Journal of Geometry and Physics 8 (1992) 243–271
North-Holland

JOURNAL OF
GEOMETRY AND
PHYSICS

Twistor correspondences for the soliton hierarchies

L.J. Mason *,**
New College, Oxford OX1 3BN, UK

G.A.J. Sparling *
Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA

Received 7 June 1991

In this article we propose a new overview on the theory of integrable systems based on symmetry reduction of the anti-self-dual Yang–Mills equations and its twistor correspondence. First, the non-linear Schrödinger (NS) equations and the Korteweg de Vries (KdV) equations are shown to be symmetry reductions of the anti-self-dual Yang–Mills (ASDYM) equation with real forms of SL(2, $\mathbb{C}$) as gauge groups.

We obtain a twistor correspondence between solutions of the NS and KdV equations and certain holomorphic vector bundles with a symmetry on the total space of the complex line bundle of Chern class two on the Riemann sphere. Remarkably, when the Chern class is increased, the correspondence extends to the NS and KdV hierarchies. If the symmetry condition is dropped we obtain a twistor correspondence for a hierarchy for the Bogomolny equations, which yields the KdV and NS hierarchies when the symmetry is imposed.

The inverse scattering transform is shown to be a coordinate realization of the twistor correspondence. Both the pure solitons and the solitonless cases are treated. The k-soliton solutions arise from the kth "Ward ansatze" in an analogous fashion to the monopole solutions.

Keywords: *twistors, soliton hierarchies, Yang–Mills equations*
1991 MSC: 32L25, 81T13

0. Introduction

It is a special pleasure for us to have this opportunity to pay our respects to Roger Penrose on the occasion of his sixtieth birthday. Roger Penrose has been a constant source of inspiration to us and indeed has generated most of the ideas on which we work. We hope that this paper will serve as yet another example of how Roger Penrose's ideas have had unexpected (and we hope significant) ap-

* Supported in part by the National Science Foundation.
** SERC Advanced Fellow. Much of this work was completed while this author was the Esmée Fairbairn Junior Research Fellow at New College and Andrew Mellon Fellow at the University of Pittsburgh.

plication in fields far from those for which they were originally intended. Fundamental to this article are twistor constructions for the solution of non-linear partial differential equations. The prototype for these constructions is Roger Penrose's non-linear graviton construction [1].

This paper is part of a programme that aims to reduce the theory of integrable systems to the study of symmetry reductions of the self-dual Yang–Mills equations and its twistor correspondence. There are obstructions to this programme as it has not yet been possible to obtain the Kadomtsev–Petviashvili equations or the Davey–Stewartson equations in $2+1$ dimensions as symmetry reductions of the self-duality equations. However, if we restrict ourselves to a working programme of reformulating the theory of integrable systems in one or two dimensions there is, as yet, no serious obstruction *.

There are two parts to this programme:

(A) *Classification.* One would like to show that all integrable systems in one and two dimensions can be obtained from the self-dual Yang–Mills equations in four dimensions. More precisely we would like to see that all integrable systems in one and two dimensions arise from imposing symmetries on the self-dual Yang–Mills equations and then putting certain constants of integration and the residual gauge freedom into a normal form. For many systems in one and two dimensions, it is sufficient to consider reductions of the Bogomolny hierarchy, a system of equations derived in section 2 of this paper related to the reduction of the self-dual Yang–Mills equations in four dimensions by one non-null translation.

(B) *Unification of the theory.* One would like to see that the diverse and intricate techniques that have been brought to bear on integrable systems can be understood as specializations or extensions of the twistor constructions that play such a significant role in the theory of the self-duality equations.

In a series of articles Ward [2–4] assembled a collection of integrable systems that arise as symmetry reductions of the self-dual Yang–Mills equations. In two dimensions these include the sine-Gordon equations, the chiral model, the non-linear sigma model, the Ernst equations and the Toda field theory. Subsequently, in ref. [5] we showed that the Korteweg–de Vries (KdV) equations and the non-linear Schrödinger (NS) equations, perhaps the most basic integrable systems, are also reductions of the self-dual Yang–Mills equations. Nevertheless, much work remains to fulfill part A (above) of the programme. A major difficulty is the question of the definition of an integrable system, and one will probably have to be satisfied with merely showing that certain large classes (such as the classes discovered by AKNS, AKS and Drinfeld and Sokolov, for example) arise from

* Even here there are difficulties — the Landau–Lifschitz model's spectral parameter lies on an elliptic curve, whereas spectral parameters of reductions of the self-dual Yang–Mills equations generally lie on the Riemann sphere. This example is less problematic, however, as it nevertheless has a twistor construction with a minitwistor space that is the total space of a line bundle over an elliptic curve rather than a sphere as in this paper (see Carey, Mason and Singer, preprint).

the Bogomolny hierarchy — it will presumably always be possible to find a system that one might like to call integrable but that will fall outside the scope of any given scheme.

The twistor constructions are correspondences between solutions of the field equations and certain holomorphic vector bundles on complex manifolds, twistor spaces. These correspondences are geometric generalizations of the inverse scattering transform. Such constructions have played a central role in the theory of the self-dual Yang–Mills equations [6–9]. Reductions of the self-dual Yang–Mills equations have a reduced twistor correspondence (see ref. [10] for the reduced twistor correspondence for the Ernst equations). In this article we present the twistor construction for the Bogomolny hierarchy and its reductions to that for the non-linear Schrödinger and Korteweg–de Vries equations.

This paper is a long version of ref. [5] with a more detailed investigation of the various structures and their relation to twistor theory. We first show that the KdV and NS equations are reductions of the anti-self-dual Yang–Mills (ASDYM) equations on $\mathbb{R}^4$ with a metric of signature (2,2). The gauge group is taken to be one of the real forms of $SL(2,\mathbb{C})$ and the solutions are required to be symmetric under a timelike translation and a null translation that is orthogonal to the time direction. Modified KdV arises from a different gauge choice in the reduction. We then describe how the *standard* twistor correspondence for full anti-self-dual Yang–Mills fields can be reduced to give a correspondence for solutions of the NS and KdV equations. This is a correspondence between solutions of the equations and holomorphic vector bundles, satisfying certain symmetry and reality conditions, on an auxiliary complex manifold, $\mathcal{O}(2)$. The complex manifold, $\mathcal{O}(2)$, is the total space of a complex line bundle, the holomorphic tangent bundle, on the Riemann sphere $\mathbb{CP}^1$ [$\mathcal{O}(n)$ denotes the complex line bundle on $\mathbb{CP}^1$ with Chern class n].

One of the basic features of an integrable (Hamiltonian) system is the existence of a sequence of constants of the motion in involution with respect to the system's poisson bracket structure. These provide Hamiltonians whose flows commute and can be realized as an infinite sequence of non-linear evolution equations that are symmetries of the original equations. In the case of the NS and KdV equations this leads to the NS and KdV hierarchies. The reduced twistor correspondence for the NS and KdV equations extends naturally to solutions of the NS and the KdV hierarchies. Solutions of the first r equations of one of the hierarchies correspond to certain holomorphic vector bundles on $\mathcal{O}(1+r)$. In the case of the 1 (non-null) translation reduction of the self-dual Yang–Mills equations, this extension leads to the "Bogomolny hierarchy", which reduces to the NS and KdV hierarchies after the imposition of a (null) symmetry and a specialization of certain constants of integration. Solutions of the Bogomolny hierarchy correspond to certain holomorphic vector bundles over regions in $\mathcal{O}(n)$.

Finally we show that the inverse scattering transform in both the reflectionless and solitonless cases can be understood as a transform from solutions of the equations to patching data for the vector bundles on twistor space. We also see that the soliton solutions correspond to a class of vector bundles on twistor space that are limiting cases of those that arise in Hitchin's study of monopoles using twistor methods.

1. Korteweg–de Vries and non-linear Schrödinger equations as reductions of the self-dual Yang–Mills equations

We work on $\mathbb{R}^4$ with co-ordinates $x^a=(x,y,v,t)$ and metric $ds^2=dx^2-dy^2+2dv\cdot dt$ of signature $(2,2)$ and a totally skew orientation tensor $\epsilon_{abcd}=\epsilon_{[abcd]}$. We will consider a Yang–Mills connection $D_a=\partial_a-A_a$, where the A_a are, for the moment, elements of the Lie algebra of $SL(2,\mathbb{C})$. The A_a are defined up to gauge transformations, $A_a\to hA_ah^{-1}-(\partial_a h)h^{-1}$, where $h\equiv h(x^a)\in SL(2,\mathbb{C})$. The connection is said to be anti-self-dual when

$$\tfrac{1}{2}\epsilon_{ab}{}^{cd}[D_c,D_d]=-[D_a,D_b]\ .$$

This is equivalent to the following three commutator equations:

$$[D_x+D_y,D_v]=0, \tag{1.1a}$$

$$[D_x-D_y,D_x+D_y]+[D_v,D_t]=0, \tag{1.1b}$$

$$[D_x-D_y,D_t]=0\ . \tag{1.1c}$$

These follow from the integrability condition on the following linear system of equations:

$$L_0 s=\{D_x-D_y+\lambda D_v\}s=0,\qquad L_1 s=\{D_t+\lambda(D_x+D_y)\}s=0\ , \tag{1.2}$$

where λ is an affine complex co-ordinate on the Riemann sphere $\mathbb{CP}^1$ (the "spectral parameter") and s is a two-component column vector.

Let us put $D_x=\partial_x-A$, $D_v=\partial_v-B$, $D_t=\partial_t-C$ and $D_y=\partial_y-D$.

Now require that the bundle and its connection possess two commuting symmetries which project to a pair of orthogonal spacetime translations, one timelike and one null. In our coordinates these are along $\partial/\partial y$ and $\partial/\partial v$. We now restrict ourselves to gauges in which the components of the connection, (A,B,C,D), are independent of v and y. We also impose the gauge condition $A+D=0$. The gauge transformations are now restricted to $SL(2,\mathbb{C})$-valued functions of t alone under which A and B transform by conjugation, $B\to hBh^{-1}$ etc. With these assumptions the equations reduce to

$$\partial_x B=0,\quad [\partial_x-2A,\partial_t-C]=0,\quad 2\partial_x A-[B,C]=\partial_t B\ . \tag{1.3a,b,c}$$

These equations follow from the integrability conditions on the reduction of the linear system (1.2),

$$L_0 s = (\partial_x - 2A + \lambda B)s = 0, \qquad L_1 s = (\partial_t - C + \lambda \partial_x)s = 0 . \qquad (1.2'\text{a,b})$$

When eq. (1.3a) holds, B depends only on the variable t, so the gauge freedom may be used to reduce B to a normal form. The reduction of eqs. (1.3) are partially classified by the available normal forms.

When B vanishes, the equations are trivially soluble, with the result that the connection may be put in the form $A_a dx^a = A(t)\, d(x+y)$. Eq. (1.1) is then automatically satisfied. Henceforth we assume that B is everywhere non-vanishing. The matrix B then has just two normal forms:

$$(\alpha) \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad (\beta) \quad B = \kappa \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We assume that the type of B is constant. In the case of type β, as t varies, κ becomes a non-zero function of t. When B is in the Lie algebra of SU(2), SU(1, 1) or SL(2, ℝ), κ is non-zero and is either real or pure imaginary. Case α leads to the KdV equation, and case β leads to the non-linear Schrödinger equation.

An analysis of the reduction of the remaining equations leads to the following theorems (this analysis is contained in the appendices).

