J. DIFFERENTIAL GEOMETRY
30 (1989) 1-50

HARMONIC MAPS INTO LIE GROUPS (CLASSICAL SOLUTIONS OF THE CHIRAL MODEL)

KAREN UHLENBECK

Dedicated to R. F. Williams

Table of Contents

This paper deals with two aspects of the algebraic structure of $\mathscr{M}$, the space of harmonic maps from a simply-connected 2-dimensional domain (either Riemannian or Lorentzian) into a real Lie group $G_{\mathbf{R}}$, the real form of a complex group G. In the language of theoretical physics, we study the classical solutions of the chiral model. In the first part of the paper (§§1–8) we construct a representation of the loop group $\mathscr{A}(S^1, G_{\mathbf{R}})$ on $\mathscr{M}$ corresponding to the Kac-Moody Lie algebra of infinitesimal deformations observed by Dolan [8]. Here the main theorems are the description of the action on $\mathscr{M}$ (Theorem 6.1) and the description of the action of a subgroup on the space of harmonic maps into Grassmannians (Theorem 8.3). In the second part of the paper (§§9–15) we restrict to a theory which applies only when Ω is a 2-dimensional, simply-connected, *Riemannian* domain and

Received February 25, 1985, and in revised form, June 6, 1988.

the group is $G_{\mathbb{R}} = \mathrm{U}(N)$. We associate with every solution a nonnegative integer n, which we call the uniton number, and show that every harmonic map from S^2 into $\mathrm{U}(N)$ has finite uniton number. Furthermore, for any simply-connected $\Omega \subseteq S^2$ and $n < \infty$, we find $0 \leq n \leq N-1$. We also give a construction for obtaining every harmonic map of uniton number n in a unique way from a harmonic map of uniton number $n-1$ (Theorem 1.46). Moreover, this construction applies also to harmonic maps into complex Grassmannians (Theorem 15.3). The moduli spaces of harmonic maps are described in terms of holomorphic sub-bundles of holomorphic bundles over Ω satisfying certain additional conditions.

The last section (§16) contains questions which are left unanswered and might profitably be investigated further.

The results described in this paper are the product of what was originally a joint research project with Steve Smith, who was interested in the group theoretic aspects of the Kac-Moody representations. Unfortunately, the actual representation found of the affine Kac-Moody Lie algebra is rather dull from the group theoretic viewpoint (and its importance in differential geometry is as yet unclear). However, the results in this paper owe a lot to our joint discussions.

Discussions with Louis Crane and Dan Freed were very encouraging and helpful. In particular, Louis found many Russian references, some of which are listed in the references. He also analyzed the action of the loop group on Yang-Mills [6]. Iz Singer directed the author to references [3] and [4]. The existence of the S^1 action of §7 was pointed out by Chuu-lian Terng.

After the fact, the author discovered that there is a very large physics literature on the subject. The bibliography of this paper is not complete. By and large articles have been included which will be useful to the mathematical audience for which this paper is written. (More comments on the literature can be found in §16.) Also, the author cannot claim that the first part of the paper is original, since much of it can now be found in many versions in the physics literature. However, the present exposition should still be of use to mathematicians, and the author apologizes where it duplicates the existing literature.

PART I

1. Basic formulas and characteristic notation.

In the first part of this paper (§§1–8) we discuss the theory of harmonic maps from a one-connected complex domain $\Omega \subset \mathbb{R}^2 \simeq \mathbb{C}$ into a compact

Lie group $G_{\mathbf{R}}$, which is the real form of a complex group G. We actually assume $G_{\mathbf{R}} = \mathrm{U}(N)$ and $G = \mathrm{GL}(N, \mathbb{C})$, although by constructing a representation $\rho\colon G \to \mathrm{GL}(N, \mathbb{C})$ and $\rho|G_{\mathbf{R}} \to \mathrm{U}(N)$, one can check our proofs are valid for many other G. In particular, our reasons for using $\mathrm{U}(N)$ instead of $\mathrm{SU}(N)$ are more fundamental than we can explain, since the "degree" in $\mathrm{U}(N)$ does describe instanton number. Determinant and trace properties trivially factor out of the solutions, but not out of the constructions. In this first part, the theory is also valid for domains $\Omega \subset E^{1,1}$. One needs only to use the characteristic coordinates $\xi = x - t$ and $\eta = x + t$ instead of $x - iy$ and $x + iy$ to make all the computations valid. Since the global theory is completely different for $E^{1,1}$, and from the point of view of differential geometry we are ultimately interested in maps $S^2 \to \mathrm{U}(N)$, we restrict to $\Omega \subset \mathbb{R}^2$.

