3$$

we find that , if $V = V_+ + V_-$ is the decomposition of V in $\tilde{L}_+ \oplus \tilde{L}_-$, then the equation

$$[\partial_x + q - z\Lambda \ , \ V_-] = [V_+ \ , \ \partial_x + q - z\Lambda] \qquad 1.3.4$$

implies that this expression is an element of $\tilde{L}(\mathfrak{g},\theta)_0$. Consequently we may define a zero curvature equation by

$$\partial_V q = [V_+ \ , \ \partial_x + q - z\Lambda] \qquad 1.3.5$$

A crucial part of the construction of V is that Λ must be a semisimple element of $\mathfrak{g}_1$ i.e. $ad\Lambda$: $x \mapsto [\Lambda, x]$ is a diagonalizable matrix in $\mathfrak{gl}(\mathfrak{g})$. It follows that

$$\mathfrak{g} = \text{Kernel}(ad\Lambda) \oplus \text{Image}(ad\Lambda) \qquad 1.3.6$$

where $ad\Lambda$ is invertible on the latter subspace. The kernel of $ad\Lambda$ is covered by the centralizer $\mathfrak{z}(z\Lambda)$ of $z\Lambda$ in the loop algebra. The centre of the centralizer will be denoted by $c(\mathfrak{z}(z\Lambda))$. The centralizer of $z\Lambda$ in $\tilde{L}$ is slightly larger than $\mathfrak{z}(z\Lambda)\otimes\mathcal{B}$, it contains some infinite series. We will denote it by $\mathfrak{Z}(z\Lambda)$.

1.3.7 <u>Proposition.</u> *For each element* $z^k v_k \in c(\mathfrak{z}(z\Lambda))$ *which lies in a homogeneous subspace* $L(\mathfrak{g},\theta)_k$ $(k\in\mathbb{Z})$, *there exists a unique series of the form*

$$V = z^k v_k + z^{k-1} v_{k-1} + \dots \in \tilde{L}\ , \quad z^j v_j \in L(\mathfrak{g},\theta)_j \otimes \mathcal{B}$$

such that

$$[\partial_x + q - z\Lambda\ ,\ V] = 0$$

and each v_j *is comprised of homogeneous differential polynomials.*

Remark. We must be careful about the use of the symbol ∂_x . When it is found in an operator, it acts on an element $f \in \mathcal{B}$ as $\partial_x f = f_x + f\partial_x$. This is the correct expression for the action of ∂_x as an element of the ring $\mathcal{B}[\partial_x]$ of differential operators over $\mathcal{B}$. So for any $V \in \tilde{L}$, $[\partial_x\ ,\ V] = V_x$.

The proof of 1.3.7 relies on the next lemma , which is the basis of the dressing method. The idea is to conjugate the operator 1.3.2 into an operator whose centralizer is easier to determine.

1.3.8 Lemma. *There exists a unique series*

$$\chi = \sum_{j=1}^{\infty} z^{-j}\chi_{-j} \in \tilde{L}_- \cap \text{Image}(\ ad z\Lambda\ )$$

such that

$$e^{\chi}\ (\ \partial_x + q - z\Lambda\ )\ e^{-\chi} = \partial_x + Z(z\Lambda) \qquad 1.3.9$$

where $Z(z\Lambda) \in \mathfrak{Z}(z\Lambda)$ *(i.e.* $Z(z\Lambda)$ *commutes with* $z\Lambda$*).*

Remark. It is best to interpret the operation of conjugation $e^{\chi}\ x\ e^{-\chi}$ as notation for the formal series

$$\exp\ ad\chi\ (x) = x + [\chi,x] + \tfrac{1}{2}[\chi,[\chi,x]] + \dots \qquad 1.3.10$$

It is easily shown that this is a Lie algebra homomorphism. This avoids having to make sense of the object e^{χ} . Moreover, we extend this operation to include any derivation ∂_t on $\tilde{L}$. We write

$$\exp\ ad\chi\ (\partial_t) = \partial_t + [\chi,\partial_t] + \dots$$
$$= \partial_t - \chi_t + \dots$$

Proof of lemma 1.3.8 We expand the formal series on the left hand side of 1.3.7 in terms of its homogeneous components in the $\mathbb{Z}$-grading on $\tilde{L}$, i.e. we collect the coefficients of z^k.

$$\begin{aligned} \text{l.h.s.} &= q - z\Lambda + [\chi, q - z\Lambda] + \ldots + \partial_x + [\chi, \partial_x] + \ldots \\ &= -z\Lambda + (q - [\chi_{-1}, \Lambda]) + \ldots \qquad 1.3.11 \\ &\quad \ldots + \partial_x - z^{-1}((\chi_{-1})_x + \ldots) + \ldots \end{aligned}$$

We wish to find χ_{-1}, $\chi_{-2}, \ldots$, elements of Image($ad\Lambda$), uniquely such that each coefficient commutes with Λ. Since $ad\Lambda$ is invertible on its image, χ_{-1} can be uniquely chosen to cancel the image-component of q, leaving the component in the centralizer $\mathfrak{z}(\Lambda)$. But now we find that each term of lower degree 1-m < 0 has χ_{-m} occurring only in the expression $[\chi_{-m}, \Lambda]$; all other terms of degree 1-m will depend upon the known quantities q and χ_{-i}, i<m.

Thus each χ_{-m} can be chosen uniquely in image($ad\Lambda$) (by virtue of the invertibility of $ad\Lambda$ on its image) to annihilate the image-components of each coefficient of z^{1-m}. Moreover, since this argument has not involved integration, each χ_{-m} is comprised of elements of the differential algebra $\mathcal{B}$. ∎

<u>Proof of prop. 1.3.7</u>. For each $z^k v_k$ given in the proposition set

$$V = e^{-\chi} z^k v_k e^{\chi}$$

Since $(v_k)_x = 0$ and $z^k v_k$ commutes with all elements of $\mathfrak{z}(z\Lambda)$ we know that

$$[\partial_x + Z(z\Lambda) \ , \ z^k v_k] = 0$$

Conjugating this by $e^{-\chi}$ gives the equation 1.3.3.

We demonstrate the uniqueness of V by examining the homogeneous components of the equation 1.3.3, which have the form

$$[\partial_x + q \ , \ v_j] = [\Lambda \ , \ v_{j-1}] \qquad 1.3.12$$

Let us define the projections

$$i : \mathfrak{g} \longrightarrow \text{Image}(ad\Lambda)$$
$$k : \mathfrak{g} \longrightarrow \text{Kernel}(ad\Lambda)$$

To determine v_{k-1} we first notice that $i(v_{k-1})$ is uniquely given by $[\partial_x + q, v_k]$, since $ad\Lambda$ is invertible on its image.

The component $k(v_{k-1})$ is determined by the equation

$$k([\partial_x + q , v_{k-1}]) = 0$$

This is a differential equation for $k(v_{k-1})$ whose solution is guaranteed by the existence of the series V. The solution is unique up to a constant, which must be zero for each v_j to be comprised of homogeneous differential polynomials from $\mathcal{B}$. (This uniqueness proof is due to Wilson (1981)) ■

We will refer to the process of conjugating an object by $e^{-\chi}$ as the dressing operation, after the terminology of Zakharov & Shabat.

Let us now write $V = V_+ + V_-$ as before. We assign a derivation over $\mathbb{C}$ on $\mathcal{B}$, commuting with ∂_x, to the element $v = z^k v_k$.

1.3.13 <u>Definition.</u> *For each homogeneous element* $v \in c(\mathfrak{z}(z\Lambda))$ *define a derivation* ∂_v *on* $\mathcal{B}$ *by*

$$\begin{aligned}\partial_v q &= [V_+ , \partial_x + q - z\Lambda] \\ &= [\partial_x + q - z\Lambda , V_-]\end{aligned}$$

Accordingly, this gives us the zero curvature equation

$$[\partial_x + q - z\Lambda , \partial_v - V_+] = 0 \qquad 1.3.14$$

In fact, now that the dressing operation has been established, it is possible to assign to each element $u \in c(\mathfrak{z}(z\Lambda))$ a derivation ∂_u defined by 1.3.13 using the series $U = \exp(ad\chi)(u)$. This defines a linear map from $c(\mathfrak{z}(z\Lambda))$ to the algebra of derivations on $\mathcal{B}$ commuting with ∂_x . The next proposition demonstrates that all these derivations commute with each other.

1.3.15. <u>Proposition.</u> *Let* $u, v \in c(\mathfrak{z}(z\Lambda))$ *have respective derivations*

$$\begin{aligned}\partial_u q &= [U_+ , \mathcal{L}] \\ \partial_v q &= [V_+ , \mathcal{L}]\end{aligned}$$

where we set $\mathcal{L} = \partial_x + q - z\Lambda$. *Then* $[\partial_u , \partial_v] = 0$.

Proof. (c.f Wilson (1979))

$$\partial_u(\partial_v q) = [\partial_u V_+ , \mathcal{L}] + [V_+ , [U_+,\mathcal{L}]]$$

$$= [\partial_u V_+ , \mathcal{L}] + [[V_+,U_+], \mathcal{L}] + [U_+ ,[V_+,\mathcal{L}]]$$

by Jacobi's identity.

Subtracting a similar expression for $\partial_v(\partial_u q)$ we find

$$(\partial_u\partial_v - \partial_v\partial_u)q = [[\partial_u - U_+,\partial_v - V_+] , \mathcal{L}]$$

Thus $$[\partial_u , \partial_v]q = [\partial_u - U_+ , [\partial_v - V_+,\mathcal{L}]] + [\partial_v - V_+ , [\mathcal{L},\partial_u - U_+]] = 0 \quad \blacksquare$$

It is now apparent that by choosing a loop algebra $L(\mathfrak{g},\theta)$ and a semisimple element $z\Lambda \in \mathfrak{g}_1$ we can construct a hierarchy of commuting flows, with zero curvature representations, which are indexed by the centre of the centralizer of $z\Lambda$.

We now wish to show that this hierarchy possesses a number of integral invariants common to all flows. Let

$$K : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{C} \quad , \quad K(x,y) = \mathrm{Trace}(adx.ady).$$

be the Killing form on $\mathfrak{g}$. There exists a symmetric bilinear form on the loop algebra

$$K : L(\mathfrak{g},\theta) \times L(\mathfrak{g},\theta) \longrightarrow \mathbb{C}$$

defined by

$$K : L(\mathfrak{g},\theta)_j \times L(\mathfrak{g},\theta)_{-j} \longrightarrow \mathbb{C}$$

$$(u , v) \longmapsto K(\pi(u),\pi(v))$$

and all other pairs of homogeneous subspaces are orthogonal under K.

Let us define , for each $v \in c(\mathfrak{z}(z\Lambda))$, a differential polynomial

$$H_v(z\Lambda) = K(Z(z\Lambda),v) \in \mathcal{B} \qquad 1.3.16$$

where $Z(z\Lambda)$ is the series in 1.3.9.

1.3.17 Proposition. *For any $v \in c(\mathfrak{z}(z\Lambda))$, the $H_v(z\Lambda)$ are such that $\partial_u H_v(z\Lambda) \in \partial_x\mathcal{B}$, i.e. they are conserved densities for the "flows" ∂_u defined above.*

Proof. Once again let $\mathcal{L} = \partial_x + q - z\Lambda$. Then

$$\partial_u H_v(z\Lambda) = K(\partial_u Z(z\Lambda)\ ,\ v)$$
$$= K(\partial_u(e^{\chi}\ \mathscr{L}\ e^{-\chi})\ ,\ v)$$
$$= K(e^{\chi}\ ([U_+,\mathscr{L}] + [e^{-\chi}\partial_u e^{\chi},\mathscr{L}])\ e^{-\chi}\ ,\ v)$$
$$= K([\text{something in } \tilde{L}\ ,\ \partial_x + Z(z\Lambda)]\ ,\ v)$$
$$= -\partial_x\ K(\text{something},\ v) + (\text{something}\ ,\ [Z(z\Lambda),v])$$

which belongs to $\partial_x\mathscr{B}$ since v commutes with $Z(z\Lambda)$. ∎

§1.4 The conservation laws.

An expression of the form

$$\partial_v H = \partial_x F \qquad ,\ H,F \in \mathscr{B} \tag{1.4.1}$$

is called a conservation law for the flow ∂_v . If H, F are realized as smooth functions of x and a parameter v along the flow ∂_v , it follows that the definite integral of H over the domain of x (with suitable boundary conditions) is independent of v. We will see that there are at least as many non-trivial, independent conserved quantities as there are non-trivial, independent equations. This will be done by examining the connection between the conserved densities H_v and the derivations ∂_v .

The evolution equation associated with v in $c(\mathfrak{z}(z\Lambda))$ is

$$-\partial_v q = \partial_x v_0 + [q\ ,\ v_0] \tag{1.4.2}$$
$$= [v_{-1}\ ,\ \Lambda]$$

where v_i are the coefficients of z^i in the expression for V, the series obtained by dressing v. It will be shown that this equation can be written in the form

$$\partial_v q = \mathscr{J}\ \frac{\delta}{\delta q} H_v \tag{1.4.3}$$

where $\mathscr{J}$ is a certain operator and

$$\frac{\delta}{\delta q}\ :\ \mathscr{B} \longrightarrow L(\mathfrak{g},\theta)_0^* \otimes \mathscr{B}$$

is the formal Euler - Lagrange operator, whose components in the basis dual to $\{\varepsilon_i\}$, where $q = \sum q_i\varepsilon_i$, are

$$\frac{\delta}{\delta q_i} = \sum_{m=0}^{\infty} (-\partial_X)^m \frac{\partial}{\partial q_i^{(m)}} \qquad 1.4.4$$

This operator is characterized as follows. Let $\Omega_{\mathcal{B}}$ be the $\mathcal{B}$-module of 1-forms (or "Kähler differentials", see for example Matsumura (1962, Benjamin)) with universal derivation (exterior derivative)

$$\delta : \mathcal{B} \longrightarrow \Omega_{\mathcal{B}}$$

This is defined for the differential algebra $(\mathcal{B},\partial_X)$ so that δ commutes with the lift of ∂_X to $\Omega_{\mathcal{B}}$. $\Omega_{\mathcal{B}}$ is freely generated by the symbols $\delta q_i^{(m)}$, with $\partial_X\delta q_i^{(m)} = \delta q_i^{(m+1)}$. For any $H\in\mathcal{B}$ it is a straightforward calculation to check that

$$\delta H = \sum_{i,m} (\partial H/\partial q_i^{(m)}) \, \delta q_i^{(m)} \equiv \sum_i \frac{\delta H}{\delta q_i}\delta q_i \mod \partial_X\Omega_{\mathcal{B}}$$

In fact this relation fixes $\delta/\delta q$ as the component of δ on the $\mathcal{B}$-module $\mathcal{B}\langle\langle\delta q_i\rangle\rangle$ complementary to $\partial_X\Omega_{\mathcal{B}}$. It follows that $\partial_X\Omega_{\mathcal{B}}$ belongs to the kernel of $\delta/\delta q$. In particular notice that the equation 1.4.3 is only dependent on the choice of the conserved densities up to exact derivatives i.e. up to the freedom allowed by the conservation law 1.4.1.

1.4.5 <u>Lemma.</u> $\dfrac{\delta}{\delta q_i}H_v = K(v_0 , \dfrac{\delta q}{\delta q_i})$

<u>Proof.</u>

$$\delta H_v = \delta\, K(v,Z(\Lambda))$$
$$= K(v , \delta\{\exp(ad\chi)(\partial_X+q-z\Lambda) - \partial_X\})$$

By expanding $\delta\{\exp(ad\chi)(\partial_X+q-z\Lambda) - \partial_X\}$ we can write it in the form

$$[\mathcal{P} , \exp(ad\chi)(\partial_X+q-z\Lambda)] + \exp(ad\chi)\delta q$$
$$= [\mathcal{P},\partial_X + Z(\Lambda)] + \exp(ad\chi)\delta q$$
$$= -\partial_X\mathcal{P} + [\mathcal{P},Z(\Lambda)] + \exp(ad\chi)\delta q$$

where $\mathcal{P}$ is a series of terms from $L(\mathfrak{g},\theta)\otimes\Omega_{\mathcal{B}}$. Hence

$$\delta H_v = K(v,\exp(ad\chi)\delta q) - \partial_X K(v,\mathcal{P})$$

since

$$K([y,Z(\Lambda)],v) = K(y,[Z(\Lambda),v]) = 0$$

for any $y \in L(\mathfrak{g},\theta)$. Therefore $\delta H_V \equiv K(V,\delta q) \mod \partial_x \Omega_{\mathcal{B}}$.

The choice of the generators δq_i gives a unique splitting for $\Omega_{\mathcal{B}}$ as the direct sum $\mathcal{B}\langle\langle \delta q_i \rangle\rangle \oplus \partial_x \Omega_{\mathcal{B}}$. Therefore

$$\frac{\delta}{\delta q_i} H_V = K(\ V\ ,\ \frac{\delta q}{\delta q_i}) = K(\ V\ ,\ \frac{\delta q}{\delta q_i})$$

Moreover, it is clear that $\frac{\delta q}{\delta q_i} \in L(\mathfrak{g},\theta)_0$ from which the lemma follows. ■

The lemma allows us to find the element $\frac{\delta}{\delta q} H_V$ of $L(\mathfrak{g},\theta)_0^* \otimes \mathcal{B}$ by

$$\frac{\delta}{\delta q} H_V = K(v_0,\ .\ ) \qquad 1.4.6$$

Using the invertibility of K: $L(\mathfrak{g},\theta)_0 \longrightarrow L(\mathfrak{g},\theta)_0^*$, we conclude

$$v_0 = K^{-1}(\frac{\delta}{\delta q} H_V) \qquad 1.4.7$$

If we insert this into the evolution equation 1.4.2 we obtain

$$\partial_V q = -[\partial_x + q\ ,\ K^{-1}(\frac{\delta}{\delta q} H_V)] \qquad 1.4.8$$

Thus we write the equation in the form 1.4.3 by choosing the operator $\mathcal{J} = -ad(\partial_x + q) \circ K^{-1}$.

It follows that there at least as many independent, non-trivial conservation laws as there are independent (i.e. commuting) equations.

§1.5 An example of a zero curvature equation for a loop algebra over $\mathfrak{sl}(3,\mathbb{C})$

The simplest examples of zero curvature equations come from loop algebras over $\mathfrak{sl}(2,\mathbb{C})$; these are the equations in the mKdV hierarchy, corresponding to the principal grading (c.f §1.1), and the so-called AKNS hierarchy (Ablowitz *et al.* (1974)) which uses the standard grading. We will see later that these two hierarchies exhaust the possibilities for $\mathfrak{a}_1$, so we will look at $\mathfrak{a}_2$.

The simple Lie algebra $\mathfrak{a}_2$ is isomorphic to the algebra

$\mathfrak{sl}(3,\mathbb{C})$ of traceless 3×3 matrices of complex numbers. We choose a basis for $\mathfrak{sl}(3,\mathbb{C})$ which consists of the matrices

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

E_{ij} = the matrix with 1 in the i^{th} row, j^{th} column, 0 elsewhere.

Let $\mathfrak{h} = \langle\langle H_1, H_2 \rangle\rangle$ be the $\mathbb{C}$-vector space spanned by H_1 and H_2. Then $\mathfrak{h}$ is a Cartan subalgebra and each E_{ij} is a root vector (see, for example, Helgason (1978, A.P.)).

We define a finite order (inner) automorphism on $\mathfrak{sl}(3,\mathbb{C})$ by choosing the element

$$T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SL_3(\mathbb{C})$$

and defining

$$\begin{aligned} \theta : \mathfrak{sl}(3,\mathbb{C}) &\longrightarrow \mathfrak{sl}(3,\mathbb{C}) \\ X &\longmapsto TXT^{-1} \end{aligned}$$

This corresponds to the $\mathbb{Z}_2$-grading on $\mathfrak{sl}(3,\mathbb{C})$

$$\begin{aligned} \mathfrak{sl}(3,\mathbb{C}) &= \mathfrak{g}_0 \oplus \mathfrak{g}_1 \\ &= \langle\langle H_1, H_2, E_{12}, E_{21} \rangle\rangle \oplus \langle\langle E_{13}, E_{31}, E_{23}, E_{32} \rangle\rangle \end{aligned}$$

The loop algebra $L(\mathfrak{a}_2,\theta)$ is isomorphic to

$$\sum_{j\in 2\mathbb{Z}} (\mathfrak{g}_0 z^j \oplus \mathfrak{g}_1 z^{j+1})$$

In the classification of Kac (1969) T corresponds to an automorphism of type (1,0,1;1) on $\mathfrak{a}_2$ where root vectors for α_0, α_1, α_2 are respectively zE_{31}, E_{12} and zE_{23} in this representation.

We choose the semisimple element

$$\Lambda = E_{13} + E_{31} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \in \mathfrak{g}_1$$

and write the potential q for the spectral operator in the coordinate form

$$q = \begin{bmatrix} q+s & p & 0 \\ r & -2s & 0 \\ 0 & 0 & s-q \end{bmatrix}$$

The spectral operator is therefore represented by

$$\partial_x + \begin{bmatrix} q+s & p & -z \\ r & -2s & 0 \\ -z & 0 & s-q \end{bmatrix} \qquad 1.5.1$$

A simple computation shows that the centralizer of Λ is commutative, i.e. Λ is regular semisimple, so that

$$\mathfrak{z}(z\Lambda) = c(\mathfrak{z}(z\Lambda)) = \sum_{j\in 2\mathbb{Z}} (\langle\!\langle H_1 - H_2 \rangle\!\rangle z^j \oplus \langle\!\langle E_{13} + E_{31} \rangle\!\rangle z^{j+1})$$

We will construct a zero curvature equation by choosing the (homogeneous) element $z^2 v_2 = z^2(H_1 - H_2)$ of this centralizer and solving the equation

$$[\partial_x + q - z\Lambda \,,\ z^2 v_2 + z v_1 + v_0 + \dots\ ] = 0$$

for v_1 and v_0. The zero curvature equation will be

$$\partial_t q = [\ v_0 \,,\ q\ ] - \partial_x v_0 \qquad 1.5.2$$

The computation yields

$$V_+ = z^2 v_2 + z v_1 + v_0$$

$$= \begin{bmatrix} z^2 & 3(pq - p_x + 3ps) & 0 \\ 3(rq + r_x + 3rs) & 3pr - 2z^2 & -3zr \\ 0 & -3zp & z^2 \end{bmatrix}$$

The equations 1.5.2 are

$$\begin{aligned} q_t &= -3(pr)_x \\ p_t &= 3(p_{xx} - pq_x - pq^2 - p^2 r) \\ &\qquad - 9s(pq - p_x + 3ps) - 3(ps)_x \\ r_t &= -3(r_{xx} + rq_x - rq^2 - r^2 p) \\ &\qquad + 9s(rq + r_x + 3rs) - 3(rs)_x \\ s_t &= 0 \end{aligned} \qquad 1.5.3$$

Notice that the variable s is stationary with respect to this flow , so that it only plays the part of a parameter. It will be shown later that this is always the case for variables corresponding to the centralizer of $z\Lambda$. Consequently we may set $s = 0$ to obtain the slightly simpler system

$$\begin{aligned} q_t &= -3(pr)_x \\ p_t &= 3(p_{xx} - pq_x - pq^2 - p^2 r) \\ r_t &= -3(r_{xx} + rq_x - rq^2 - r^2 p) \end{aligned} \qquad 1.5.4$$

Notice that q is a conserved density for this system. In general it is difficult to compute the conserved densities defined in 1.3.16, but Wilson (1981) derived other conserved densities given by

$K(z^{-1}v_{-1} , z\Lambda)$ for $V = v + \ldots + v_{-1} + \ldots$ 1.5.5

This is necessarily congruent to $H_V(z\Lambda)$ modulo exact derivatives since Wilson has shown that 1.5.5 is also a "Hamiltonian" for the equation characterized by v, i.e. we can substitute 1.5.5 for H_V in 1.4.3. For the system 1.5.4 the conserved density given by 1.5.5, for $s = 0$, is

$$\frac{3}{2}(pr_x - rp_x) - 3pqr$$

Remark. It should be explained that we do not expect to be able to write the equations 1.5.4 in the form 1.4.3 since, by setting s equal to zero, the potential q belongs to a subspace of $L(\mathfrak{g},\theta)_0 \otimes \mathcal{B}$ on which the form K is degenerate. Therefore we cannot repeat the construction given in 1.4.9. Nevertheless the quantity $\frac{3}{2}(pr_x - rp_x) - 3pqr$ will still be a conserved density since it is obtained by setting $s = 0$.

An interesting specialization of these equations (or more correctly, a slight modification of 1.5.4) is obtained by restricting the potential q to take values in a real form of $\mathfrak{g}$, that is, in a Lie subalgebra $\mathfrak{u}$ of the $\mathbb{R}$-Lie algebra generated by $\mathfrak{g}$ whose complexification is $\mathfrak{g}$. A trivial example of this is to take the real form $\mathfrak{sl}(3,\mathbb{R})$ of $\mathfrak{sl}(3,\mathbb{C})$. Then the equations 1.5.4 specialize to the identical equations with q, p and r $\mathbb{R}$-valued functions.

A slightly more interesting example is to take the subalgebra of skew Hemitian matrices, $\mathfrak{su}(3,\mathbb{C})$. Then the potential q will be restricted to be

$$q_R = \begin{bmatrix} iq & p & 0 \\ -\bar{p} & 0 & 0 \\ 0 & 0 & -iq \end{bmatrix}$$

where q is now an $\mathbb{R}$-valued function. Of course we must also require Λ and v_2 to be skew Hermitian, which can be done by replacing the current choices by $\Lambda = i(E_{13}+E_{31})$ and $v_2 = i(H_1 - H_2)$. These choices modify the equations in $\mathfrak{sl}(3,\mathbb{C})$ by multiplying the right hand side of 1.5.4 by i. The specialization to $\mathfrak{su}(3,\mathbb{C})$ then gives

$$q_t = -3(|p|^2)_x$$
$$ip_t = 3(p_{xx} + p|p|^2 + p(q^2 - iq_x))$$

A similar specialization can be obtained using the real form $\mathfrak{su}(2,1)$.

We will see later (in §3.6) that the equations 1.5.4 also possess a transformation of Miura type.

CHAPTER 2

§2.1 The determination of equivalent hierarchies of zero curvature equations.

A hierarchy of zero curvature equations is determined by a choice of the spectral operator

$$\partial_x + q - z\Lambda \qquad\qquad q \in \tilde{\mathfrak{g}}_0$$

or, more correctly, by a choice of loop algebra $L(\mathfrak{g},\theta)$ (which determines the potential q and therefore the differential algebra $\mathcal{B}$) and a choice of non-zero constant semisimple element $z\Lambda \in L(\mathfrak{g},\theta)_1$. Equivalently, we will consider a hierarchy to be given by a triple $(\mathfrak{g},\theta,\Lambda)$ consisting of a $\mathbb{Z}_m$-graded semisimple Lie algebra $(\mathfrak{g};\theta)$ and a non-zero semisimple element $\Lambda\in\mathfrak{g}_1$.