Theorem α. *The self-dual Yang–Mills equations are solved with B of type α by the ansätze:*

$$2A = \begin{pmatrix} q & 1 \\ q_x - q^2 & -q \end{pmatrix}, \qquad 2C = \begin{pmatrix} (q_x - q^2)_x & -2q_x \\ 2w & -(q_x - q^2)_x \end{pmatrix},$$

where $4w = q_{xxx} - 4qq_{xx} - 2q_x^2 + 4q^2 q_x$, *a subscript x or t denotes differentiation with respect to that variable, and provided that q satisfies*

$$4q_t = q_{xxx} - 6(q_x)^2 .$$

With the definition $u = -q_x = \frac{1}{2}\mathrm{Tr}(BC)$ *we obtain the Korteweg–de Vries equation*

$$4u_t = u_{xxx} + 12uu_x .$$

Conversely, every solution of the equations for type α, with Tr(*AB*) *everywhere non-zero, may be reduced to this form, at worst after suitable co-ordinate and gauge transformations. [When* Tr(*AB*) *is identically zero, the equations are explicitly soluble.]* □

Remarks.

(1) The modified KdV equations can be obtained by a different choice of gauge. We still require that B be in the above normal form, but instead of setting $A+D=0$, we require that $A+D$ is strictly lower triangular and that A be upper triangular.

This can be implemented with a gauge transformation of the form

$$G=\begin{pmatrix}1 & 0\\ g & 1\end{pmatrix}.$$

The equations can then be reduced in such a way that the basic dependent variable is the top left entry of A, which satisfies the modified KdV equation as a consequence of the self-dual Yang–Mills equations.

(2) The Drinfeld–Sokolov form of the linear system can be obtained by a different gauge transformation of the above form in which the function g is chosen so as to eliminate the diagonal components of A.

Theorem β. *The self-dual Yang–Mills equations are solved, with B of type β, by the ansatz:*

$$2A=\begin{pmatrix}0 & \psi\\ -\tilde{\psi} & 0\end{pmatrix},\qquad 2\kappa C=\begin{pmatrix}\psi\tilde{\psi} & \psi_x\\ \tilde{\psi}_x & -\psi\tilde{\psi}\end{pmatrix},$$

provided ψ *and* $\tilde{\psi}$ *satisfy*

$$2\kappa\psi_t=\psi_{xx}+2\psi^2\tilde{\psi},\qquad 2\kappa\tilde{\psi}_t=-\tilde{\psi}_{xx}-2\tilde{\psi}^2\psi\,, \tag{1.4}$$

and $2\kappa=1$ *or* $-\mathrm{i}$. *Conversely, every solution of the equations for type β may be reduced to this form, at worst after suitable gauge and co-ordinate transformations.* □

When A, B and C are in the Lie algebra of $\mathrm{SL}(2,\mathbb{R})$, all quantities may be taken to be real, $2\kappa=1$ and eq. (1.4) becomes the pair of coupled equations for real $p=\psi$ and $r=-\tilde{\psi}$,

$$p_t=p_{xx}-2p^2r,\qquad r_t=-r_{xx}+2r^2p\,.$$

When A, B and C are in the Lie algebra of SU(2), one has $2\kappa=-\mathrm{i}$, and $\tilde{\psi}$ is the complex conjugate of ψ. Equation (1.4) then becomes just the following:

$$\mathrm{i}\psi_t=-\psi_{xx}-2|\psi|^2\psi\,. \tag{1.5}$$

Equation (1.5) is the non-linear Schrödinger equation for the unknown ψ, with an attractive self-interaction.

For the SU(1, 1) case, one has $2\kappa=-\mathrm{i}$ and $\tilde{\psi}$ is the negative complex conjugate of ψ. Then eq. (1.4) becomes the following:

$$\mathrm{i}\psi_t=-\psi_{xx}+2|\psi|^2\psi\,. \tag{1.6}$$

Equation (1.6) is the non-linear Schrödinger equation for the unknown ψ, with a repulsive self-interaction.

The proof of these results in the forward direction follows from direct calculation. The converses require more work, which is presented in appendices A and B.

2. The twistor correspondences and the Bogomolny hierarchies

The information of the general local analytic solution to the anti-self-dual Yang–Mills equations may be encoded in the global analytic structure of certain holomorphic complex vector bundles over suitable domains in twistor space $\mathbb{PT}$, which is complex projective three-space $\mathbb{CP}^3$ [6]. Rather than describe the original correspondence, we shall derive a generalization of its reduction by a non-null symmetry. This will give the correspondence for the Bogomolny hierarchy.

In this section all spaces and structures will be taken to be holomorphic, thus ψ will be taken to be a holomorphic function of the complex variables x and t etc.

When a single non-null translational symmetry (i.e. along ∂_y as in section 1) is imposed on the anti-self-dual Yang–Mills system, the resulting equations are called the Bogomolny equations. Solutions of these equations will be seen to correspond to bundles invariant under the corresponding symmetry of the twistor space. The action of the symmetry on the relevant domain in $\mathbb{PT}$ has no fixed points so that invariant bundles are the pullback of bundles on the quotient of the domains by the symmetry. We shall denote the largest such quotient by $\mathcal{O}(2)$; this has become known as minitwistor space. It is a complex line bundle of Chern class two fibred over the Riemann sphere, $\mathbb{CP}^1$ (strictly speaking, it is an affine bundle — the choice of a zero section corresponds to a choice of origin in $\mathbb{C}^3$).

Solutions of the Korteweg–de Vries and non-linear Schrödinger equations are obtained when we impose the further null symmetry (along ∂_v in section 1) and so correspond to certain (rank-2) holomorphic vector bundles over $\mathcal{O}(2)$, on which we have the action of an additional symmetry, corresponding to this extra symmetry. It is not possible to divide out directly by the action of this symmetry for it completely fixes a fibre of $\mathcal{O}(2)$. Indeed, we shall see later that important features of the correspondence reside in the structure around this fibre.

In order to describe the hierarchies we shall take our twistor spaces to be the complex line bundles of Chern class n, $\mathcal{O}(n)$. We shall see that, for $\mathcal{O}(2)$, the $\mathbb{C}^3$ with coordinates (v, x, t) on which the (complexified) Bogomolny equations are defined is the space of global holomorphic sections of $\mathcal{O}(2)$ over the Riemann sphere, $\mathbb{CP}^1$. If $\mathcal{O}(2)$ is replaced by $\mathcal{O}(n)$, the affine bundle of Chern class $n \geqslant 2$ over the Riemann sphere, the corresponding space of such holomorphic sections, $\Gamma(\mathcal{O}(n))$, is $\mathbb{C}^{n+1}$. If one now studies holomorphic vector bundles over $\mathcal{O}(n)$, we will see that they encode the information of gauge equivalence classes of solutions to a system $B(n)$ of non-linear partial differential equations on $\mathbb{C}^{n+1}$ written out explicitly in eq. (2.5) below. The system $B(n)$ is included (non-canonically) in the system $B(n+1)$, so it is natural to let n go to infinity. This gives, by definition, $B(\infty)$, the Bogomolny hierarchy. We will now "derive" the Bogomolny hierarchy from holomorphic vector bundles on $\mathcal{O}(n)$.

2.1. THE TWISTOR CORRESPONDENCE FOR THE BOGOMOLNY HIERARCHY

Let $\mathcal{O}(n)$ denote the twistor space, the complex line bundle of Chern class $n \geq 1$, over the Riemann sphere, $\mathbb{CP}^1$. The Riemann sphere, $\mathbb{CP}^1$, can be represented by means of stereographic projection as $\mathbb{C} \cup \infty$ with coordinate λ on $\mathbb{C}$ and $\lambda' = 1/\lambda$ a coordinate on $\mathbb{C}' = \{\mathbb{C} \cup \infty\} - \{0\}$. The line bundles, $\mathcal{O}(n)$, can then be given coordinates (μ, λ) in the region fibred over $\mathbb{C}$ and by $(\mu', \lambda') = (\mu/\lambda^n, 1/\lambda)$ in the region fibred over $\mathbb{C}'$. The projection $p: \mathcal{O}(n) \to \mathbb{CP}^1$ is given by $(\mu, \lambda) \to \lambda$ in these coordinates.

Let $\Gamma(n)$ be the space of global holomorphic sections of $\mathcal{O}(n)$. It can be seen that $\Gamma(n) \equiv \mathbb{C}^{n+1}$ as follows. A point σ of $\Gamma(n)$, $\sigma: \mathbb{CP}^1 \to \mathcal{O}(n)$, can be represented by $\mu = \sigma(t_i, \lambda)$ is a polynomial of degree n in λ with coefficients t_i, $i = 0, \ldots, n$:

$$\sigma(\lambda) = \sum_{k=0}^{n} t_k \lambda^k . \tag{2.1}$$

On $\mathbb{C}'$, $\mu' = \sum_{k=0}^{n} t_k (\lambda')^{n-k}$, which is regular at $\lambda = \infty$, so these sections are indeed global. These give all the global sections (further powers of λ in (2.1) would give either poles at $\lambda = 0$ or at $\lambda = \infty$). The coefficients $(t_0, \ldots, t_n) = (v, x, t, t_3, \ldots, t_n)$ are linear co-ordinates for $\Gamma(n)$.

A point Z of $\mathcal{O}(n)$ with coordinates (μ, λ) corresponds in $\Gamma(n)$ to the twistor hyperplane Σ_Z of all holomorphic sections passing through Z: Σ_Z consists of those t_i satisfying (2.1) for fixed $(\sigma(\lambda), \lambda) = (\mu, \lambda)$.

Definition. For an open region, R, of $\Gamma(n)$, denote by $\mathcal{O}(n)_R$ the region in $\mathcal{O}(n)$ consisting of all Z such that Σ_Z intersects R. The domain R will be said to be suitable if each $\Sigma_Z \cap R$ is homotopically and analytically trivial and if the subset of $\Gamma(n)$, consisting of sections lying in $\mathcal{O}(n)_R$, coincides with R itself.

We now prove the following theorem:

Theorem 2.1. *For R suitable, let E be a vector bundle over $\mathcal{O}(n)_R$ with structure group* $\mathrm{SL}(m, \mathbb{C})$ *such that its restriction to any holomorphic section of $\mathcal{O}(n)_R$ over $\mathbb{CP}^1$ is trivial. Each such E determines and is determined by a gauge equivalence class of solutions of a system of non-linear differential equations on R, the nth level of the Bogomolny hierarchy, $B(n)$, as defined in eqs. (2.4) and (2.5) below.*

Remark. The assumption that E be trivial on a holomorphic section is satisfied generically in the sense that for a general choice of E, E will be trivial on every section, σ, corresponding to points in all but a complex codimension-one subset of R. This follows from the fact that holomorphic bundles on $\mathbb{CP}^1$ with zero Chern class are generically trivial. The fields satisfying $B(n)$ will have rational singularities on this codimension-one subset in a suitable gauge. One need only move the region R away from this subset in order to satisfy the requirements of the theorem.

Proof. We shall use the Čech description of the bundle: choose an open Stein cover, $\{U_\alpha\}$, of $\mathcal{O}(n)_R$, where the numerical index α ranges over the number of sets in the cover (for example, the cover $U_0=\mathcal{O}(n)_R\cap\{\lambda\neq\infty\}$ and $U_1=\mathcal{O}(n)\cap\{\lambda\neq 0\}$ is often adequate so that $\alpha=0, 1$).