If $s\colon \Omega \to \mathrm{U}(N) = G_{\mathbf{R}}$, the energy is

$$\begin{aligned} E(s) &= \frac{1}{2} \iint_\Omega \left(\left| s^{-1} \frac{\partial s}{\partial x} \right|^2 + \left| s^{-1} \frac{\partial s}{\partial y} \right|^2 \right) dx\, dy \\ &= -\frac{1}{2} \iint_\Omega \operatorname{tr} \left(\frac{\partial s}{\partial x} \frac{\partial s}{\partial x}^{-1} + \frac{\partial s}{\partial y} \frac{\partial s}{\partial y}^{-1} \right) dx\, dy. \end{aligned}$$

For $a \in \mathfrak{g}_{\mathbf{R}}$, $a^* = \bar{a}^T = -a$ and $|a|^2 = \operatorname{tr} aa^*$. Since $s^{-1}(q) = s^*(q)$ for $q \in \Omega$, we can replace s^{-1} by s^*. However, we shall need formulas for the complex group $G = \mathrm{GL}(N, \mathbb{C})$, where s^* must be replaced by s^{-1}. We therefore consistently use the expression above.

The Euler-Lagrange equations for the integral E are the equations

$$(1) \qquad \frac{\partial}{\partial x} \left(s^{-1} \frac{\partial s}{\partial x} \right) + \frac{\partial}{\partial y} \left(s^{-1} \frac{\partial s}{\partial y} \right) = 0.$$

Solutions to such equations are called *harmonic maps* from $\Omega \subset \mathbb{R}^2$ into $\mathrm{GL}(N, \mathbb{C}) = G$. When $s^{-1}(q) = s^*(q)$, then s is a harmonic map into $\mathrm{U}(N) = G_{\mathbf{R}}$. Recall that if $s\colon S^1 \to \mathrm{U}(N)$, then s is a geodesic if

$$\frac{d}{d\theta} \left(s^{-1} \frac{d}{d\theta} s \right) = 0, \quad \text{or} \quad s^{-1} \frac{d}{d\theta} s \equiv a, \quad \text{or} \quad s(\theta) = s(0) e^{a\theta},$$

$a \in \mathfrak{g}_{\mathbf{R}} = \mathfrak{u}(N)$. The harmonic map equations are therefore very natural extensions of the geodesic equation. Let $A_x = \frac{1}{2} s^{-1}\, \partial s / \partial x$, $A_y = \frac{1}{2} s^{-1}\, \partial s / \partial y$ and define the one-form

$$(2) \qquad A = A_x\, dx + A_y\, dy = \tfrac{1}{2} s^{-1}\, ds.$$

Note both A_x and A_y are skew Hermitian matrices. In terms of A, we can write (1) as

$$(3) \qquad d^*A = \frac{\partial}{\partial x}(A_x) + \frac{\partial}{\partial y}(A_y) = 0.$$

As a consequence of the definition of A, we have the identity

$$(4) \qquad dA + [A, A] = \frac{\partial}{\partial x}A_y - \frac{\partial}{\partial y}A_x + 2[A_x, A_y] = 0.$$

In a simply connected region, equations (3) and (4) and a specification of the basepoint $s(p)$ at $p \in \Omega$ are equivalent to (1). To see this, either construct s by line integrals from A and $s(p)$, or note $2(dA + [A, A])$ is the curvature of the connection $d + 2A$ in the bundle $\Omega \times \mathbb{C}^N$. Then $s\colon \Omega\times\mathbb{C}^N \to \Omega\times\mathbb{C}^N$ is the change in gauge (trivialization) which transforms the flat connection $d + 2A$ into the trivial connection d. We also record the second variation formula, or the linearization of (1) at a solution s. This is expressed in terms of $\Lambda = s^{-1}\delta s$, where $\Lambda\colon \Omega \to \mathrm{u}(N)$. We call Λ a Jacobi field if

(5)

$$0 = d^*(d\Lambda + 2[A, \Lambda]) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\Lambda + 2[A_x, \Lambda]\right) + \frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}\Lambda + 2[A_y, \Lambda]\right).$$

(3) and (4) can be rewritten in a number of forms which are very reminiscent of classical integrable systems. For example, if we let y be time, $\partial/\partial y =^\bullet$, $A_x - iA_y = P$ and $A_x + iA_y$, then

$$(6) \qquad \dot{P} = i\frac{\partial}{\partial x}P + i[Q, P], \qquad \dot{Q} = i\frac{\partial}{\partial x}Q - i[P, Q].$$

The conditions that A_x and A_y are skew-Hermitian translate to $P^* = -Q$ (which is of course a reality condition). We are in fact discussing exactly the theory of these equations in this paper. The choice of time t was however arbitrary, and the Hamiltonian formulations available are not satisfactory to the author. For our theory, the correct change of coordinates seems to be *characteristic* coordinates. Since (1) can be written

$$\Delta s - \frac{\partial s}{\partial x}s^{-1}\frac{\partial s}{\partial x} - \frac{\partial s}{\partial y}s^{-1}\frac{\partial s}{\partial y} = 0,$$

its top order term is Δ and characteristic coordinates for the system are $(z = x + iy, \bar{z} = x - iy)$. Recall that $\partial = \partial/\partial z = \frac{1}{2}(\partial/\partial x - i\partial/\partial y)$, $\bar{\partial} = \partial/\partial z = \partial/\partial x + i\partial/\partial y$, $dz = dx + idy$, and $d\bar{z} = dx - idy$. The harmonic map equations (1) are now (we consistently use $\partial/\partial z = \partial$, $\partial/\partial\bar{z} = \bar{\partial}$)

$$(7) \qquad \bar{\partial}(s^{-1}\partial s) + \partial(s^{-1}\bar{\partial}s) = 0.$$