The question of the existence of such an element Λ for a given $(\mathfrak{g},\theta)$ will be left until later; it is a difficult problem to classify these cases. The purpose of this section is to define a notion of equivalence between hierarchies and then determine when two hierarchies are equivalent in this sense.

2.1.1. Definition. *Let $\mathcal{B}$ and $\mathcal{B}'$ be two differential algebras and let $\mathcal{D}$ and $\mathcal{D}'$ be two collections of derivations, on $\mathcal{B}$ and $\mathcal{B}'$ respectively. We will say that $\mathcal{D}$ and $\mathcal{D}'$ are equivalent if there exists an isomorphism between $\mathcal{B}$ and $\mathcal{B}'$ that induces a bijection between $\mathcal{D}$ and $\mathcal{D}'$.*

Remark. For any isomorphism $\varphi:\mathcal{B} \longrightarrow \mathcal{B}'$ of the algebras , a derivation ∂ on $\mathcal{B}$ induces a derivation $\varphi\circ\partial\circ\varphi^{-1}$ on $\mathcal{B}'$.

Using this definition we will consider the equivalence of hierarchies as collections of derivations on the differential algebra $\mathcal{B}$.

An obvious case where two hierarchies are equivalent is the following. Let G_0 denote the adjoint group of transformations on $\mathfrak{g}_0$; it is a subgroup of the adjoint group G of $\mathfrak{g}$. Since $[\mathfrak{g}_0, \mathfrak{g}_i] \subseteq \mathfrak{g}_i$ for each homogeneous subspace $\mathfrak{g}_i \subseteq \mathfrak{g}$, G_0 has a representation on each $\mathfrak{g}_i$ which can be lifted to $L(\mathfrak{g}, \theta)_i$. In particular, G_0 has a representation as a subgroup of $GL(\mathfrak{g}_1)$. It is a straightforward conclusion from the definition of semisimple elements that every element in the G_0-orbit of a semisimple element is semisimple. It follows that the hierarchy $(\mathfrak{g}, \theta, \Lambda)$ belongs to a collection of hierarchies $\{(\mathfrak{g}, \theta, g.\Lambda) \mid g \in G_0\}$.

2.1.2. <u>Lemma.</u> *All the hierarchies in the collection* $\{(\mathfrak{g}, \theta, g.\Lambda) \mid g \in G_0\}$ *are equivalent.*

<u>Proof.</u> Each equation in the hierarchy $(\mathfrak{g}, \theta, \Lambda)$ is given by a constant element $v \in c(\mathfrak{z}(z\Lambda))$ and characterized by the unique series $V = v + \ldots \in \tilde{L}$ satisfying

$$[\ \partial_x + \ q - z\Lambda\ ,\ V\ ] = 0$$

For any $g \in G_0$ the adjoint action is a Lie algebra homomorphism, so that

$$[\ \partial_x + g.q - g.z\Lambda\ ,\ g.V\ ] = 0 \qquad 2.1.3$$

But clearly this series $g.V = g.v + \ldots$ characterizes the equation for the derivation $\partial_{g.v}$ in the hierarchy $(\mathfrak{g}, \theta, g.\Lambda)$. We can write

$$g.q = \sum q_i(g.\varepsilon_i) = \sum p_i \varepsilon_i \qquad p_i \in \mathcal{B}$$

and identify this with the potential p in the operator

$$\partial_x + p - z(g.\Lambda)$$

associated with the hierarchy $(\mathfrak{g}, \theta, g.\Lambda)$. Therefore the differential algebra $\mathcal{B}$ for $(\mathfrak{g}, \theta, \Lambda)$ is isomorphic to the algebra $\mathcal{B}'$ for $(\mathfrak{g}, \theta, g.\Lambda)$. This isomorphism induces the bijection $\partial_v \longmapsto \partial_{g.v}$ between the hierachies $(\mathfrak{g}, \theta, \Lambda)$ and $(\mathfrak{g}, \theta, g.\Lambda)$. ∎

Recall that the equations of the hierarchy $(\mathfrak{g}, \theta, \Lambda)$ are indexed by the centre $c(\mathfrak{z}(z\Lambda))$ of the centralizer of $z\Lambda$ in the loop algebra. Now suppose $\Lambda' \in \mathfrak{g}_1$ is another semisimple element with the same centralizer as Λ; $\mathfrak{z}(\Lambda') = \mathfrak{z}(\Lambda)$. Necessarily $c(\mathfrak{z}(z\Lambda')) = c(\mathfrak{z}(z\Lambda')$. We intend to show that in this case the hierarchies $(\mathfrak{g},\theta,\Lambda')$ and $(\mathfrak{g},\theta,\Lambda)$ are equivalent after a specialization, namely, after setting some of the dependent variables equal to zero.

Recall from the previous chapter that the equation

$$[\ \partial_x + q - z\Lambda\ ,\ \partial_v - V_+\ ] = 0$$

can also be written as

$$-[\ \partial_x + q\ ,\ v_0\ ] = \partial_v q = [\ \Lambda\ ,\ v_{-1}\ ] \qquad 2.1.4$$

where $V = v + \ldots + v_0 + v_{-1} + \ldots$. If we recall also the projections i, k of $\mathfrak{g}$ onto the image and kernel of $ad\Lambda$ respectively, we notice that 2.1.4 implies

$$\partial_v k(q) = k([\ \Lambda\ ,\ v_{-1}\ ]) = 0$$

Consequently we may set to zero the coordinates of $k(q)$ without disturbing the consistency of the equations 2.1.4 (we choose a basis of $\mathfrak{g}_0$ which is compatible with the splitting $\mathfrak{g}_0 = k(\mathfrak{g}_0) \oplus i(\mathfrak{g}_0))$. Moreover, it can readily be seen that the effect of this specialization is equivalent to constructing the hierarchy based on the operator

$$\partial_x + i(q) - z\Lambda \qquad 2.1.5$$

The whole mechanism of the dressing method is still valid; for each $v \in c(\mathfrak{z}(z\Lambda))$ there exists a unique series $\dot{V} = v + \ldots$ whose homogeneous terms belong to $L(\mathfrak{g},\theta)_j \otimes \dot{\mathcal{B}}$, where $(\dot{\mathcal{B}},\partial_x)$ is the differential algebra of differential polynomials in the coordinates of $i(q)$. The equation

$$[\ \partial_x + i(q) - z\Lambda\ ,\ \partial_v - \dot{V}\ ] = 0 \qquad 2.1.6$$

is precisely the specialization of 2.1.4 above. We will denote the hierarchy of specialized equations taken from $(\mathfrak{g},\theta,\Lambda)$ by $\{\mathfrak{g},\theta,\Lambda\}$.

<u>Remark.</u> In practice these specialized equations are of equal significance to the original equations. The coordinates for $k(q)$ are stationary with respect to all the derivations in the hierarchy, therefore they only play the part of parameters in the equations. For example, in §1.5 we found the equations 1.5.4 much easier to handle than the full equations 1.5.3.

Note also that, formally, the algebra $\dot{\mathcal{B}}$ is described as the quotient of $\mathcal{B}$ by the differential ideal $\mathcal{K}$ generated by the coordinates of $k(q)$. The derivation ∂_v on $\mathcal{B}$ preserves $\mathcal{K}$, for if q_i is a coordinate of $k(q)$ and $r\in\mathcal{K}$ then $\partial_v(q_i r) = q_i\partial_v r \in \mathcal{K}$. Thus we can push ∂_v onto $\mathcal{B}/\mathcal{K}$ and its defining equation will be 2.1.6, since $i(q)\equiv q \bmod \mathcal{K}$.

The next proposition is a key result about the equivalence of these specialized hierarchies. The proof will need to be developed in a series of lemmas.

2.1.7. <u>Proposition.</u> *Let* Λ, Λ' *be two semisimple elements satisfying* $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$. *Then the hierarchies* $\{\mathfrak{g},\theta,\Lambda\}$ *and* $\{\mathfrak{g},\theta,\Lambda'\}$ *are equivalent.*

The idea behind the proof is quite simple. The next lemma shows that there is an isomorphism between the abelian subalgebra $c(\mathfrak{z}(z\Lambda))$ and the algebra of operators commuting with $\partial_x+i(q)-z\Lambda$. Similarly, there is an algebra of commuting operators isomorphic to $c(\mathfrak{z}(z\Lambda')$ corresponding to the hierarchy $\{\mathfrak{g},\theta,\Lambda'\}$. We will see that when $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$ we can identify these two algebras of operators and thereby identify the zero curvature equations of one hierarchy with those of the other.

2.1.8. <u>Lemma.</u> *For any* u, v $\in c(\mathfrak{z}(z\Lambda))$ *the two equations*

$$\begin{aligned} [\ \partial_x + i(q) - z\Lambda\ ,\ \partial_u - \dot{U}_+\ ] &= 0 \\ [\ \partial_x + i(q) - z\Lambda\ ,\ \partial_v - \dot{V}_+\ ] &= 0 \end{aligned} \qquad 2.1.9$$

imply

$$[\ \partial_u - \dot{U}_+\ ,\ \partial_v - \dot{V}_+\ ] = 0$$

<u>Proof.</u> Using 2.1.9 and the Jacobi identity we find that

$$[\ [\ \partial_u-\dot{U}_+\ ,\ \partial_v-\dot{V}_+\ ]\ ,\ \partial_x + i(q) - z\Lambda\ ] = 0 \qquad 2.1.10$$

Now let $f = [\ \partial_u-\dot{U}_+\ ,\ \partial_v-\dot{V}_+\ ] \in L_+\otimes\dot{\mathcal{B}}$. In the $\mathbb{Z}$-grading on $L_+\otimes\dot{\mathcal{B}}$ we can write $f = f_0 + zf_1 + \ldots + z^n f_n$. The component of degree zero in 2.1.10 is

$$\partial_x f_0 + [\ i(q)\ ,\ f_0\ ] = 0$$

Since $\partial_x f_0$ has one more derivative than $[i(q), f_0]$ can have, f_0 does not belong to the differential algebra unless it is zero. If we repeat this argument for each f_i, $i=1,\dots,n$ successively, we discover that $f \equiv 0$. ∎

In fact it was not necessary to specialize the hierarchy to prove this lemma; the reason for the specialization is the following.

If we choose Λ' such that $\mathfrak{z}(\Lambda') = \mathfrak{z}(\Lambda)$ then there exists an operator

$$\partial_y + p - z\Lambda' \qquad 2.1.11$$

such that

$$[\ \partial_y + p - z\Lambda'\ ,\ \partial_v - \dot{V}_+\ ] = 0 \qquad 2.1.12$$

for all the operators in the hierarchy $\{\mathfrak{g}, \theta, \Lambda\}$. In particular

$$[\ \partial_y + p - z\Lambda'\ ,\ \partial_x + i(q) - z\Lambda\ ] = 0 \qquad 2.1.13$$

where we have written ∂_y for $\partial_{z\Lambda'}$. With a change of perspective we intend to view the hierarchy $\{\mathfrak{g}, \theta, \Lambda\}$ as the hierarchy $\{\mathfrak{g}, \theta, \Lambda'\}$ using 2.1.11 as the spectral operator. The specialization is necessary so that the object p can be taken to be a "potential" for a spectral operator; we will see that $p = i(p)$, which forces us to consider only spectral operators having values in $i(\mathfrak{g}_0)$.

In order to show that the hierarchies for Λ and Λ' are equivalent we need to find an isomorphism between their respective differential algebras. To do this we need the following lemma.

2.1.14. Lemma. *The maps $ad\Lambda$, $ad\Lambda'$ induce the same splitting*

$$\mathfrak{g} = k(\mathfrak{g}) \oplus i(\mathfrak{g})$$

if $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$.

Proof. Of course $k(\mathfrak{g}) = \mathfrak{z}(\Lambda)$, so we want to show that the images of $ad\Lambda$ and $ad\Lambda'$ are the same. Let $c = c(\mathfrak{z}(\Lambda)) =$

$c(\mathfrak{z}(\Lambda'))$. This is an abelian subalgebra of semisimple elements of $\mathfrak{g}$, so its adjoint representation provides a weight space decomposition

$$\mathfrak{g} = \sum_{\lambda\in W} \mathfrak{g}^{\lambda} \qquad\qquad W \subset c^{*}$$

where

$$\mathfrak{g}^{\lambda} = \{\ x\in\mathfrak{g}\ |\ [c,x] = \lambda(c)x \quad \forall\ c\in c\ \}$$

and $\qquad W = \{\ \lambda\in c^{*}\ |\ \mathfrak{g}^{\lambda} \neq \{0\}\ \}$

It follows that for any $c\in c$

$$\mathrm{Image}(adc) = \sum (\ \mathfrak{g}^{\lambda}\ |\ \lambda(c) \neq 0\ )$$

Both Λ, $\Lambda' \in c$, and since $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$ we conclude that $\lambda(\Lambda) = 0$ if and only if $\lambda(\Lambda') = 0$, when $\lambda\in W$. Therefore the images of $ad\Lambda$ and $ad\Lambda'$ are the same. ∎

2.1.15. Lemma. *There exists an invertible linear transformation* $C\colon i(\mathfrak{g}) \longrightarrow i(\mathfrak{g})$ *such that* $p = C(i(q))$.

Proof. Firstly, $p = i(p)$ by virtue of its construction; p is the component of degree zero in the infinite series $\exp ad(-\chi)(z\Lambda')$ from the dressing operation (there is no difficulty in "specializing" the entries of χ). Therefore $p = -[\chi_{-1},\Lambda']$, which belongs to $i(\mathfrak{g})\otimes\dot{\mathcal{B}}$ by the previous lemma.

From the equation 2.1.13 we have the identity

$$[\Lambda'\ ,\ i(q)] = [\Lambda\ ,\ p]$$

Therefore the lemma follows if we set $C = (ad\Lambda)^{-1}\circ(ad\Lambda')$, which is invertible on $i(\mathfrak{g})$. ∎

The result of this is that we may use the coordinates of p, call them p_i, as generators for the differential algebra $(\dot{\mathcal{B}},\partial_x)$. The equation 2.1.13 gives the relationship between ∂_x and ∂_y as

$$\partial_x p + [\ i(q)\ ,\ p\ ] = \partial_y i(q) \qquad 2.1.16$$

It follows that if we replace $i(q)$ by $C^{-1}(p)$ every differential polynomial in $\dot{\mathcal{B}}$ can be written as a differential polynomial in the y-derivatives of the coordinates p_i.

Now let us write the spectral operator associated with the hierarchy $\{\mathfrak{g},\theta,\Lambda'\}$ as

$$d_y + \rho - z\Lambda' \qquad 2.1.17$$

so that the differential algebra for this hierarchy is the algebra $(\mathcal{P},d_y)$ of all polynomials in the coordinates ρ_i and their formal derivatives with respect to d_y .

There is an obvious isomorphism between $(\mathcal{P},d_y)$ and $(\mathring{\mathcal{B}},\partial_x)$ where we map ρ_i it q_i and d_y to ∂_x . However, we wish to relate the two operators 2.1.17 and 2.1.11.

2.1.18. Lemma. *Let $\ell:(\mathcal{P},d_y) \longrightarrow (\mathring{\mathcal{B}},\partial_x)$ be the homomorphism of differential algebras characterized by $\ell(\rho_i) = p_i = C(i(q))_i$ and $\ell\circ d_y = \partial_y\circ\ell$. Then this is a isomorphism.*

Proof. The image of ℓ will clearly be the differential algebra generated by all the p_i and their y-derivatives, for which the isomorphism with $(\mathcal{P},d_y)$ is clear. However, the differential algebra $(\mathring{\mathcal{B}},\partial_x)$ is identical to the differential algebra generated by all p_i and their y-derivatives due to the lemma 2.1.15 and the relation 2.1.16. ■

We are now in a position to prove proposition 2.1.7. It will be shown that the zero curvature equations in $\{\mathfrak{g},\theta,\Lambda'\}$ constructed using the dressing method on the operator 2.1.17 map, under ℓ, to the equations 2.1.12 in $\{\mathfrak{g},\theta,\Lambda\}$.

Proof of proposition 2.1.7. It suffices to exhibit this mapping for the equations given by homogeneous elements $v = z^k v_k \in c(\mathfrak{z}(z\Lambda))$. By proposition 1.3.7 there exists a unique series

$$\mathcal{V} = z^k v_k + z^{k-1}\omega_{k-1} + \ldots + \omega_0 + \ldots$$

with homogeneous terms in $L(\mathfrak{g},\theta)\otimes\mathcal{P}$, satisfying

$$[\ d_y + \rho - z\Lambda'\ ,\ \mathcal{V}\ ] = 0$$

We define a derivation d_v on $(\mathcal{P},d_y)$ by

$$[\ d_y + \rho - z\Lambda'\ ,\ d_v - \mathcal{V}_+\ ] = 0$$

This equation is founded on the identities

$$[\, d_y + \rho \,,\, \omega_j \,] = [\, \Lambda' \,,\, \omega_{j-1} \,] \qquad j\geq 1 \qquad 2.1.19$$
$$d_y k(\omega_0) + k([p,\omega_0]) = 0$$

which uniquely determine ω_0 given v_k .

However, clearly the equation

$$[\, \partial_y + p - z\Lambda' \,,\, \partial_v - \bar{V}_+ \,] = 0$$

in $\{\mathfrak{g},\theta,\Lambda\}$, where $\dot{V}_+ = z^k v_k + z^{k-1} v_{k-1} + \dots + v_0$, implies that the terms v_j, $j\geq 1$, are the unique solutions to the image under ℓ of the equations 2.1.19. It follows that $\ell(\omega_0) = v_0$. Therefore the differential polynomials $d_v\rho_1 \in \mathcal{P}$ given by

$$d_v\rho = -[\, d_y + \rho \,, \omega_0 \,]$$

map to $\partial_v p_1$ from

$$\partial_v p = -[\, \partial_y + p \,, v_0 \,] \qquad \blacksquare$$

If we combine the results of 2.1.2 and 2.1.7 we arrive at a larger equivalence class for the specialized hierarchy $\{\mathfrak{g},\theta,\Lambda\}$.

2.1.20. Proposition. *For any two semisimple elements* Λ, $\Lambda' \in \mathfrak{g}_1$ *the specialized hierarchies* $\{\mathfrak{g},\theta,\Lambda\}$ *and* $\{\mathfrak{g},\theta,\Lambda'\}$ *are equivalent if, for some* $g \in G_0$, $\mathfrak{z}(g.\Lambda) = \mathfrak{z}(\Lambda')$.

Remark. It is interesting to note that $c(\mathfrak{z}(z\Lambda)) = c(\mathfrak{z}(z\Lambda'))$ if and only if $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$. We know the latter implies the former, and the former implies $c(\mathfrak{z}(\Lambda)) = c(\mathfrak{z}(\Lambda'))$ in $\mathfrak{g}$. But this means $[\Lambda,\mathfrak{z}(\Lambda')]=0$, therefore $\mathfrak{z}(\Lambda') \subseteq \mathfrak{z}(\Lambda)$. If we interchange Λ and Λ' in this argument we see $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$.

In order to determine the equivalence class of a specialized hierarchy $\{\mathfrak{g},\theta,\Lambda\}$ we need to examine the G_0-conjugacy class of $\mathfrak{z}(\Lambda)$ in $\mathfrak{g}$. In the next section we will describe briefly how to do this and then provide a few examples.

§2.2 <u>The Cartan subspace and equivalent specialized hierarchies.</u>

In order to understand the G_0-orbit of $\mathfrak{z}(\Lambda)$ we must understand the G_0-orbits of semisimple elements of $\mathfrak{g}_1$.

It is well-known (see e.g. Helgason (1978,A.P.)) that the G-orbits of semisimple elements are characterized by any Cartan subalgebra of $\mathfrak{g}$. Every semisimple element lies in a Cartan subalgebra and all Cartan subalgebras are G-conjugate. Moreover, two elements of the same Cartan subalgebra $\mathfrak{h}$ are G-conjugate if and only if they are conjugate under the action of the Weyl group; this is the finite group isomorphic to the quotient group

$$\text{Normalizer}(\mathfrak{h})/\text{Stabilizer}(\mathfrak{h})$$

Vinberg (1976) showed that the characterization of G_0-orbits of semisimple elements in $\mathfrak{g}_1$ can be done in exactly the same manner, replacing the notion of a Cartan subalgebra by a Cartan subspace of $\mathfrak{g}_1$.

Given $(\mathfrak{g},\theta)$, Vinberg (1976) defined a Cartan subspace $\mathfrak{s} \subseteq \mathfrak{g}_1$ to be a maximal subspace of semisimple elements which is also an abelian subalgebra of $\mathfrak{g}$. He showed that the G_0-orbit of any semisimple element in $\mathfrak{g}_1$ intersects $\mathfrak{s}$. Consequently, all Cartan subspaces are G_0-conjugate and the union of all Cartan subspaces contains all semisimple elements in $\mathfrak{g}_1$. Vinberg defined the Weyl group of a graded Lie algebra to be the finite group isomorphic to

$$W(\mathfrak{s}) = N(\mathfrak{s})/S(\mathfrak{s})$$

where

$$N(\mathfrak{s}) = \{ g \in G_0 \mid g.\mathfrak{s} = \mathfrak{s} \}$$

$$S(\mathfrak{s}) = \{ g \in G_0 \mid g.c = c \ \forall c \in \mathfrak{s} \}$$

are the normalizer and stabilizer of $\mathfrak{s}$ in G_0. Vinberg also proved that two elements of $\mathfrak{s}$ are G_0-conjugate if and only if they are conjugate under the action of this group $W(\mathfrak{s})$.

We are interested in finding the classes of semisimple elements of $\mathfrak{g}_1$ which have the same centralizer up to G_0-conjugacy. We begin by examining when two elements of $\mathfrak{s}$ have the same centralizer.

Since $\mathfrak{s}$ is an abelian subalgebra of semisimple elements of $\mathfrak{g}$ its adjoint representation on $\mathfrak{g}$ yields a weight space decomposition

$$\mathfrak{g} = \sum_{\mu\in\mathfrak{S}} \mathfrak{g}^{\mu} \quad , \qquad \mathfrak{S} \subseteq \mathfrak{s}^{*} \qquad 2.2.1$$

where

$\mathfrak{g}^{\mu} = \{ x \in \mathfrak{g} \mid [c , x] = \mu(c)x \quad \forall c \in \mathfrak{s} \}$

and $\mathfrak{S}$ is the collection of weights $\mu\in\mathfrak{s}^{*}$ such that $\mathfrak{g}^{\mu}$ is non-trivial. Notice that we have allowed $0 \in \mathfrak{S}$; its weight space is the centralizer $\mathfrak{z}(\mathfrak{s})$ of $\mathfrak{s}$.

The centralizer of any $\Lambda\in\mathfrak{s}$ is given by

$$\mathfrak{z}(\Lambda) = \sum (\mathfrak{g}^{\mu} \mid \mu(\Lambda)=0) \qquad 2.2.2$$

Let us denote by $\mathfrak{S}_{\Lambda}$ the subset of $\mathfrak{S}$ containg all those μ which annihilate Λ. Then all elements with common centralizer $\mathfrak{z}(\Lambda)$ are contained in the intersection of the hyperplanes of $\mathfrak{s}$ which are the kernels of each $\mu \in \mathfrak{S}_{\Lambda}$. This intersection will be called

$$\mathcal{H}_{\Lambda} = \{ c \in \mathfrak{s} \mid \mu(c) = 0 \ \forall \mu \in \mathfrak{S}_{\Lambda} \}$$

Remark. In the case where $\mathfrak{s}$ is actually a Cartan subalgebra of $\mathfrak{g}$, an example of which will soon be given, the weight space $\mathfrak{S}$ is a root space, each μ is a root and the hyperplanes $\{c \mid \mu(c)=0\}$, for each non-zero $\mu\in\mathfrak{S}$, are dual to the walls of the Weyl chambers in $\mathfrak{s}^{*}$. The group $W(\mathfrak{s})$ must then be isomorphic to the Weyl group of the Lie algebra.

However, each $\mathcal{H}_{\Lambda}$ also contains elements of $\mathfrak{s}$ with a larger centralizer than $\mathfrak{z}(\Lambda)$, since $\mathfrak{z}(\Lambda) \subseteq \mathfrak{z}(\Lambda')$ if $\mathfrak{S}_{\Lambda} \subseteq \mathfrak{S}_{\Lambda'}$ i.e. if $\mathcal{H}_{\Lambda'} \subseteq \mathcal{H}_{\Lambda}$. Therefore it is necessary (and sufficient) to have $\mathcal{H}_{\Lambda'} = \mathcal{H}_{\Lambda}$ in order to say that $\mathfrak{z}(\Lambda) = \mathfrak{z}(\Lambda')$.

In most cases it will be easier to work with the sets of weights $\mathfrak{S}_{\Lambda}$. It is easy to establish that the Weyl group $W(\mathfrak{s})$ acts on the set of weights $\mathfrak{S}$ by

$$w^{*}\mu = \mu\circ w^{-1} \qquad w \in W(\mathfrak{s}),\ \mu \in \mathfrak{S}$$

Therefore we can say that two elements of $\mathfrak{s}$ have the same centralizer if and only if the set of weights annihilating one is conjugate, under this action, to the set of weights annihilating the other.

Treating the general case beyond this brief discussion is difficult and requires a better understanding of the Weyl group of a graded Lie algebra. We will finish this discussion by looking at a few simple examples.

2.2.3 <u>Example.</u> Examine the periodic grading corresponding to the automorphism of type (1,0,1;1) on $\mathfrak{a}_2$ presented at the end of chapter 1. It was shown that, for the choice of semisimple element $\Lambda = E_{13}+E_{31}$ the intersection of $\mathfrak{z}(\Lambda)$ with $\mathfrak{g}_1$ is one dimensional (see 1.5.2). Therefore the dimension of a Cartan subspace for this graded Lie algebra must be one, since $\Lambda \in \mathfrak{s}$ implies $\mathfrak{s} \subseteq \mathfrak{z}(\Lambda)$. We conclude that there is only one specialized hierarchy of equations for this choice of grading.

2.2.4 <u>Example.</u> An interesting example is provided by the $\mathbb{Z}_2$-grading given by an automorphism of type (1,0,1;1) on c_2. The diagram $c_2^{(1)}$ is

$$\overset{\alpha_0}{\circ} \Longrightarrow \overset{\alpha_1}{\circ} \Longleftarrow \overset{\alpha_2}{\circ}$$

In this case each Cartan subspace is actually a Cartan subalgebra for c_2. We will compute a Cartan subspace and classify the specialized hierarchies up to equivalence. We use the representation $\mathfrak{sp}(2,\mathbb{C})$ of c_2, consisting of the 4×4 complex matrices of the form

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & -A_1^t \end{bmatrix}$$

where A_1, A_2 and A_3 are 2×2 blocks, with A_2, A_3 both symmetric. To obtain a root space decomposition we fix the Cartan subalgebra $\mathfrak{h}$ of diagonal matrices, spanned by

$$H_1 = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{bmatrix} \qquad H_2 = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & -1 \end{bmatrix}$$

H_1 and H_2 are proportional to the coroots of α_1 and α_2

respectively, where $\alpha_1 = (a_1, 0)$, $\alpha_2 = (a_2, 1)$.