By definition, a holomorphic rank-m vector bundle, E, when restricted to a Stein open set U_α, is holomorphically trivialized by a frame f_α of E over U_α,

$$f_\alpha : E|_{U_\alpha}\simeq U_\alpha\times\mathbb{C}^m .$$

On each overlap, $U_\alpha\cap U_\beta$, one must also have $f_\alpha=P_{\alpha\beta}f_\beta$, where $P_{\alpha\beta}(\mu,\lambda)$ is a $SL(m,\mathbb{C})$-valued holomorphic function on $U_\alpha\cap U_\beta$, often referred to as a *patching function*. The ensemble of all the patching functions, $P_{\alpha\beta}$, obey on any triple intersection, $U_\alpha\cap U_\beta\cap U_\gamma$, the compatibility (a non-linear version of the Čech cocycle) condition $P_{\alpha\beta}P_{\beta\gamma}=P_{\alpha\gamma}$ and conversely any such system of functions $P_{\alpha\beta}$ satisfying the compatibility condition, determines a bundle E.

If we were to choose different trivializations, $f'_\alpha=g_\alpha f_\alpha$, where g_α are $SL(m,\mathbb{C})$-valued functions on U_α, then $P_{\alpha\beta}$ would be replaced by $P'_{\alpha\beta}=g_\alpha P_{\alpha\beta}g_\beta^{-1}$. Thus $P_{\alpha\beta}$ and $P'_{\alpha\beta}$ determine the same bundle. A bundle E is trivial when $E\simeq\mathcal{O}(n)_R\times\mathbb{C}^m$, i.e., there exists a global frame, f. This implies that there exist g_α with $f_\alpha=g_\alpha f$ so that $P_{\alpha\beta}=g_\alpha g_\beta^{-1}$.

In order to construct the solutions of the Bogomolny hierarchy on $\mathbb{C}^{n+1}$, restrict the patching functions $P_{\alpha\beta}$ to any section $\mu=\sigma(t_i,\lambda)$, corresponding to a point t_i of R. By assumption, E is trivial on $\mu=\sigma$, so that we have $SL(n,\mathbb{C})$-valued functions $s_\alpha(t_i,\lambda)$ (depending on λ and t_i), satisfying the following relations on each $\{\mu=\sigma\}\cap U_\alpha\cap U_\beta$:

$$s_\alpha(t_i,\lambda)=P_{\alpha\beta}(\sigma(t_i,\lambda),\lambda)s_\beta(t_i,\lambda) . \tag{2.2}$$

Now the polynomials $\sigma(t_i,\lambda)$ are annihilated by the differential operators $V_k=\partial_k-\lambda\partial_{k-1}$, $1\leqslant k\leqslant n$, where ∂_k is the co-ordinate derivative $\partial/\partial t_k$. Thus the $P_{\alpha\beta}(\sigma(t_i,\lambda),\lambda)$ are annihilated by each operator V_k. Define the quantities $\gamma_k^\alpha(t_i,\lambda)$ by the formula

$$V_k s_\alpha(t_i,\lambda)=s_\alpha(t_i,\lambda)\gamma_k^\alpha(t_i,\lambda) . \tag{2.3}$$

Then applying V_k to eq. (2.2) one obtains immediately $\gamma_k^\alpha(t_i,\lambda)=\gamma_k^\beta(t_i,\lambda)$ on each overlap region $\sigma\cap U_\alpha\cap U_\beta$. Thus there is a function, $\gamma_k(t_i,\lambda)$, defined for all λ such that restricted to any $\sigma\cap U_\alpha$, $\gamma_k(t_i,\lambda,\mu)$ agrees with $\gamma_k^\alpha(t_i,\lambda,\mu)$. The quantity $\gamma_k(t_i,\lambda)$ is regular for all $\lambda\in\mathbb{C}$, and $(1/\lambda)\gamma_k$ is regular as $\lambda\to\infty$ (the only difficulty with γ_k as $\lambda \to \infty$ is the linear dependence of V_k on λ), so that by a minor extension of Liouville's theorem, it is *linear* in λ, so that one has $\gamma_k(t_i, \lambda) = A_k + \lambda B_{k-1}$, for some A_k and B_{k-1}, which depend only on t_i. Putting s equal to $s_\alpha(t_i, \lambda)$, for some α, eq. (2.3) shows that s is a solution to the following linear system:

$$L_k s \equiv (\Delta_k - \lambda D_{k-1})s = 0, \quad 1 \leqslant k \leqslant n\,, \tag{2.4}$$

where

$$\Delta_k = \partial_k - A_k, \qquad D_{k-1} = \partial_{k-1} - B_{k-1}, \quad 1 \leqslant k \leqslant n\,.$$

Note that in this section the matrices A_k, B_{k-1}, etc. are understood to be acting on the right. Equation (2.4) defines the linear system for the nth level of the Bogomolny hierarchy. The integrability conditions for this linear system define the field equations for $B(n)$, the nth level of the Bogomolny hierarchy. Thus $B(n)$ consists of the following commutator equations:

$$[\Delta_k, \Delta_j] = 0, \tag{2.5a}$$

$$[D_{j-1}, D_{k-1}] = 0\,, \tag{2.5b}$$

$$[\Delta_k, D_{j-1}] - [\Delta_j, D_{k-1}] = 0, \quad 1 \leqslant k \leqslant n\,. \tag{2.5c}$$

The equations are gauge invariant under the transformation $A_k \to h^{-1}A_k h - h^{-1}\partial_k h$ and $B_k \to h^{-1}B_k h - h^{-1}\partial_k h$. This corresponds to the freedom in the choice of s_α, $s_\alpha \to s'_\alpha = s_\alpha h$, with h an $\mathrm{SL}(m, \mathbb{C})$-valued function of the t_i alone (h must be global and holomorphic in λ and is thus independent of λ by Liouville's theorem). The differences $A_k - B_k$, $1 \leqslant k \leqslant n-1$, are canonical, transforming according to the adjoint representation.

From E, therefore, we have obtained a gauge equivalence class of solutions of the system $B(n)$. It can be seen that this solution is independent, modulo gauge transformations, of all choices made.

Conversely, given an analytic solution of $B(n)$ on R for R suitable, one may construct such a bundle E on $\mathcal{O}(n)_R$ as follows.

For any fixed twistor, $Z = (\mu, \lambda)$, the equation of Σ_Z is $\mu = \sigma(t_i, \lambda)$, where $\sigma(t_i, \lambda)$ is given in eq. (2.1). Since the vector fields V_k kill $\sigma(t_i, \lambda)$, they are tangent to any such twistor plane, for any fixed λ and μ; there are n such (independent) vector fields which span the tangent space of the hyperplane Σ_Z. Given a solution of eq. (2.5) on R, construct a vector bundle E on $\mathcal{O}(n)_R$ by defining the fibre at $Z \in \mathcal{O}(n)_R$, E_Z, to be

$$E_Z = \{\text{solutions of } L_k s = (V_k - A_k + \lambda B_{k-1})s = 0 \text{ on } \Sigma_Z \cap R\}\,;$$

the linear system, eq. (2.4), defines a linear connection on every $\Sigma_Z \cap R$, which has zero curvature, by virtue of eqs. (2.5). Since $\Sigma_Z \cap R$ is homotopically trivial, there is no holonomy, so eq. (2.4), considered globally on each $\Sigma_Z \cap R$, becomes the defining equation for the fibre, at each Z, of a holomorphic vector bundle, E, over Z_R. It is easily seen that E is trivial when restricted to any section of $\mathcal{O}(n)_R$ over $\mathbb{CP}^1$ and that this construction of E from the solution of eq. (2.5) is the inverse of the construction of the solution of eq. (2.5) from a given E, given above. □

The case n=1. $B(1)$ is vacuous, so the theorem amounts to the encoding of the information of an arbitrary analytic vector bundle with analytic connection in two dimensions, into the structure of a vector bundle over a domain in complex projective two-space.

The case n=2. $B(2)$ gives the standard Bogomolny system. Equations (2.5) for $B(2)$ are

$$[\Delta_1, \Delta_2]=0, \qquad [D_0, D_1]=0, \qquad [\Delta_2, D_0]-[\Delta_1, D_1]=0 .$$

With the co-ordinate redefinitions $t_0=v$, $t_1=x$, $t_2=t$, we can identify D_0 with D_v, Δ_2 with D_t, D_1 with D_x-D and Δ_1 with D_x+D. Then eqs. (2.5), for the case $n=2$, becomes eq. (1.1) in which the operator D_y is replaced by the Lie algebra element, $-D$. This corresponds to imposing the symmetry along ∂_y to (1.1) and requiring that A, B, C and D depend only on (v, x, t). This is, by definition, the standard Bogomolny system.

To reduce further to the systems corresponding to the Korteweg–de Vries and non-linear Schrödinger equations, we had to specialize to the case where the group is $SL(2,\mathbb{C})$, impose a symmetry in the v (or t_0) direction and then impose a reality structure to bring $SL(2,\mathbb{C})$ down to one of its real forms. Clearly, the bundle E must be of rank 2 to give rise to an $SL(2,\mathbb{C})$ solution. The symmetry and reality conditions are implemented on E as follows.

The symmetry condition on E. The action $(t_0, t_1, t_2, ..., t_n) \to (t_0+a, t_1, t_2, t_2, ..., t_n)$ induces the action $(\mu, \lambda) \to (\mu+a, \lambda)$ on $\mathcal{O}(2)$. Denote this action by K. This action K fixes the fibre of $\mathcal{O}(2)$ over $\lambda=\infty$ $(\lambda'=0)$. We are therefore unable to divide out by this symmetry action since $\mathcal{O}(2)/K$ is not a manifold. (In the co-ordinates (μ', λ') K acts by $(\mu', \lambda') \to (\mu'+a\lambda'^n, \lambda')$ so that the action fixes the fibre $\lambda'=0$ to nth order.)

A holomorphic vector bundle E corresponds to a symmetric solution of $B(n)$ if we can lift K to act as a holomorphic one-parameter symmetry group, $\tilde{K}$, of E. If this is the case $\tilde{K}$ induces an automorphism of each fibre of the vector bundle over the fixed points of K in $\mathcal{O}(2)$, $\lambda=\infty$. The Lie algebra generators of these automorphisms give rise to the canonical matrices B_0, which distinguish the non-linear Schrödinger and Korteweg–de Vries equations (as discussed in section 1).

Remark. Provided that B_0 is always non-zero, if we lift $\tilde{K}$ to the principal bundle $P(E)$ corresponding to E, it acts without fixed points, so that the quotient, $P(E)/\sim$ of $P(E)$ by the lifted action of $\tilde{K}$ is a well-defined four complex dimensional manifold, although it is no longer an $\mathrm{SL}(2,\mathbb{C})$ bundle. The space $P(E)/\sim$ fibres over $\mathbb{CP}^1$ with fibre $\mathrm{SL}(2,\mathbb{C})$ over $\lambda\neq\infty$, but with fibre $\mathrm{SL}(2,\mathbb{C})/\{\exp a B_0\}$ over $\lambda=\infty$ where $\{\exp a B_0\}$ denotes the subgroup of $\mathrm{SL}(2,\mathbb{C})$ generated by B_0. $\mathrm{SL}(2,\mathbb{C})/\{\exp a B_0\}$ is the complex line bundle of Chern class minus one [the Hopf bundle or $\mathcal{O}(-1)$] over the Riemann sphere, if B_0 is of type α, or the two-dimensional complex manifold consisting of ordered pairs of non-coincident points on the Riemann sphere, if B_0 is of type β (equivalently this latter fibre is the affine line bundle of Chern class minus one over the Riemann sphere). Away from the manifold $\lambda=\infty$, $P(E)/\sim$ is simply a trivial vector bundle over the complex plane. Thus the entire structure of the solution is built into the complex manifold $(P(E)/\sim)|_{\lambda=\infty}$ and into its glueing into this trivial bundle. Similar remarks apply to the case of reducing $\mathrm{SL}(m,\mathbb{C})$ by a symmetry in the t_0 direction. This discussion will be pursued further, elsewhere.