The one-form A is, of course, still

$$A = A_{\bar z}\,d\bar z + A_z\,dz = \tfrac{1}{2}s^{-1}\,ds, \tag{8}$$

so $A_{\bar z} = \frac{1}{2}(Ax + iA_y)$ and $A_z = \frac{1}{2}(A_x - iA_y)$. We write the pair (3) and (4) together as

$$\begin{aligned} &\text{(a)}\quad d^*A = \bar\partial A_z + \partial A_{\bar z} = 0,\\ &\text{(b)}\quad dA + [A,A] = \bar\partial A_z - \partial A_{\bar z} + 2[A_{\bar z}, A_z] = 0. \end{aligned} \tag{9}$$

The second variation formula is now

$$\bar\partial(\partial\Lambda + 2[A_{\bar z},\Lambda]) + \partial(\bar\partial\Lambda + [2A_{\bar z},\Lambda]) = 0. \tag{10}$$

Also, the pair (6) makes quite a bit of sense as a pair equivalent to (9):

$$\begin{aligned} &\text{(a)}\quad \bar\partial A_z + [A_{\bar z}, A_z] = 0,\\ &\text{(b)}\quad \partial A_{\bar z} + [A_z, A_{\bar z}] = 0. \end{aligned} \tag{11}$$

We shall use (7)–(11) throughout the paper. All the formulas can be written in terms of (x,y) also, but they are far less illuminating.

As one final comment, note that out of habit, everything is pulled back on the left ($\Lambda = s^{-1}\delta s$, $A = \frac{1}{2}s^{-1}\,ds$ and so forth). Formulas (up to sign changes) are valid for the right pull-backs. This is explained fully as we proceed, especially in the explanation of the S^1 action, where the representation of (-1) acts to take s to s^{-1}. If s is harmonic, then s^{-1} is also (see Theorem 8.3 for an explicit statement of this fact).

$$\begin{gathered} \delta(s^{-1}) = -s^{-1}\delta s s^{-1} = -\Lambda s^{-1},\\ A = -\tfrac{1}{2}d(s^{-1})s. \end{gathered}$$

The switch from s to s^{-1} clearly reverses my habit of using the left pull-back to the right convention. There is no preferred choice and both are used in the literature of differential geometry.

2. The canonical associated linear problem or Lax pair.

We recall that in considering the equations (9), the second is the statement that the curvature of the connection $(\bar\partial + 2A_{\bar z}, \partial + 2A_z)$ in $\Omega \times \mathbb{C}^N$ vanishes. In seeking reformulations of (9) which will give us insight, we discover an interesting fact which is the basis for our entire theory. This is very well known although it is not known to the author who deserves the credit for the discovery.[1]

[1] *Added in proof.* This is due to K. Pohlmeyer, *Integrable Hamiltonian systems and interactions through constraints*, Comm. Math. Phys. **46** (1976) 207–221.

Theorem 2.1. *Let Ω be simply connected and $A\colon \Omega \to T^*(\Omega) \otimes \mathfrak{g}$. Then $2A = s^{-1}\,ds$, where s is harmonic if and only if the curvature of the connection in $\Omega \times \mathbb{C}^N$*

$$D_\lambda = (\bar\partial + (1-\lambda)A_{\bar z}, \partial + (1-\lambda^{-1})A_x z)$$

vanishes for all $\lambda \in \mathbb{C}^* = \mathbb{C} - 0$.

Proof. We write out the curvature equations and expand in λ:

$$\begin{aligned}\bar\partial(1-\lambda^{-1})A_x &- \partial(1-\lambda)A_{\bar z} + [(1-\lambda)A_{\bar z}, (1-\lambda^{-1})A_x] \\ &= \lambda(\partial A_{\bar z} - [A_{\bar z}, A_z]) + \lambda^0(\bar\partial A_z - \partial A_{\bar z} + 2[A_{\bar z}, A_z]) \\ &\quad + \lambda^{-1}(-\bar\partial A_z - [A_{\bar z}, A_z]).\end{aligned}$$

Clearly if $A = \frac{1}{2}s^{-1}\,ds$ for s harmonic, this curvature vanishes since the coefficients of λ^α, $\alpha = -1, 0, 1$, are (9)(b) and (11)(a), (b). On the other hand, if all the curvatures vanish, these coefficients vanish and we get (11)(a), (b) and (9)(b) to hold. However, (9)(a) is the sum of (11)(a) and (11)(b), and we have the existence of our harmonic s.

Naturally, again since Ω is simply-connected, our next step is to trivialize the connections

$$D_\lambda = (\bar\partial + (1-\lambda)A_{\bar z}, \partial + (1-\lambda^{-1})A_z).$$

This involves solving simultaneously the linear equations

$$\bar\partial E_\lambda = (1-\lambda)E_\lambda A_{\bar z}, \qquad \partial E_\lambda = (1-\lambda^{-1})E_\lambda A_z \tag{12}$$

for $\lambda \in \mathbb{C}^*$. Since the curvature vanishes, we can do this, and a solution is uniquely determined by prescribing $E_\lambda(p)$ for $p \in \Omega$ any base point. For the purposes of the first part of this paper, we will need to do this, and we assume $s(p) = I$ at the basepoint $p \in \Omega$ and choose $E_\lambda(p) = I$ (I is the identity of the group). Since $\tilde s = s^{-1}(p)s$ is harmonic if s is, and $\tilde s(p) = I$, we lose nothing by assuming $s(p) = I$. In Part II, and possibly in general, this normalization is awkward and we abandon it. However, note that if $\lambda = 1$, $dE_1 \equiv 0$ and $E_1 \equiv E(p)$, which is I under our convention. The condition $E_1 \equiv I$ we find useful to preserve under all circumstances. At this point we have very nearly given a proof of the next theorem.