Rather than explicitly list the root vectors which belong to $\mathfrak{g}_0$ and $\mathfrak{g}_1$ in this grading let us simply note that the subalgebra $\mathfrak{g}_0$ consists of all matrices of the form

$$\begin{bmatrix} A_1 & \underline{0} \\ \underline{0} & -A_1^t \end{bmatrix} \qquad \underline{0} \text{ is the } 2\times 2 \text{ block of } 0\text{'s}$$

and $\mathfrak{g}_1$ is the subspace of all matrices of the form

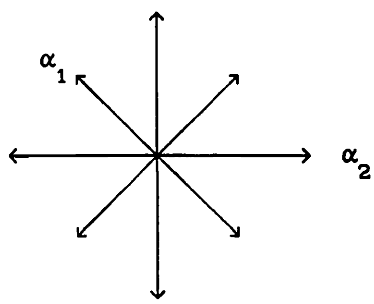
$$\begin{bmatrix} \underline{0} & A_2 \\ A_3 & \underline{0} \end{bmatrix}$$

In particular, the elements

$$\Lambda_1 = \begin{bmatrix} \underline{0} & \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \\ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \underline{0} \end{bmatrix} \qquad \Lambda_2 = \begin{bmatrix} \underline{0} & \begin{matrix} 1 & -1 \\ -1 & 1 \end{matrix} \\ \begin{matrix} 1 & -1 \\ -1 & 1 \end{matrix} & \underline{0} \end{bmatrix}$$

span a Cartan subalgebra $\mathfrak{s}$ in $\mathfrak{g}_1$. A quick computation shows that these can be (simultaneously) diagonalized, to H_1 and H_2 respectively.

As mentioned in the remark earlier, we can use our knowledge of the root system for c_2 to determine the conjugacy classes of centralizers of elements of $\mathfrak{s}$. The root system is a set of vectors in $\mathfrak{s}^* \cong \mathbb{C}^2$, but it can be represented schematically by the diagram



There are two types of roots, long and short. The Weyl group is generated by the reflections along each root. We see that all long roots are conjugate and all short roots are conjugate under combinations of these reflections (the group is, of course, the group of symmetries of the square, D_8).

The diagram above indicates the regions corresponding to elements of $\mathfrak{s}$ with centralizers of different dimension.

The regular elements, those whose centralizer is $\mathfrak{s}$, correspond to the regions between the rays generated by the roots; these are the Weyl chambers. The set of all regular elements forms one class yielding one specialized hierarchy, since every regular element has the same centralizer.

A feature of the geometry of this root system is that each long (respectively, short) root is orthogonal to another long (short) root. Therefore the kernel of any root can be identified with the subspace generated by another root of the same length. Any non-zero element of $\mathfrak{s}$ lying in the kernel of a root has a three dimensional centralizer. Since all roots of the same length are conjugate, there are precisely two conjugacy classes of centralizers of non-regular elements.

This simple analysis of the root system has determined that there are three distinct equivalence classes of specialized hierarchies for this choice of grading on c_2. They correspond to choosing $\Lambda \in \mathfrak{s}$ to be either

(i) any regular element,

or (ii) any element annihilated by a short root

or (iii) any element annihilated by a long root.

Each specialized hierarchy is characterized by a specialized spectral operator. We may choose representative spectral operators corresponding to the three cases listed above to be

$$\text{(i)} \qquad \partial_x + \begin{pmatrix} \begin{matrix} 2q & 3p+r \\ 3p-r & s \end{matrix} & \underline{0} \\ \underline{0} & \begin{matrix} -2q & r-3p \\ -3p-r & -s \end{matrix} \end{pmatrix} - z \begin{pmatrix} \underline{0} & \begin{matrix} 2 & 0 \\ 0 & 1 \end{matrix} \\ \begin{matrix} 2 & 0 \\ 0 & 1 \end{matrix} & \underline{0} \end{pmatrix}$$

$$\text{(ii)} \qquad \partial_x + \begin{pmatrix} \begin{matrix} q & p \\ r & q \end{matrix} & \underline{0} \\ \underline{0} & \begin{matrix} -q & -r \\ -p & -q \end{matrix} \end{pmatrix} - z \begin{pmatrix} \underline{0} & \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \\ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \underline{0} \end{pmatrix}$$

$$\text{(iii)} \qquad \partial_x + \begin{pmatrix} \begin{matrix} 0 & p \\ r & q \end{matrix} & \underline{0} \\ \underline{0} & \begin{matrix} 0 & -r \\ -p & -q \end{matrix} \end{pmatrix} - z \begin{pmatrix} \underline{0} & \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} & \underline{0} \end{pmatrix}$$

Notice that only in case (ii) has the element Λ been chosen to lie in the Cartan subalgebra $\mathfrak{s}$. In cases (i) and

(iii) the semisimple elements belong to the Cartan subalgebra generated by $E_{13}+E_{31}$ and $E_{24}+E_{42}$ (once again E_{ij} is the matrix with a 1 in the i-th row, j-th column and 0's elsewhere). This choice has been made because it is easier to construct zero curvature equations from these operators.

For example, in each case above an equation belonging to the specialized hierarchy $\{c_2, \theta, \Lambda\}$ corresponds to the element $z^3\Lambda$ in $c(\mathfrak{z}(z\Lambda))$. These equations are

(i)
$$
\begin{aligned}
2q_t &= f_x + 6pk + 2rh \\
s_t &= g_x - 6pk - 2rh \\
3p_t &= h_x - (2q-s)k + r(g-f) \\
r_t &= -k_x + (2q-s)h + 3p(g-f)
\end{aligned}
$$

where
$$
\begin{aligned}
f &= \tfrac{1}{8}\, q_{xx} + \tfrac{1}{2}\, \{\, 2(rp_x-pr_x) - (pr)_x \,\} \\
g &= \tfrac{1}{4}\, s_{xx} - \{\, 2(pr)_x - (rp_x-pr_x) \,\} \\
h &= \tfrac{1}{3}\, p_{xx} - \tfrac{1}{3}\, (2q+s)r_x - \tfrac{1}{6}\, (q_x-s_x)r \\
k &= r_{xx} + (2q+s)p_x + \tfrac{3}{2}\, (q_x+s_x)p
\end{aligned}
$$

(ii)
$$
\begin{aligned}
q_t &= (\, -\tfrac{1}{2}\, q_{xx} + q^3 + 3pqr \,)_x \\
p_t &= \tfrac{1}{4}\, p_{xxx} - pq_{xx} - \tfrac{3}{2}\, \{(pq^2)_x + rpp_x\} \\
r_t &= \tfrac{1}{4}\, r_{xxx} - rq_{xx} - \tfrac{3}{2}\, \{(rq^2)_x + prr_x\}
\end{aligned}
$$

(iii)
$$
\begin{aligned}
q_t &= \tfrac{1}{4}\, (\, q_{xx} + 2(pr_x-rp_x) - 8pqr - 2q^3 \,)_x \\
p_t &= p_{xxx} + \tfrac{3}{4}\, (\, pq_{xx} + 2q_x(p_x-pq) - 2p_x(q^2+3pr) - 2p^2r_x \,) \\
r_t &= r_{xxx} + \tfrac{3}{4}\, (\, rq_{xx} + 2q_x(r_x+rq) + 2r_x(q^2+3pr) + 2r^2p_x \,)
\end{aligned}
$$

2.2.5. Example. For each affine diagram $\Gamma(A)$ the grading of type $(1,1,..,1)$ on $L\mathfrak{g}(A)$ provides a unique equivalence class of hierarchies, for the following reasons.

Recall that the root space of each simple affine root α_i is generated by the single vector e_i. The subspace $\mathfrak{g}_1$ of $\mathfrak{g}$ elements of degree one for this grading is identified with

$$L(\mathfrak{g},\theta)_1 = \{ x = \sum_{i=0}^{\ell} x_i e_i \mid x_i \in \mathbb{C} \}$$

We will see, at the end of the section, that an element of $\mathfrak{g}_1$ is semisimple if and only if every $x_i \neq 0$. When $\Gamma(A)$ is a diagram from table 1 (so $\Gamma(A)$ is an extended Dynkin diagram) this result has already been proved by Kostant (1959).

However, every such semisimple element has the same centralizer, up to conjugacy by G_0 (which is isomorphic to $H_0 = \exp \mathfrak{h}_0$ for this grading). This is not difficult to see. Any element h of $\mathfrak{h}_0$ can be written as

$$h = \sum_{i=0}^{\ell} k_i h_i \quad , \qquad k_i \in \mathbb{C}$$

and by definition

$$\exp(k_i \, ad \; h_i) e_j = \exp(k_i a_{ji}) . e_j \qquad A = (a_{ij})$$

Although the vectors h_i are linearly dependent they have corank 1, therefore it is possible to conjugate an element x in $L(\mathfrak{g},\theta)_1$, with $x_i \neq 0$ for all i, into any element

$$x' = \sum_{i=0}^{\ell} x'_i e_i \quad , \quad x_i \neq 0 \text{ for all } i$$

with the condition, say, $x'_0 = x_0$. Clearly then, any semisimple element of $L(\mathfrak{g},\theta)_1$ is conjugate to kx for some $k \neq 0$. But kx and x have the same centralizer, therefore all semisimple elements in $\mathfrak{g}_1$ possess the same G_0-conjugacy class of centralizers.

Finally let us note that, in this grading, no specialization is required for the equivalence relation 2.1.20 to apply. Recall from lemma 2.1.2 that unspecialized hierarchies are equivalent if their semisimple elements are G_0-conjugate. We have just shown that in this grading every semisimple element of $\mathfrak{g}_1$ is $G_0 (\cong H_0)$-conjugate up to a scalar multiple. However, it is clear that the hierarchies (g,θ,Λ) and $(\mathfrak{g},\theta,k\Lambda)$ are equivalent; we merely rescale the equations. By combining these properties we obtain our result. We could also have proved this by showing that $\mathfrak{g}_0$ $(\cong \mathfrak{h}_0)$ equals $i(\mathfrak{g}_0)$ i.e. no semisimple element of $L(\mathfrak{g},\theta)_1$ commutes with an element of $\mathfrak{h}_0$.

§2.3 A condition for the existence of a non-trivial hierarchy.

It is time to address the question of when the graded Lie algebra $(\mathfrak{g},\theta)$ has a non-zero semisimple element $\Lambda \in \mathfrak{g}_1$.

It is well-known that the semisimple elements of $\mathfrak{g}$ are characterized by the property that their orbits, under the action of the adjoint group G, are Zariski-closed (later we will prove this)

It has already been mentioned that the adjoint group G_0 of $\mathfrak{g}_0$ is a subgroup of G, and that its action preserves $\mathfrak{g}_1$. It was shown by Vinberg (1976) that the semisimple elements of $\mathfrak{g}_1$ can also be characterized intrinsically by their G_0-orbits. We state without proof:

2.3.1 Propostion. (Vinberg, 1976) *An element of $\mathfrak{g}_1$ is semisimple if and only if its G_0-orbit is Zariski-closed.*

In future closure will be taken with respect to the Zariski topology unless otherwise stated.

In this section it will be shown that non-trivial elements of $\mathfrak{g}_1$ with closed G_0-orbits exist if a certain property of the weights for the representation of $\mathfrak{g}_0$ on $\mathfrak{g}_1$ is satisfied. Recall that the representation of $\mathfrak{g}_0$ on $\mathfrak{g}_1$ is equivalent to that of $L(\mathfrak{g},\theta)_0$ on $L(\mathfrak{g},\theta)_1$. It follows that the weights are precisely the set

$$\{ \alpha \in \mathfrak{h}_0^* \mid (\alpha,1) \in \Delta \}$$

where Δ is the affine root system for $L(\mathfrak{g},\theta)$. The main result we intend to prove is the following, which is a particular application of a result proved by Dadok & Kac (1985).

2.3.2 Proposition. *Suppose the set Δ_1 of affine roots of degree 1 contains a collection* B *of roots with the properties:*

(i) $\alpha-\beta \notin \Delta_0$ *for any* α, $\beta \in$ B,

(ii) there exists a set of strictly positive integers $\{ \mu_\beta \mid \beta \in B \}$ *satisfying* $\sum_{\beta \in B} \mu_\beta \mathfrak{b} = 0$, $\beta = (\mathfrak{b}, 1)$.
Then $\mathfrak{g}_1$ *contains a non-trivial element with a closed* G_0*-orbit.*

Before we prove this proposition we present some preliminary results. The first of these, again due to Vinberg (1976), says that $\mathfrak{g}_1$ contains no non-trivial semisimple elements unless each of the integers $s_0, \ldots, s_\ell$ is either 0 or 1. It is for this reason that we expressed interest only in these cases earlier. Vinberg's proof is based upon an examination of the action of the centre of the reductive algebra $\mathfrak{g}_0$ on the irreducible subspaces of $\mathfrak{g}_1$. A simpler proof will be presented.

Let $\Pi = \{\alpha_0, \ldots, \alpha_\ell\}$ be a basis of simple roots for Δ and define

$$\Pi_0 = \{ \alpha_i \in \Pi \mid s_i = 0 \}$$
$$\Pi_1 = \{ \alpha_i \in \Pi \mid s_i = 1 \}$$
$$\Pi_2 = \{ \alpha_i \in \Pi \mid s_i \geq 2 \}$$

2.3.3 Proposition. $\mathfrak{g}_1$ *has no non-trivial semisimple elements if* Π_2 *is non-empty.*

Proof. Let Γ be the affine diagram corresponding to Δ and let $\Delta_1 = \{(a,1) \in \Delta\}$ be the collection of affine roots of degree one. Clearly all of these come from the sub-diagram of Γ obtained by deleting all the vertices corresponding to Π_2 . However, every *proper* sub-diagram of Γ is the Dynkin diagram of some semisimple Lie subalgebra $\mathfrak{a}$ of $\mathfrak{g}$. So the roots in Δ_1 correspond to positive roots of $\mathfrak{a}$. Moreover there exists a basis of roots for $\mathfrak{g}$ for which the positive roots of $\mathfrak{a}$ are also positive for $\mathfrak{g}$. Consequently $\mathfrak{g}_1$ belongs to the nilpotent subalgebra of $\mathfrak{g}$ comprised of the root spaces for its positive roots. Each element of this subalgebra is nilpotent (in the adjoint representation each element is represented by an upper triangular matrix); it follows that the only semisimple element in $\mathfrak{g}_1$ is 0. ∎

Remark. It is not such a strong result to prove that every element of $\mathfrak{g}_1$ is nilpotent. A very simple argument shows that if $\mathfrak{g}_1$ does not contain a non-zero semisimple element then necessarily every element is nilpotent. Any $x \in \mathfrak{g}_1$ is characterized by $\theta(x) = \omega x$, where ω is a primitive root of unity with the same order as θ. Every $x \in \mathfrak{g}$ has a unique Jordan decomposition $x = x_s + x_n$, x_s semisimple, x_n nilpotent, $[x_s, x_n] = 0$. Thus

$$\begin{aligned} \theta(x) &= \theta(x_s) + \theta(x_n) \\ = \omega x &= \omega x_s + \omega x_n \end{aligned}$$

However, for any automorphism θ of $\mathfrak{g}$, $\theta(x_s)$ is semisimple, $\theta(x_n)$ is nilpotent and $[\theta(x_s), \theta(x_n)] = 0$. Thus $\theta(x_s) = \omega x_s$ by the uniqueness of the Jordan decomposition. So we have proved that if $x \in \mathfrak{g}_1$ then $x_s \in \mathfrak{g}_1$.

Now let us assume that Π_2 is empty. For convenience write $L(\mathfrak{g}, \theta)_1 = \mathfrak{L}$, then we have the root space decomposition

$$\mathfrak{L} = \bigoplus_{\beta \in \Delta_1} L(\mathfrak{g}, \theta)^{\beta}$$

It is not always true that each affine root space is one-dimensional. However if we let Δ^0 denote the collection of singular (or imaginary) roots,

$$\Delta^0 = \{(0, j) \in \Delta\}$$

it is true that $L(\mathfrak{g}, \theta)^{\beta}$ is one-dimensional for each $\beta \in \Delta - \Delta^0$ (see e.g Helgason (1978, A.P.)). The next lemma demonstrates that the presence of a singular root in Δ_1 implies the existence of a non-trivial semisimple element in $\mathfrak{g}_1$.

2.3.4 Lemma. *Let $(0,1) \in \Delta_1$. Then $\mathfrak{L}$ contains semisimple elements which are non-zero.*

Proof. Recall that an affine root $\alpha = \sum_{i=0}^{\ell} m_i \alpha_i$ belongs to Δ_1 if $\sum_{i=0}^{\ell} m_i s_i = 1$. So if $\alpha = (0,1)$ it must be that $\alpha = \sum_{i=0}^{\ell} n_i \alpha_i$, where the integers n_i are the unique (normalized) positive integers satisfying

$$\sum_{i=0}^{\ell} n_i a_i = 0 \ , \ \alpha_i = (a_i, s_i).$$

Consequently $\sum_{i=0}^{\ell} n_i s_i = 1$, so that an imaginary root has degree one only if all the s_i are zero except for one $s_i=1$ for which $n_i=1$. However, by inspection we quickly see that the only automorphisms θ with this property have order 1, 2 or 3 and correspond to symmetries of the Dynkin diagram for $\mathfrak{g}$ which have the same order. If the automorphism is trivial then $\mathfrak{k} \cong \mathfrak{g}$. The non-trivial symmetries induce automorphisms of a Cartan subalgebra of $\mathfrak{g}$ and, being of order 2 or 3, must have a non-zero eigenspace of degree one. Thus $\mathfrak{k}$ contains a non-zero subspace of a Cartan subalgebra. ∎

From now on we may assume that Δ_1 does not contain a singular root. We intend to show that, in this case, we can exploit the following theorem of Kempf & Ness (1978) to find closed orbits.

2.3.5 Theorem. (Kempf & Ness, 1978) *Let* W *be a finite dimensional* $\mathbb{C}$*-vector space,* **G** *a connected reductive subgroup of* GL(W) *with a maximal compact subgroup* **U** < **G**. *Let* $\|\cdot\|$ *be a Hermitian norm on* W *which is* **U***-invariant, and for each* x ∈ W *define a length function*

$$p_x : \mathbf{G} \longrightarrow \mathbb{R}^+ \cup \{0\}$$
$$\mathbf{g} \mapsto \|\mathbf{g}.x\|^2$$

Then the **G***-orbit* **G**.x *is closed if and only if* p_x *has a critical point on* **G**.

Since this is a rather important theorem it deserves some explanation. The first important step in the proof was to show that the only critical points of the function p_x are minima. It follows that if G.x is closed then p_x must attain a minimum and therefore have a critical point. If G.x is not closed, an extension of the Hilbert-Mumford theorem (see Birkes (1971)) says that there exists a one-parameter subgroup $T = \{\gamma(s) \mid s \in \mathbb{C}\}$, of G such that $\lim_{s\to\infty} \gamma(s).x$ lies in W but not in G.x . Kempf & Ness showed that on the quotient space

$$\text{(max. compact subgroup of T)}\backslash T \cong S^1\backslash\mathbb{C}^* \cong \mathbb{R}^+$$

the function $p_x(s)$ is strictly decreasing as $|s|\to\infty$, thus p_x never attains a minimum for $s\in\mathbb{C}$.

2.3.6 <u>Example.</u> We will use this theorem to show that x is semisimple if and only if G.x is closed. First we must describe the compact real form of $\mathfrak{g}$ we wish to use, and the Hermitian norm.

Let us fix a root space decomposition of $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha\in R} \mathfrak{g}^\alpha$$

We can choose root vectors $e_\alpha \in \mathfrak{g}^\alpha$ such that $K(e_\alpha, e_{-\alpha})=1$ for all $\alpha\in R$, where K is the Killing form on $\mathfrak{g}$. It follows that $[e_\alpha, e_{-\alpha}]=h_\alpha$, where $K(h_\alpha, \cdot)=\alpha$ (notice that h_α is not in general the coroot of α; this definition has been made for later convenience). Treating $\mathfrak{g}$ as a Lie algebra over $\mathbb{R}$, we define the real form

$$\mathfrak{u} = \sum_{\alpha\in R} i\mathbb{R}\langle\langle h_\alpha\rangle\rangle + \sum_{\alpha\in R} \mathbb{R}\langle\langle e_\alpha - e_{-\alpha}\rangle\rangle + \sum_{\alpha\in R} i\mathbb{R}\langle\langle e_\alpha + e_{-\alpha}\rangle\rangle$$

where $\mathbb{R}\langle\langle h_\alpha\rangle\rangle$ is the real vector space generated by h_α. Clearly $\mathfrak{u}$ is a real form, since $\mathfrak{g} = i\mathfrak{u} + \mathfrak{u}$. In fact $\mathfrak{u}$ is a compact form since the Killing form is negative definite on $\mathfrak{u}$. With respect to this choice of real form we have the complex conjugation

$$\hat{} : \mathfrak{g} \longrightarrow \mathfrak{g}$$

$$x+iy \mapsto x-iy \qquad x,y \in i\mathfrak{u}$$

Consequently we obtain a Hermitian norm $\|z\|^2 = K(z,\hat{z})$. Moreover this norm is invariant under the compact group of "unitary transformations" $U < G$ whose Lie algebra is $\mathfrak{u}$.

We will now show that, for any $x\in\mathfrak{g}$, the function

$$p_x : G \longrightarrow \mathbb{R}^+\cup\{0\}$$

$$g \mapsto \|g.x\|^2$$

has a critical point if and only if x is semisimple. It is clear that for any critical point g of p_x the function $p_{g.x}$ has a critical point at e, the identity of G. So we may assume that x is such that e is a critical point. The identity e is a critical point if and only if

$$\frac{d}{ds}p_x(\exp sy)|_0 = 0 \text{ for all } y \in \mathfrak{g} \qquad 2.3.7$$

where we have used the fact that $\exp: \mathfrak{g} \longrightarrow G$ is onto in a neighbourhood of the identity. If we compute the derivative we find that 2.3.7 is equivalent to

$$K([y,x] , \hat{x}) = 0 \quad \text{for all } y \in \mathfrak{g}$$

i.e.

$$K(y , [x,\hat{x}]) = 0 \quad \text{for all } y \in \mathfrak{g}$$

Thus e is a critical point of p_x if and only if $[x,\hat{x}] = 0$, since K is non-degenerate on $\mathfrak{g}$. However, it is a standard result of linear algebra that a matrix is diagonalizable if and only if it is normal, i.e. it commutes with its Hermitian transpose. In the adjoint representation, the Hermitian transpose of adx is $ad\hat{x}$. We conclude that normality is equivalent to $[x,\hat{x}] = 0$. ■

Now let us apply Theorem 2.3.5 to the representation of G_0 on $\mathfrak{L}$. As in the example, there is a canonical choice of compact form for $L(\mathfrak{g},\theta)_0$ given by

$$\mathfrak{u}_0 = \sum_{\alpha\in\Delta_0} i\mathbb{R}\langle\langle h_\alpha\rangle\rangle + \sum_{\alpha\in\Delta_0} \mathbb{R}\langle\langle e_\alpha - e_{-\alpha}\rangle\rangle + \sum_{\alpha\in\Delta_0} i\mathbb{R}\langle\langle e_\alpha + e_{-\alpha}\rangle\rangle$$

where $h_\alpha \in \mathfrak{h}_0$, $e_\alpha \in L(\mathfrak{g},\theta)^\alpha$ are generators for $L(\mathfrak{g},\theta)_0$ satisfying $K(e_\alpha, e_{-\alpha})=1$, $K(h_\alpha,\cdot)=a$, $\alpha=(a,0)$. This is the Lie algebra for a maximal compact subgroup U_0 of G_0.

We have assumed that Δ_1 does not contain singular roots, therefore the space

$$\mathfrak{M} = L(\mathfrak{g},\theta)_1 + L(\mathfrak{g},\theta)_{-1}$$

can be spanned by root vectors e_β, $e_{-\beta}$, $\beta\in\Delta_1$, chosen to satisfy $K(e_\beta, e_{-\beta})=1$. Consequently $\mathfrak{M}$ can be decomposed, as an $\mathbb{R}$-vector space, into $\mathfrak{M} = \mathfrak{N} + i\mathfrak{N}$ where

$$\mathfrak{K} = \sum_{\beta\in\Delta_1} \mathbb{R}\langle\langle e_\beta - e_{-\beta}\rangle\rangle + \sum_{\beta\in\Delta_1} i\mathbb{R}\langle\langle e_\beta + e_{-\beta}\rangle\rangle$$

The complex conjugation $\bar{\ }: \mathfrak{M} \longrightarrow \mathfrak{M}$ of $\mathfrak{M}$ with respect to $\mathfrak{K}$ is characterized by $\bar{e}_\beta = -e_{-\beta}$. Clearly, conjugation maps $L(\mathfrak{g},\theta)_1$ onto $L(\mathfrak{g},\theta)_{-1}$ and therefore we may define a Hermitian norm

$$|\cdot| : \mathfrak{L} \longrightarrow \mathbb{R}^+\cup\{0\}$$
$$x \mapsto K(x,\bar{x})^{1/2}$$

Moreover this norm is U_0-invariant. U_0 acts along $\mathfrak{K}$ (since $[\mathfrak{u}_0,\mathfrak{K}] \subseteq \mathfrak{K}$) so, given any $g \in U_0$, $a+ib \in \mathfrak{M}$,

$$\overline{g.a+ib} = \overline{g.a + ig.b} = g.a - ig.b = g.\overline{(a+ib)}$$

Therefore

$$K(\ g.x\ ,\ \overline{g.x}\ ) = K(\ g.x\ ,\ g.\bar{x}\ ) = K(\ x\ ,\ \bar{x}\ )$$

by the G_0-invariance of the bilinear form K. Hence we have established the existence of a function

$$p_x : G_0 \longrightarrow \mathbb{R}^+\cup\{0\} \quad x \in \mathfrak{L}$$
$$g \mapsto |g.x|^2$$

satisfying the conditions of Theorem 2.3.5.

2.3.8 <u>Lemma.</u> *p_x has a critical point at e if and only if* $[x,\bar{x}] = 0$.

<u>Proof.</u> As in example 2.3.6, the total derivative of p_x at e is zero if and only if

$$K(\ [y,x]\ ,\ \bar{x}\ ) = 0 \qquad \text{for all } y \in L(\mathfrak{g},\theta)_0 \qquad 2.3.9$$

However, $K([y,x],\bar{x}) = K(y,[x,\bar{x}])$, and K is non-degenerate on $L(\mathfrak{g},\theta)_0$ (see e.g. Helgason (1978,A.P.)). Therefore 2.3.9 is equivalent to $[x,\bar{x}] = 0$. ∎

We want to equate the property "p_x has a critical point" with the existence of a collection of roots with the properties *(i)* and *(ii)* in 2.3.2. For convenience let us make the following definition.