The reality condition on E. $(t_0, t_1, t_2, ..., t_n)\to(\bar{t}_0, \bar{t}_1, \bar{t}_2, ..., \bar{t}_n)$, complex conjugation, maps Σ_Z to $\Sigma_{\bar{Z}}$, where in the co-ordinates above, if $Z=(\mu,\lambda)$, then $\bar{Z}=(\bar{\mu},\bar{\lambda})$. If the gauge group is unitary for real t_i, then the complex conjugate of a solution of the linear system (2.4) on Σ_Z is a solution to the linear system associated to E^* on $\Sigma_{\bar{Z}}$. The complex conjugation therefore lifts to an anti-holomorphic map from E to E^*. Similarly for gauge group $\mathrm{SL}(2,\mathbb{R})$ the complex conjugation lifts to an anti-holomorphic map from E to E.

3. Symmetry reduction to the soliton hierarchies

The soliton hierarchies are a sequence of non-linear partial differential equations representing the (commuting) Hamiltonian flows of the infinite sequence of conserved quantities on the infinite-dimensional Poisson manifold of rapidly decreasing initial data. The first such hierarchy, the Korteweg–de Vries hierarchy, was discovered by Gardner, Green, Kruskal and Miura [11]. Its equations have the form $\partial_k u=P_k(u)$, $k\geqslant 1$, where the real variable u is a function of co-ordinates t_j, $j\geqslant 1$, $\partial_k=\partial/\partial t_k$ and $P_k(u)$ is a certain polynomial in the derivatives of u with respect to the variable t_1 (where t_1 is often referred to as the "x" coordinate) up to order $2k-1$. When $k=2$, one has $P_2(u)=\partial_1^3 u+6u\,\partial_1 u$, so the corresponding evolution is that of the Korteweg–de Vries equation. The non-linear Schrödinger hierarchy was found next [12]. Its equations read $\partial_k\psi=Q_k(\psi)$, where ψ is now complex and $Q_k(\psi)$ is a polynomial in the derivatives of ψ with respect to the variable t_1 up to order k. When $k=2$, one has $\psi_2=\mathrm{i}\partial_1^2\psi+2\mathrm{i}\epsilon|\psi|^2\psi$ $(\epsilon=\pm1)$ giving the evolution of the non-linear Schrödinger equation for the cases of the attractive $(\epsilon=1)$ and repulsive $(\epsilon=-1)$ potentials. These hierarchies are usually analyzed under the assumption that all quantities rapidly decrease as $|t_1|\to\infty$. They will then be referred to as the standard hierarchies.

It is possible to give a simple expression for the equations of these hierarchies. Let g_k, for $k\geqslant 0$, be a collection of functions of real variables t_j, $j\geqslant 1$, with values in the Lie algebra g of a Lie group G. Then define the hierarchy $g(\infty)$ to be the system of partial differential equations

$$\partial_j g_k = \sum_{m=0}^{j} [g_m, g_{j+k-m}], \quad j\geqslant 1, k\geqslant 0 ,$$

where [,] is the Lie algebra commutator. These equations are invariant under the adjoint action of G on g, applied simultaneously to all the g_k. When these equations hold one finds that quantities h_k are constant, where the h_k are defined by the formula

$$2h_k = \sum_{r=0}^{k} h(g_m, g_{k-m}), \quad k\geqslant 0 .$$

Here $h(\ ,\ )$ is any symmetric bilinear form on g, invariant under the adjoint action of G on g. When G is semi-simple, we can take h to be the Killing form. We have the following theorem.

Theorem 3.1. *Let G be a real form of* $\mathrm{SL}(2,\mathbb{C})$ *and h the Killing form.*

If $h_0<0$ *and* $h_k=0$ $\forall k\geqslant 1$, *then* $g(\infty)$ *is the standard non-linear Schrödinger hierarchy for the attractive case when* $G=\mathrm{SU}(2)$ *and for the repulsive case when G is* $\mathrm{SU}(1,1)$. *If the* h_k, *for* $k\geqslant 0$, *are not zero, we obtain variations of these hierarchies appropriate to varying boundary conditions on the system.*

If $h_0>0$, $g(\infty)$ *yields a hierarchy of evolutions for two real-valued functions, whose simplest evolution is the real analogue of the NS equation, eq. (1.4a).*

If $h_0=0$, *but* $h_1\neq 0$ *and all the other* h_k, $k\geqslant 2$, *are zero, we find the standard Korteweg–de Vries hierarchy. If the* h_k *with* $k\geqslant 2$ *are not zero, we obtain variations of the KdV hierarchy, appropriate to varying boundary conditions on the system.*

The special case where both h_0 and h_1 vanish has not yet been examined in detail.

Theorem 3.2. *The reduced Bogomolny hierarchy,* $B(\infty)$, *where all quantities are independent of* t_0, *coincides with* $g(\infty)$.

Results analogous to theorem 3.1 appear in many places in the literature (see, for example, refs. [13–15]). We will prove theorem 3.2 first and then theorem 3.1.

Proof of theorem 3.2. We consider the Bogomolny hierarchy equations at level n, $B(n)$, subject to the requirement that the system possess a symmetry in the t_0 direction. Choose a gauge such that all the quantities A_k and B_{k-1}, with $k\geqslant 1$, are independent of the variable t_0. Equation (2.5a) implies that one may then take a gauge in which all the quantities B_k, with $k\geqslant 1$, vanish. Define $A_0=-B_0$. The Bogomolny hierarchy of eqs. (2.5) now becomes the following system:

$$\partial_k A_0=0, \quad \partial_k A_1=[A_0,A_{k+1}]+\partial_{k+1}A_0, \qquad 1\leqslant k\leqslant n-1\,, \tag{3.1a,b}$$

$$\partial_j A_k-\partial_k A_j-[A_j,A_k]=0, \quad \partial_k A_{j+1}-\partial_j A_{k+1}=0, \quad 1\leqslant j,k\leqslant n\,. \tag{3.2a,b}$$

Denote this system by $S(n)$. For simplicity, we begin by analyzing the limiting system, denoted by $S(\infty)$, where n is taken to infinity. Explicitly, $S(\infty)$ is the following system:

$$\partial_k g_0=0, \quad \partial_k g_1=[g_0,g_{k+1}], \qquad k\geqslant 1\,, \tag{3.3a,b}$$

$$\partial_j g_k-\partial_k g_j-[g_j,g_k]=0, \quad \partial_k g_{j+1}-\partial_j g_{k+1}=0, \quad j,k\geqslant 1\,. \tag{3.4a,b}$$

Note that, if one puts $A_k=g_k$, for $0\leqslant k\leqslant n$, one obtains a solution of $S(n)$, with the property that A_0 is constant. Equations (3.3) and (3.4) may be directly solved for the quantities $\partial_j g_k$, with the result

$$\partial_j g_k=\sum_{m=0}^{j}[g_m,g_{j+k-m}], \quad j\geqslant 1, k\geqslant 0\,. \tag{3.5}$$

Indeed, eqs. (3.4a) and (3.4b) give the recursion relation $\partial_j g_k=\partial_{j-1}g_{k+1}+[g_j,g_k]$, for all $j\geqslant 2$ and $k\geqslant 1$. Then eq. (3.5) follows immediately [after using eqs. (3.4a) and (3.3) at the last step to eliminate the quantity $\partial_1 g_{j+k-1}$]. Thus eqs. (3.3) and (3.4) yield eq. (3.5). Conversely, trivial algebraic identities show that, given quantities g_k, for all $0\leqslant k$, obeying eq. (3.5), then eqs. (3.3) and (3.4) follow. Hence $S(\infty)$ is completely equivalent to the system satisfying just eq. (3.5), which is precisely the system $g(\infty)$ described above. This proves theorem 3.2. □

Note that $S(2)$ yields eqs. (1.3) when one identifies $t_1=x$, $t_2=t$, $A_0=B$, $A_1=A$ and $A_2=C$.

Proof of theorem 3.1. We will proceed by using eq. (3.5) to obtain the recursion operators for the NS and KdV hierarchies.

Firstly some preliminaries. It is a straightforward algebraic identity that eq. (3.5) determines a consistent system of equations. We obtain $\partial_j\partial_k g_m=\partial_k\partial_j g_m$, $\forall j,k\geqslant 1$ and $m\geqslant 0$ as a consequence of eq. (3.5). One may write down immediately an infinite collection of constants of the motion for the evolution given by eq. (3.5) — define the quantities h_k by the formula

$$2h_k=\sum_{m=0}^{k}\mathrm{Tr}(g_m g_{k-m}), \quad k\geqslant 0\,. \tag{3.6}$$

Here Tr denotes the matrix trace (or more generally some invariant bilinear form on the Lie algebra). It is a simple algebraic identity that $\partial_j h_k$ vanishes, for each

$j \geqslant 1$ and $k \geqslant 0$, when eq. (3.5) holds.

When $j=1$, eq. (3.5) gives a partial recursion relation for the quantities g_k:

$$\partial_1 g_k - [g_1, g_k] = [g_0, g_{k+1}], \quad k \geqslant 0 . \tag{3.7}$$

When $g_0=0$, one sees that, after a gauge transformation, the dependence of all quantities on the variable t_1 drops out and the system collapses to itself, with the role of the variable t_k now played by t_{k+1}. So henceforth we assume that g_0 is non-zero.

We do not yet have a general description of the construction, by recursion, of the quantities g_k for a general choice of group, G, so from now onwards, we shall again restrict ourselves to the case where G has the complexification $\mathrm{SL}(2, \mathbb{C})$.

When $G_{\mathbb{C}} = \mathrm{SL}(2, \mathbb{C})$, g_0 has the characteristic features that, whenever the equation $[g_0, Y]=0$ holds, then Y is proportional to g_0. The equation $[g_0, Y]=Z$ is solvable for Y, given Z if and only if $\mathrm{Tr}(g_0 Z)=0$. Furthermore the matrix g_0 may be taken to be either of type α or of type β. The matrix g_0 is of type β iff $\mathrm{Tr}(g_0^2) \neq 0$ and of type α iff $\mathrm{Tr}(g_0)^2=0$. We discuss these cases separately.

Type β. We shall use eqs. (3.6) and (3.7) to determine the g_k inductively in terms of the constants h_k and g_1.

Let h_j for $j \geqslant 2$ be a given set of constants. Let g_1 be given arbitrarily except that $h_1 \equiv \mathrm{Tr}(g_0 g_1)$ is required to be constant and define $2h_0 = \mathrm{Tr}(g_0^2)$. Suppose that the quantities g_k have been found for all $k \leqslant m$, such that eq. (3.7) holds for all $k<m$ and such that eq. (3.6) holds for all $k<m$. Then, by induction, using eq. (3.6) one shows that the quantity $\mathrm{Tr}(g_0\{\partial_1 g_m - [g_1, g_m]\})$ vanishes. Thus eq. (3.7), with $k=m$, may be solved for the unknown g_{m+1}, uniquely modulo scalar multiples of g_0. This ambiguity in g_{m+1} is removed by imposing eq. (3.7) with $k=m+1$. By induction, therefore, g_k is uniquely determined, for all $k \geqslant 0$, given g_0, g_1 and all the h_j for $j \geqslant 2$.