Theorem 2.2. *If s is harmonic and $s(p) \equiv I$, then there exists a unique $E\colon \mathbb{C}^* \times \Omega \to G$ satisfying equations* (12) *with*

(a) $E_1 \equiv I$,

(b) $E_{-1} = s$,

(c) $E_\lambda(p) = I$.

Moreover, E is analytic and holomorphic in $\lambda \in \mathbb{C}^*$. Finally, if s is unitary, E_λ is unitary for $|\lambda| = 1$.

Proof. The existence and uniqueness of E_λ follow from understanding the curvature condition of (2.1) (in whatever fashion the reader has already chosen). The regularity follows from the same conditions and the analyticity of s. We check that from (12) we get

$$-\bar\partial E_\lambda^{-1} = (1-\lambda)A_{\bar z}E_\lambda^{-1}, \qquad -\partial E_\lambda^{-1} = (1-\lambda^{-1})A_z E_\lambda^{-1}.$$

The adjoint of this pair of equations is

$$-\partial(E_\lambda^{-1})^* = (1-\bar\lambda)(E_\lambda^{-1})^*(A_{\bar z})^*, \qquad -\bar\partial(E_\lambda^{-1})^* = (1-(\lambda^{-1})^*)(E_\lambda^{-1})^*(A_z)^*.$$

When s is unitary, $(A_z)^* = -A_{\bar z}$. If $|\lambda| = 1$ or $\bar\lambda = \lambda^{-1}$, the uniqueness gives $(E_\lambda^{-1})^* = E_\lambda$. This finishes the proof.

The converse now reduces the original harmonic map equation (1) to an equation with an entirely different flavor.

Theorem 2.3. *Suppose $E\colon \mathbb{C}^* \times \Omega \to G$ is analytic and holomorphic in the first variable, $E_1 \equiv I$ and the expressions*

$$\frac{E_\lambda^{-1}\bar\partial E_\lambda}{1-\lambda}, \qquad \frac{E_\lambda^{-1}\partial E_\lambda}{1-\lambda^{-1}}$$

are constant in λ. Then $s = E_{-1}$ is harmonic.

Proof. We let

$$A = \frac{1}{2}s^{-1}\,ds = E_\lambda^{-1}\left(\left(\frac{1}{1-\lambda}\right)\bar\partial E_\lambda\,d\bar z + \left(\frac{1}{1-\lambda^{-1}}\right)\partial E_\lambda\,dz\right),$$

$$\bar\partial E_\lambda = (1-\lambda)E_\lambda A_{\bar z}, \qquad \partial E_\lambda = (1-\lambda^{-1})E_\lambda A_z.$$

It follows that the connections $\bar\partial + (1-\lambda)A_{\bar z}$, $\partial + (1-\lambda^{-1})A_z$ have zero curvature, which implies $E_{-1} = s$ is harmonic.

From now on we replace the equations for s by the equations for E_λ. In other words, our variables are

$$E\colon C^* \times \Omega \to \mathrm{GL}(N, \mathbb{C})$$

with $E_1 \equiv I$. The holomorphic maps $Q\colon C^* \to \mathrm{GL}(N,\mathbb{C})$ with $Q_1 = I$ act by left multiplication, since if E satisfies the equations of Theorem 2.3, QE does also. The old harmonic map E_{-1} becomes $Q_{-1}E_{-1}$, which is also harmonic. The *unitary condition* is the requirement that E_λ be unitary for $|\lambda| = 1$. Note that we may expand E_λ in a Laurent series,

$$E_\lambda = \sum_{\alpha=-\infty}^{\infty} T_\alpha \lambda^\alpha,$$

where $T_\alpha\colon \Omega \to N \times N$ matrices ($= \mathfrak{g}$). The unitarity condition is

$$E_\lambda^{-1} = \sum_{\alpha=-\infty}^{\infty} T_\alpha^* \lambda^{-\alpha}.$$

The form of the unitarity or reality condition which we shall use most often is one of the equations

$$E_\lambda E^*_{\sigma(\lambda)} = I \quad \text{or} \quad E_\lambda^{-1} = E^*_{\sigma(\lambda)}. \tag{13}$$

Here $\sigma\colon C^* \to C^*$ is the anti-conformal involution $\sigma(\lambda) = (\bar{\lambda})^{-1}$. In Yang-Mills theory, the involution σ is replaced by $\lambda \to -(\bar{\lambda})^{-1}$, which is a great deal more difficult to treat [6].