2.3.10 <u>Definition.</u> *Let* $B \subset \Delta_1$ *be a non-empty collection of* *roots with the two properties:*

(i) $\alpha-\beta \notin \Delta_0$ *for any* α, $\beta \in B$,

(ii) there exists a set of strictly positive integers $\{ \mu_\beta \mid \beta \in B \}$ *satisfying* $\sum_{\beta\in A} \mu_\beta \mathfrak{b} = 0$, $\beta = (\mathfrak{b}, 1)$.

We will say that such a collection is of affine type.

Remark. This name for the set B is prompted by the observation that if B is an indecomposable set then its elements are represented by Cartan matrix of affine type (see Kac (1985 C.U.P) for an explanation of generalized Cartan matrices).

2.3.11 Proposition. *Suppose* Δ_1 *contains no singular root and there exists a collection* $B \subset \Delta_1$ *of affine type. Then there exist* $x_\beta \in \mathbb{C}-\{0\}$, $\beta \in B$, *such that the element*

$$x_B = \sum_{\beta\in B} x_\beta e_\beta \in \mathfrak{L}$$

is semisimple.

Remark. This result is a special case of a result of Kac & Dadok (1985) which provides a sufficient condition for the existence of closed orbits for a rational representation of a reductive algebraic group on a $\mathbb{C}$-vector space.

Proof. We will show that, if we choose the values x_β such that $|x_\beta|^2 \equiv x_\beta \bar{x}_\beta = \mu_\beta$ (the strictly positive integers in 2.3.9 *(ii)*), then $[x_B, \bar{x}_B] = 0$. We conclude, by lemma 2.3.7 and the application of the Theorem 2.3.5 of Kempf & Ness, that the G_0-orbit of x_B is closed. Now,

$$[x_B, \bar{x}_B] = [\sum_{\alpha\in B} x_\alpha e_\alpha , -\sum_{\beta\in B} \bar{x}_\beta e_{-\beta}]$$

$$= - \sum_{\beta\in B} |x_\beta|^2 [e_\beta, e_{-\beta}]$$

since $\alpha-\beta \notin \Delta$ for α, $\beta \in B$. By definition, $[e_\beta, e_{-\beta}] = h_\beta$. If we apply the bilinear form K we have

$$K([x_B,\bar{x}_B],\ \cdot\ ) = -\sum_{\beta\in B} |x_\beta|^2 K(\ h_\beta,\ \cdot\ )$$

$$= -\sum_{\beta\in B} |x_\beta|^2 \alpha \quad , \quad \beta = (\alpha, 1)$$

Thus directly from our choice $|x_\beta|^2 = \mu_\beta$ we have

$$K([x_B,\bar{x}_B],\cdot) = 0$$

As we mentioned earlier, K is non-degenerate on $L(\mathfrak{g},\theta)_0$, to which $[x_B,\bar{x}_B]$ belongs, therefore $[x_B,\bar{x}_B] = 0$. ∎

Remark. It should be pointed out that in general it is *not* possible to conclude that x_B is semisimple directly from the property $[x_B,\bar{x}_B] = 0$. In general we cannot find a compact form for the algebra $\mathfrak{g}$ for which the complex conjugation maps $\pi(e_\beta)$ to $-\pi(e_{-\beta})$, where $\pi: L(\mathfrak{g},\theta) \longrightarrow \mathfrak{g}$. However, in most cases this can be done. For any finite order inner automorphism, or outer automorphism of index 2, the graded Lie algebra $(\mathfrak{g},\theta)$ does have a compact form characterized by the mapping $e_\beta \mapsto -e_{-\beta}$ pushed down onto $\mathfrak{g}$. Thus we could prove proposition 2.3.11 for all the loop algebras classified by the diagrams in tables 1 & 2 (see chapter 1, §1.2) simply by invoking the elementary theorem about normal matrices.

Unfortunately, for the few remaining cases (those pertaining to the diagram $\delta_4^{(3)}$) it is not so elementary to prove 2.3.11, therefore it was necessary to invoke the result 2.3.5. In any case, the proof 2.3.11 is more elegant than a case by case proof.

It could be shown that for *any* non-zero values $y_\beta \in \mathbb{C}$, $\beta\in B$, the element $y_B = \sum_{\beta\in B} y_\beta e_\beta$ is semisimple. It is possible to show that, if H_0 is the maximal abelian subgroup of G_0 whose Lie algebra is $\mathfrak{h}_0$, then y_B is always a scalar multiple of some element of the orbit $H_0.x_B$.

Remark. Recall in the previous section, example 2.2.5, it was claimed that the element

$$\sum_{i=0}^{\ell} x_i e_i$$

is semisimple if and only if $x_i \neq 0$ for all i. This is evident now from two facts:

(a) the set of simple affine roots is an affine collection (indeed this is where the name comes from)

(b) any proper subset of $\{e_i \mid i=0,..\ell\}$ generates a subalgebra consisting entirely of nilpotents.

Notice that if Δ_1 does contain the singular root (0,1) then the singleton {(0,1)} can be considered to be a collection of affine type. So, between the results 2.3.4 and 2.3.11 we have the result 2.3.2, which we may restate as:

A sufficient condition for the existence of a semisimple element in $\mathfrak{g}_1$ *is that the set of roots* Δ_1 *contains a collection of affine type.*

§2.4 Graded Lie algebras with a non-trivial Cartan subspace.

Using the results of the previous section we will compile a list of affine diagrams together with the gradings of type $(s_0,...,s_\ell)$ for which $(\mathfrak{g},\theta)$ admits a collection of roots in Δ_1 of affine type. In general it does not appear possible to classify these cases completely without performing an almost case by case analysis. Our main aim in this section is to construct a table of the simple Lie algebras with rank ≤ 4 together with the gradings which admit a non-trivial Cartan subspace. It will be shown later that this table lists all the graded simple Lie algebras, rank($\mathfrak{g}$) ≤ 4, possessing this property.

We begin by proving a few results which hold for arbitrary rank.

2.4.1. Lemma. *If θ is an involution (i.e. has order 2) then $(\mathfrak{g},\theta)$ has a non-trivial Cartan subspace.*

Proof. Let α_i be a simple affine root corresponding to a vertex labelled with $s_i=1$; at least one exists. Let $\alpha_i=(a_i,1) \in \mathfrak{h}_0^*\times\mathbb{Z}$. Then $(-a_i,2n-1)$ is a root for any integer n since $-\alpha_i$ is a root and the order of θ is 2. The set $\{(a_i,1),(-a_i,1)\}$ is clearly a collection of affine type in Δ_1, so the lemma follows. ∎

2.4.2. Lemma. *Given any non-trivial grading of type $(s_0,\ldots,s_\ell)$, $s_i\in\{0,1\}$, on the diagram $\mathfrak{a}_\ell^{(1)}$ or, for $\ell > 2$, on $\mathfrak{d}_\ell^{(2)}$, the corresponding graded Lie algebra has a non-trivial Cartan subspace.*

Proof. For each of these diagrams the coefficients of linear dependence $\{n_i|\ i=0,\ldots,\ell\}$ of the columns of the Cartan matrix satisfy $n_i=n_j$ for all i, j. Therefore the simple roots $\{\alpha_i=(a_i,s_i)\}$ have the property

$$\sum_{i=0}^{\ell} a_i = 0$$

Given the grading of type $(s_0,\ldots,s_\ell)$ we can decompose the affine diagram $\mathfrak{a}_\ell^{(1)}$ or $\mathfrak{d}_\ell^{(2)}$ into a collection of disjoint, connected subdiagrams $\{\Gamma_k\}$ satisfying

(a) every vertex of the original diagram lies in some Γ_k,

(b) for each Γ_k one and only one vertex corresponds to an $s_i=1$.

It can be shown (see, for example, Helgason (1978, A.P.)) that for each connected subdiagram Γ_k the simple sum

$$\alpha_{(k)} = \sum (\alpha_j \mid \alpha_j\in\Gamma_k)$$

is an affine root. By property (b) $\{\alpha_{(k)}\} \subseteq \Delta_1$. Since no two subdiagrams intersect we conclude that $\alpha_{(j)}-\alpha_{(k)}$ is not a root for any j,k. If we write $\alpha_{(k)}=(a_{(k)},1)$ then

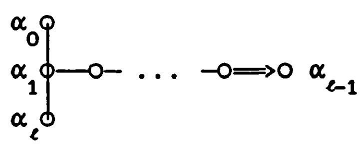
$$\sum a_{(k)} = \sum_{i=0}^{\ell} a_i = 0$$

Therefore $\{\alpha_{(k)}\}$ is a collection of affine type, whereupon the lemma follows. ■

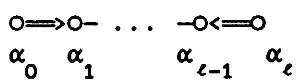
Remark. According to this lemma any finite order *inner* automorphism on $\mathfrak{sl}(\ell+1,\mathbb{C})$ (for which $(s_0,\ldots,s_\ell)$ is a sequence of 0's and 1's) admits a semisimple element in its eigenspace of degree one.

2.4.3. Lemma. *Given any grading of type* $(1,s_1,\ldots,s_{\ell-1},1)$, $s_i\in\{0,1\}$, *on either* $\mathfrak{b}_\ell^{(1)}$ *or* $\mathfrak{c}_\ell^{(1)}$, $\ell \geq 2$, *with the vertices labelled as*

$\mathfrak{b}_\ell^{(1)}$

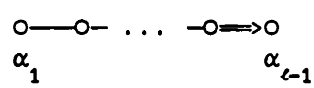


$\mathfrak{c}_n^{(1)}$



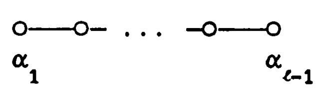
the corresponding graded Lie algebra has a non-trivial Cartan subspace.

Proof. The coefficients $\{n_i \mid i=0,\ldots\ell\}$ are $n_0=1$, $n_\ell=1$, $n_i=2$ for $i= 1,\ldots,\ell-1$. Therefore if all $s_i=0$ for $i=1,\ldots,\ell-1$ the corresponding automorphism is an involution and the lemma is true by 2.4.1. So we may now assume $s_i=1$ for some $i\in\{1,\ldots,\ell-1\}$. In this case it is possible to decompose either of the subdiagrams



of $\mathfrak{b}_\ell^{(1)}$ 2.4.4

or



of $\mathfrak{c}_\ell^{(1)}$ 2.4.5

into disjoint, connected subdiagrams Γ_k with the properties (a) and (b) listed in the proof of the previous lemma. We proceed as before to define

$$\alpha_{(k)} = \sum (\alpha_j \mid \alpha_j \in \Gamma_k) = (a_{(k)}, 1)$$

The collection $\{\alpha_0, \alpha_\ell, \alpha_{(k)}\} \subseteq \Delta_1$ is of affine type since

(i) no pair of these roots has a difference which is a root (this follows from the disjointedness of the Γ_k)

$$\text{(ii)}\quad a_0 + a_\ell + \sum 2a_{(k)} = a_0 + a_\ell + 2\sum_{i=1}^{\ell-1} a_i = 0 \quad ■$$

2.4.6. <u>Lemma</u> *Given any grading of type* $(0, s_1, \ldots, s_{\ell-1}, 0)$, $s_i \in \{0,1\}$, *on* $\mathfrak{b}_\ell^{(1)}$ *the corresponding graded Lie algebra has a non-trivial Cartan subspace.*

<u>Proof.</u> Let s_k be the first integer in the sequence $(s_1, \ldots, s_{\ell-1})$ which equals 1. Define

$$\beta_0 = \alpha_0 + \alpha_1 + \ldots + \alpha_k$$
$$\beta_\ell = \alpha_\ell + \alpha_1 + \ldots + \alpha_k$$

Then β_0 and β_ℓ are affine roots in Δ_1. Moreover $\beta_0 - \beta_\ell$ is not a root.

If $k = \ell-1$ then $\{\beta_0, \beta_\ell\}$ is a collection of affine type whereupon the lemma is true. If $k < \ell-1$ there is a non-trivial diagram

$$\underset{\alpha_{k+1}}{\circ} - \circ - \ldots - \underset{\alpha_{\ell-1}}{\circ} \Longrightarrow \circ \qquad 2.4.7$$

If none of $s_{k+1}, \ldots, s_{\ell-1}$ equals 1 then the corresponding automorphism is an involution, so the lemma follows by 2.4.1. Otherwise it is possible to decompose the graph 2.4.7 into a collection of disjoint, connected diagrams $\{\Gamma_m\}$ which have the properties (a) and (b) listed in the proof of lemma 2.4.2 As before we let

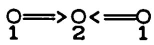
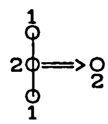
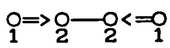
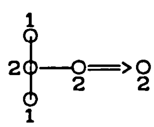
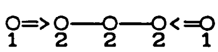
$$\alpha_{(m)} = \sum (\alpha_j \mid \alpha_j \in \Gamma_m)$$

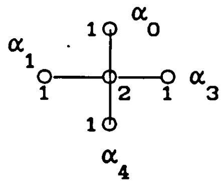
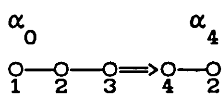
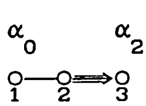
The collection $\{\beta_0, \beta_\ell, \alpha_{(m)}\}$ is a collection of affine type in Δ_1 for the same reasons (i) and (ii) as in the proof of the previous lemma. ■

The purpose of these results 2.4.2, 2.4.3 and 2.4.6 is to reduce the number of individual cases which must be treated in the table to come. The table takes the following form. Each affine diagram corresponding to a simple Lie algebra of rank ≤ 4 is given in the left-hand column, together with the scheme for labelling the vertices by $\alpha_0,\ldots,\alpha_\ell$ (in the case of $b_\ell^{(1)}$ and $c_\ell^{(1)}$ the scheme will be the same as that given in proposition 2.4.3). Next to that appears a sequence of integers $(s_0,\ldots,s_\ell)$ corresponding to a periodic grading for which there exists a non-trivial Cartan subspace. If this fact is the result of one of the previous lemmas then it is referred to in the right-hand column. Otherwise the right-hand column will list, for each case, a collection of roots of affine type in Δ_1.

One should keep in mind that many gradings of type $(s_0,\ldots,s_\ell)$ are equivalent due to the symmetries of a diagram and only one representative need be listed. Moreover, in general the involutions will usually not be listed since they are easy to determine and all such cases are covered by lemma 2.4.1.

Some of the cases which appear in this table have already been investigated by Vinberg (1976) where he lists, for those graded exceptional Lie algebras for which $\mathfrak{g}_0$ is semisimple, the dimension of the Cartan subspace and the group $W(\mathfrak{s})$. Some of the information presented in the paper by Kac (1980) can also be used to determine the existence of a non-trivial Cartan subspace, although one needs to know how to describe the representation of G_0 on $\mathfrak{g}_1$ in his terms.

Affine diagram	Grading $(s_0,\ldots,s_\ell)$	Collection of roots of affine type.
$c_2^{(1)}$ 	(1, 0, 1) (1, 1, 1)	lemma 2.4.3
$b_3^{(1)}$ 	(1, 1, 0, 0)	α_0, $\alpha_1+2\alpha_2$, $\alpha_1+\alpha_3$
	(1, 0, 0, 1)	lemma 2.4.3
	(0, 1, 1, 0)	lemma 2.4.6
	(1, 1, 0, 1) (1, 0, 1, 1) (1, 1, 1, 1)	} lemma 2.4.3
$c_3^{(1)}$ 	$(1, s_1, s_2, 1)$ $s_1 = 0$ or 1	lemma 2.4.3
$b_4^{(1)}$ 	(1, 1, 0, 0, 0)	$\alpha_0, \alpha_1+2\alpha_2+2\alpha_3, \alpha_1+\alpha_4$
	(1, 0, 1, 0, 0)	$\alpha_0+\alpha_1+\alpha_4, \alpha_1+\alpha_2+2\alpha_3, \alpha_2$
	$(1,s_1,s_2,s_3,1)$	lemma 2.4.3
	$(0,s_1,s_2,s_3,0)$ $s_1 = 0$ or 1	lemma 2.4.6
$c_4^{(1)}$ 	(1, 0, 1, 0, 0)	$\alpha_0+\alpha_1$, $\alpha_1+\alpha_2+\alpha_3$ $\alpha_2+\alpha_3+\alpha_4$
	(1, 0, 0, 1, 0)	$\alpha_0+\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$ $\alpha_3+\alpha_4$
	(0, 1, 0, 1, 0)	$\alpha_0+\alpha_1$, $\alpha_1+\alpha_2$, $\alpha_2+\alpha_3$ $\alpha_3+\alpha_4$
	$(1,s_1,s_2,s_3,1)$ $s_1 = 0$ or 1	lemma 2.4.3

$\mathfrak{d}_4^{(1)}$ 	(1, 1, 0, 0, 0)	$\alpha_0+\alpha_2+\alpha_4$, $\alpha_1+\alpha_2+\alpha_3$
	(1, 0, 1, 0, 0)	α_0, $\alpha_1+\alpha_2+\alpha_3+\alpha_4$, α_2
	(1, 0, 0, 0, 1)	$\alpha_0+\alpha_1+\alpha_2$, $\alpha_2+\alpha_3+\alpha_4$
	(1, 1, 1, 0, 0)	α_0, α_1, α_2, $\alpha_2+\alpha_3+\alpha_4$
	(1, 0, 1, 0, 1)	α_0, α_2, α_4, $\alpha_1+\alpha_2+\alpha_3$
	(1, 1, 0, 1, 1)	α_0, α_1, $\alpha_2+\alpha_3$, $\alpha_2+\alpha_4$
	(1, 1, 1, 1, 1)	α_0, α_1, α_2, α_3, α_4
$\mathfrak{f}_4^{(1)}$ 	(0, 0, 1, 0, 0)	$\alpha_0+..+\alpha_3$, $\alpha_1+..+\alpha_4$ $\alpha_2+2\alpha_3+\alpha_4$
	(1, 1, 0, 0, 0)	α_0, α_1 $\alpha_1+3\alpha_2+4\alpha_3+2\alpha_4$
	(1, 0, 1, 0, 0) (0, 1, 0, 0, 1)	$\alpha_0+\alpha_1$, $\alpha_1+\alpha_2$ $\alpha_2+2\alpha_3+\alpha_4$
	(1, 1, 1, 0, 0)	α_0, α_1, $\alpha_2+\alpha_3+\alpha_4$ $\alpha_2+2\alpha_3$
	(0, 1, 0, 1, 0)	$\alpha_0+\alpha_1+\alpha_2$, α_1 $\alpha_2+\alpha_3+\alpha_4$, α_3
	(1, 0, 1, 0, 1)	$\alpha_0+\alpha_1$, $\alpha_1+\alpha_2+2\alpha_3$ α_2, $\alpha_3+\alpha_4$
	(1, 1, 1, 0, 1)	α_0, α_1, $\alpha_2+2\alpha_3$ α_2, α_4
	(1, 1, 1, 1, 1)	$\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4$
$\mathfrak{g}_2^{(1)}$ 	(1, 1, 0)	α_0, $\alpha_1+3\alpha_2$, α_1
	(1, 1, 1)	α_0, α_1, α_2

$a_2^{(2)}$	(1, 0)	involution
1 ⇛ 2 α_0 α_1	(1, 1)	α_0, α_1
$a_4^{(2)}$	(1, 0, 0)	involution
α_0 α_2	(0, 1, 0)	$\alpha_0+\alpha_1$, $\alpha_1+2\alpha_2$
1 ⇒ 2 ⇒ 2	(1, 1, 0)	α_0, $\alpha_1+2\alpha_2$, α_1
	(1, 0, 1)	$\alpha_0+2\alpha_1$, α_2
	(1, 1, 1)	α_0, α_1, α_2
$\delta_4^{(3)}$	(s_0, s_1, s_2)	
α_1 α_2 α_0	(0, 1, 0) (1, 0, 0)	$\alpha_1+2\alpha_2+\alpha_0$
1 ⇛ 2 — 1	(0, 0, 1) (1, 1, 0)	$\alpha_1+\alpha_2$, $\alpha_2+\alpha_0$
	(1, 1, 1)	α_0, α_1, α_2

§2.5 Graded Lie algebras with a trivial Cartan subspace, rank($\mathfrak{g}$) ≤ 4.

At present there does not appear to be a direct proof that the sufficient condition, 2.3.2, for the existence of a non-trivial Cartan subspace is also necessary. Nevertheless this is true for all graded Lie algebras $(\mathfrak{g},\theta)$ with rank($\mathfrak{g}$) ≤ 4. In the last section all those cases where the condition 2.3.2 holds were listed. In this section we will see that for every other grading of those algebras there are no non-trivial semisimple elements in $\mathfrak{g}_1$. This will be done case by case, mainly using a result due to Vinberg (1976).

It is well known in the literature that an element of $\mathfrak{g}$ is nilpotent if and only if the Zariski closure of its G-orbit contains 0. A similar characterization holds for the representation of G_0 on $\mathfrak{g}_1$. We will use without proof the result:

2.5.1. Proposition. (Vinberg, 1976) *An element* $x\in\mathfrak{g}_1$ *is nilpotent if and only if clos*$(G_0.x)$ *contains* 0.

Once again, closure is taken with respect to the Zariski topology.

Remark. I have not found, in the literature, a simple proof of the statement that $x\in\mathfrak{g}$ is nilpotent if and only if clos(G.x) contains 0. It seems worth providing one here.

Suppose $x\in\mathfrak{g}$ is nilpotent, then by the Morozov embedding theorem (see, for example, Jacobson (1962, Wiley)) there exists $h\in\mathfrak{g}$ such that $[h,x] = -x$. Therefore

$$\lim_{t\to\infty} \exp(adth).x = (\lim_{t\to\infty} \exp(-t)).x = 0$$

Hence $0 \in$ clos(G.x).

Conversely, if $0 \in$ clos(G.x) then for any G-invariant polynomial f on $\mathfrak{g}$ for which $f(0)=0$ we have $f(x)=0$. The adjoint representation pulls invariant polynomials on $\mathfrak{sl}(\mathfrak{g})$ back to G-invariant polynomials on $\mathfrak{g}$. It follows that $\mathrm{tr}[(adx)^n] = 0$ for all positive integers n, therefore adx is nilpotent.

The orbit characterizations of semisimple and nilpotent elements (propositions 2.3.1 and 2.5.1) lead to the next result, the proof of which can be found in the paper by Vinberg (1976).

2.5.2. Proposition. (Vinberg, 1976) *The closure of the G_0-orbit of any element $x \in \mathfrak{g}_1$ contains precisely one closed orbit, namely, the orbit of the semisimple part of* x *in its Jordan decomposition.*

As a corollary we obtain a condition which guarantees that $\mathfrak{g}_1$ only contains nilpotents.

2.5.3. Corollary. *If, for some $x \in \mathfrak{g}_1$, $clos(G_0.x) = \mathfrak{g}_1$ then $\mathfrak{g}_1$ only contains nilpotents.*

Proof. Let $y \in \mathfrak{g}_1$, then the G_0-orbit of its semisimple part, y_s, belongs to $\mathfrak{g}_1 = \mathrm{clos}(G_0.x)$. But 0 also has the closed orbit {0} and so by the previous proposition we must have $G_0.y_s = \{0\}$. Therefore $y_s=0$, thus y is nilpotent. ∎

Consequently, to prove that $\mathfrak{g}_1$ contains only nilpotents we need only show that there exists an element of $\mathfrak{g}_1$ satisfying the next condition. Vinberg (1976) mentions this result without proof.

2.5.4. Proposition. *Suppose, for some $x \in \mathfrak{g}_1$ that $[\mathfrak{g}_0, x] = \mathfrak{g}_1$. Then $clos(G_0.x) = \mathfrak{g}_1$, therefore $\mathfrak{g}_1$ contains nilpotents only.*

Proof. It is not difficult to see that $\mathrm{clos}(G_0.x)$ is a closed, irreducible variety with the same dimension as $G_0.x$.

Since $G_0.x$ is non-singular its dimension equals the dimension of $[\mathfrak{g}_0, x]$. Thus we have a closed irreducible subvariety $\text{clos}(G_0.x) \subseteq \mathfrak{g}_1$ whose dimension equals that of $\mathfrak{g}_1$. An elementary theorem of algebraic geometry states that $\text{clos}(G_0.x)$ cannot be a proper subset (see, for example, Humphreys (1975, Springer)) i.e. $\text{clos}(G_0.x) = \mathfrak{g}_1$. ∎

This result will be the main tool used to show that all the gradings ignored by the earlier table correspond to graded Lie algebras with a trivial Cartan subspace. In practice showing that $[\mathfrak{g}_0, x] = \mathfrak{g}_1$ for some $x \in \mathfrak{g}_1$ is quite tedious; sometimes it is possible to use a simpler test.

2.5.5. <u>Lemma.</u> *Let θ be an inner automorphism of $\mathfrak{g}$ (these correspond to the affine diagrams in table 1 of chapter 1). If the order of θ does not divide the degree of any homogeneous G-invariant polynomial belonging to a basis for the ring of G-invariant polynomials on $\mathfrak{g}$ then $\mathfrak{g}_1$ contains nilpotents only.*

It is well-known that the ring $\mathbb{C}[\mathfrak{g}]^G$ of G-invariant polynomials on $\mathfrak{g}$ is finitely generated. Any two bases of homogeneous G-invariant polynomials for this ring have the same list of the degrees of the polynomials in the basis, that is to say, these degrees are invariants of $\mathfrak{g}$. A list of these numbers can be found in, for example, Bourbaki (1968, Hermann).