From eqs. (3.6) and (3.7) we see that, in fact, g_k is a polynomial in the derivatives with respect to t_1 up to order $k-1$, of the entries of the matrix g_1, which consist of two free functions. The evolution of these quantities with respect to t_j is given by eq. (3.5) with $k=1$, giving a hierarchy of evolution equations. One may present this argument more explicitly, as follows. Write out the matrices g_k as follows:

$$g_k = \begin{pmatrix} q_k & p_k \\ r_k & -q_k \end{pmatrix}, \quad k \geqslant 0 ,$$

so one has $p_0 = r_0 = 0$ and $q_0 \neq 0$ is a constant. Then eq. (3.5) gives the following system:

$$\partial_j p_k = \sum_{m=0}^{j} 2(q_m p_{j+k-m} - p_m q_{j+k-m}), \quad j \geqslant 1, k \geqslant 0 , \tag{3.8a}$$

$$\partial_j q_k = \sum_{m=0}^{j} (p_m r_{j+k-m} - r_m p_{j+k-m}), \quad j \geqslant 1, k \geqslant 0 , \tag{3.8b}$$

$$\partial_j r_k = \sum_{m=0}^{j} 2(r_m q_{j+k-m} - q_m r_{j+k-m}), \quad j \geqslant 1, k \geqslant 0 . \tag{3.8c}$$

In particular the evolution of the quantities p_1, q_1 and r_1 is found by putting $k=1$ in eqs. (3.8):

$$\partial_k q_1 = 0, \quad \partial_k p_1 = 2q_0 p_{k+1}, \quad \partial_k r_1 = -2q_0 r_{k+1}, \quad k \geqslant 1 . \tag{3.9}$$

To determine the quantities p_k and r_k, iteratively, we first write out eq. (3.7):

$$2q_0 p_{k+1} = \partial_1 p_k - 2q_1 p_k + 2p_1 q_k,$$

$$2q_0 r_{k+1} = -\partial_1 r_k - 2q_1 r_k + 2r_1 q_k, \quad k \geqslant 0 , \tag{3.10}$$

$$\partial_1 q_k = p_1 r_k - r_1 p_k, \quad k \geqslant 0 . \tag{3.11}$$

Equations (3.10) and (3.11) give the required recursion relation: given p_k and r_k, one first solves eq. (3.11) to determine q_k and then (3.10) for p_{k+1} and r_{k+1} and iterates.

This procedure is conveniently summarized by the following matrix recursion operator, for the quantities p_{k+1} and r_{k+1}, which is obtained by eliminating q_k between eqs. (3.10) and (3.11) using the inverse derivative operator ∂_1^{-1}:

$$2q_0 \begin{pmatrix} p_{k+1} \\ r_{k+1} \end{pmatrix} = \begin{pmatrix} \partial_1 - 2q_1 - 2p_1 \partial_1^{-1} r_1 & 2p_1 \partial_1^{-1} p_1 \\ -2r_1 \partial_1^{-1} r_1 & -\partial_1 - 2q_1 + 2r_1 \partial_1^{-1} p_1 \end{pmatrix} \begin{pmatrix} p_k \\ r_k \end{pmatrix} . \tag{3.12}$$

Equation (3.12) holds for $k \geqslant 1$. This expression is somewhat unsatisfactory, because a boundary condition is hidden in the definition of the inverse derivative operator. It is thus better to use eq. (3.6), which immediately gives q_{k+1}, given p_j, q_j and r_j, for all $j \leqslant k$. Writing out eq. (3.6) explicitly, one has:

$$2q_0 q_{k+1} = h_{k+1} - \sum_{m=0}^{k-1} (r_{m+1} p_{k-m} + q_{m+1} q_{k-m}), \quad k \geqslant 0 . \tag{3.13}$$

[In eq. (3.13), the summation is understood to be vacuous when $k=0$.] Equations (3.10) and (3.13) give a complete recursion formula for the quantities p_k, q_k and r_k.

In the case that the constants h_k, for $k \geqslant 1$, are taken to be zero, and G is respectively either the group SU(2) or SU(1,1), these equations give the standard nonlinear Schrödinger hierarchy, for the attractive and repulsive couplings, respectively [15]. Note that these constants automatically become zero if it is required that all the matrices g_k, for $k \geqslant 1$, vanish at infinity.

Type α. Next we turn to the case where g_0 is of type α. Then eq. (3.8) is to be analyzed, subject to the condition that $p_0 = q_0 = 0$ and $r_0 = 1$. Also eq. (3.6) gives the following equations:

$$h_0 = 0, \qquad h_1 = p_1, \qquad h_2 = p_2 + q_1^2 + h_1 r_1, \tag{3.14}$$

$$p_{k+1} + h_1 r_k + 2q_1 q_k + r_1 p_k = h_{k+1} - \sum_{m=1}^{k-2} (r_{m+1} p_{k-m} + q_{m+1} q_{k-m}), \quad k \geqslant 2. \tag{3.15}$$

In eq. (3.15), the summation is understood to be vacuous when $k=2$. We will assume that the constant h_1 is non-zero. To obtain the recursion formula, first write out eq. (3.7):

$$2h_1 q_k = -\partial_1 p_k + 2q_1 p_k, \tag{3.16a}$$

$$2q_{k+1} = \partial_1 r_k + 2q_1 r_k - 2r_1 q_k, \quad k \geqslant 0, \tag{3.16b}$$

$$p_{k+1} - h_1 r_k = -\partial_1 q_k - r_1 p_k, \qquad k \geqslant 0. \tag{3.17}$$

Using eqs. (3.16) and (3.17), we derive the following identity:

$$\partial_1 (p_{k+1} + h_1 r_k) = -2q_1 \partial_1 q_k - 2r_1 (q_1 p_k - h_1 q_k). \tag{3.18}$$

Combining eqs. (3.17) and (3.18), we find the following expression for the quantity p_{k+1}:

$$2\partial_1 p_{k+1} = -\partial_1 (r_1 p_k) - 2q_1 r_1 p_k - (\partial_1^2 + 2q_1 \partial_1 - 2h_1 r_1) q_k. \tag{3.19}$$

The quantity q_k may be eliminated from eq. (3.19), using eq. (3.16a). Again employing the inverse derivative operator, we obtain the following compact recursion relation for the quantity p_k:

$$4h_1 p_{k+1} = (\partial_1^2 - 4x_1 + 2\partial_1^{-1} y_1) p_k, \quad k \geqslant 1,$$

$$y_1 = \partial_1 x_1, \qquad x_1 = \partial_1 q_1 + q_1^2 + h_1 r_1 = 2\partial_1 q_1 + h_2. \tag{3.20}$$

Equation (3.20) must be supplemented with boundary conditions to fix the inverse derivative operator. Once p_k has been found, q_k may be determined from eq. (3.16a) and then r_{k-1} from eq. (3.17). Note that $p_2 = -\partial_1 q_1$, so that p_k is a polynomial in the derivatives of q_1 with the highest derivative term proportional to the $(2k-3)$th derivative of q_1. As before, we may avoid the use of the inverse derivative operator if, instead of eq. (3.18), we use eq. (3.15). By adding eqs. (3.15) and (3.17), we obtain the formula:

$$2p_{k+1} = h_{k+1} - \partial_1 q_k - 2q_1 q_k - 2r_1 p_k - \sum_{m=1}^{k-2} (r_{m+1} p_{k-s} + q_{m+1} q_{k-s}), \quad k \geqslant 2. \tag{3.21}$$

Note that this expression only involves the quantities r_j, for $j \leqslant k-1$. Thus the recursion procedure goes as follows: if all the quantities p_j, q_j and r_{j-1} are given, for $j \leqslant k$, $k \geqslant 2$, first p_{k+1} is determined by eq. (3.21); then r_k is found from eq. (3.17) (provided that h_1 is a non-zero constant); finally q_{k+1} is obtained from eq. (3.16b) and so on. This recursion is provided with the initial conditions

$$p_0 = 0, \quad p_1 = h_1, \quad p_2 = -\partial_1 q_1,$$
$$q_0 = 0, \quad q_1 = q_1, \quad q_2 = (2h_1)^{-1}(\partial_1^2 q_1 - 2q_1\,\partial_1 q_1),$$
$$r_0 = 1, \quad r_1 = h_1^{-1}(h_2 + \partial_1 q_1 - q_1^2). \tag{3.22}$$

Thus every quantity is a polynomial in the derivatives of q_1. We do not analyze the case where h_1 vanishes, here. Finally the evolution of q_1 is obtained from eq. (3.8b), by putting $k=1$. One finds the simple equation

$$\partial_k q_1 = -p_{k+1}, \quad k \geqslant 1. \tag{3.23}$$

Equations (3.20) and (3.23) agree with the standard recursion equations for the Korteweg–de Vries hierarchy (see, for example, ref. [15]) establishing that the system $g(\infty)$, for g_0 of type α, with the constant h_1 non-zero, gives the standard Korteweg–de Vries hierarchy. Note that, if the boundary condition is that q_1 and all its derivatives are required to rapidly decrease as t_1 goes to infinity, then all the constants h_k with $k \geqslant 3$ must vanish. Also, eq. (3.23) for $k=2$ is the following equation:

$$\partial_2 q_1 = -h_3/2 + h_1^{-1}[\partial_1^3 q_1 - 6(\partial_1 q_1)^2 - 4h_2\,\partial_1 q_1]. \tag{3.24}$$

Differentiate this equation with respect to t_1 and put $u = -\partial_1 q_1$. One obtains the equation

$$\partial_2 u = h_1^{-1}[\partial_1^3 u + 12u\,\partial_1 u - 4h_2\,\partial_1 u]. \tag{3.25}$$

So the standard Korteweg–de Vries equation and its hierarchy is obtained in the rapidly decreasing case if all the constants h_k, with $k \geqslant 2$, vanish. This completes the proof of theorem 3.1. □

Remark. Note that eq. (3.5) in many cases describes a system preserving a Poisson structure. Specifically, let lower case Greek indices be used for the Lie algebra g. Let g have structure tensor $C_{\beta\gamma}{}^{\alpha}$. Define the Poisson brackets

$$\{g_j^{\alpha}, g_k^{\beta}\} = D_{\gamma}{}^{\alpha\beta} g_{j+k}{}^{\gamma} \tag{3.26}$$

(repeated Greek indices are summed over). Equation (3.26) defines a consistent Poisson bracket structure, provided that the tensor $D_{\alpha}{}^{\beta\gamma}$ is skew and obeys the Jacobi identity, so that it gives a Lie algebra structure to g^*, the dual to g. Then one finds that the flows given by eq. (3.26) preserve the Poisson structure if and only if $D_{\alpha}{}^{\beta\gamma}$ is invariant under the adjoint action of g and $C_{\beta\gamma}{}^{\alpha}$ is invariant under the adjoint action of g^*. Further, if $h_{\alpha\beta}$ is a symmetric tensor such that $D_\gamma{}^{\alpha\beta}h_{\delta\beta}=C_{\gamma\delta}{}^\alpha$, then the flows are Hamiltonian, with Hamiltonians h_k given by the formula

$$h_k=\sum_{m=0}^{k} h_{\alpha\beta}g_m^\alpha g_{k-m}^\beta, \quad k\geqslant 0. \tag{3.27}$$

In particular, if G is semi-simple, $h_{\alpha\beta}$ may be taken to be the Killing form and $D_\gamma{}^{\alpha\beta}=C_{\gamma\delta}{}^\alpha h^{\beta\delta}$, where $h^{\alpha\beta}$ is the inverse of $h_{\alpha\beta}$ and all the conditions are satisfied. Equation (3.27) defines the same h_k as eq. (3.6).