3. The Dolan representation

In this section we review the infinitesimal construction of the representation of the affine Kac-Moody Lie algebra based on $G_{\mathbf{R}}$ due to Dolan [8], [9] and Chau, Lin and Shi [3] for both the harmonic map equation and the anti-self-dual Yang-Mills fields. These papers are based on the assumption that the solution space of an equation is a differentiable manifold $\mathscr{M}$ with the tangent space given by the linearized equations.

Definition 3.1. A *conservation law* for an equation is a vector-field along the solution manifold $\mathscr{M}$. This is actually the same thing as an *infinitesimal symmetry.*

The Lie algebra structure of the conservation laws, or infinitesimal symmetries, is just the Lie bracket of the vector fields. So to verify the representation, it is necessary to do some very messy computations with the Lie brackets. Since we obtain directly a representation of a group, we can avoid these computations. We also transform the "hidden" symmetry into a directly constructable one.

What we actually describe in this section is how to canonically obtain a large number of Jacobi fields, or solutions to the linearized harmonic map equation (5) or equivalently (10). In the appropriate context, these solutions would represent vector fields, and could be integrated to give flows. We avoid this by going to the extended or so-called finite problem and solving in another fashion.

Our first construction of solutions to the linearized equations follows that of Chau, Lin and Shi [3] for anti-self-dual Yang-Mills equations. Small elements a, b, c, etc. are reserved for elements of the Lie algebra, either $\mathfrak{g}_{\mathbf{R}}$ or $\mathfrak{g} = \mathrm{gl}(N, \mathbb{C})$ in our general case.

Let us review this linearized equation. For every local harmonic map $s\colon \Omega \to \mathrm{U}(N)$, we obtain $A = \frac{1}{2}s^{-1}\,ds$. Associated with s is the linear equation (10) for $\Lambda = s^{-1}\delta s$. Since $d^*A = 0$, this equation can be written as either

$$d^*(d\Lambda + 2[A, \Lambda]) = 0 \tag{14}$$

or

$$d^*d\Lambda + 2[A; d\Lambda] = 0. \tag{15}$$

(14) and (15) are equivalent.

Theorem 3.2. *If* $\lambda \in \mathbb{C}^*$, *then* $\Lambda_\lambda(a) = E_\lambda^{-1} a E_\lambda$ *solves the linearized equation* (14). *If* $|\lambda| = 1$ *and* a *is skew, then* $\Lambda_\lambda(a)\colon \Omega \to \mathrm{u}(N)$ *is also skew.*

Proof. The skew-Hermitian condition follows from the fact that E_λ is unitary when $|\lambda| = 1$. To do the computation, $P_{\bar z} = (1-\lambda)A_{\bar z}$, $P_z = (1-\lambda^{-1})A_z$, and $\Lambda = \Lambda - \lambda(a)$ temporarily. For $\lambda = 1$, $\Lambda = a$ is automatically a solution. For $\Lambda \neq a$

$$\bar\partial \Lambda = [\Lambda, P_{\bar z}], \qquad \partial \Lambda = [\Lambda, P_z].$$

To show Λ satisfies the linearized equation, we verify the form (15) of the equation:

$$\begin{aligned}
&(\partial\bar\partial + \bar\partial\partial)\Lambda + 2[A_{\bar z}, \bar\partial\Lambda] + 2[A_z, \partial\Lambda] \\
&\quad = \partial[\Lambda, P_{\bar z}] + \bar\partial[\Lambda, P_z] + [2A_{\bar z}, \bar\partial\Lambda] + [2A_{\bar z}, \partial\Lambda] \\
&\quad = [2A_{\bar z} - P_{\bar z}, [\Lambda, P_z]] + [2A_z - P_z, [\Lambda, P_{\bar z}]] + [\Lambda, \partial P_{\bar z} + \bar\partial P_z] \\
&\qquad + [\Lambda, (1-\lambda)\bar\partial A_z + (1-\lambda^{-1})\partial A_{\bar z}] \\
&\qquad + (\lambda - \lambda^{-1})([A_{\bar z}, [\Lambda, A_z]] - [A - z, [\Lambda, A_{\bar z}]]) \\
&\quad = [\Lambda, (1-\lambda)\bar\partial A_z + (1-\lambda^{-1})\partial A_{\bar z} + (\lambda - \lambda^{-1})[A_z, A_{\bar z}]].
\end{aligned}$$

This vanishes according to (9) and (11).

Since $\Lambda_\lambda(a)$ is linear in a and holomorphic in λ, a large number of solutions of the linearized harmonic map equation can be found by expanding $\Lambda_\lambda(a)$ appropriately in power series about various points in λ. The points which are in fact used are $\lambda = 1$, where $E_1 \equiv I$, and $\Lambda_1(a) = a$ and $\lambda = -1$, where $E_{-1} \equiv s$ and $\Lambda_{-1}(a) = sas^{-1}$. Note that the first term in the expansion about $\lambda = 1$, or $\delta s = s\Lambda = sa$ represents the infinitesimal action of the group by multiplication of solutions on the right by constants, whereas the first term at $\lambda = -1$ is $\delta s = s\Lambda = as$, or the infinitesimal representative of the action of the group on the left by multiplication by constants. Both actions clearly preserve the harmonic map equation. To preserve the condition $s(p) = I$, $\Lambda_\lambda(a)$ evaluated at $p \in \Omega$ is to be zero. The allowable variation is $\delta s = sa - as$ or $\Lambda = \Lambda_1(a) - \Lambda_{-1}(a)$.