<u>Proof.</u> Let $\{\mathcal{J}_i\}$ be a basis of the ring $\mathbb{C}[\mathfrak{g}]^G$ where $\mathcal{J}_i$ has degree d_i. The subvariety of $\mathfrak{g}$

$$\{ x \in \mathfrak{g} \mid \mathcal{J}_i(x) = 0, \ \forall i \}$$

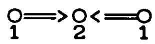
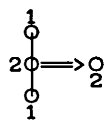
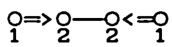
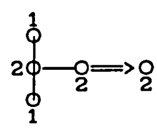
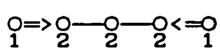
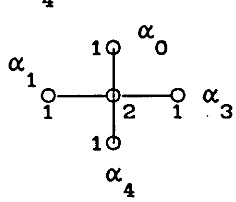
is called the null-cone of $\mathfrak{g}$; it is well known that this set only contains nilpotents (in fact we can deduce this from the proof given earlier that an element is nilpotent if the closure of its G-orbit contains 0). Now let m = order(θ) and let ω be a primitive m-th root of unity. If θ is an inner automorphism (and therefore belongs to G) for any element

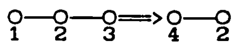
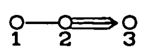
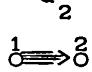
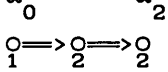
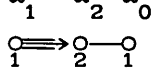
$x \in \mathfrak{g}_1$ we have

$$\mathcal{J}_i(x) = \mathcal{J}_i(\theta.x) = \mathcal{J}_i(\omega x) = \omega^{d_i}\mathcal{J}_i(x) \quad \text{for all } i$$

If m does not divide d_i then $\omega^{d_i} \neq 1$ for any i, therefore $\mathcal{J}_i(x) = 0$ for all i. Hence $\mathfrak{g}_1$ belongs to the null-cone. ■

Between the results 2.5.4 and 2.5.5 we will be able to show that each grading not appearing in the table in the previous section pertains to a case where $\mathfrak{g}_1$ contains only nilpotents. The following table is arranged similarly to the previous one, with the affine diagrams in the left-hand column and the grading in the central column. In the column on the right there will appear either a set of roots from Δ_1 or the order of the grading. The former designates that there exists a vector x, which is a sum of non-zero root vectors for the roots given, with the property $[\mathfrak{g}_0, x] = \mathfrak{g}_1$. To actually demonstrate this in each case would be a tiresome procedure; one example will be computed afterwards to give an indication of how to proceed in general. When the order of an inner automorphism appears in the right-hand column this indicates that it does not divide any of the d_i. These degrees will be listed beside the name of the relevant affine diagram.

Affine diagram	Grading $(s_0,\dots,s_\ell)$	$[\mathfrak{g}_0, x] = \mathfrak{g}_1$ $x = \sum (x_\alpha e_\alpha \mid \alpha$ listed)
$c_2^{(1)}$ 2,4 	(1, 1, 0)	order = 3
$b_3^{(1)}$ 2,4,6 	(1, 0, 1, 0)	α_0, $\alpha_1+\alpha_2+\alpha_3$
	(1, 1, 1, 0)	order = 5
$c_3^{(1)}$ 2,4,6 	(1, 1, 0, 0)	α_1
	(0, 1, 1, 0)	α_1, α_2
	(1, 1, 1, 0)	order = 5
$b_4^{(1)}$ 2,4,6,8 	(1, 0, 0, 1, 0)	$\alpha_0, \alpha_1+\alpha_2+\alpha_3+\alpha_4$
	(1, 1, 1, 0, 0)	order = 5
	(1, 1, 0, 1, 0)	order = 5
	(1, 0, 1, 1, 0)	order = 5
	(1, 1, 1, 1, 0)	order = 7
$c_4^{(1)}$ 2,4,6,8 	(0, 1, 1, 0, 0)	$\alpha_0+\alpha_1$, α_2
	(1, 1, 0, 0, 0)	α_0 α_1
	(1, 1, 1, 0, 0)	order = 5
	(1, 0, 1, 1, 0)	order = 5
	(1, 1, 0, 1, 0)	order = 5
	(1, 1, 1, 1, 0)	order = 7
$\mathfrak{d}_4^{(1)}$ 2,4,6,4 	(1, 1, 1, 1, 0)	order = 5

$f_4^{(1)}$ 2,6,8,12	(0, 0, 0, 1, 0)	$\alpha_3,\ \alpha_3+\alpha_4$
α_0 α_4	(0, 0, 0, 1, 1)	$\alpha_3,\ \alpha_4$
	(1, 0, 0, 0, 1)	$\alpha_0+\alpha_1+\alpha_2+2\alpha_3\ ,\ \alpha_4$
	(1, 0, 1, 1, 0)	$\alpha_0+\alpha_1,\ \alpha_2,\ \alpha_3$
	(0, 1, 0, 1, 1)	$\alpha_1,\ \alpha_2+\alpha_3,\ \alpha_4$
	(1, 0, 0, 1, 0) (0, 1, 1, 0, 0) (0, 0, 1, 0, 1) (1, 1, 0, 0, 1)	order = 5
	(0, 0, 1, 1, 0) (1, 1, 0, 1, 0) (1, 0, 0, 1, 1) (0, 1, 1, 0, 1)	order = 7
	(0, 1, 1, 1, 0) (0, 0, 1, 1, 1) (1, 1, 0, 1, 1)	order = 9
	(1, 1, 1, 1, 0) (1, 0, 1, 1, 0)	order = 10
	(0, 1, 1, 1, 1)	order = 11
$g_2^{(1)}$ 2,6	(0, 0, 1)	α_2
α_0 α_2	(0, 1, 1)	order = 5
	(1, 0, 1)	order = 4
$a_2^{(2)}$	(0 ,1)	α_1
α_0  α_1		
$a_4^{(2)}$	(0, 0, 1)	α_2
α_0 α_2 	(0, 1, 1)	α_1,α_2
$d_4^{(3)}$	(s_0,s_1,s_2)	
α_1 α_2 α_0	(0, 1, 1)	$\alpha_1,\ \alpha_2$
	(1, 0, 1)	$\alpha_0,\ \alpha_2$

2.5.6. Example. One example is sufficient to demonstrate the method of proving the existence of a dense orbit. Take the example of the grading of type (1,0,1,0) for $\mathfrak{b}_3^{(1)}$. We will see that there exists an element x in the subspace

$$L(\mathfrak{g},\theta)^{\alpha} + L(\mathfrak{g},\theta)^{\beta} \qquad \alpha = \alpha_0,\ \beta = \alpha_1+\alpha_2+\alpha_3$$

such that

$$[\ L(\mathfrak{g},\theta)_0\ ,\ x\ ] = L(\mathfrak{g},\theta)_1$$

The space $L(\mathfrak{g},\theta)_1$ is the sum of root spaces for the roots

$$\alpha_0\ ,\ \alpha_0+\alpha_1\ ,\ \alpha_0+\alpha_1+\alpha_2\ ,\ \alpha_2\ ,\ \alpha_1+\alpha_2\ ,\ \alpha_1+\alpha_2+\alpha_3$$

The roots corresponding to $L(\mathfrak{g},\theta)_0$ lie in

$$\Delta_0 = \{\ \pm\alpha_1\ ,\ \pm\alpha_3\ ,\ \pm(\alpha_1+\alpha_3)\ \}$$

Therefore if x_α , x_β are non-zero root vectors for α, β respectively it is clear that

$$[\ \sum_{\gamma\in\Delta_0} L(\mathfrak{g},\theta)^{\gamma}\ ,\ x_\alpha + x_\beta\ ]$$

contains the root spaces corresponding to

$$\alpha_2 = \beta - (\alpha_1+\alpha_3) \quad , \quad \alpha_0+\alpha_1 = \alpha + \alpha_1$$
$$\alpha_0+\alpha_1+\alpha_3 = \alpha + (\alpha_1+\alpha_3) \quad , \quad \alpha_1+\alpha_2 = \beta - \alpha_3$$

Moreover

$$L(\mathfrak{g},\theta)_0 = \mathfrak{h}_0 \oplus \sum_{\gamma\in\Delta_0} L(\mathfrak{g},\theta)^{\gamma}$$

where $\mathfrak{h}_0$ is a 3-dimensional abelian subalgebra which acts diagonally on the root spaces of $L(\mathfrak{g},\theta)$. Thus

$$[\ \mathfrak{h}_0\ ,\ x_\alpha + x_\beta\ ] = L(\mathfrak{g},\theta)^{\alpha} + L(\mathfrak{g},\theta)^{\beta}$$

since α and β are linearly independent. The result follows.

Finally let us note that, in principle, the paper by Kimura *et al.* (1986) contains all the information we seek. They report in that paper the classification of all the rational representations of reductive algebraic groups which admit only a finite number of orbits. In particular, if the representation of G_0 on $\mathfrak{g}_1$ admits only a finite number of orbits then $\mathfrak{g}_1$ must contain a dense orbit, therefore $\mathfrak{g}_1$ contains only nilpotents, by 2.5.3 (if $\mathfrak{g}_1$ is the union of a finite number of orbits then it is the union of their closures, but $\mathfrak{g}_1$ is irreducible therefore one of these closures is not a proper subset).

Consequently if I could translate each representation of G_0 on $\mathfrak{g}_1$ into the language used by Kimura *et al.* then I could make use of their results. Unfortunately it is not at all clear to me what the relationship is between their notation and the description of the representation of G_0 on $\mathfrak{g}_1$.

CHAPTER 3

§3.1 Miura's transformation.

The basic observation of Miura (1968) was that the mKdV equation

$$q_t = q_{xxx} - 6q^2 q_x \qquad 3.1.1$$

could be transformed into the KdV equation

$$u_t = u_{xxx} + 6uu_x \qquad 3.1.2$$

by the substitution

$$u = q_x - q^2 \qquad 3.1.3$$

We wish to find a context in which this transformation can be understood, for it is not obvious that the differential polynomial $u_t = q_{xt} - 2qq_t$ should be able to be expressed purely as a differential polynomial in the variable u. More formally, recall from chapter 1 that $\mathbb{C}\{q\}$ denotes the differential algebra of polynomials in $q, q_x, \ldots$. Then we wish to discover why the derivation ∂_t on $\mathbb{C}\{q\}$ defined by 3.1.1 preserves the subalgebra $\mathbb{C}\{u\}$.

In this section we will describe how the Miura transformation is explained as an example of some general machinery developed by Drinfel'd & Sokolov (1981,1985). The following section will show the contrast between this and a more elegant explanation due to Wilson (to be published). Before anything can be done a little preliminary discussion is necessary.

It had been known for some time before 1981 that 3.1.3 is the prescription for factorizing the Schrödinger operator,

$$\partial_x^2 + u = (\partial_x - q)(\partial_x + q) \qquad 3.1.4$$

This operator plays a pivotal role in the theory of the KdV equation (see Gardner *et al.* (1968) for example), however let us examine this factorization at face value.

This second order operator has a two dimensional null-space spanned by functions ψ, ϕ satisfying

$$(\partial_x^2 + u)\psi = 0 \qquad (\partial_x^2 + u)\phi = 0 \qquad W(\psi,\phi) = k \qquad 3.1.5$$

where $W(\psi,\phi)$ is the Wronskian

$$W(\psi,\phi) = \psi\phi_x - \phi\psi_x$$

and k is a non-zero constant. In particular we may choose ψ, ϕ to be compatible with the factorization, that is to say

$$(\partial_x + q)\psi = 0 \qquad (\partial_x + q)\psi = k\psi^{-1} \qquad 3.1.6$$

where the second relation follows from the fact that ψ^{-1} spans the null-space of $\partial_x - q$.

If we choose ϕ, ψ such that k=1 then the relations in 3.1.6 can be rewritten as

$$\left[\partial_x + \begin{pmatrix} q & -1 \\ 0 & -q \end{pmatrix} \right] \begin{pmatrix} \psi & \phi \\ 0 & 1/\psi \end{pmatrix} = 0 \qquad 3.1.7$$

whereas the relations in 3.1.5 are equivalent to

$$\left[\partial_x + \begin{pmatrix} 0 & -1 \\ u & 0 \end{pmatrix} \right] \begin{pmatrix} \psi & \phi \\ \psi_x & (\phi\psi_x - 1)/\psi \end{pmatrix} = 0 \qquad 3.1.8$$

Drinfel'd & Sokolov pointed out that the operators in 3.1.7 and 3.1.8 are equivalent under a unique gauge transformation

$$\partial_x + \begin{pmatrix} 0 & -1 \\ u & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix} \left[\partial_x + \begin{pmatrix} q & -1 \\ 0 & -q \end{pmatrix} \right] \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix} \qquad 3.1.9$$

Moreover they showed that there is an analogous result for the factorization of the operator

$$\partial_x^{\ell+1} + u_{\ell-1}\partial_x^{\ell-1} + \dots + u_0 = (\partial_x + q_\ell)\dots(\partial_x + q_0) \qquad 3.1.10$$

with

$$\sum_{i=0}^{\ell} q_i = 0$$

In this case the matrix operators

$$\partial_x + \begin{pmatrix} q_0 & -1 & 0 \dots & & \\ 0 & q_1 & -1 & 0 \dots & \\ & & \ddots & \ddots & -1 \\ & & & 0 & q_\ell \end{pmatrix} \qquad 3.1.11$$

and

$$\partial_x + \begin{pmatrix} 0 & -1 & 0 \dots & & \\ & 0 & -1 & 0 \dots & \\ & & \ddots & \ddots & -1 \\ u_0 & \dots & \dots & u_{\ell-1} & 0 \end{pmatrix} \qquad 3.1.12$$

are gauge equivalent by the action of a lower triangular unipotent matrix with unique entries which belong to $\mathbb{C}\{q_i\}$. The factorization 3.1.10 has been used by Sokolov & Shabat (1980) and Fordy & Gibbons (1980) to construct zero curvature equations, for the variables q_i, given a Lax equation with the scalar spectral operator 3.1.10 (see also Kupershmidt & Wilson (1981)).

Of course it is not obvious that this transformation of matrix operators allows to pass from one equation to another. For the Miura transformation the argument presented by Drinfel'd & Sokolov (1985) goes briefly as follows.

In zero curvature form the mKdV equation has the spectral operator with representation

$$\partial_x + \begin{pmatrix} q & -z \\ -z & -q \end{pmatrix} \tag{3.1.13}$$

Recall from chapter 1, §1.2, that this uses the principal realization of the loop algebra. In the standard realization this operator has the form

$$\partial_x + \begin{pmatrix} q & -1 \\ -z & -q \end{pmatrix} \tag{3.1.14}$$

Under the adjoint action of lower triangular unipotent matrices in **SL(2,$\mathbb{C}$)** the element

$$\begin{pmatrix} 0 & 0 \\ -z & 0 \end{pmatrix} \tag{3.1.15}$$

is fixed. Consequently

$$\begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix}\left[\partial_x + \begin{pmatrix} q & -1 \\ -z & -q \end{pmatrix}\right]\begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix} = \partial_x + \begin{pmatrix} 0 & -1 \\ u-z & 0 \end{pmatrix} \tag{3.1.16}$$

A more general statement is that the operator

$$\mathcal{L} = \partial_x + \begin{pmatrix} q & -1 \\ r-z & -q \end{pmatrix} \tag{3.1.17}$$

can be transformed into

$$M(\mathcal{L}) = \partial_x + \begin{pmatrix} 0 & -1 \\ \mu-z & 0 \end{pmatrix} \quad , \ \mu = q_x - q^2 + r \tag{3.1.18}$$

by a gauge transformation using a unique lower triangular unipotent matrix (it is of course the same lower triangular matrix used in 3.1.16).

Drinfel'd & Sokolov were forced to introduce the operator $\mathcal{L}$ because gauge action of the group $\mathcal{N}_-$ of matrix "functions" of the form

$$\begin{bmatrix} 1 & 0 \\ a(x) & 1 \end{bmatrix}$$

does not preserve the set of operators having diagonal potentials. $\mathcal{L}$ can be thought of as a generic element of the set of operators with lower triangular potentials, on which $\mathcal{N}_-$ acts by gauge transformations.

Having noticed 3.1.16 we would like to be able to say that the Miura transformation is the result of "dividing out" by the action of $\mathcal{N}_-$. We think of the variable μ as being an $\mathcal{N}_-$-gauge invariant since the map $\mathcal{L} \mapsto M(\mathcal{L})$ assigns a unique $M(\mathcal{L})$ to the $\mathcal{N}_-$-gauge orbit of $\mathcal{L}$. Unfortunately this idea does not go through, but Drinfel'd & Sokolov managed to salvage some of this concept ingeniously.

They defined *two* derivations

$$\begin{aligned} \partial_t \mathcal{L} &= [\ V(r)_+\ ,\ \mathcal{L}\ ] \\ \nabla_t \mathcal{L} &= [\ V(r)^+\ ,\ \mathcal{L}\ ] \end{aligned} \qquad 3.1.19$$

where $[V(r),\mathcal{L}] = 0$ and $V(r)_+$ (respectively, $V(r)^+$) is the finite series of terms in $V(r)$ of non-negative degree in the principal (respectively, standard) grading. The derivation ∂_t yields the mKdV equation when $r\equiv 0$ (clearly $V(0)$ is the series V commuting with the operator 3.1.13), while the derivation ∇_t gives the KdV equation for the variable μ (it is this latter derivation which is, so to speak, $\mathcal{N}_-$-invariant).

According to Drinfel'd & Sokolov we look at

$$(\nabla_t - \partial_t)\mathcal{L} = [\ V(r)^+ - V(r)_+\ ,\ \mathcal{L}\ ] \qquad 3.1.20$$

The matrix $V(r)^+ - V(r)_+$ is lower triangular (recall that in the principal grading the generator

$$f_1 = z^{-1}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

has degree -1, while in the standard grading

$$f_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

and therefore has degree 0). The idea of Drinfel'd & Sokolov is that $\nabla_{t_+}-\partial_t$ must be tangent to the one-parameter subgroup $\exp(s(V(r)^+ - V(r)_+))$ of N_-. Since μ is an invariant of this group they conclude that $(\nabla_t-\partial_t)\mu = 0$. Therefore ∂_t also gives the KdV equation for u when $r\equiv 0$.

The remarkable achievement of Drinfel'd & Sokolov was to prove that a similar procedure can be followed for any zero curvature equation associated with the grading of type (1,1,...,1) (the principal grading) on the Lie algebra $L\mathfrak{g}(A)$ with affine Cartan matrix A. Their general result is this: let $(\mathfrak{g},\nu,\Lambda)$ be the hierarchy of zero curvature equations corresponding to the principal grading on $L\mathfrak{g}(A)$, where

$$z\Lambda = \sum_{i=0}^{\ell} e_i$$

in terms of the canonical generators for $L\mathfrak{g}(A)$. To any vertex α_i of the affine diagram $\Gamma(A)$ there corresponds another hierarchy of equations, usually referred to as the generalized KdV equations, which are obtained from the hierarchy $(\mathfrak{g},\nu,\Lambda)$ by a transformation of Miura type.

When the diagram is $a_\ell^{(1)}$ the generalised KdV hierarchy obtained for any choice of vertex is precisely the hierarchy of Lax equations associated with the scalar operator 3.1.10 (see Drinfel'd & Sokolov (1985)).

A similar result can be obtained for some gradings other than the principal grading. Later we will look at the conditions necessary to obtain this result; this entails a rigorous development of the proof outlined by Drinfel'd & Sokolov.

For the meanwhile we will return to Miura's transformation. The explanation just given appears rather ungainly; in the next section we will examine a more elegant explanation due to Wilson (to be published).

§3.2 Miura's transformation and SL(2,$\mathbb{C}$)

Let us reconsider the factorization 3.1.4 in the light of differential Galois theory. The simple results we will use in this section can be found in the book by Kaplansky (1957, Hermann, see especially chapter vi).

Let $\mathbb{C}\langle u\rangle$ denote the differential field of rational functions of u, u_x, u_{xx} and so on, i.e. the quotient field of the differential algebra $\mathbb{C}\{u\}$. We view $\mathbb{C}\langle q\rangle$ as the field extension of $\mathbb{C}\langle u\rangle$ necessary to factorize the Schrödinger operator. As Wilson (to be published) points out, it would be very convenient if this was a Galois extension, that is, if $\mathbb{C}\langle u\rangle$ was the fixed field of some group G of (differential) automorphisms on $\mathbb{C}\langle q\rangle$. For then any G-equivariant derivation on $\mathbb{C}\langle q\rangle$ would preserve $\mathbb{C}\langle u\rangle$.

However, this is not the case, but the reason why implies a more interesting explanation of Miura's transformation.

The field extension $\mathbb{C}\langle u\rangle \subset \mathbb{C}\langle \psi,\phi\rangle$ implicitly used in the previous section is the extension containing the solution space to the equation

$$(\partial_x^2 + u)\psi = 0$$

It is a Galois extension with Galois group SL(2,$\mathbb{C}$), where the action is

$$\begin{aligned}(\ \psi \quad \phi\ ) \longmapsto (\ \psi \quad \phi\ )\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}\\ = (\ \alpha\psi+\gamma\phi \quad \beta\psi+\delta\phi\ ) \qquad \alpha\delta-\beta\gamma = 1\end{aligned} \qquad 3.2.1$$

It is not difficult to see why. Consider the transformations 3.2.1 for arbitrary $\alpha,\beta,\gamma,\delta$. Provided $\alpha\delta-\beta\gamma\neq 0$ this is a change of basis for the null-space of $\partial_x^2 + u$, therefore u is invariant. However we also have the relation $W(\psi,\phi)=1$, which is not preserved unless $\alpha\delta-\beta\gamma=1$.

At this point we are reminded that the field $\mathbb{C}\langle \psi,\phi\rangle$ is not freely generated; it implicitly contains the relation $\psi\phi_x-\phi\psi_x=1$. We prefer to work with a free field since then we have independent indeterminates in which to write our equations. Fortunately this can be remedied by letting $\eta = \phi/\psi$. The differential field $\mathbb{C}\langle\eta\rangle$ is the subfield of $\mathbb{C}\langle \psi,\phi\rangle$ fixed by the centre $\{\pm I\}$ of SL(2,$\mathbb{C}$). The centre is a normal subgroup, therefore we conclude that $\mathbb{C}\langle u\rangle \subset \mathbb{C}\langle\eta\rangle$ is also a Galois extension, with Galois group PSL(2,$\mathbb{C}$) =

SL(2,$\mathbb{C}$)/centre.

Indeed we have a sequence of extensions

$$\mathbb{C}\langle u\rangle \subset \mathbb{C}\langle q\rangle \subset \mathbb{C}\langle \eta\rangle \qquad 3.2.2$$

where

$$u = q_x - q^2 \quad , \quad q = \frac{1}{2}\,\eta_{xx}\eta_x^{-1} \qquad 3.2.3$$

The action of PSL(2,$\mathbb{C}$) on $\mathbb{C}\langle\eta\rangle$ is by linear fractional transformations

$$\eta \longmapsto \frac{\beta + \delta\eta}{\alpha + \gamma\eta} \qquad \alpha\delta-\beta\gamma=1 \qquad 3.2.4$$

We see from this that $\mathbb{C}\langle q\rangle \subset \mathbb{C}\langle\eta\rangle$ is the Galois extension corresponding to the solvable subgroup PB_+ of upper triangular matrices in SL(2,$\mathbb{C}$) modulo {±I}. This subgroup acts via the affine transformations

$$\eta \longmapsto (\beta\alpha + \eta)\alpha^{-2} \qquad 3.2.5$$

However, PB_+ is not a normal subgroup of PSL(2,$\mathbb{C}$), therefore $\mathbb{C}\langle q\rangle$ cannot be a Galois extension of $\mathbb{C}\langle u\rangle$. Nevertheless, Wilson points out that we have all we need to explain Miura's transformation.

The equation

$$\eta_t = \eta_{xxx} - \frac{3}{2}\,\eta_{xx}^2\eta_x^{-1} \qquad 3.2.6$$

on $\mathbb{C}\langle\eta\rangle$ is PSL(2,$\mathbb{C}$)-invariant. This can be seen by writing this equation as

$$\eta_t = S(\eta)\eta_x \qquad 3.2.7$$

The expression $S(\eta)$ is called the Schwartzian derivative of η; it is well known to be invariant under linear fractional transformations (indeed $S(\eta) = 2u$). Therefore the derivation ∂_t defined by 3.2.6 preserves both the subfields $\mathbb{C}\langle q\rangle$ and $\mathbb{C}\langle u\rangle$, and of course any subfield fixed by a subgroup of PSL(2,$\mathbb{C}$). This equation yields the mKdV and KdV equations for the variables q and u respectively. Wilson (to be published) calls 3.2.6 the "Ur-KdV equation".

Now it would be very pleasing if we could say that this explanation is the blueprint for a conceptually neater proof of the result obtained by Drinfel'd & Sokolov described at the end of the previous section. However, it is not clear, except in the case of $\mathfrak{sl}(\ell+1,\mathbb{C})$, what the correct analogue of the top field $\mathbb{C}<\phi,\psi>$ is. So first let us look at the case of $\mathfrak{sl}(\ell+1,\mathbb{C})$ and later on we will attempt to understand the more difficult cases.

§3.3 The case of the principal grading on $\mathfrak{sl}(\ell+1,\mathbb{C})$.

We will look at how the Galois group approach can be used for zero curvature equations with the spectral operator

$$\partial_x + \begin{bmatrix} q_0 & -z & 0 \dots & & \\ 0 & q_1 & -z & 0 \dots & \\ & & \dots & \dots & -z \\ -z & & & 0 & q_\ell \end{bmatrix} \qquad \sum_{i=0}^{\ell} q_i = 0 \qquad 3.3.1$$

These come from the principal grading on the Lie algebra with diagram $a_\ell^{(1)}$. The zero curvature equation defines a derivation on the differential field $\mathbb{C}<q_0,\dots,q_\ell>$ which lies in the middle of the sequence of field extensions

$$\mathbb{C}<u_0,\dots,u_\ell> \subset \mathbb{C}<q_0,\dots,q_\ell> \subset \mathbb{C}<\psi_0,\dots,\psi_\ell> \qquad 3.3.2$$

where

$$L_{\ell+1} = \partial_x^{\ell+1} + u_\ell\partial_x^\ell + \dots + u_0 = (\partial_x+q_\ell)\dots(\partial_x+q_0) \qquad 3.3.3$$

and

$$(\partial_x+q_i)\dots(\partial_x+q_0)\psi_i = 0 \quad , \quad i=0,\dots,n \qquad 3.3.4$$

The constraint

$$\sum_{i=0}^{\ell} q_i = 0$$

is equivalent to $u_\ell \equiv 0$ and therefore

$$W(\psi_0,\dots,\psi_\ell) = \text{constant} \qquad 3.3.5$$

We will fix the Wronskian to be 1. Consequently the differential fields in the sequence 3.3.2 are not freely generated (although it is trivial to remedy this for the bottom two fields). We will look at this problem later. First we will see how to construct an $SL(\ell+1,\mathbb{C})$-invariant equation on the top field which induces our given zero curvature equation on $\mathbb{C}<q_0,\dots,q_\ell>$.