4. The relationship between the inverse scattering transform and the Penrose–Ward transform

The inverse scattering transform proceeds, both for the NS and the KdV equations, by studying the scattering problem for solutions of the linear system, eq. (1.2′). For the NS equation, eq. (1.2′) is precisely the normal Lax pair as in ref. [14]. For the KdV equation, the first component, s_1, of s is a solution of the Schrödinger equation $(\partial_x^2-2q_x)s_1=\lambda s_1$ that forms the basis of the normal Lax pair for KdV [the second component is given by $s_2=(\partial_x-q)s_1$]. We have seen, however, in the proof of theorem 2.1 that (local) solutions of (1.2) correspond to (local) sections of the vector bundle E on $\mathcal{O}(2)$. Solutions of (1.2′) correspond to sections of the vector bundle that are invariant under the symmetry $\partial_0=\partial_v$ in terms of their representation on spacetime, and hence are invariant under the lift to the bundle of the corresponding symmetry ∂_μ on twistor space. Therefore, solutions of the Lax pairs for the NS and KdV equations determine invariant local sections of the holomorphic vector bundle on twistor space. In this section we will see that the scattering data provide the patching data for the bundle in terms of local sections represented as solutions of (1.2′). Thus the inverse scattering transform is a concrete realization of the Ward transform appropriate to rapidly decreasing boundary conditions.

We shall restrict ourselves to the non-linear Schrödinger equation in order to illustrate the basic ideas. The details for the KdV equation are more complicated and will be treated elsewhere. In this section we will consider the case where the scattering data have no discrete spectrum, just a continuous spectrum. In the next section we will consider the soliton solutions for which the scattering data are zero, but which have a non-trivial discrete spectrum.

First we will summarize some facts from the standard theory of inverse scattering; see ref. [14] for the proof of these results.

We work with solutions of eq. (1.2′a) of the linear system

$$\left[\partial_x - \begin{pmatrix} 0 & \psi \\ \pm\bar{\psi} & 0 \end{pmatrix} + \tfrac{1}{2}\lambda \mathrm{i} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = 0 , \qquad (1.2'\mathrm{a})$$

where $s = \binom{s_1}{s_2}$ is a two-component column vector and $\pm$ is $-$ for the attractive case and $+$ for the repulsive case. This is the basic equation in the linear system for the non-linear Schrödinger equation as used in ref. [14].

If $s = \binom{s_1}{s_2}$ is a solution of eqs. (1.2′), then so is

$$\hat{s} = \begin{pmatrix} \bar{s}_2 \\ \pm\bar{s}_1 \end{pmatrix} ,$$

where by $\bar{f}(\lambda)$, the complex conjugate function associated to $f(\lambda)$, we mean the holomorphic function $\bar{f}(\lambda) = \overline{f(\bar{\lambda})}$. Note that $\hat{\hat{s}} = \pm s$.

For the purpose of the inverse scattering transform assume that $\psi \to 0$ "rapidly" as $|x| \to \infty$ and consider the solutions of eq. (1.2′a) that have the asymptotic forms

$$s_+ \to \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\lambda x/2} \\ 0 \end{pmatrix}, \qquad \hat{s}_+ \to \begin{pmatrix} 0 \\ \pm\mathrm{e}^{\mathrm{i}\lambda x/2} \end{pmatrix}, \quad \text{as } x \to +\infty ,$$

$$s_- \to \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\lambda x/2} \\ 0 \end{pmatrix}, \qquad \hat{s}_- \to \begin{pmatrix} 0 \\ \pm\mathrm{e}^{\mathrm{i}\lambda x/2} \end{pmatrix}, \quad \text{as } x \to -\infty .$$

The scattering data consist of the functions $a(\lambda)$, $b(\lambda)$, $\bar{a}(\lambda)$ and $\bar{b}(\lambda)$ defined by

$$s_- = a(\lambda) s_+ + b(\lambda)\hat{s}_+ , \qquad \hat{s}_- = \pm\bar{b}(\lambda) s_+ + \bar{a}(\lambda)\hat{s}_+ . \qquad (4.1)$$

The determinant of a pair of solutions is independent of (x, t) as the connection matrices are trace free. We therefore have the identity $\det(s_\pm, \hat{s}_\pm) = \pm 1$. This yields

$$a\bar{a} \mp b\bar{b} = 1 . \qquad (4.2)$$

These solutions enjoy the following analytic continuation properties:

$$\mathrm{e}^{\mathrm{i}\lambda x/2} s_+ \text{ and } \mathrm{e}^{-\mathrm{i}\lambda x/2}\hat{s}_- \text{ are analytic in } \lambda \text{ for } \operatorname{im}\lambda \leqslant 0 \text{ and all } x ,$$

$$\mathrm{e}^{\mathrm{i}\lambda x/2} s_- \text{ and } \mathrm{e}^{-\mathrm{i}\lambda x/2}\hat{s}_+ \text{ are analytic in } \lambda \text{ for } \operatorname{im}\lambda \geqslant 0 \text{ and all } x , \qquad (4.3)$$

and as $|\lambda| \to \infty$ we have

$$\mathrm{e}^{\mathrm{i}\lambda x/2} s_+ \to \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathrm{o}(1), \qquad \mathrm{e}^{-\mathrm{i}\lambda x/2}\hat{s}_- \to \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mathrm{o}(1) ,$$

for $\operatorname{im}\lambda \leqslant 0$ and all x,

$$\mathrm{e}^{\mathrm{i}\lambda x/2} s_- \to \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathrm{o}(1), \qquad \mathrm{e}^{-\mathrm{i}\lambda x/2}\hat{s}_+ \to \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathrm{o}(1) , \qquad (4.4)$$

for $\operatorname{im}\lambda \geqslant 0$ and all x.

These properties follow by integrating eq. (1.2′a) to obtain a Volterra integral equation and then examining its convergence properties under iteration for different values of λ. See section 1.5 of ref. [14] for full details.

Note that $a(\lambda)=\det(e^{i\lambda x/2}s_-, e^{-i\lambda x/2}\hat{s}_+)$ is therefore analytic in the upper half-plane, and $\bar{a}(\lambda)$ is analytic in the lower half-plane.

In the case where the potential is attractive [gauge group SU(2)], it is possible that there are "bound state" solutions to (1.2′a), so that s_+ and $\hat{s}_-$ are proportional at a collection of complex values of λ with im $\lambda<0$ (this implies that s_+ is exponentially decaying both as $x\to+\infty$ and as $x\to-\infty$ at that value of λ). For such values of λ, $\bar{a}(\lambda)=0$. [Similarly it is possible that s_- and $\hat{s}_+$ are linearly dependent for some λ with im $\lambda>0$, at which values of λ, $a(\lambda)=0$.] In the next subsection we treat the case where this does not happen, $a(\lambda)\neq 0$; in the subsequent one we will treat the case where $b(\lambda)=0$, but where there are values of λ at which $a(\lambda)$ does vanish.

4.1. NON-SOLITONIC SOLUTIONS

Since s_- and $\hat{s}_+$ are linearly independent solutions of (1.2′a) for im $\lambda\geqslant 0$, they determine a ∂_μ-invariant frame f_u for the vector bundle E on the twistor space $\mathcal{O}(2)$ defined on the intersection of im $\lambda\geqslant 0$ with the union of the sections corresponding to real points in $\mathbb{C}^3$. Similarly s_+ and $\hat{s}_-$ determine a frame f_ℓ on the region in $\mathcal{O}(2)$ given by the intersection of im $\lambda\leqslant 0$ with the union of the sections corresponding to real points in $\mathbb{C}^3$. However, the fall-off conditions (4.3) show that, for this frame to be regular as $|\lambda|\to\infty$ in the relevant domains, we must include exponential factors to cancel the essential (exponential) singularity at $\lambda=\infty$. As they stand, the factors of $e^{i\lambda x/2}$ are not functions on $\mathcal{O}(2)$, but the function $e^{i\mu/2}=\exp[i(v+\lambda x+\lambda^2 t)/2]$ is a function on $\mathcal{O}(2)$ and agrees with $e^{i\lambda x/2}$ when v and t are held constant.

Thus there exist frames

$$F_u=\frac{1}{\sqrt{a}}(e^{i\mu/2}s_-, e^{-i\mu/2}\hat{s}_+), \qquad F_\ell=\frac{1}{\sqrt{\bar{a}}}(e^{i\mu/2}s_+, e^{-i\mu/2}\hat{s}_-),$$

that are regular on

$$U_u=\{(\mu,\lambda)\in\mathcal{O}(2)\mid \text{im }\lambda\geqslant 0, \mu=v+x\lambda+t^2\lambda^2 \text{ for some real } (v,x,t)\},$$

$$U_\ell=\{(\mu,\lambda)\in\mathcal{O}(2)\mid \text{im }\lambda\leqslant 0, \mu=v+x\lambda+t^2\lambda^2 \text{ for some real } (v,x,t)\},$$

respectively. (The factors of $1/\sqrt{a}$ etc. have been included in order to give F_u and F_ℓ unit determinants; recall that a was assumed to be non-vanishing so there is no serious ambiguity.) These sections are no longer invariant under the symmetry due to the insertion of the factors of $e^{i\mu/2}$. On twistor space, the action of the symmetry is just ∂_μ, and the action of the symmetry on a section s of E in the frames F_u and F_ℓ is

$$\mathscr{L}_{\partial_\mu} s = \partial_\mu s + \tfrac{1}{2}\mathrm{i}s\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

We have therefore produced frames of E which, although not invariant, have their μ-dependence in normal form.

On $\lambda=\bar{\lambda}$ the frames F_{u} and F_{ℓ} are related by $F_{\mathrm{u}}=F_{\ell}P$, where P is given by the formula

$$P(\mu,\lambda)=\frac{1}{\sqrt{1\pm b\bar{b}}}\begin{pmatrix}1 & \mp \mathrm{e}^{-\mathrm{i}\mu}\bar{b}\\ \mathrm{e}^{\mathrm{i}\mu}b & 1\end{pmatrix}.$$

Thus we see that P is a patching function for the bundle.

The transform from $P(\mu,\lambda)$ to $\psi(x,t)$ is the inverse scattering transform. From the point of view of both the inverse scattering and the Ward transform (cf. section 2) one reconstructs ψ from P by finding F_{u} and F_{ℓ}, matrix functions of (x,t,λ), such that

$$F_{\mathrm{u}}(x,t,\lambda)=F_{\ell}(x,t,\lambda)P(\lambda x+\lambda^2 t,\lambda)\ .$$

This is a Riemann–Hilbert problem for each fixed (x,t) with $F_{\mathrm{u}}(x,t,\lambda)$ holomorphic on im $\lambda\geqslant 0$ and $F_{\ell}(x,t,\lambda)$ holomorphic on im $\lambda\leqslant 0$. Such parametrized Riemann–Hilbert problems, in general, are known to have a solution for all but a complex codimension-one subset of values of x and t. The solution is unique up to left multiplication of both F_{u} and F_{ℓ} by a matrix depending only on (x,t). This freedom can be fixed by requiring that, as $\lambda\to\infty$, $F_{\mathrm{u}}=1$.

Remark. In the above, we have avoided discussing the precise nature of the domain and cover in $\mathcal{O}(2)$ with respect to which we are obtaining the Čech description of the bundle. There is some novelty due to the fact that the sets as defined above are closed, not open as usually required for a Čech description, and, indeed, the intersection of the two sets is of real codimension-two and is the common Shilov boundary of the two sets. This reflects the geometry of the subset of $\mathcal{O}(2)$ swept out by sections corresponding to real points in spacetime. This does not make any practical difference to the correspondence, but in the case of non-analytic solutions global in x and t, it is essential as the domain on which the bundle E is defined includes all values of μ for λ complex but only real values of μ when λ is real.