The change of variables used to construct the representation due to the physicists is completely mysterious without the following rather tedious development (which is actually completely unnecessary for understanding our final development but leads to some beautiful looking contour integration formulas).

Theorem 3.3. *Suppose Λ is a solution of the linear equation* (14). *Then there exist two additional solutions Λ^+ and Λ^- to* (14) *satisfying*

$$d\Lambda^+ = *\,(d\Lambda + 2[jA, \Lambda]), \tag{16}$$

$$d\Lambda = *\,(d\Lambda^- + 2[A, \Lambda^-]). \tag{17}$$

If $\Lambda^+(0) = 0$ and $\Lambda^-(0) = 0$, the solutions are unique.

Proof. Since the second variation equations read $d(*(d\Lambda + 2[A, \Lambda])) = 0$, and Ω is simply connected, the existence of Λ^+ is assured. To see that Λ^+ satisfies the equation, note we can write it as

$$d(*d\Lambda^+) + [2A, *d\Lambda^+] = 0.$$

However, $d+2A$ is the flat connection, and $[(d+2A), (d+2A)] =$ curvature $= 0$. It follows that

$$d(*d\Lambda^+) + [2A, *d\Lambda^+] = [d + 2A, d + 2A]\Lambda^+ = 0.$$

The existence of Λ^- and its properties follow exactly in the same way. It is clear that the equation for Λ^- is implied by $d^2A = 0$, whereas existence follows from the fact that $d + 2A$ is the flat connection in another gauge. So if $(d + 2A)Q = 0$, $Q = d\Lambda^- + 2[A, \Lambda^{-1}]$. Another interpretation is that we obtain the lowering operation $\Lambda \to \Lambda^-$ by raising $s\Lambda s^{-1}$, a solution to the linearized equation for s^{-1}, and conjugating back $s^{-1}(s\Lambda s^{-1})^+ s = \Lambda^-$ to a solution of the linearized equation for s.

In order for this operation to be unique, we emphasize that the normalization $\Lambda(p) = 0$ at $p \in \Omega$ is necessary. Also, it follows that if Λ is skew-Hermitian, then $(\Lambda^+)^* = -\Lambda^+$ and $(\Lambda^-)^* = -\Lambda$ from the uniqueness. For the same reason $(\Lambda^+)^- = (\Lambda^-)^+ = \Lambda$ if $\Lambda(p) = 0$.

We now describe the representation of the affine Kac-Moody Lie algebra based on $\mathfrak{g}_{\mathbf{R}}$. First let us recall the definition

Definition 3.4. If $\mathfrak{g}_{\mathbf{R}}$ is a lie algebra, the (uncompleted) affine Kac-Moody Lie algebra $\sum(\mathfrak{g}_{\mathbf{R}})$ based on $\mathfrak{g}_{\mathbf{R}}$ is the ring of finite power series $\sum_\alpha a_\alpha t^\alpha$, $a_\alpha = 0$ for $\alpha < k$ and $\alpha > k'$ with $a_\alpha \in \mathfrak{g}_{\mathbf{R}}$. The Lie bracket operation is

$$\left[\sum_\alpha a_\alpha t^\alpha, \sum_\beta b_\beta t^\beta\right] = \sum_\gamma \left(\sum_\alpha [a_\alpha, b_{\gamma-\alpha}]\right) t^\gamma.$$

We do not actually attempt to discuss the completion (the ring of all formal, infinite power series). However, from our point of view as analysts we see

$$\sum(\mathfrak{g}_{\mathbf{R}}) \subset C^\infty(S^1, \mathfrak{g}_{\mathbf{R}}).$$

By setting $t = e^{i\theta}$, we identify

$$\sum_\alpha a_\alpha t^\alpha \to f(\theta) = \sum_\alpha a_\alpha e^{i\alpha\theta},$$

where f is a map from S^1 into $\mathfrak{g}$ with finite Fourier series. The Lie algebra structure is $[f, g](\theta) = [f(\theta), g(\theta)]$. Another point of view is that we identify the formal element $\sum_\alpha a_\alpha t^\alpha$ with the holomorphic map from $\mathbb{C}^*$ to $\mathfrak{g}$ with finite poles at 0 and ∞ which is given by considering $t \in \mathbb{C}^*$. (Careful, $t \neq \lambda$!) This last approach is the one we will take.

We now describe the representation. We fix a harmonic map s. At s, we define the representation on a basis at^α by $at^\alpha \to \Lambda^\alpha(a)$. The elements $\Lambda^\alpha(a)$ are defined used the raising and lowering operators of (16) and (17):

$$\Lambda^\alpha(a) = (\Lambda^{\alpha-1}(a))^+, \quad \Lambda^{-\alpha}(a) = (\Lambda^{-\alpha+1}(a))^-, \quad \Lambda^0(a) = a - s^{-1}as.$$

By our discussion above, at each solution s this is well defined. However, it does not show that the space of solutions is a manifold. If the space of solutions is a manifold $\mathscr{M}$ with tangent space given by solutions to the linearized problem, then $\Lambda^\alpha(a)$ is a vector field along the solution space. This does not begin to deal with the calculation of Lie brackets, however.