In the principal realization the zero curvature equation has the form

$$[\ \partial_x + q - z\Lambda\ ,\ \partial_t - (v_k z^k + \ldots + v_0)\ ] = 0 \qquad 3.3.6$$

where $\partial_x + q - z\Lambda$ is the operator 3.3.1. However, we may also write these two commuting operators in the standard realization, so that we have the commutator

$$[\ \partial_x + q - (z\Lambda_- + \Lambda_+)\ ,\ \partial_t - (\nu_k z^k + \ldots + \nu_0)\ ] = 0 \qquad 3.3.7$$

Here Λ_+ is the matrix with 1's along the super diagonal and 0's elsewhere, while Λ_- has a single 1 in the bottom left corner. The matrix ν_0 is upper triangular with entries belonging to the differential algebra $\mathbb{C}\{q_0, \ldots, q_\ell\}$. We know that the equation 3.3.7 is identical to

$$[\ \partial_x + q - \Lambda_+\ ,\ \partial_t - \nu_0\ ] = 0 \qquad 3.3.8$$

since these operators are the components independent of z (more formally, of degree zero in the standard grading) from the pair of commuting operators above. For example, the mKdV equation can be represented by the commuting pair of upper triangular matrices

$$\left[\ \partial_x + \begin{pmatrix} q & -1 \\ 0 & -q \end{pmatrix}\ ,\ \partial_t - \begin{pmatrix} 2q^3 - q_{xx} & 2(q_x - q^2) \\ 0 & q_{xx} - 2q^3 \end{pmatrix}\ \right] = 0 \qquad 3.3.9$$

The operator $\partial_x + q - \Lambda_+$ is the operator 3.1.11. As with the $\mathfrak{sl}(2,\mathbb{C})$ case, it is possible to write the relations 3.3.4 in the form

$$(\partial_x + q - \Lambda_+)\Psi = 0 \qquad 3.3.10$$

where Ψ is upper triangular and has top row

$$(\ \psi_0\ \ \psi_1\ \ \ldots\ \ \psi_\ell\ )$$

All the other entries of Ψ are elements of the field $\mathbb{C}\langle\psi_0, \ldots, \psi_\ell\rangle$ and are uniquely determined by the condition 3.3.10.

It is not difficult to see that the equation 3.3.8 is the self-consistency condition for defining the derivation ∂_t on $\mathbb{C}\langle\psi_0, \ldots, \psi_\ell\rangle$ by

$$\partial_t \Psi = \nu_0 \Psi \qquad 3.3.11$$

Needless to say, on the subfield $\mathbb{C}\langle q_0, \ldots, q_\ell\rangle$ this derivation is identical to the zero curvature derivation.

We want to show that the action of $SL(\ell+1,\mathbb{C})$ on

$\mathbb{C}<\psi_0,\dots,\psi_\ell>$ given by

$$(\ \psi_0\ \dots\ \psi_\ell\ ) \longmapsto (\ \psi_0\ \dots\ \psi_\ell\ )g \qquad g\in SL(\ell+1,\mathbb{C}) \qquad 3.3.12$$

leaves the equation 3.3.11 invariant. First we translate the action into an action on Ψ.

For any $g \in SL(\ell+1,\mathbb{C})$ we can factorize the matrix Ψg into the product

$$\Psi g = (\Psi g)_-(\Psi g)_+ \qquad 3.3.13$$

where $(\Psi g)_-$ is a lower triangular unipotent matrix, $(\Psi g)_+$ is upper triangular, both of whose entries are rational expressions in the coordinates of Ψ. This is a manifestation of the fact that the big cell N_-B_+ is open dense in $SL(\ell+1,\mathbb{C})$, where N_- is the group of lower triangular unipotent matrices (see, for example, Humphreys (1975, Springer))

3.3.14. Lemma. *The action* 3.3.12 *induces the mapping*

$$\Psi \longmapsto (\Psi g)_+$$

Proof. We can easily check that the top row of $(\Psi g)_+$ is precisely the image of 3.3.12; notice that it is equal to the top row Ψg. To prove that the other entries are correct we show that they bear the same relation to this top row as the entries of Ψ do to its top row.

Recall Ψ is characterized by the equation 3.3.10. So clearly

$$(\Psi g)_x(\Psi g)^{-1} = -q + \Lambda_+$$

The left hand side is

$$[(\Psi g)_-]_x(\Psi g)_-^{-1} + (\Psi g)_-[(\Psi g)_+]_x(\Psi g)_+^{-1}(\Psi g)_-^{-1}$$

Therefore

$$\partial_x - [(\Psi g)_+]_x(\Psi g)_+^{-1} = (\Psi g)_-^{-1}[\partial_x + q - \Lambda_+](\Psi g)_-$$

The left hand side of this is upper triangular, whereas the right hand side only has Λ_+ above the diagonal, therefore it has the form

$$\partial_x + q' - \Lambda_+ \qquad 3.3.15$$

where q' is a diagonal matrix of trace zero. The entries of q' are characterized by the equation

$$(\partial_x + q' - \Lambda_+)(\Psi g)_+ = 0$$

and therefore are completely determined by the top row of $(\Psi g)_+$. Consequently the other entries of $(\Psi g)_+$ are determined by the top row using equations identical to those in 3.3.10. ∎

We already know from the Galois theory that the Galois group of $\mathbb{C}\langle\psi_0,\ldots,\psi_\ell\rangle$ over $\mathbb{C}\langle q_0,\ldots,q_\ell\rangle$ is B_+ and we recover part of this from the observation that if $g \in B_+$ then $(\Psi g)_+ = \Psi g$, so

$$(\Psi g)_x(\Psi g)^{-1} = \Psi_x g.g^{-1}\Psi^{-1} = -q + \Lambda_+ \qquad 3.3.16$$

Therefore the entries of q are certainly B_+-invariant.

Remark. Let us look at the reason why B_+ is the Galois group of $\mathbb{C}\langle\psi_0,\ldots,\psi_\ell\rangle$ over $\mathbb{C}\langle q_0,\ldots,q_\ell\rangle$. The factorization 3.3.3 of $L_{\ell+1}$ uniquely determines a flag of subspaces of the linear span Y of $\{\psi_0,\ldots,\psi_\ell\}$. This flag is $Y_0\subset Y_1\subset\ldots\subset Y_\ell=Y$ where Y_i is the kernel of the operator in 3.3.4. The action of $SL(\ell+1,\mathbb{C})$ on Y defined by 3.3.12 is such that g fixes the flag above if and only if $g \in B_+$. This is because $\{\psi_0,\ldots,\psi_i\}$ spans Y_i.

The entries of the matrix q′ in 3.3.15 show how $SL(\ell+1,\mathbb{C})$ acts on the elements $q_0,\ldots,q_\ell$.

3.3.17. Corollary. *For any* $g \in SL(\ell+1,\mathbb{C})$ *we may write*

$$\partial_x + g\circ q - \Lambda_+ = (\Psi g)_-^{-1}[\partial_x + q - \Lambda_+](\Psi g)_- \qquad 3.3.18$$

where

$$g\circ q = diag(g\circ q_0,\ldots,g\circ q_\ell)$$

and $g\circ q_i$ *is the expression for the action of* g *on* q_i *induced by 3.3.12.*

A more explicit expression comes from the expansion of 3.3.18

$$g\circ q - \Lambda_+ = (\Psi g)_-^{-1}\partial_x(\Psi g)_- + (\Psi g)_-^{-1}(q - \Lambda_+)(\Psi g)_-$$

Since the right hand side is upper triangular the lower triangular terms on the left hand side must cancel so that

$$g \circ q = q - [(\Psi g)_-^{-1} \Lambda_+ (\Psi g)_-]_d \qquad 3.3.19$$

where the subscript "d" denotes the diagonal component of the matrix in brackets.

It will now be shown that the equation 3.3.11 is $SL(\ell+1,\mathbb{C})$-invariant.

3.3.20. Proposition. *Let* $g \circ \Psi$, $g \circ \nu_0$ *denote the transformation by* $g \in SL(\ell+1,\mathbb{C})$ *of the elements of* Ψ *and* ν_0 *induced by the action* 3.3.12. *Then*

$$\partial_t(g \circ \Psi) = (g \circ \nu_0)(g \circ \Psi)$$

Thus the derivation ∂_t *is* $SL(\ell+1,\mathbb{C})$*-equivariant, so the equation is* $SL(\ell+1,\mathbb{C})$*-invariant.*

Proof. We have already established that $g \circ \Psi = (\Psi g)_+$. Now we must prove that

$$\partial_t(g \circ \Psi)(g \circ \Psi)^{-1} = g \circ \nu_0 \qquad 3.3.21$$

To establish the right hand side we return to the definition of ν_0 as the component of degree zero, in the standard grading, of V_+. Recall that V_+ is defined as the series of terms of non-negative degree (in the principal grading) from the series V commuting with the spectral operator. Therefore ν_0 is determined by the equation

$$[\ \partial_x + q - (z\Lambda_- + \Lambda_+)\ ,\ V\ ] = 0$$

where we have chosen to use the standard realization. This equation implies

$$[\ (\Psi g)_-^{-1}(\partial_x + q - \Lambda_+)(\Psi g)_- + z\Lambda_-\ ,\ (\Psi g)_-^{-1} V (\Psi g)_-\ ] = 0$$

since Λ_- is fixed by the group N_- (it is a lowest weight vector for $SL(\ell+1,\mathbb{C})$). We can write this as

$$[\ \partial_x + g \circ q - \Lambda_+\ ,\ g \circ V\ ] = 0$$

by virtue of lemma 3.3.17. The series $g \circ V$ is therefore precisely the series V with the differential polynomials in q_i replaced by the same polynomials in $g \circ q_i$. This follows from the fact that we can conjugate V and $g \circ V$ into the same element of $c(\mathfrak{z}(z\Lambda))$ in the manner of the method of dressing. Thus the series have the same construction using different potentials in the spectral operator.

By definition $g\circ\nu_0$ is the matrix with the differential polynomials in q_i replaced by the same polynomials in $g\circ q_i$. Therefore $g\circ\nu_0$ is the term of degree zero (in the standard grading) of $(g\circ V)_+$, the series of non-negative terms (in the principal grading) of $g\circ V$. Denote this term of degree zero by $(g\circ V)^0$. Then

$$g\circ\nu_0 = (g\circ V)^0 = [(\Psi g)_-^{-1}V(\Psi g)_-]^0$$

$$= [(\Psi g)_-^{-1}V(\Psi g)_-]_+ \qquad 3.3.22$$

The latter identity is a result of $(\Psi g)_-$ being lower unipotent. Once again the subscript "+" denotes the terms of non-negative degree in the principal grading. In particular this term will be upper triangular since the object inside the brackets has degree zero in the standard grading.

Finally we will prove the proposition by showing that the left-hand side of 3.3.21 equals 3.3.22.

$$\begin{aligned}[(\Psi g)_+]_t(\Psi g)_+^{-1} &= [(\Psi g)_-^{-1}(\Psi g)]_t[(\Psi g)_-^{-1}(\Psi g)]^{-1}\\ &= \partial_t(\Psi g)_-^{-1}.(\Psi g)_- + (\Psi g)_-^{-1}.\Psi_t\Psi^{-1}.(\Psi g)_-\\ &= \partial_t(\Psi g)_-^{-1}.(\Psi g)_- + (\Psi g)_-^{-1}.\nu_0.(\Psi g)_-\end{aligned}$$

This must be upper triangular, therefore the first term must cancel the strictly lower triangular part of the second term, leaving 3.3.22. ■

Consequently the derivation ∂_t defined by 3.3.11 preserves both the subfields $\mathbb{C}\langle q_0,\dots q_\ell\rangle$ and $\mathbb{C}\langle u_0,\dots u_\ell\rangle$. Drinfel'd & Sokolov (1985) have shown that the zero curvature equation 3.3.6 induces a scalar Lax equation

$$\partial_t L_{\ell+1} = [P, L_{\ell+1}]$$

where P is a scalar differential operator with coefficients from $\mathbb{C}\langle u_0,\dots u_\ell\rangle$. We may conclude from their result that the equation 3.3.11 implies this Lax equation for the variables $u_0,\dots u_\ell$.

To write the equation 3.3.11 explicitly in terms of the variables $\psi_0,\dots,\psi_\ell$ we must include the constraint

$W(\psi_0,\dots,\psi_\ell) = 1$. However, as with the Ur-KdV equation, if we choose to define $\eta_i = \psi_i\psi_0^{-1}$ we find that the variables $\eta_1,\dots,\eta_\ell$ are differentially independent, that is, the algebra $\mathbb{C}\{\eta_1,\dots,\eta_\ell\}$ is free. We can see this by returning to the definitions 3.3.3 and 3.3.4.

Let us assume for the moment that u_ℓ is a free variable. Then all the fields in the sequence 3.3.2 are free. The Galois group of the top field over the bottom field is now $GL(\ell+1,\mathbb{C})$ (we have a "Picard-Vessiot" extension). In particular

$$W(g\circ\psi_0,\dots,g\circ\psi_\ell) = \det(g).W(\psi_0,\dots,\psi_\ell)$$

(see Kaplansky (Hermann, 1957)). The subfield $\mathbb{C}\langle\eta_1,\dots,\eta_\ell\rangle$ is clearly free in this case; it is the fixed field of the centre $\{kI \mid k\in\mathbb{C}^*\}$ of $GL(\ell+1,\mathbb{C})$. It follows that the Wronskian (and its derivatives) cannot belong to the field $\mathbb{C}\langle\eta_1,\dots,\eta_\ell\rangle$ since it is not fixed by this subgroup. Therefore the elements $\eta_1,\dots,\eta_\ell$ and $W(\psi_0,\dots,\psi_\ell)$ are differentially independent. As a result, constraining the Wronskian has no effect on the independence of the variables $\eta_1,\dots,\eta_\ell$.

The orbit of solutions under the Galois group.

As well as explaining the transformations of Miura type, the Galois group induces a group of transformations on the solution space of a zero curvature equation. This transformation is given by 3.3.17; since $g\circ\Psi = (\Psi g)_+$ is a solution of 3.3.11 it follows that $g\circ q$ must be a solution to the equation 3.3.8. This transformation will be non-trivial if $g \notin B_+$ (the orbit of a solution will be identifiable with the homogeneous space $SL(\ell+1,\mathbb{C})/B_+$).

3.3.23 Example. For the mKdV equation the $SL(2,\mathbb{C})$ action 3.2.1 induces the transformation

$$q \longmapsto q' = \tfrac{1}{2}\eta_{xx}\eta_x^{-1} - (\alpha + \gamma\eta)_x(\alpha + \gamma\eta)^{-1} \qquad 3.3.24$$

where we recall $q = \frac{1}{2}\eta_{xx}\eta_x^{-1}$.

For example the trivial solution $q=0$ transforms into the almost equally trivial rational solution

$$q' = \frac{a}{x + b} \quad \text{a, b constants}$$

A less trivial example is the transformation of the one-soliton solution

$$q(x,t) = z\operatorname{sech}\omega \quad \text{where } \omega = zx + z^3t$$

For arbitrary parameters α, γ in 3.3.24 this solution is transformed into

$$q'(x,t) = z\operatorname{sech}\omega - [\frac{\alpha}{\gamma}e^{-2\rho} - \frac{1}{4z}e^{2\rho}\ln(\operatorname{cosec}2\rho + \cot 2\rho)]^{-1}$$

where $\rho = \operatorname{Arctan}(e^{\omega})$. This new solution still vanishes as x tends towards $\pm\infty$ but, unfortunately, its behaviour is far from nice since the function

$$\frac{\alpha}{\gamma}e^{-2\rho} - \frac{1}{4z}e^{2\rho}\ln(\operatorname{cosec}2\rho + \cot 2\rho)$$

has a zero in the range of ρ, $0 < \rho < \frac{\pi}{2}$.

The examples demonstrate that the transformation 3.3.24 is significantly different from the "Bäcklund transformation" constructed by Wadati (1974). This is also evident from the fact that the transformation 3.2.12 does not explicitly involve the time variable t, whereas in the Bäcklund transformation it is necessary to include an equation governing the t-dependence of the new solution.

§3.4 G-invariant equations in general.

The special form of the operator 3.3.1 allowed us to associate the variables $q_0, \ldots q_\ell$ with the factorization of a scalar linear operator. In general the spectral operator for a zero curvature equation will not have a form which suggests any such associated linear operator. Without this it is not clear that there is any sequence of field extensions analogous to 3.3.2.

Nevertheless, it is possible to approach the construction of G-invariant equations, at least for $G = PSL(\ell+1,\mathbb{C})$, from another point of view which does not depend upon the existence of the associated scalar differential operator. This holds some hope of being useful in the general case. It will be illustrated by returning to the case $G = PSL(2,\mathbb{C})$; the Ur-KdV - mKdV - KdV sequence introduced in §3.2.

The principle, due to Wilson (private communication), is to regard the Ur-KdV equation as defining a local flow on a space of functions $\mathfrak{X}$ on which G acts on the right. The field extensions 3.2.3 correspond (in reverse order) to the quotient spaces in the sequence

$$\mathfrak{X} \longrightarrow \mathfrak{X}/PB_+ \longrightarrow \mathfrak{X}/G \qquad 3.4.1$$

Intuitively each differential field represents the "field of functions" over one of these spaces.

<u>Remark.</u> Unfortunately this analogy between finite dimensional varieties where the field of functions is the field of fractions of the coordinate ring, and infinite dimensional manifolds is not a good one. One of the problems is that we cannot write every functional on a function space as the (integral of) a rational expression involving the functions and their derivatives.

For our purposes the analytic properties of the space $\mathfrak{X}$ are irrelevant provided we end up with a sensible coordinate description. Therefore, to avoid questions about global structure, we will deal with germs of functions.

We define $\mathfrak{X}$ to be the space of germs (at 0) of holomorphic functions

$$\Psi : \mathbb{C} \underset{0}{\longrightarrow} N_-\backslash G \qquad 3.4.2$$

with the property that, for any lift

$$\Phi : \mathbb{C} \underset{0}{\longrightarrow} G \qquad 3.4.3$$

of Ψ (i.e $\Psi(x) = N_-\Phi(x)$ in some neighbourhood of 0) the matrix germ $\Phi_x\Phi^{-1}$ has the form

$$\begin{bmatrix} * & 1 \\ * & * \end{bmatrix} \qquad 3.4.4$$

where the unspecified entries are germs of $\mathbb{C}$-valued functions. This property can be stated more formally as

$$\Phi_x\Phi^{-1} : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{b}_- + \Lambda_+ \qquad 3.4.5$$

where $\mathfrak{b}_- \subset \mathfrak{sl}(2,\mathbb{C})$ is the subalgebra of lower triangular matrices, and Λ_+ is the matrix 3.4.4 with 0's in place of the unspecified entries.

In coordinates we write any element of $G = PSL(2,\mathbb{C})$ as the equivalence class

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{mod}\{\pm I\} \mid ad-bc=1 \right\}$$

The quotient $N_-\backslash G$ is the space

$$\{ [\, a \quad b \,] \mid (\, a \quad b \,) \in \mathbb{C}^2-\{(0\ 0)\} \}$$

If we write

$$\Phi = \begin{bmatrix} a(x) & b(x) \\ c(x) & d(x) \end{bmatrix}$$

the condition 3.4.5 reduces to $ab_x-a_xb = 1$. Therefore $\mathfrak{X}$ is the space

$$\{ [\, a(x) \quad b(x) \,] \mid W(a,b) = 1 \}$$

The Wronskian condition can be rewritten as

$$\left(\frac{b}{a}\right)_x = a^{-2} \qquad \text{or} \qquad \left(\frac{a}{b}\right)_x = -b^{-2}$$

therefore $\mathfrak{X}$ can be identified with

$$\{ \text{ certain germs of holomorphic functions } \eta : \mathbb{C} \underset{0}{\longrightarrow} \mathbb{P}^1 \}$$

Previously we fixed $\eta = \frac{b}{a}$ and ignored the point at infinity. However, this point is very important; the value of η passes through the point at infinity precisely when any lift Φ ceases to lie in the big cell N_-PB_+ of $PSL(2,\mathbb{C})$. We see this in the factorization

$$\Phi = \begin{bmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & \frac{1}{a} \end{bmatrix}\begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix}$$

This leads us back to the previous description of the coordinate η; in the previous section Ψ was the matrix of indeterminates representing

$$\begin{bmatrix} a & b \\ 0 & \frac{1}{a} \end{bmatrix} \qquad W(a,b) = 1 \qquad 3.4.6$$

3.4.7. Remark. It is worth noting that the matrix germ Ψ in 3.4.6 is a meromorphic germ. The condition $ab_x - a_x b = 1$ implies that any zero of a is a simple zero (otherwise b would not be holomorphic). Therefore Ψ has isolated (simple) poles. Equivalently, any lift Φ of Ψ takes values in the big cell except at isolated points.

3.4.8. Lemma. *$\mathfrak{X}/PB_+$ is in bijection with a set of meromorphic germs*

$$q : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{h} \cong \mathbb{C}$$

where $\mathfrak{h}$ is the Cartan subalgebra of diagonal matrices in $\mathfrak{sl}(2,\mathbb{C})$.

Proof. It is clear that we can identify each element Ψ of $\mathfrak{X}$ with a meromorphic matrix germ $\Psi:\mathbb{C} \underset{0}{\longrightarrow} PB_+$; it is given by 3.4.6. This germ satisfies

$$\Psi_x\Psi^{-1} : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{h} + \Lambda_+$$

therefore the diagonal $-q$ of $\Psi_x\Psi^{-1}$ is a meromorphic germ and is also PB_+-invariant (c.f. 3.3.16).

If two germs Ψ, Ψ' are such that $\Psi_x\Psi^{-1} = \Psi'_x(\Psi')^{-1}$ then $\Psi = \Psi' g$ for some $g \in PB_+$ therefore the map $\Psi_x\Psi^{-1} \longrightarrow q$ is a bijection. ■

The space $\mathfrak{X}/G$ is described using the gauge transformation due to Drinfel'd & Sokolov presented in §3.1. We can state as a fact that for each germ q there exists a germ

$$n(x) = \mathbb{C} \underset{0}{\longrightarrow} N_-$$

such that

$$n(\partial_x + q - \Lambda_+)n^{-1}$$

has the form

$$\partial_x + \begin{bmatrix} 0 & -1 \\ * & 0 \end{bmatrix} \qquad 3.4.9$$

where the unspecified entry is a meromorphic germ. This operator is the unique operator of this form in the gauge orbit of the group of lower triangular matrices. Given this we can prove the following lemma.

3.4.10. Lemma. *$\mathcal{X}/G$ is in bijection with the set of meromorphic germs $u : \mathbb{C} \underset{0}{\longrightarrow} \mathbb{C}$ obtained by the gauge transformation above on the operator $\partial_x - \Psi_x\Psi^{-1}$.*

Proof. Let Φ be any lift of Ψ. We will show first that all germs Φg, $g \in G$, lead to the same operator of the form 3.4.9.

The operator $\partial_x - \Phi_x\Phi^{-1}$ is the same as $\partial_x - (\Phi g)_x(\Phi g)^{-1}$ for every element g of G. It is equivalent to $\partial_x - \Psi_x\Psi^{-1}$ by the gauge transformation

$$m(x)(\partial_x - \Psi_x\Psi^{-1})m(x)^{-1} \qquad m(x) = \Phi\Psi^{-1}$$

This can be seen from

$$(\Phi\Psi^{-1}\Psi)_x(\Phi\Psi^{-1}\Psi)^{-1} = \Phi_x\Phi^{-1}$$

Therefore each element of ΦG, for any lift Φ of Ψ, leads to the same entry u in the operator 3.4.9.

Now suppose that Φ, Φ' are two lifts of elements of $\mathcal{X}$ which both map to the same operator of the form 3.4.8. Then

$$\partial_x - \Phi_x\Phi^{-1} = \partial_x - (l(x)\Phi')_x(l(x)\Phi')^{-1}$$

for some meromorphic germ $l(x)$ with values in N_- (this follows from both n, m having values in N_-). We may conclude that $(l\Phi')_x(l\Phi')^{-1}$ is holomorphic since $\Phi_x\Phi^{-1}$ is. Therefore

$$\Phi = l(x)\Phi' g$$

and $l(x)\Phi'$ must be holomorphic in a neighbourhood of 0. Consequently for each x in this neighbourhood

$$N_-\Phi(x) = N_-\Phi'(x)g$$

Therefore Φ, Φ' are both lifts of the same element of $\mathcal{X}/G$. ∎

So finally we have the structure 3.4.1 we require. The most important thing is that we have obtained decent coordinates η, q and u for $\mathcal{X}$, $\mathcal{X}/PB_+$ and $\mathcal{X}/G$ (q is the entry in the diagonal matrix q). If we ignore the details of the analysis we can begin to formulate an approach to generalizing this construction.

The essential ingredient for the generalization is the presence of a second $\mathbb{Z}$-grading on $L(\mathfrak{g},\theta)$. We replace **G** by the adjoint group **A** of a Lie subalgebra $\mathfrak{a} \subset \mathfrak{g}$ where $\mathfrak{a}$ is (isomorphic to) the subalgebra $L(\mathfrak{g},\theta)^0$ of elements of $L(\mathfrak{g},\theta)$ with degree zero in the second $\mathbb{Z}$-grading. This second grading is coarser than the original in the sense that

$$L(\mathfrak{g},\theta)_0 \subset L(\mathfrak{g},\theta)^0$$

Remark. The loop algebra $L(\mathfrak{g},e)$ with the standard grading has the most coarse of all gradings since the Lie algebra of elements of degree zero is isomorphic to $\mathfrak{g}$. Thus we define $\mathfrak{a}$ to be the Lie subalgebra of $\mathfrak{g}$ corresponding to the embedding of $L(\mathfrak{g},\theta)^0$ in $L(\mathfrak{g},e)_0$.

For example, the case dealt with in the previous section compares the principal grading with the standard grading. In this case $L(\mathfrak{g},\theta)_0 \cong \mathfrak{h}$.

In general the original grading on $L(\mathfrak{g},\theta)$ induces a non-trivial $\mathbb{Z}$-grading on $\mathfrak{a}$. We use this to obtain the "triangular" decomposition

$$\mathfrak{a} = \mathfrak{m}_- + \mathfrak{g}_0 + \mathfrak{m}_+$$

where $\mathfrak{g}_0 \cong L(\mathfrak{g},\theta)_0$ and $\mathfrak{m}_-$ ($\mathfrak{m}_+$) is the nilpotent subalgebra of elements of negative (positive) degree. In the case of the principal-vs-standard grading this is

$$\mathfrak{g} = \mathfrak{n}_- + \mathfrak{h} + \mathfrak{n}_+$$

We would like to take the space $\mathcal{X}$ to be the subspace of

$$\{ \text{ holomorphic germs } \Psi:\mathbb{C} \underset{0}{\longrightarrow} M_-\backslash A \} \qquad M_- = \exp \mathfrak{m}_-$$

such that for each lift Φ of Ψ

$$\Phi_x\Phi^{-1} : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{p}_- + \Lambda_+ \quad , \ \mathfrak{p}_- = \mathfrak{m}_- + \mathfrak{g}_0 \qquad 3.4.11$$

Here Λ_+ is the component of Λ ($\in \mathfrak{g}_1$) lying in $\mathfrak{m}_+$. We will see later that the properties desired of Λ_+ strongly affect the choice that is made for the second grading.