Two distinct cases that one can consider are, firstly, the case where $b(\lambda)$ decreases faster than $\mathrm{e}^{-c\lambda^2}$ as $|\lambda|\to\infty\ \forall c$, and, secondly the case where $b(\lambda)$ is analytic along $\lambda=\bar{\lambda}$ including $\lambda=\infty$. At least in the linearized limit, $b(\lambda)$ rapidly decreasing corresponds to ψ analytic as a function of x, and $b(\lambda)$ analytic corresponds to ψ rapidly decreasing since, at least in the linearized limit, $b(\lambda)$ is the Fourier transform of ψ.

4.2. THE RAPIDLY DECREASING SOLITON SOLUTIONS

The soliton solutions are characterized by $b(\lambda)=0$. However, the coefficient $a(\lambda)$ is now allowed to have zeros at a selection of values of λ, $\{\lambda_i\}$, $i=1, ..., k$ in the upper half-plane. We will assume for simplicity that a has no multiple zeros or poles. This requirement, together with the fact that $a\bar{a}=1$ on $\lambda=\bar{\lambda}$ and the condition that a be analytic on the upper half λ-plane, determines $a(\lambda)$ as

$$a(\lambda)=\prod_{i=1}^{k}\frac{(\lambda-\lambda_i)}{(\lambda-\bar{\lambda}_i)} .$$

In particular $\bar{a}(\lambda)=1/a(\lambda)$.

The condition $b=0$ implies that $P=\text{Identity}$ so that the frames $\sqrt{a}\,F_{\text{u}}$ and $(1/\sqrt{a})F_{\ell}$ are analytic continuations of each other and are therefore the same frame,

$$F=(e^{i\mu/2}s_-, e^{-i\mu/2}\hat{s}_+)=(1/\bar{a})(e^{i\mu/2}s_+, e^{-i\mu/2}\hat{s}_-) .$$

The columns of F are solutions of the Lax pair that are globally meromorphic in λ due to the factor of $1/\bar{a}=a$. Furthermore, $\det(F)=a(\lambda)$ so that the columns of F are proportional at $\lambda=\lambda_i$:

$$e^{i\mu/2}s_-(\lambda_i)=\gamma_i e^{-i\mu/2}\hat{s}_+(\lambda_i) , \tag{4.5}$$

where the γ_i are non-zero constants. Similarly at $\lambda=\bar{\lambda}_i$,

$$e^{-i\mu/2}\hat{s}_-(\bar{\lambda}_i)=-\bar{\gamma}_i e^{i\mu/2}s_+(\bar{\lambda}_i) .$$

The data $\{(\lambda_i, \gamma_i)\}$, $i=1, ..., k$, determine the soliton solution. In section 2.5 of ref. [14] the soliton solutions are constructed explicitly by solving a "Riemann–Hilbert problem with zeros" which involves expressing $F(x, t, \lambda)$ as an ordered product of "Blaschke–Potapov factors".

4.3. THE TWISTOR DESCRIPTION

The bundle, E, on $\mathcal{O}(2)$ corresponding to the soliton solution has a pair of holomorphic line subbundles, L and $\hat{L}$. These are determined by the span of the first and second columns of F, respectively; L is determined by the span of $e^{i\mu/2}s_-$ on $\text{im}\,\lambda\geq 0$ and $e^{i\mu/2}s_+$ on $\text{im}\,\lambda\leq 0$, and similarly for $\hat{L}$. The line bundles L and $\hat{L}$ are linearly dependent at $\lambda=\lambda_i$.

The line subbundles are both isomorphic to $\tilde{\mathcal{O}}(-k)$, the pull-back to $\mathcal{O}(2)$ of $\mathcal{O}(-k)$ on $\mathbb{CP}^1$, the line bundle of Chern class $-k$. This follows from the fact that the patching function for L is a on $\text{im}\,\lambda=0$ (since $e^{i\mu/2}s_-=ae^{i\mu/2}s_+$) and a has degree k. We have then

$$0\to\tilde{\mathcal{O}}(-k)\to E\to\tilde{\mathcal{O}}(k)\to 0 ,$$

where the quotient $E/\tilde{\mathcal{O}}(-k)\simeq\tilde{\mathcal{O}}(k)=\tilde{\mathcal{O}}(-k)^*$ because sections of $E/\tilde{\mathcal{O}}(-k)$ are dual to those of $\tilde{\mathcal{O}}(-k)$ by taking the determinant of the representative sections in E [recall that E is an $\mathrm{SL}(2,\mathbb{C})$ bundle].

The bundle E can be represented as a deformation of $\tilde{\mathcal{O}}(-k)\oplus\tilde{\mathcal{O}}(k)$ that preserves the inclusion $\tilde{\mathcal{O}}(-k)\to E$. Such bundles are completely characterized by their extension class. This can be represented as follows.

Choose a cover $\{U_\alpha\}$ for $\mathcal{O}(2)$ and a pair of sections, (s_α^1, s_α^2) constituting a frame F_α with unit determinant for E on each U_α such that the s_α^1 are a fixed collection of trivializations of $\tilde{\mathcal{O}}(-k)$ with patching functions $a_{\alpha\beta}$. Then the patching data for E are

$$(s_\alpha^1, s_\alpha^2)=(s_\beta^1, s_\beta^2)\begin{pmatrix} a_{\alpha\beta} & a_{\alpha\beta}h_{\alpha\beta}\\ 0 & a_{\alpha\beta}^{-1}\end{pmatrix}$$

(no summation is implied over repeated indices), where the freedom $s_\alpha^2\to s_\alpha^2+g_\alpha s_\alpha^1$ with g_α holomorphic functions on U_α leads to

$$h_{\alpha\beta}\to h_{\alpha\beta}+g_\alpha-a_{\alpha\beta}^{-2}g_\beta\ .$$

Such a $h_{\alpha\beta}$ up to this equivalence constitutes an element h of $H^1(\mathcal{O}(2),\tilde{\mathcal{O}}(-2k))$.

For the purposes of the soliton solutions, h must be chosen so that E is invariant and satisfies the reality condition. With this end in mind, we cover $\mathcal{O}(2)$ by the open sets

$$U_{\mathrm{u}}=\{(\mu,\lambda)\in\mathcal{O}(2)\,|\,\mathrm{im}\,\lambda>-\epsilon\},\qquad U_{\ell}=\{(\mu,\lambda)\in\mathcal{O}(2)\,|\,\mathrm{im}\,\lambda<-\epsilon\}\ ,$$

for some $\epsilon>0$ chosen so that a does not vanish on the intersection. Then put $s_{\mathrm{u}}^1=\mathrm{e}^{\mathrm{i}\mu/2}s_-$ and $s_\ell^1=\mathrm{e}^{\mathrm{i}\mu/2}s_+$. For s_{u}^2 we cannot use $\mathrm{e}^{-\mathrm{i}\mu/2}\hat{s}_+$ since this is proportional to s_{u}^1 at $\lambda=\lambda_i$. However, at $\lambda=\lambda_i$, $a=\det(F)$ vanishes only to first order (by assumption), so that $(\lambda-\lambda_i)^{-1}(\mathrm{e}^{-\mathrm{i}\mu/2}\hat{s}_+-\gamma_i^{-1}\mathrm{e}^{-\mathrm{i}\mu/2}s_-)$ is regular and transverse to s_{u}^1 except where $\lambda=\lambda_j$, $j\neq i$. Extending this principle, we put

$$s_{\mathrm{u}}^2=\frac{1}{a}\left(\mathrm{e}^{-\mathrm{i}\mu/2}\hat{s}_+-\sum_{i=1}^{k}\frac{a(\lambda)}{\gamma_i(\lambda-\lambda_i)a'(\lambda_i)}\mathrm{e}^{\mathrm{i}\mu/2}s_-\right),$$

where $a'(\lambda_i)$ is the derivative of a with respect to λ evaluated at λ_i. It can be checked that this section is regular and linearly independent of s_{u}^1 on U_{u} and furthermore that $\det(s_{\mathrm{u}}^1, s_{\mathrm{u}}^2)=1$. Similarly put

$$s_\ell^2=a\left(\mathrm{e}^{-\mathrm{i}\mu/2}\hat{s}_-+\sum_{i=1}^{k}\frac{\bar{a}(\lambda)}{\bar{\gamma}_i(\lambda-\bar{\lambda}_i)\bar{a}'(\bar{\lambda}_i)}\mathrm{e}^{\mathrm{i}\mu/2}s_+\right).$$

This choice of frames leads to the patching data, $a_{\mathrm{u}\ell}=a$, and

$$h_{\mathrm{u}\ell}=-\sum_{i=1}^{k}\left(\frac{1}{a^2\bar{\gamma}_i(\lambda-\bar{\lambda}_i)\bar{a}'(\bar{\lambda}_i)}+\frac{1}{\gamma_i(\lambda-\lambda_i)a'(\lambda_i)}\right).$$

This is the representative of $H^1(\mathcal{O}(2), \bar{\mathcal{O}}(-2k))$ that determines E.

5. Conclusion

There are many more techniques for studying solutions of the non-linear Schrödinger and KdV equations such as the Krichever construction, Bäcklund transformations etc. These generally amount to finding some way to present the bundle on twistor space in such a way that one can perform the Riemann–Hilbert problem explicitly. These will be discussed in a subsequent work.

Appendix A. Proof of theorem α

We proceed by reducing eqs. (1.3a,b,c):

$$\partial_x B=0, \quad [\partial_x-2A, \partial_t-C]=0, \quad 2\partial_x A-[B,C]=\partial_t B . \qquad (1.3\text{a,b,c})$$

We assume B is of type α so the matrices A, B and C are given as follows:

$$A=\begin{pmatrix} q & p \\ r & -q \end{pmatrix}, \quad B=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad C=\begin{pmatrix} v & u \\ w & -v \end{pmatrix}.$$

Then eq. (1.3a) follows immediately, and eq. (1.3c) implies: $p_x=0$, $u=-q_x$ and $v=\frac{1}{2}r_x$. In particular, therefore, p is a function only of t. The final equation, (1.3b), when written out yields

$$p_t=-q_{xx}+2qq_x+pr_x, \qquad 2q_t=r_{xx}-2rq_x-2pw,$$
$$r_t=w_x+2qw-rr_x . \qquad (\text{A.1a,b,c})$$

The remaining gauge freedom takes $(p,q,r)\to(p,q-pg,r+2qg-g^2p)$ where $q\equiv g(t)$.

When p vanishes identically the equations are completely soluble with solution (after using the remaining gauge freedom)

$$p=0, \qquad q=\epsilon a \tan X,$$

$$r=\frac{1}{(\epsilon a)^2}(\epsilon a_t X+c)(X\tan X+1)-b_t,$$

$$u=-q_x, \qquad 2v=r_x,$$

$$4a^4 w=X^2(\epsilon a_t X+c)^2 \sec^2 X+2X(2\epsilon a_t X+c)(\epsilon a_t X+c)\tan X$$
$$+(\epsilon a c_t-2c\epsilon a_t-2\epsilon^3 a^3 b_{tt})\sin(2X)$$
$$+d\cos^2 X+2\epsilon^2 a a_{tt} X^2+2\epsilon a c_t X+c^2,$$

where $\epsilon^4=1$, $X=\epsilon a(x+b)$ and a, b, c and d are arbitrary real functions of t, except that a is positive. The limiting case when the function a becomes identically zero gives the solution:

$$q=-X^{-1}, \qquad r=-b_t+cX^2\,,$$
$$u=X^{-2}, \qquad v=cX\,,$$
$$w=b_{tt}X+dX^2+c_tX^3+c^2X^4, \quad X=x+c\,,$$

where b, c and d are arbitrary real functions of t. We do not discuss here the behaviour of the solutions when either p or a approaches zero.