We now relate this discussion to the earlier construction. Note that $E_\lambda(p) \equiv I$ does determine E_λ uniquely. The expression from Theorem 3.2, $E_\lambda^{-1} a E_\lambda = \Lambda_\gamma(a)$, is the unique map $\Lambda_\lambda(a)\colon \Omega \to \mathfrak{g}_{\mathbb{C}}$ satisfying

$$\begin{aligned}
\Lambda_\lambda(a)(p) &= E_\lambda^{-1}(p) a E_\lambda(p),\\
\bar\partial(\Lambda_\lambda(a)) &= (\lambda - 1)[A_{\bar z}, \Lambda_\lambda(a)],\\
\partial(\Lambda_\lambda(a)) &= (\lambda^{-1} - 1)[A_z, \Lambda_\lambda(a)].
\end{aligned}$$

A simple manipulation gives

$$\begin{aligned}
\bar\partial\Lambda_\lambda(a) &= \left(\frac{\lambda-1}{\lambda+1}\right)(\bar\partial\Lambda_\lambda(a) + 2[A_{\bar z}, \Lambda_\lambda(a)]),\\
\partial\Lambda_\lambda(a) &= -\left(\frac{\lambda-1}{\lambda+1}\right)(\partial\Lambda_\lambda(a) + 2[A_z, \Lambda_\lambda(a)]).
\end{aligned}$$

We compare with our previous construction, and set $i\tau = (\lambda - 1)/(\lambda + 1)$. The singularities at $\lambda = (0, \infty)$ go to singularities at $\pm i$, $\lambda = 1$ becomes $\tau = 0$, and $\lambda = -1$ becomes $\tau = \infty$. In our new variable τ

$$\Lambda_\tau(a) - a = \tau(\Lambda_\tau(a))^+.$$

Theorem 3.5. *An equivalent description of the representation of $\sum(\mathfrak{g}_{\mathbb{R}})$ is*

$$\Lambda^\alpha(a) = \frac{1}{\alpha!}\left(\frac{\partial}{\partial\tau}\right)^\alpha |_{\tau=0}\Lambda_\tau(a), \qquad \Lambda^{-\alpha}(a) = -\frac{1}{\alpha!}\left(\frac{\partial}{\partial\tau}\right)^\alpha |_{\tau=0}\Lambda_{\tau^{-1}}(a)$$

for $a > 0$ and

$$\Lambda^0(a) = a - s^{-1}as = \Lambda_0(a) - \Lambda_\infty(a).$$

Proof. The theorem is straightforward, modulo the normalization problem. Formally the power series

$$R_\tau(a) = \sum_{\alpha=1}^{\infty} \Lambda^\alpha(a)\tau^\alpha, \quad |\tau| < \varepsilon, \qquad L_\tau(a) = \sum_{\alpha=1}^{\infty} \Lambda^{-\alpha}(a)\tau^{-\alpha}, \quad |\tau| > \varepsilon,$$

solve

$$R_\tau(a) = \tau(R_\tau(a) + \Lambda_a^0)^+, \qquad L_\tau(a) = \tau^{-1}(L_\tau(a) + \Lambda_a^0)^-.$$

Note $\Lambda^0(0) = \Lambda_0(a) - \Lambda_\infty(a)$ is correct, and it is useful to see that $(s^{-1}as)^+ = 0$, $a^- = 0$. A little calculation shows $R_\tau(a) = \Lambda_\tau(a) - \Lambda_0(a)$ for $|\tau| < \varepsilon$, and $L_\tau(a) = -\Lambda_\tau(a) + \Lambda_\infty(a)$ for $|\tau| > \varepsilon$.

4. The variation formulas for the extended solution

This section is needed to connect the discussion of the previous section with the vector space representations given in §§5 and 6. We already know that for $\gamma \in \mathbb{C}^*$ and $a \in \mathfrak{g}$

$$\delta s = \delta E_1 = s\Lambda_\gamma(a) = E_{-1}E_\gamma^{-1}aE_\gamma$$

gives an infinitesimal variation of the harmonic maps. Morally we know this must determine δE_γ for $\lambda \in \mathbb{C}^*$. However, the actual expression for δE_λ can be calculated. Note we actually use $\delta s = s(E_\gamma^{-1}aE_\gamma - s^{-1}as)$ to preserve the normalization $E_\lambda(p) = I$.