We propose to replace the sequence of projections 3.4.1 by

$$\mathcal{X} \longrightarrow \mathcal{X}/P_+ \longrightarrow \mathcal{X}/A \qquad 3.4.12$$

where $P_+ = \exp \mathfrak{p}_+$, $\mathfrak{p}_+ = \mathfrak{g}_0 + \mathfrak{m}_+$. For this to make any sense at all we must at least be able to show that

$$\mathcal{X}/P_+ \cong \{ \text{ certain germs } q : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{g}_0 \} \qquad 3.4.13$$

This is identifiable with a set of potentials for the spectral operator. This can be done, as earlier, by identifying each element of $\mathcal{X}$ with a matrix germ

$$\Psi : \mathbb{C} \underset{0}{\longrightarrow} P_+$$

Then

$$\Psi_x \Psi^{-1} : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{g}_0 + \Lambda_+$$

Conversely, Ψ is the unique solution, up to right multiplication by an element of P_+, to the equation

$$(\partial_x + q - \Lambda_+)\Psi = 0 \qquad 3.4.14$$

Given a zero curvature equation

$$[\partial_x + q - z\Lambda \ , \ \partial_t - V_+] = 0$$

we have a local flow ∂_t for the coordinates, given by 3.4.13, on $\mathcal{X}/P_+$. In principle we can define a flow on $\mathcal{X}$ by

$$\partial_t \Psi = \nu_0 \Psi \qquad 3.4.15$$

where ν_0 is the component of degree zero of V_+ in the second grading on $L(\mathfrak{g},\theta)$. This induces the flow

$$[\partial_x + q - \Lambda_+ \ , \ \partial_t - \nu_0] = 0$$

on $\mathcal{X}/P_+$, which is precisely the zero curvature equation we started with. This is in direct analogy to the observations 3.3.10 and 3.3.11 in the previous section.

In fact symbolically we can repeat the steps of the proof of proposition 3.3.20 that the equation 3.4.15, in the case of $\mathfrak{sl}(\ell+1,\mathbb{C})$ with the principal-vs-standard grading, is A-invariant. However this is meaningless unless the equations can be written in coordinates. This is an unresolved problem.

Curiously, it is easier to describe the space $\mathcal{X}/A$ in coordinates. The basic idea is contained in the proof of lemma 3.4.10. We identify each element of $\mathcal{X}/A$ with a double coset $\mathcal{M}_-\Phi A$, where

$$\mathcal{M}_- = \{ \text{ germs } m : \mathbb{C} \underset{0}{\longrightarrow} M_- \}$$

and Φ satisfies 3.4.11. This is in turn identified with a gauge-equivalence class of operators

$$\{ m(\partial_x - \Phi_x\Phi^{-1})m^{-1} \mid m \in \mathcal{M}_- \} \qquad 3.4.16$$

which contains the operator in 3.4.14. In the next section we will effectively show that, for certain choices of Λ_+, there exists a unique operator in this class which belongs to the set

$$\{ \partial_x + \mu - \Lambda_+ \mid \mu : \mathbb{C} \underset{0}{\longrightarrow} \mathfrak{l} \}$$

where $\mathfrak{l}$ is a particular subspace of $\mathfrak{a}$ (in the case of $\mathfrak{sl}(\ell+1,\mathbb{C})$ dealt with in the previous section the unique operator is given by 3.1.12).

This brings us the full circle round to the method of Drinfel'd & Sokolov. We saw in §3.1 that they managed to describe the transformation

$$\mathcal{X}/PB_+ \longrightarrow \mathcal{X}/SL(2,\mathbb{C})$$

without any knowledge of the space $\mathcal{X}$. In the next section we will use the ideas of Drinfel'd & Sokolov to, effectively, construct the map

$$\mathcal{X}/P_+ \longrightarrow \mathcal{X}/A$$

although we will return to an entirely algebraic point of view. A caveat on this construction is that the element Λ_+ must satisfy certain conditions, so only certain hierarchies of zero curvature equations admit this "transformation of Miura type".

§3.5 The Miura-Drinfel'd-Sokolov transformation.

We will replace the map

$$\mathcal{X}/P_+ \longrightarrow \mathcal{X}/A \qquad 3.5.1$$

with the "dual" mapping (inclusion) of differential algebras

$$\mathbb{C}\{u_i\} \hookrightarrow \mathbb{C}\{q_i\} = \mathcal{B} \qquad 3.5.2$$

This will be done by formalizing the method used by Drinfel'd & Sokolov (1985) to describe the coordinates of the gauge-equivalence classes introduced at the end of the previous section. The setting for our main result will be as follows.

Let $(\mathfrak{g},\theta,\Lambda)$ be a hierarchy of zero curvature equations and let $L(\mathfrak{g},\theta)$ have the grading of type $(s_0,\dots,s_\ell)$ (this is a sequence of 0's and 1's). Choose another $\mathbb{Z}$-grading, of type $(\sigma_0,\dots,\sigma_\ell)$, which is coarser than the first in the sense: $\sigma_j = 0$ if $s_j = 0$. We denote the homogeneous subspace of elements of degree j in the latter grading by $L(\mathfrak{g},\theta)^j$. Then $L(\mathfrak{g},\theta)_0 \subset L(\mathfrak{g},\theta)^0$ and we can decompose the latter reductive subalgebra into

$$L(\mathfrak{g},\theta)^0 = \mathfrak{m}_- \oplus L(\mathfrak{g},\theta)_0 \oplus \mathfrak{m}_+$$

where $\mathfrak{m}_-$ $(\mathfrak{m}_+)$ is the nilpotent subalgebra of elements of negative (positive) degree in $L(\mathfrak{g},\theta)$. As before we define $\mathfrak{p}_- = \mathfrak{m}_- + L(\mathfrak{g},\theta)_0$. Let $\lambda \in L(\mathfrak{g},\theta)_1$ be the semisimple element covering $\Lambda \in \mathfrak{g}_1$ (i.e $\lambda \equiv z\Lambda$). It is clear that

$$L(\mathfrak{g},\theta)_1 \subset L(\mathfrak{g},\theta)^0 \oplus L(\mathfrak{g},\theta)^1$$

We split $\lambda = \lambda_0 + \lambda_1$ accordingly (it follows that λ_1 is the lift of Λ_+ to $L(\mathfrak{g},\theta)_1$).

Our aim is to prove the following result, which is a generalization of the result obtained by Drinfel'd & Sokolov (1985) (described at the end of §3.1).

3.5.3. <u>Proposition.</u> *Suppose the hierarchy $(\mathfrak{g},\theta,\Lambda)$ has been chosen together with a coarse grading of type $(\sigma_0,\dots,\sigma_\ell)$ such that the semisimple element λ in the loop algebra satisfies:*

(i) $[\lambda_1 , \mathfrak{m}_-] = 0$,

(ii) $ad\lambda_0 : \mathfrak{m}_- \longrightarrow \mathfrak{p}_-$ *is injective.*

Then there exists a free subalgebra $\mathbb{C}\{u_i\}$ of $\mathbb{C}\{q_i\}$ such that any derivation on $\mathbb{C}\{q_i\}$, defined by a zero curvature equation from $(\mathfrak{g},\theta,\Lambda)$, maps $\mathbb{C}\{u_i\}$ into itself. The variables u_i are obtained by a transformation of Drinfel'd-Sokolov type, in other words, by a gauge transformation of the operator $\partial_x+q-\lambda$.

3.5.4. <u>Example.</u> Let us show that the cases treated by Drinfel'd & Sokolov satisfy the conditions of the proposition. When $L(\mathfrak{g},\theta)$ corresponds to the principal grading we may assume, without loss of generality,

$$\lambda = \sum_{i=0}^{\ell} e_i$$

where $\{e_i,h_i,f_i\}$ is the set of canonical generators for $L\mathfrak{g}(A)$. Drinfel'd & Sokolov choose the second grading to be of type $(0,\ldots,s_i,\ldots,0)$, $s_i=1$. In this case the algebra $\mathfrak{m}_-$ is generated by $\{f_j|j\neq i\}$, $\mathfrak{p}_-$ is generated by $\{f_j,h_j|j\neq i\}$ and

$$\lambda_0 = \sum_{j\neq i} e_j \quad , \quad \lambda_1 = e_i$$

Here property *(i)* is a consequence of the relations $[e_i,f_j]=\delta_{ij}h_i$.

Property *(ii)* follows from a result of Kostant's (1959), referred to earlier, which says that $\lambda_0+\rho$ is regular semisimple if ρ is a lowest weight vector for $\mathfrak{p}_-$. A regular semisimple element only has semisimple elements in its centralizer, whereas $\mathfrak{m}_-$ only consists of nilpotents. Therefore the kernel of $ad(\lambda_0+\rho)$ on $\mathfrak{m}_-$ is trivial. However, $[\rho,\mathfrak{m}_-] = 0$ since ρ is a lowest weight vector, therefore $ad\lambda_0$ is injective on $\mathfrak{m}_-$. More generally, this argument shows that if λ is regular semisimple then *(ii)* of 3.5.3 follows from *(i)*.

The principle behind the transformation developed by Drinfel'd & Sokolov is the notion of "dividing out" by the gauge action of the group $\mathcal{M}_-$ on the space of operators

$$\{ \partial_x - \Phi_x\Phi^{-1} \mid \Phi \text{ a lift of an element of } \mathfrak{X} \}$$

described above. In the algebraic formulation we replace this set by a "generic element"; the operator

$$\begin{aligned}\mathcal{L} &= \partial_x + \sum_i q_i\varepsilon_i + \sum_m r_m\zeta_m - \lambda \\ &= \partial_x + q + r - \lambda \qquad 3.5.4\end{aligned}$$

where $\{\varepsilon_i\}$ is a basis for $L(\mathfrak{g},\theta)_0$,$\{\zeta_m\}$ is a basis for $\mathfrak{m}_-$ and $\{r_m\}$ is a set of indeterminates independent of $\mathbb{C}\{q_i\}$. This operator belongs to the class

$$\partial_x + \mathfrak{p}_- \otimes \mathcal{A} - \lambda \qquad 3.5.5$$

where $\mathcal{A} = \mathbb{C}\{q_i, r_m\}$. For example, for the mKdV equation $\mathcal{L}$ has been given in 3.1.17.

The group $\mathcal{M}_-$ must be replaced by the group $\mathcal{M}_{\mathcal{A}} = \exp(\mathfrak{m}_- \otimes \mathcal{A})$. As promised at the end of the previous section, it will be shown that if λ satisfies the conditions *(i)* and *(ii)* in 3.5.3 then each $\mathcal{M}_{\mathcal{A}}$-gauge orbit in

$$\partial_x + \mathfrak{p}_- \otimes \mathcal{A} - \lambda$$

contains a unique element of the class of operators

$$\partial_x + \Gamma \otimes \mathcal{A} - \lambda$$

where the subspace $\Gamma \subset \mathfrak{p}_-$ is given by the next lemma.

3.5.6. Lemma. *Suppose*

$$ad\lambda_0 : \mathfrak{m}_- \longrightarrow \mathfrak{p}_-$$

is injective; then there exists a homogeneous subspace $\Gamma \subset \mathfrak{p}_-$ *satisfying*

$$\mathfrak{p}_- = \Gamma \oplus [\ \lambda_0,\ \mathfrak{m}_-\ ]$$

where $dim\ \Gamma = dim\ L(\mathfrak{g},\theta)_0$ *and* $\Gamma = \oplus_j \Gamma_j$, $\Gamma_j \subset L(\mathfrak{g},\theta)_j$.

Proof. Define $\mathfrak{p}_j = \mathfrak{p}_- \cap L(\mathfrak{g},\theta)_j$. For each $j<0$ the map

$$ad\lambda_0 : \mathfrak{p}_j \longrightarrow \mathfrak{p}_{j+1}$$

is injective by property *(i)* of 3.5.3. We define Γ_{j+1} to be a complementary subspace to $[\lambda_0, \mathfrak{p}_j]$ in $\mathfrak{p}_{j+1}$, and $\Gamma_k = \mathfrak{p}_k$ where k is the lowest degree. It follows that $dim\ \Gamma = dim\ L(\mathfrak{g},\theta)_0$ since $ad\lambda_0$ is injective on $\mathfrak{m}_-$. ∎

3.5.7. Remark. In the case of $\mathfrak{sl}(\ell+1,\mathbb{C})$ with the principal-vs-standard grading (as in sections 3.2 and 3.3) the subalgebra $\mathfrak{p}_-$ is the algebra of lower triangular matrices and we may choose Γ to corresond to space of "companion" matrices, which have non-zero entries only in bottom row (excluding the diagonal position).

From now on we will assume that λ satisfies the conditions of proposition 3.5.3.

3.5.8. Lemma. *Let* $\mathcal{E}$ *be a differential algebra,* $p \in \mathfrak{p}_- \otimes \mathcal{E}$.

Then there exists a unique $y \in m_- \otimes \mathcal{E}$ *such that*

$$\exp ady\, (\partial_x + p - \lambda) \in \partial_x + \Gamma \otimes \mathcal{E} - \lambda$$

<u>Proof.</u> We decompose p into $\sum_{j\le 0} p_j$, y into $\sum_{j<0} y_j$ where $p_j, y_j \in \mathfrak{p}_j$. Then the series above expands into

$$-\lambda_1 - \lambda_0 + \{p_0 + [\ \lambda_0,\ y_{-1}\ ]\}$$
$$+ \{p_{-1} + [\lambda_0, y_{-2}] + \tfrac{1}{2}[y_{-1}, [\lambda_0, y_{-1}]] - \partial_x y_{-1}\} + \ldots$$

where we have used the property $[\lambda_1, y] = 0$ from 3.5.3. Here we have gathered all the terms of the same degree in $L(\mathfrak{g},\theta)$ (note that λ_0 has degree 1). Since $ad\lambda_0$ is injective on m_- we see that there exists a unique y_{-1} such that the first term in braces belongs to Γ_0. Similarly there exists a unique y_{-2}, given y_{-1}, such that the second term in braces belongs to Γ_{-1}. The series is finite, therefore y is uniquely determined by this process. ∎

As a result of this lemma it is possible to define a map

$$M_{\varepsilon} : \partial_x + \mathfrak{p}_- \otimes \mathcal{E} - \lambda \longrightarrow \partial_x + \Gamma \otimes \mathcal{E} - \lambda$$

which assigns to each element of the left hand side the unique operator given by this lemma.

3.5.10. <u>Definition.</u> *We will call the transformation*

$$\mathcal{L} \longmapsto M_{\mathcal{A}}(\mathcal{L})$$

the Miura-Drinfel'd-Sokolov transformation of the operator $\mathcal{L}$. *Fix a basis* $\{\xi_i\}$ *of* Γ, *then we define* $\{\mu_i\}$ *to be the set of coordinates for* $M_{\mathcal{A}}(\mathcal{L})$:

$$M_{\mathcal{A}}(\mathcal{L}) = \partial_x + \sum \mu_i \xi_i - \lambda$$

We denote by $\mathcal{C}$ *the differential subalgebra* $\mathbb{C}\{\mu_i\}$ *of* $\mathcal{A}$.

Now we are going to prove proposition 3.5.3 after choosing the variables u_i to be given by setting all $r_m \equiv 0$ in the expressions for $\mu_i \in \mathcal{A}$. The algebra $\mathcal{A}$ then collapses to $\mathbb{C}\{q_i\} = \mathcal{B}$, so we have the inclusion 3.5.2. Notice also that with all $r_m \equiv 0$ the operator $\mathcal{L}$ becomes $\partial_x + q - \lambda$ and we still have a gauge transformation

$$\exp\, ad\; y(r_m \equiv 0)\{\partial_x + q - \lambda\} \in \partial_x + \Gamma \otimes \mathbb{C}\{u_i\} - \lambda$$

where $y(r_m \equiv 0)$ denotes that all $r_m \equiv 0$ in y. The variables u_i are, of course, the coordinates of this operator.

To prove 3.5.3 we must show that each zero curvature derivation preserves $\mathbb{C}\{u_i\}$. To do this we define two derivations, ∂_V and ∇_V, on $\mathcal{A}$. The former will yield the zero curvature equation when all $r_m \equiv 0$.

Both derivations are defined using the same series V(r) commuting with the operator $\mathcal{L}$. V(r) is characterized by the element v in $c(\mathfrak{z}(\lambda))$; we can repeat the dressing construction used in chapter 1, §1.3. The important point is that V(r) reduces to the unique series V corresponding to the derivation ∂_V on $\mathcal{B}$ when all $r_m \equiv 0$. The proof that this can be done will be delayed until the end of the section; it is a straightforward extension of proposition 1.3.7.

Given this, we can write

$$V(r) = V(r)_+ + V(r)_-$$
$$= V(r)^+ + V(r)^-$$

where $V(r)_+$ ($V(r)^+$) is the finite series of terms of non-negative degree in the grading of type $(s_0,\dots,s_\ell)$ (respectively, of type $(\sigma_0,\dots,\sigma_\ell)$). Now define

$$\partial_V\mathcal{L} = [\ V(r)_+\ ,\ \mathcal{L}\ ] \qquad 3.5.11$$
$$\nabla_V\mathcal{L} = [\ V(r)^+\ ,\ \mathcal{L}\ ]$$

In each case the derivations are well-defined since in both cases the right-hand side is an element of $\mathfrak{p}_-\otimes\mathcal{A}$. This is evident from the identities

$$[\ V(r)_+\ ,\ \mathcal{L}\ ] = [\ \mathcal{L}\ ,\ V(r)_-\ ]$$
$$[\ V(r)^+\ ,\ \mathcal{L}\ ] = [\ \mathcal{L}\ ,\ V(r)^-\ ]$$

It was explained earlier, in §3.1, that the derivation ∇_V is needed because ∂_V is not equivariant with respect to the transformations which leave μ_1 fixed. So we cannot restrict ∂_V to $\mathcal{C}$ by "dividing out by the action of $\mathcal{M}_\mathcal{A}$". Unfortunately, neither can we demonstrate that ∇_V is "$\mathcal{M}_\mathcal{A}$-equivariant" on $\mathcal{A}$. In fact the gauge action of $\mathcal{M}_\mathcal{A}$ does not induce a group of transformations on $\mathcal{A}$. An example clarifies this point.

Recall the original Miura-Drinfel'd-Sokolov transformation from §3.1

$$\mathcal{L} \longmapsto M(\mathcal{L})$$

$$\partial_x + \begin{bmatrix} q & -1 \\ r-z & -q \end{bmatrix} \longmapsto \partial_x + \begin{bmatrix} 0 & -1 \\ \mu-z & 0 \end{bmatrix} \quad \mu = q_x - q^2 + r$$

It is a quick computation to show that conjugating $\mathcal{L}$ by

$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \qquad c \in \mathbb{C}\{q,r\} \tag{3.5.12}$$

induces the transformation

$$\begin{aligned} q &\longmapsto q + c \\ r &\longmapsto -c_x + 2cq + c^2 + r \end{aligned} \tag{3.5.13}$$

on $\mathbb{C}\{q,r\}$. However, unless c is independent of both q and r a second application of this transformation does not give the same result as replacing c by $2c$ (which is the result of squaring 3.5.13).

To rectify this we must define c to be another indeterminate which we adjoin to $\mathbb{C}\{q,r\}$. We can show that 3.5.13 does define a (one parameter) group of automorphisms of $\mathbb{C}\{q,r,c\}$ and that ∇_V (which is extended to $\mathbb{C}\{q,r,c\}$ by defining $\nabla_V c \equiv 0$) is equivariant with respect to the action of this group. I claim that $\mathbb{C}\{\mu,c\}$ is the subalgebra fixed by this group, therefore ∇_V preserves it. After we set $c\equiv 0$ it follows that ∇_V maps $\mathbb{C}\{\mu\}$ into itself.

By following this principle it will be shown that, in general whenever proposition 3.5.3 applies, ∇_V preserves $\mathscr{C}$. It remains then to show that $(\nabla_V - \partial_V)$ is identically zero on $\mathscr{C}$. This is done in much the same way as described earlier, in §3.1. Finally we will see that setting all $r_m \equiv 0$ gives us the result we desire.

Thus the crux of the proof of 3.5.3 lies in describing a group of automorphisms whose invariants are the "coordinates" μ_i of $M(\mathcal{L})$.

Recall that the set $\{\zeta_m\}$ is a basis for m_-. We let $c = \sum c_m \zeta_m$ where $\{c_m\}$ is a set of indeterminates which we adjoin to $\mathcal{A}$ to give the differential algebra $\mathcal{A}\{c_m\} = \mathbb{C}\{q_i, r_m, c_m\}$. We define $X(s) = \exp(ad\, sc)$, for a complex parameter s, and equate this with a one parameter group of gauge transformations

$$X(s) : \partial_X + p_-\otimes\mathcal{A}\{c_m\} - \lambda \longrightarrow \partial_X + p_-\otimes\mathcal{A}\{c_m\} - \lambda \qquad 3.5.14$$
$$\mathcal{L} \longmapsto \exp(s\,adc)[\mathcal{L}]$$

In coordinates

$$X(s)\circ\mathcal{L} = \partial_X + \sum q_i(s)\varepsilon_i + \sum r_m(s)\zeta_m - \lambda$$

for some $q_i(s)$, $r_m(s) \in \mathcal{A}\{c_m\}$ depending upon s.

Corresponding to this we define a one parameter family of automorphisms of $\mathcal{A}\{c_m\}$:

$$\varphi(s) : \mathcal{A}\{c_m\} \longrightarrow \mathcal{A}\{c_m\}$$
$$\begin{bmatrix} q_i \\ r_m \\ c_m \end{bmatrix} \mapsto \begin{bmatrix} q_i(s) \\ r_m(s) \\ c_m \end{bmatrix}$$

In fact this family forms a group under composition of maps.

3.5.15. <u>Lemma.</u> $\varphi(t)\circ\varphi(s) = \varphi(s+t)$ *where* s,t $\in \mathbb{C}$.

<u>Proof.</u> We prove that this holds for each generator q_i, r_m of $\mathcal{A}$; trivially it is true for each c_m.

Let $q_i(s) = Q_i(q_i, r_m, sc_m)$ describe the differential polynomial given by $\varphi(s)q_i$. It is the ε_i-coordinate of $X(s)\mathcal{L}$. It follows that $Q_i(q_i(t), r_m(t), sc_m)$ is the ε_i-coordinate of $X(s)\circ X(t)\mathcal{L}$, since X(t) merely replaces q_i, r_m by $q_i(t)$, $r_m(t)$. Clearly $X(s)\circ X(t)\mathcal{L} = X(s+t)\mathcal{L}$. Thus

$$\varphi(s+t)q_i = Q_i(q_i(t), r_m(t), sc_m)$$

However

$$\varphi(t)q_i(s) = Q_i(q_i(t), r_m(t), sc_m)$$

by definition, therefore $\varphi(s+t)q_i = \varphi(t)\circ\varphi(s)q_i$. The same argument applies to each r_m. ∎

3.5.16. <u>Lemma.</u> *The subalgebra* $\mathcal{C}\{c_m\}$ *of* $\mathcal{A}\{c_m\}$ *is precisely the subalgebra of invariants of the one-parameter group* $S = \{ \varphi(s) \mid s\in\mathbb{C} \}$.

<u>Proof.</u> If we set $\mathcal{E} = \mathcal{A}\{c_m\}$ in lemma 3.5.8 we see that

$$M_{\mathcal{A}}(\mathcal{L}) = M_{\mathcal{E}}(\mathcal{L}) = M_{\mathcal{E}}(X(s)\mathcal{L})$$

since $c \in \mathfrak{m}_{-}\otimes\mathcal{A}\{c_m\}$. Therefore each μ_i is S-invariant and, of course, so is each c_m. Thus every element of $\mathcal{C}\{c_m\}$ is S-invariant.

To show that these are the only S-invariants we look at the identity

$$\mathcal{L} = \exp(-ady)M_{\mathcal{A}}(\mathcal{L})$$

where $y = \sum y_m\zeta_m$ is given by lemma 3.5.8. This tells us that the set $\{q_i, r_m\}$ belongs to the differential algebra $\mathbb{C}\{\mu_i, y_m, c_m\}$, therefore this is equal to $\mathcal{A}\{c_m\}$. However, the action of $\varphi(s)$ on y_m is characterized by

$$X(s)\mathcal{L} = \exp(s\,adc)\circ\exp(-ady)M_{\mathcal{A}}(\mathcal{L})$$

$$= \exp(ad\ y(s))M_{\mathcal{A}}(\mathcal{L})$$

for some $y(s) = \sum y_m(s)\zeta_m$ in $\mathfrak{m}_{-}\otimes\mathcal{A}\{c_m\}$ (the existence of $y(s)$ is a corollary to lemma 3.5.8 since $X(s)\mathcal{L}$ belongs to the class $\partial_x + \mathfrak{p}_{-}\otimes\mathcal{A}\{c_m\} - \lambda$).

Now suppose $f(y_m)$ is a S-invariant differential polynomial, then $f(y_m(s))$ belongs to $\mathbb{C}\{y_m\}$. In particular it must be invariant under a specialization of $\varphi(s)$ where we set $c_m = y_m$ for all m. In this case

$$\exp(ad\ y(s)) = \exp((s-1)ad\ y)$$

Therefore $f(y_m) = f((s-1)y_m)$ for all s. But f cannot be invariant under scaling of all coordinates unless it is a constant. Therefore only $\mathcal{C}\{c_m\}$ contains S-invariants. ∎

We extend ∇_V to $\mathcal{A}\{c_m\}$ by defining $\nabla_V c_m \equiv 0$ for all c_m. Then ∇_V is S-equivariant.

3.5.17. <u>Lemma.</u> $\varphi(s)^{-1}\circ\nabla_V\circ\varphi(s) = \nabla_V$.

<u>Proof.</u> It suffices to show this for the generators of $\mathcal{A}$; it is trivially true for each c_m.

We want to show that

$$\nabla_V[X(s)\mathcal{L}] = \nabla_V\mathcal{L}$$

By definition

$$\nabla_V[X(s)\mathcal{L}] = \exp(s\,adc)\ [V(r)^+,\mathcal{L}]$$
$$= [\ \exp(s\,adc)V(r)^+\ ,\ \exp(s\,adc)\mathcal{L}\ ]$$

Notice that c belongs to $L(\mathfrak{g},\theta)^0\otimes\mathcal{A}\{c_m\}$, therefore

$$\exp(s\,adc)V(r)^+ = [\exp(s\,adc)V(r)]^+$$

However

$$[\ \exp(s\,adc)V(r)\ ,\ X(s)\mathcal{L}\ ] = 0$$

We will prove later that exp(s*adc*)V(r) is the unique series commuting with $X(s)\mathcal{L}$ corresponding to $v \in c(\mathfrak{z}(\lambda))$, hence it must be the series obtained from V(r) by replacing q_i, r_m with $q_i(s)$, $r_m(s)$. Therefore, with a slight abuse of the notation

$$\nabla_V[X(s)\mathcal{L}] = \varphi(s)[\ V(r)^+,\ \mathcal{L}\ ]$$
$$= \varphi(s)(\nabla_V\mathcal{L})$$

More correctly

$$\varphi(s)^{-1}\nabla_V\varphi(s)\circ\mathcal{L} = \nabla_V\mathcal{L} \quad ■$$

3.5.18. <u>Corollary.</u> ∇_V *maps* $\mathscr{C}$ *into itself.*

<u>Proof.</u> Let $f \in \mathscr{C} \subset \mathscr{C}\{c_m\}$; then f is S-invariant by lemma 3.5.16. Thus

$$\nabla_V f = \nabla_V(\varphi(s)f) = \varphi(s)\circ\nabla_V f$$

by the previous lemma. So $\nabla_V f$ is S-invariant and must lie in $\mathscr{C}\{c_m\}$. However, the definition of ∇_V is independent of the indeterminates c_m, therefore $\nabla_V f$ belongs to $\mathscr{C}$. ■

3.5.19. <u>Lemma.</u> *The derivation* $\nabla_V-\partial_V$ *is identically zero on the algebra* $\mathscr{C}$.