We now consider the generic case where p is assumed to be everywhere nonzero. Equation (A.1a), when integrated with respect to x, gives the relation $rp=q_x-q^2+xp_t+m_t$, for some gauge invariant function $m(t)$. Equation (A.1c) can be manipulated until every term is an exact derivative with respect to x as follows (we first multiply by p^2 and take the first and last terms of the r.h.s. as the subject):

$$\begin{aligned}[p^2(2w-r^2)]_x &= 2[p(pr)_t-p_t(pr)-2p^2qw]\\ &=2[pq_{xt}+p(p_{tt}x+m_{tt})-p_tpr-pqr_{xx}+2pqq_xr]\\ &=2[pq_{tx}+p(p_{tt}x+m_{tt})\\ &\quad+(pq_xr-q\{pr\}_x)_x-pr(q_x-q^2+p_tx+m_t)\\ &=[2pq_t+pp_{tt}x^2+2pm_{tt}x+2pq_xr-2pqr_x-p^2r^2]_x\,.\end{aligned}$$

Integrating we obtain the following expression:

$$2p^2w=2pq_t+2pq_xr-2pqr_x+pp_{tt}x^2+2pm_{tt}x+n\,, \tag{A.2}$$

where n is an arbitrary function of t. A gauge transformation takes n to $n-4p^2g_t-2pp_tg$, so g may be chosen so as to reduce n to zero. The residual gauge freedom has $g=g_0|p|^{-1/2}$, where g_0 is a constant. Comparing eqs. (A.1b) and (A.2) we obtain the following evolution equation for q:

$$\begin{aligned}pq_t &= (pr)_{xx}-4(pr)q_x+2(pr)_xq-pp_{tt}x^2-2pm_{tt}x\\ &=q_{xxx}-6q_x^2-4q_x(p_tx+m_t)+2p_tq-pp_{tt}x^2-2pm_{tt}x\,.\end{aligned} \tag{A.3}$$

Make the redefinition $|p|^{1/2}q'=q+\frac{1}{4}(p_tx^2+2m_tx+k)$, where $k(t)$ is a solution of the ODE

$$2k_t=(p_tk+m_t^2)/p\,.$$

The freedom in choice of k and the gauge freedom in q both transform q' to $q'+c$, where c is a constant. In terms of q', eq. (A.3) reads

$$4pq'_t=q'_{xxx}-6|p|^{1/2}q'^2_x+2q'_x(p_tx+m_t)\,. \tag{A.4}$$

Next define vector fields E and F by the formulae

$$2E=|p|^{-1/2}\frac{\partial}{\partial x}, \qquad 2F=\frac{|p|^{1/2}}{p}\left(\frac{\partial}{\partial t}-\frac{p_t x+m_t}{2p}\frac{\partial}{\partial x}\right).$$

Note that E and F commute, so they are co-ordinate vector fields, $E=\partial/\partial x'$ and $F=\partial/\partial t'$, for some new co-ordinate system (x', t'); in fact (x', t') are given by the following formulae:

$$x'=2|p|^{1/2}x+f(t), \quad f_t(t)=\frac{|p|^{1/2}m_t}{p},$$

$$t'=h(t), \quad h_t=2p|p|^{-1/2}. \tag{A.5}$$

In the new co-ordinate system eq. (A.4) now reads

$$q'_{t'}=q'_{x'x'x'}-3(q'_{x'})^2. \tag{A.6}$$

Define $u'=-q'_{x'}$ and differentiate eq. (A.6) with respect to x', to obtain an equation for u'. Dropping primes, one obtains the equation

$$u_t=u_{xxx}+6uu_x. \tag{A.7}$$

This is the Korteweg–de Vries equation! The upshot of this discussion is that, modulo gauge and co-ordinate transformations, one might as well have begun with $p=1$ and all integration constants may be taken to be zero. Thus one arrives at the statement of theorem α. □

Example. Consider eq. (A.3), with $p=1/4$, $m=3t^2$, $q=w(x)-2xt+8t^3$. Then eq. (A.3) becomes just $x=w_{xxx}-6w_x^2$. Defining $y=w_x$ we obtain *Painlevé's equation of the first kind*:

$$y_{xx}=6y^2+x. \tag{A.8}$$

The co-ordinate transformation of eq. (A.5) in this case is $x'=x+6t^2$, $t'=t$. The corresponding solution of eq. (A.7) (the KdV equation) is given by $u=-2y(x-6t^2)-2t$.

Appendix B. Proof of theorem β

Consider eqs. (1.3a,b,c) when B is of type β. i.e., $B=\mathrm{diag}(\kappa, -\kappa)$. The available gauge freedom consists of diagonal elements of G, depending only on t. Then eqs. (1.3b) and (1.3c), with A and C given in terms of p, q, r, u, v, w, as in the beginning of appendix A, give the following equations:

$$\kappa_x=0, \quad q_x=\kappa_t, \quad p_x=2\kappa u, \quad r_x=-2\kappa w,$$

$$v_x-q_t-pw+ru=0, \tag{B.1}$$

$$u_x-p_t-2qu+2pv=0, \qquad w_x-r_t+2qw-2rv=0. \tag{B.2}$$

Integrating eq. (B.1), one finds expressions for the quantities q and v:

$$q=\Lambda_x, \qquad v=-\frac{pr}{2\kappa}+\Lambda_t, \quad 2\Lambda=x^2\kappa_t+2xm_t+2n_t. \tag{B.3}$$

Here m and n depend only on t. Also eq. (B.2) may now be written as follows:

$$p_{xx}-2\kappa p_t-2qp_x+\kappa pv=0=r_{xx}+2\kappa r_t+2qr_x+4\kappa rv. \tag{B.4}$$

Now make the substitution $p=\epsilon\kappa e^{2\Lambda}\Psi$, $r=\epsilon\kappa e^{-2\Lambda}\Psi'$, where Λ is given in eq. (B.3) and $\epsilon=1$ or i and is chosen so that the quantity $\epsilon\kappa$ is real. Under the remaining gauge transformations, Ψ and Ψ' transform by $\Psi\to\lambda\Psi$ and $\Psi'\to(1/\lambda)\Psi'$, where λ is a constant such that the diagonal matrix diag$(\lambda, 1/\lambda)$ belongs to G. Then in terms of Ψ and Ψ', eq. (B.4) becomes the following:

$$\Psi_{xx}+2(x\kappa_t+m_t)\Psi_x-2\kappa\Psi_t-2\epsilon^2\kappa^2\Psi^2\Psi'=0,$$

$$\Psi'_{xx}-2(x\kappa_t+m_t)\Psi'_x+2\kappa\Psi'_t-2\epsilon^2\kappa^2\Psi'^2\Psi=0. \tag{B.5}$$

Next make the co-ordinate transformation $x'=\epsilon(x\kappa+m)$, $t'=f(t)$ with $2f_t=\epsilon\kappa$. One then has

$$\partial/\partial x'=(\epsilon\kappa)^{-1}\partial/\partial x, \qquad \partial/\partial t'=2(\epsilon\kappa)^{-1}[\partial/\partial t-\kappa^{-1}(x\kappa_t+m_t)\partial/\partial x].$$

Equation (B.5) becomes in the new co-ordinates

$$\Psi_{x'x'}-\epsilon^{-1}\Psi_{t'}-2\Psi^2\Psi'=0, \qquad \Psi'_{x'x'}+\epsilon^{-1}\Psi'_{t'}-2\Psi'^2\Psi=0. \tag{B.6a,b}$$

For the SL(2,ℝ) case, all quantities are real, $\epsilon=1$ and eqs. (B.6) become the coupled system (1.4a). For the SU(2) and SU(1, 1) cases one has $\epsilon=$i, and κ, m, n and Λ are purely imaginary; also the quantities r, w and Ψ' are the negative complex conjugates of the quantities p, u, and Ψ, respectively, in the SU(2) case and are their complex conjugates in the SU(1, 1) case. Then eq. (B.6b) is the negative complex conjugate of (B.6a) in the SU(2) case is its complex conjugate equation in the SU(1, 1) case. Thus in the SU(2) case, eqs. (B.6) reduce to the single equation

$$\mathrm{i}\Psi_{t'}=-\Psi_{x'x'}-2|\Psi|^2\Psi. \tag{B.7}$$

Equation (B.7) is the attractive non-linear Schrödinger equation for the unknown $\Psi(x', t')$. For the SU(1, 1) case, the remaining equation becomes the following equation:

$$\mathrm{i}\Psi_{t'}=-\Psi_{x'x'}+2|\Psi|^2\Psi. \tag{B.8}$$

Equation (B.8) is the repulsive non-linear Schrödinger equation for the unknown $\Psi(x', t')$.

We see that, modulo gauge and co-ordinate transformations, we might just as well have taken κ to be constant a priori, and put all the integration constants equal to zero. Thus we arrive at the statement of theorem β. □

References

[1] R.P. Penrose, Nonlinear gravitons and curved twistor theory, Gen. Rel. Grav. 7 (1976) 31–52.

[2] R.S. Ward, Stationary axisymmetric space–times; a new approach, Gen. Rel. Grav. 15 (1983) no. 2.

[3] R.S. Ward, Integrable and solvable systems and relations among them, Philos. Trans. R. Soc. London A 315 (1985) 451–457.

[4] R.S. Ward, Multi-dimensional integrable systems, in: *Field Theory, Quantum Gravity and Strings*, eds. H.J. deVega and N. Sanchez, Lecture Notes in Physics, Vol. 246 (Springer, Berlin, 1986).

[5] L.J. Mason and G.A.J. Sparling, Nonlinear Schrödinger and Korteweg–de Vries are reductions of self-dual Yang–Mills, Phys. Lett. A 137 (1989) 29–33.

[6] R.S. Ward, On self-dual gauge fields, Phys. Lett. A 61 (1977) 81–82.

[7] R.S. Ward, Ansätze for self-dual Yang–Mills fields, Commun. Math. Phys. 80 (1981) 563–574.

[8] M.F. Atiyah, Geometry of Yang–Mills fields, in: Fermi Lectures (Accademia Nazionale dei Lincei Scuola Normale Superiore, Pisa, 1979).

[9] N.J. Hitchin, Monopoles and geodesics, Commun. Math. Phys. 83 (1982) 579–602.

[10] N.M.J. Woodhouse and L.J. Mason, The Geroch group and non-Hausdorff twistor spaces, Nonlinearity 1 (1988) 73–114.

[11] C.S. Gardner, J.M. Greene, M.D. Kruskal and R.M. Miura, Method for solving the Korteweg–de Vries equation, Phys. Rev. Lett. 19 (1967) 1095–1097.

[12] V.E. Zakharov and A.B. Shabat, A scheme for integrating the non-linear equations of mathematical physics by the method of inverse scattering, Funct. Anal. Appl. 8 (1974) 226–235.

[13] A.C. Newell, *Solitons in Mathematical Physics*, Conf. Ser. in Applied Mathematics, Vol. xvi (CBMS-NSF, Society for Industrial and Applied Mathematics, Philadelphia, 1985).

[14] L. Fadeev and M. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Springer, Berlin, 1987).

[15] B.G. Konopelchenko, *Non-linear Integrable Equations, Recursion Operators; Group Theoretical and Hamiltonian Structures of Soluble Equations*, Lectures Notes in Physics, Vol. 270 (Springer, Berlin, 1987).

[16] R.S. Ward, Completely solvable gauge-field equations in dimensions greater than four, Nucl. Phys. B 236 (1984) 381–396.

[17] R.S. Ward, The Nahm equations, finite gap potentials and Lamé functions, J. Phys. A 20 (1987) 2679–2683.

[18] N.M.J. Woodhouse, Twistor description of the symmetries of Einstein's equation for stationary axisymmetric space–times, Class. Quantum Grav. 4 (1987) 799–814.