<u>Proof.</u> Compare the derivation

$$(\nabla_V-\partial_V)\mathcal{L} = [\ V(r)^+-V(r)_+\ ,\ \mathcal{L}\ ]$$

with the derivation ∂_S from $\mathcal{A}$ into $\mathcal{A}\{c_m\}$ defined by

$$\partial_S\mathcal{L} = [c,\mathcal{L}] = \frac{d}{ds}\,X(s)\mathcal{L}\ \Big|_{s=0} \qquad 3.5.19$$

Since the generators μ_i of $\mathscr{C}$ are S-invariants we find

$$\partial_S\mu_i = \frac{d}{ds}\,\varphi(s)\mu_i\ \Big|_{s=0} = 0$$

Consequently we have the following algebraic property of the differential polynomials μ_i: the derivation

$$\begin{aligned} \partial_s \,:\, \mathcal{A} &\longrightarrow \mathcal{A}\{c_m\} \\ q_i &\mapsto F_i(q_i, r_m, c_m) \\ r_m &\mapsto G_m(q_i, r_m, c_m) \end{aligned} \qquad 3.5.20$$

defined by 3.5.19 is such that

$$\partial_s \mu_i(q_i, r_m) \equiv 0 \quad \text{for all } \mu_i.$$

Now if we replace c_m in 3.5.20 by the ζ_m-coordinate of $V(r)^+ - V(r)_+ \in \mathfrak{m}_- \otimes \mathcal{A}$ we obtain the definition of $\nabla_V - \partial_V$ mapping $\mathcal{A}$ into itself. It follows that $(\nabla_V - \partial_V)\mu_i \equiv 0$ for all μ_i. ■

Finally we have proved the proposition 3.5.3, since it follows immediately from the lemma above that:

The derivation ∂_V maps $\mathscr{C}$ into itself and therefore maps $\mathbb{C}\{u_i\}$ into itself after setting all $r_m \equiv 0$.

In the next section a few examples will be produced to demonstrate the utility of the Miura-Drinfel'd-Sokolov transformation. Before ending this section we will return to the proof of the following result, which was postponed earlier.

3.5.21 <u>Proposition.</u> *For each $v \in c(\mathfrak{z}(\lambda))$ there exists a unique series $V(r)$, with coefficients $v_j(r)$ consisting of homogeneous differential polynomials, commuting with $\mathcal{L}$. When we set each $r_m \equiv 0$, $V(r)$ is the unique series given by proposition 1.3.7.*

As with proposition 1.3.7 we prove this using the dressing method.

<u>Lemma.</u> *There exists $\chi(r) \in L_- \equiv \prod_{j<0}(L(\mathfrak{g},\theta)_j \otimes \mathcal{A})$ such that*

$$\exp ad\chi(r)\,(\partial_x + q + r - \lambda) \in \partial_x + \mathfrak{Z}(\lambda)$$

where $\mathfrak{Z}(\lambda)$ is the centralizer of λ in L, where

$$L = \{\oplus_{j\geq 0} L(\mathfrak{g},\theta)_j \otimes \mathcal{A}\} \oplus L_-.$$

Proof. This is a straightforward extension of the proof of lemma 1.3.8. Since λ is semisimple $\chi(r)$ can be determined by its homogeneous components in $L(\mathfrak{g},\theta)_j \otimes \mathcal{A}$. These are successively given by the requirement that the homogeneous terms in the expansion of

$$\exp(ad\chi(r)).(\partial_x+q+r-\lambda)$$

lie in $\mathfrak{z}(\lambda)$. ∎

Corollary. *For each* $v \in c(\mathfrak{z}(\lambda)) \cap L(\mathfrak{g},\theta)_k$ *the series*

$$V(r) = \exp ad(-\chi(r)).v$$

$V(r)$ *is the unique homogeneous series with leading term* v *commuting with* $\mathcal{L}$.

Proof. Once again this is a straightforward extension of the proof of proposition 1.3.7. Certainly $V(r)$ commutes with $\mathcal{L}$, its uniqueness is verified by looking at the equations implicit in

$$[\ \partial_x + q + r - \lambda\ ,\ v + v_{k-1} + \dots\ ] = 0$$

where v_j are the components of V of degree j in the grading of type $(s_0,\dots,s_\ell)$. Looking at the homogeneous terms in the expansion of this equation we find

$$\partial_x v_j + [q,v_j] + \sum_{n>0} [r_{-n},v_{j+n}] = [\lambda,v_{j-1}]$$

where r_{-n} is the component of degree $-n$ in the grading above (we have defined $r \in \mathfrak{m}_- \otimes \mathcal{A}$ therefore it has no components of non-negative degree). This determines the element v_{j-1}, given v_j, v_{j+1}, etc, up to its component in $\mathfrak{z}(\lambda)$. This component is uniquely determined by the differential equation for v_{j-1}, since we require each coefficient to consist of homogeneous differential polynomials. ∎

§3.6 Some examples.

Example 1. An interesting example comes from looking at $\mathfrak{sl}(3,\mathbb{C})$ with the principal grading. We write the spectral operator as

$$\partial_x + \begin{bmatrix} q & -z & 0 \\ 0 & p-q & -z \\ -z & 0 & -p \end{bmatrix} \qquad 3.6.1$$

which, in terms of the canonical generators for the loop algebra, is

$$\partial_x + qh_1 + ph_2 - z(e_0 + e_1 + e_2)$$

There are, up to equivalence, two possible gradings more coarse than the principal grading on a_2. They are the standard grading (of type (1,0,0)) and the grading of type (1,0,1).

In the latter grading the spectral operator has the representation

$$\partial_x + \begin{bmatrix} q & -1 & 0 \\ 0 & p-q & -z \\ -z & 0 & -p \end{bmatrix} \qquad 3.6.2$$

Conjugating this by

$$\begin{bmatrix} 1 & 0 & 0 \\ -q & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad 3.6.3$$

yields

$$\partial_x + \begin{bmatrix} 0 & -1 & 0 \\ r & p & -z \\ -z & 0 & -p \end{bmatrix} \quad r = q_x - q^2 + qp \qquad 3.6.4$$

Therefore we expect zero curvature derivations from the hierarchy with spectral operator 3.6.1 to preserve the subalgebra $\mathbb{C}\{p,r\}$ of $\mathbb{C}\{q,p\}$. One such derivation is defined by the equations

$$\begin{aligned} 3q_t &= (q_x - q^2)_x - 2(p_x - p^2 + qp)_x \\ 3p_t &= -(p_x - p^2)_x + 2(q_x - q^2 + qp)_x \end{aligned} \qquad 3.6.5$$

On the the subalgebra $\mathbb{C}\{p,r\}$ this gives

$$3p_t = -p_{xx} + 2pp_x + 2r_x$$

$$3r_t = r_{xx} - 2(p_x - p^2)_{xx} - 2p(p_x - p^2)_x - 2(pr)_x - 4p_x r$$

In the standard grading the spectral operator 3.6.1 has the representation

$$\partial_x + \begin{bmatrix} q & -1 & 0 \\ 0 & p-q & -1 \\ -z & 0 & -p \end{bmatrix} \qquad 3.6.7$$

We already know from §3.3 that the Miura transformation in this case can be obtained from the sequence of field extensions

$$\mathbb{C}\langle u, v\rangle \subset \mathbb{C}\langle q, p\rangle \subset \mathbb{C}\langle \psi_0, \psi_1, \psi_2\rangle \qquad 3.6.8$$

where ψ_0, ψ_1, ψ_2 span the kernel of the operator

$$\partial_x^3 + v\partial_x + u \qquad 3.6.9$$

and are chosen to be compatible with its factorization into

$$(\partial_x - p)(\partial_x + p - q)(\partial_x + q) \qquad 3.6.10$$

The variables u, v are

$$\begin{aligned} u &= (q_x - q^2)_x + q(p_x - p^2 + qp) \\ v &= (p_x - p^2) + (q_x - q^2 + qp) \end{aligned} \qquad 3.6.11$$

The zero curvature equation 3.6.5 on $\mathbb{C}\langle q, p\rangle$ induces the equations

$$\begin{aligned} u_t &= u_{xx} - \tfrac{2}{3}v_{xxx} - \tfrac{2}{3}vv_x \\ v_t &= v_{xx} + 2u_x \end{aligned} \qquad 3.6.12$$

on $\mathbb{C}\langle u, v\rangle$. These equations also have a Lax pair representation, as expected

$$\begin{gathered} L_t = [P, L] \\ L = \partial_x^3 + v\partial_x + u \quad , \quad P = \partial_x^2 + \tfrac{2}{3}v \end{gathered} \qquad 3.6.13$$

Remark. The system 3.6.12 contains the Boussinesq equation (see e.g. Fordy & Gibbons (1981)). The corresponding zero curvature equation 3.6.5 has been dubbed the "modified Boussinesq" equation by Fordy & Gibbons. The change of variables

$$Q = \tfrac{1}{2}(q + p) \quad , \quad S = \tfrac{1}{2}(q - p)$$

simplifies 3.6.5 to

$$\begin{aligned} Q_t &= S_{xx} - 2(QS)_x \\ 3S_t &= -Q_{xx} + (4S^2 - 2Q^2)_x \end{aligned}$$

An equivalent system (using different choices for the variables) was constructed by Sokolov & Shabat (1980) as part of their investigations into "modifying" Lax equations.

There is a remarkable connection between the field $\mathbb{C}\langle p,r\rangle$ and the sequence 3.6.8. We find that $\mathbb{C}\langle p,r\rangle$ is the subfield of $\mathbb{C}\langle\psi_0,\psi_1,\psi_2\rangle$ fixed by the paraboloic subgroup $\mathbf{P}_1$ of $SL(3,\mathbb{C})$ of matrices of the form

$$\begin{bmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{bmatrix} \in SL(3,\mathbb{C})$$

This subgroup fixes the flag of vector spaces

$$\langle\!\langle\psi_0,\psi_1\rangle\!\rangle \subset \langle\!\langle\psi_0,\psi_1,\psi_2\rangle\!\rangle$$

This flag can be identified with the partial factorization

$$(\partial_x - p)(\partial_x^2 + p\partial_x + q_x - q^2 + qp) \qquad 3.6.14$$

of the operator 3.6.9. Therefore the invariants of the action of $\mathbf{P}_1$ are generated by p and r. Consequently the equations 3.6.6 fit into the scheme

$$\mathcal{X} \longrightarrow \mathcal{X}/\mathbf{B}_+ \longrightarrow \mathcal{X}/\mathbf{P}_1 \longrightarrow \mathcal{X}/SL(3,\mathbb{C}) \qquad 3.6.15$$

where $\mathcal{X}$ is defined much the same as the similar space described in §3.4. Using the substitution

$$u = r_x - pr \quad , \quad v = p_x - p^2 + r \qquad 3.6.16$$

the equations 3.6.6 can be transformed into the equations 3.6.12.

It is by no means clear that the Miura- Drinfel'd-Sokolov transformation between the operators 3.6.2 and 3.6.4 should produce the invariants of $\mathbf{P}_1$. In the first place there is no factorization of the scalar operator $\partial_x^3+v\partial_x+u$ implicit in the representation 3.6.2.

In the language of §3.4 the two choices of gradings of type (1,0,1) and (1,0,0) correspond to two different spaces $\mathcal{X}'$ and $\mathcal{X}$ with projections

$$\mathcal{X}' \longrightarrow \mathcal{X}'/\mathbf{P}_+ \longrightarrow \mathcal{X}'/\mathbf{A}$$

$$\mathcal{X} \longrightarrow \mathcal{X}/\mathbf{B}_+ \longrightarrow \mathcal{X}/SL(3,\mathbb{C})$$

Here $\mathbf{A} \cong SL(2,\mathbb{C})\otimes\mathbb{C}^*$, represented as the group of matrices of the form

$$\begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{bmatrix} \in SL(3,\mathbb{C})$$

By construction $\mathcal{X}'/P_+$ and $\mathcal{X}/B_+$ must both have the coordinates q and p. What is striking is that both $\mathcal{X}'/A$ and $\mathcal{X}/P_1$ have the same coordinates p and r. I will not pursue this further in this thesis, aside from pointing out that

$$\partial_x + \begin{bmatrix} 0 & -1 & 0 \\ r & p & -1 \\ -z & 0 & -p \end{bmatrix}$$

is the unique operator of the type

$$\partial_x + \begin{bmatrix} 0 & -1 & 0 \\ * & * & -1 \\ -z & 0 & * \end{bmatrix}$$

in the orbit of the spectral operator 3.6.7 under the gauge action of the group of matrices of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{where } a \in \mathbb{C}\langle q,p\rangle$$

<u>Example</u> <u>2.</u> Recall from §1.5 of chapter 1 that we derived the system of equations

$$\begin{aligned} q_t &= -3(pr)_x \\ p_t &= 3(p_{xx} - pq_x - pq^2 - p^2r) \\ r_t &= -3(r_{xx} + rq_x - rq^2 - r^2p) \end{aligned} \qquad 3.6.17$$

If we make the substitution

$$\begin{aligned} u &= q^2 - q_x + pr \\ v &= pq - p_x \\ w &= r \end{aligned} \qquad 3.6.18$$

then the equations 3.6.17 transform into

$$\begin{aligned} u_t &= -6(vw)_x \\ v_t &= 3(v_{xx} - uv) \\ w_t &= -3(w_{xx} - uw) \end{aligned} \qquad 3.6.19$$

This substitution was obtained by the following Miura-Drinfel'd-Sokolov transformation:

$$\partial_x + \begin{pmatrix} q & p & -1 \\ r & 0 & 0 \\ -z & 0 & -q \end{pmatrix} \longmapsto \partial_x + \begin{pmatrix} 0 & 0 & -1 \\ w & 0 & 0 \\ u-z & v & 0 \end{pmatrix}$$

This is possible because the matrix

$$\lambda = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -z & 0 & 0 \end{pmatrix}$$

satisfies properties *(i)* and *(ii)* of proposition 3.5.3. In this case $\mathfrak{m}_-$ is the set of matrices of the form

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ * & * & 0 \end{pmatrix} \in \mathfrak{sl}(3,\mathbb{C})$$

and $\mathfrak{p}_-$ is the set of matrices of the form

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ * & * & * \end{pmatrix} \in \mathfrak{sl}(3,\mathbb{C})$$

In this example the grading on $L(\mathfrak{a}_2,\theta)$ is the grading of type (1,0,1) and the coarser grading is the standard grading (of type (1,0,0)).

It is a surprising fact (brought to my attention by Drs. J. D. Gibbon and J. Gibbons) that if we replace x by ix, t by $-it$ and set

$$A = w = \bar{v} \quad , \quad B = u$$

(where the bar denotes the complex conjugate) then we obtain from 3.6.19 a set of equations used to model Langmuir waves

$$iA_t = 3(A_{xx} - AB)$$
$$B_t = -6(|A|^2)_x$$

for which Yajima & Oikawa (1978) discovered a curious inverse scattering problem.

<u>Example 3.</u> For our last example, let us look at an example where the conditions of proposition 3.5.3 do not hold. Recall from §2.2 in chapter 2 that three spectral operators were given, corresponding to the three distinct (specialized) hierarchies admitted by the grading of type (1,0,1) on $c_2^{(1)}$. We will look at the prospects of finding a Miura-Drinfel'd-Sokolov transformation when we fix the coarser grading to be the standard grading (of type (1,0,0)). In this case the subalgebra $L(\mathfrak{g},\theta)^0$ has the

triangular decomposition

$$L(\mathfrak{g},\theta)^0 = \mathfrak{m}_- + L(\mathfrak{g},\theta)_0 + \mathfrak{m}_+$$

where

$$\mathfrak{m}_- = \langle\!\langle E_{42},\ E_{41}+E_{32},\ E_{31}\rangle\!\rangle$$

$$\mathfrak{g}_0 = \langle\!\langle H_1,\ H_2,\ E_{12}-E_{43},\ E_{21}-E_{34}\rangle\!\rangle$$

$$\mathfrak{m}_+ = \langle\!\langle E_{24},\ E_{14}+E_{23},\ E_{13}\rangle\!\rangle$$

and E_{ij}, H_i have been defined in §2.2. This is more clearly represented as

$$\begin{bmatrix} 0 & 0 & + & + \\ 0 & 0 & + & + \\ - & - & 0 & 0 \\ - & - & 0 & 0 \end{bmatrix}$$

since $L(\mathfrak{g},\theta)^0 \cong c_2$.

Look at the operator

$$\partial_x + \begin{bmatrix} 0 & p & 0 & 0 \\ r & q & 0 & -z \\ 0 & 0 & 0 & -r \\ 0 & -z & -p & -q \end{bmatrix}$$

which corresponds to case *(iii)* in example 2.2.4. In the standard grading its semisimple element λ splits into

$$\lambda_0 + \lambda_1 = E_{24} + zE_{42}$$

A Miura-Drinfel'd-Sokolov transformation cannot be applied in this case since $[E_{24},E_{31}] = 0$ therefore λ_0 is not injective on $\mathfrak{m}_-$.

However, one can check that the other two cases, with the respective spectral operators

$$\partial_x + \begin{bmatrix} \begin{matrix} 2q & 3p+r \\ 3p-r & s \end{matrix} & \underline{0} \\ \underline{0} & \begin{matrix} -2q & r-3p \\ -3p-r & -s \end{matrix} \end{bmatrix} - z\begin{bmatrix} \underline{0} & \begin{matrix} 2 & 0 \\ 0 & 1 \end{matrix} \\ \begin{matrix} 2 & 0 \\ 0 & 1 \end{matrix} & \underline{0} \end{bmatrix}$$

for case *(i)*, and

$$\partial_x + \begin{bmatrix} \begin{matrix} q & p \\ r & q \end{matrix} & \underline{0} \\ \underline{0} & \begin{matrix} -q & -r \\ -p & -q \end{matrix} \end{bmatrix} - z\begin{bmatrix} \underline{0} & \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \\ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \underline{0} \end{bmatrix}$$

for case *(ii)*, do admit Miura - Drinfel'd - Sokolov transformations. These transformations predict the substitution

$$u = 2r - 3p$$
$$v = q_x + 2q^2 + r^2 - 6pr + 9p^2$$
$$w = 3p_x - r_x + 3pq + rq + 3ps - rs$$
$$y = s_x + 9p^2 - r^2 + s^2$$

for the hierarchy in case *(i)*, and the substituion

$$\beta = r_x - 2rq$$
$$\gamma = q_x - q^2 - pr$$
$$\delta = p_x - 2pq$$

for the equations in the hierarchy given by case *(ii)*. I leave it to the interested reader to compute the (new ?) "integrable" systems of p.d.e's obtained by applying these substitutions to the systems given in example 2.2.4.

Conclusion.

It should be clear to readers of Drinfel'd & Sokolov (1985) that this thesis has made substantial use of their ideas, which prove to be quite robust in that the fundamental principles lend themselves easily to generalization.

No significant effort was needed to prove the following result:

to each periodically graded semisimple Lie algebra $(\mathfrak{g}, \theta)$ *admitting a semisimple element in* $\mathfrak{g}_1$ *there corresponds a hierarchy of integrable equations; the equations are indexed by the abelian subalgebra* $\mathfrak{c}(\mathfrak{z}(z\Lambda))$ *of* $L(\mathfrak{g}, \theta)$.

It seems quite reasonable to refer to these equations as "integrable". In every case the equation possesses an inverse scattering problem which can, in principle, be solved for certain classes of potential. In particular, soliton solutions can always be constructed.

Naturally we wish to know when $(\mathfrak{g}, \theta)$ admits a hierarchy of equations. I would like to tender the following conjecture which is wholly consistent with the results obtained in chapter 2 (and other results I have neglected to include).

Conjecture: *A periodically graded semisimple Lie algebra* $(\mathfrak{g}, \theta)$ *possesses a non-trivial Cartan subspace (i.e. has a non-zero semisimple element in* $\mathfrak{g}_1$*) if and only if the collection of affine roots of degree one for* $L(\mathfrak{g}, \theta)$ *contains a collection of affine type.*

From a practical point of view it is unlikely that anyone will feel the need to compute zero curvature equations when the rank of $\mathfrak{g}$ is large. One must bear in mind that the number of variables is at least equal to the rank and at most equal to the dimension of $\mathfrak{g}$ (although in the latter case at least rank($\mathfrak{g}$) many variables will be stationary with respect to all flows since the dimension of $\mathfrak{z}(\Lambda)$ is at least equal to the rank of $\mathfrak{g}$)

Perhaps the most interesting result in this thesis is the description of the Miura-Drinfel'd-Sokolov (M.D.S.) transformation. One point which has not been investigated is the extent to which it can be used, that is to say, for which $(\mathfrak{g},\theta)$ does there exist an element Λ satisfying the conditions under which the transformation holds? A similar question arises if we try to find a concrete description of the scheme of Wilson's, that the M.D.S transformation is only part of a series of transformations obtained by dividing out by the action of a Lie group. Conceptually we have viewed this scheme as the sequence

$$\mathcal{X} \longrightarrow \mathcal{X}/\mathbf{P} \longrightarrow \mathcal{X}/\mathbf{A}$$

The interesting problem is to make sense of this scheme. Essentially it is a problem of "coordinates", that is to say, determining whether or not each space in this sequence corresponds to a freely generated differential field. Moreover, the space $\mathcal{X}/\mathbf{P}$ should provide the coordinates for the zero curvature equations. The M.D.S. transformation should describe the coordinates on $\mathcal{X}/\mathbf{A}$. So the problem is to find a sensible description of the "space" $\mathcal{X}$ on which a Lie group $\mathbf{A}$ acts (on the right) which fits in with the previous two conditions.

References.

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

Beals, R. & Coifman, R. R. 1984 Scattering and inverse scattering for first order systems. *Comm. Pure and Appl. Math.*, 37, *39 - 90.*

Birkes, D. 1971 Orbits of linear algebraic groups. *Ann. of Math.*, 93, *459 - 475.*

Bourbaki, N. 1968, Hermann: Paris. Éléments de mathématique, Fasc. 36, Groups et algébres de Lie.

Dadok, J. & Kac, V. 1985 Polar representations. *J. Algebra*, 92, *504 - 524.*

Drinfel'd, V. G. & Sokolov, V. V. 1981 Equations of Korteweg-de Vries type and simple Lie algebras. *Soviet Math. Dokl.*, 23, *457 - 462.*

Drinfel'd, V. G. & Sokolov, V. V. 1985 Lie algebras and equations of Korteweg-de Vries type. *J. Soviet Math.*, 30, *1975 - 2036.*

Fordy, A. P. & Gibbons, J. 1981 Factorization of operators II. *J. Math. Phys.*, 22, *1170 - 1175.*

Gardner, C. S., Greene, J. M., Kruskal, M. D. & Miura, R. M. 1967 Method for solving the Korteweg-de Vries equation. *Phys. Rev. Lett.*, 19, *1095 - 1097.*

Helgason, S. 1978, Academic Press: New York. *Differential geometry, symmetric spaces and Lie groups.*

Humphreys, J. E. 1975, Springer-Verlag (Graduate texts in math.): Berlin. *Linear algebraic groups.*

Jacobson, N. 1962, Wiley (Interscience): New York. *Lie algebras.*

Kac (Kats), V. G. 1969 Automorphisms of finite order semisimple Lie algebras. *Funct. Anal. Appl.*, 3, *252 - 254.*

Kac, V. G. 1980 Some remarks on nilpotent orbits. *J. of Algebra*, 64, *190 - 213.*

Kac, V.G. 1985, Cambridge University Press. *Infinite dimensional Lie algebras.*

Kaplansky, I. 1957, Hermann: Paris. *Introduction to differential algebra.*

Kempf, G. & Ness, L. 1978 The length of vectors in representation space, in *Algebraic geometry*, Lecture Notes in Math., No. 732, Springer-Verlag: Copenhagen.

Kimura, T., Shin-ichi, K. & Osami, Y. 1986 A classification of the representations of reductive algebraic groups which admit only a finite number of orbits. *Am. J. Math.*, 108, *643 - 692.*

Kostant, B. 1959 The principal three dimensional subgroup and the Betti numbers of a complex simple Lie group. *Am. J. Math.*, 81, *973 - 1032.*

Kostant, B. 1963 Lie group representations on polynomial rings. *Am. J. Math.*, 85, *327 - 404.*

Kupershmidt, B.A. & Wilson, G. 1981 Modifying Lax equations and the second Hamiltonian structure. *Inventiones Math.*, 62, *403 - 436.*

Lax, P.D. 1968 Integrals of non-linear equations of evolution and solitary waves. *Comm. Pure and Appl. Math.*, 21, *467 - 490.*

Matsumura, H. 1962, Benjamin: New York. *Commutative algebra.*

Miura, R.M. 1968 Korteweg-de Vries equation and generalizations I. A remarkable explicit non-linear transformation. *J. Math. Phys.*, 9, *1202 - 1204.*

Sokolov, V.V. & Shabat, A.B. 1980 (L,A)-pairs and a Ricatti type substitution. *Funct. Anal. Appl.*, 14, *79 - 80.*

Vinberg, E.B. 1976 The Weyl group of a graded Lie algebra. *Math. USSR - Izv.*, 10, *463 - 495.*

Wadati, M. 1974 Bäcklund transformation for solutions of the modified Korteweg-de Vries equation. *J. Phys. Soc. of Japan*, 36, *1498.*

Wilson, G. 1979 Commuting flows and conservation laws for Lax equations. *Math. Proc. Camb. Phil. Soc.*, 86, *131 - 143.*

Wilson, G. 1981 The modified Lax and two-dimensional Toda lattice equations associated with simple Lie algebras. *Ergod. Th. & Dynam. Sys.*, 1, *361 - 380.*

Wilson, G., to be published. On the quasi-Hamiltonian formalism of the Korteweg-de Vries equation. *Phys. Lett. A.*

Yajima, N. & Oikawa, M. 1978 Formation and interaction of sonic-Langmuir solitons. *Prog. Th. Phys.*, 56, *1719 - 1739.*

Zakharov, V.E. & Shabat, A.B 1972 Exact theory of two-dimensional self focussing and one dimensional self-modulation of waves in non-linear media. *Soviet Phys. JETP*, 34, *62 - 69.*

Zakharov, V.E. & Shabat, A.B 1974 A scheme for integrating the non-linear equations of mathematical physics by the method of the inverse scattering problem I. *Funct. Anal. Appl.*, 8, *226 - 235.*

Zakharov, V.E. & Shabat, A.B 1979 Integration of the non-linear equations of mathematical physics by the method of inverse scattering II. *Funct. Anal. Appl.*, 13, *166 - 174.*