Proposition 4.1. Given the first variation $\delta s = s(E_\gamma^{-1}aE_\gamma - s^{-1}as)$, we obtain

$$\delta E_\lambda = \frac{(\lambda - 1))(\gamma + 1)}{2(\lambda - \gamma)}(E_\lambda E_\gamma^{-1}aE_\gamma - aE_\lambda).$$

Proof. We do the calculation for $\bar\partial$. The given expression is invariant under $\lambda \to \lambda^{-1}$ and $\gamma \to \gamma^{-1}$, so the exact computation with $\lambda \to \lambda^{-1}$ and $\gamma \to \gamma^{-1}$ holds for ∂. First we note

$$\delta A_{\bar z} = \frac{1}{2}\bar\partial(E_\gamma^{-1}aE_\gamma) + [A_{\bar z}, E_\gamma^{-1}aE_\gamma] = -\frac{1}{2}(\gamma + 1)[E_\gamma^{-1}aE_\gamma, A_{\bar z}].$$

We also have

$$\bar\partial(\delta E_\lambda E_\lambda^{-1}) = (1 - \lambda)E_\lambda \delta A_{\bar z} E_\lambda^{-1} - \frac{1}{2}(\gamma + 1)(1 - \lambda)E_\lambda[E_\gamma^{-1}aE_\gamma, A_{\bar z}]E_\lambda^{-1}.$$

However

$$\bar\partial(E_\lambda E_\gamma^{-1}aE_\gamma E_\lambda^{-1}) = (\lambda - \gamma)E_\lambda[E_\gamma^{-1}aE_\gamma, A_{\bar z}]E_\lambda^{-1}.$$

Remember the normalization $E_\lambda(p) = E_\gamma(p)$ and $\delta E_\lambda(p) = 0$. These calculations yield the candidate formula

$$\delta E_\lambda E_\lambda^{-1} = -\frac{\frac{1}{2}(\gamma+1)(1-\lambda)}{(\lambda-\gamma)}(E_\lambda(E_\gamma^{-1} a E_\gamma)E_\lambda^{-1} - a).$$

Repeat the calculation for ∂ to discover that this answer is indeed the only possible correct one.

In order to connect this formula with our earlier development, we shall need to change variables to $i\tau = (\lambda-1)/(\lambda+1)$ and $i\xi = (\gamma-1)/(\gamma+1)$. We use the notation $F_\tau = E_\lambda$. These changes in variables are unfortunately terribly confusing. The singularities now are at $\tau, \xi = \pm i$. In these variables $F_0 \equiv I$, $F_\infty = s$ and

$$\delta F_\tau = \tau \frac{(F_\tau F_\xi^{-1} a F_\xi - a F_\tau)}{\tau - \xi}.$$

Consider how to obtain the representation of $\sum_{\alpha>0} a_\alpha \xi^\alpha = f(\xi)$. For $\alpha > 0$, we develop the expression in contour integrals $|\xi| < 1$ to get the appropriate representation

$$\delta F_\tau = \frac{\tau}{2\pi i} \oint\limits_{\substack{|\xi|<1 \\ \text{enclosing } 0}} \sum_\alpha \frac{F_\tau F_\xi^{-1} a_\alpha \xi^{-\alpha-1} F_\xi - a_\alpha \xi^{-\alpha-1} F_\tau}{t-\xi} d\xi.$$

On the other hand, if $f(\xi) = \sum_{\alpha>0} a_\alpha \xi^{-\alpha}$, we want minus the Laurent expansion at $\xi = \infty$:

$$\delta F_\tau = \frac{\tau}{2\pi i} \oint\limits_{\substack{|\xi|<1 \\ \text{enclosing } \infty}} \sum_\alpha \frac{F_\tau F_\xi^{-1} a_\alpha \xi^{\alpha-1} F_\xi - a_\alpha \xi^{\alpha-1} F_\tau}{\tau-\xi} d\xi.$$

Note that the expression inside the contour integrals is holomorphic at every point $\xi \neq (0,\infty)$ and $\xi \neq \pm i$. As long as the Fourier series is finite, we encounter no difficulties with convergence. One can check that the constant term is correct, also.

Theorem 4.2. *Let $f(\xi) = \sum_{\alpha=N}^{M} a_\alpha \xi^\alpha$ for $a_\alpha \in Q$. Then the variation of F_τ associated with f is just*

$$\delta F_\tau = \frac{\tau}{2\pi i} \oint \frac{(F_\tau F_\xi^{-1} f(\xi^{-1}) F_\xi - f(\xi^{-1}) F_\tau)}{\tau - \xi} \frac{d\xi}{\xi}$$

over any contour enclosing the two points $\pm i$ in a counterclockwise direction and avoiding 0.

Proof. Provided one believes the formulas, one need only note that the contour about zero and the reverse contour about infinity can be deformed into contours about $\pm i$. Another description of an allowable contour is one just below $|\xi| = 1$ follows counterclockwise and one just above $|\xi| = 1$ followed clockwise. This description we gave in the theorem has the advantage that we can extend the representation of finite Fourier series to include a representation of any function holomorphic in the neighborhoods of $\pm i$!

The striking thing about the development is that in this section we have disposed entirely of the variable $(z, \bar{z}) = q$ which looked so very important. The business end of the construction is entirely in the variable τ.

Finally, it might be useful to write the formulas back in terms of λ and γ. We make the appropriate change of variable. Now the singularities are at 0 and the contour $|\xi| = 1$ is $\gamma =$ imaginary. Also, let $v(\xi) = f(\xi^{-1})$.

Theorem 4.3. *If v is holomorphic on $\mathbb{C} \cup \{\infty\} - \{+1, -1\}$, it represents the variation*

$$\delta E_\lambda = \frac{(\lambda - 1)}{2\pi i} \oint \frac{E_\lambda E_\gamma^{-1} v(\gamma) E_\gamma - v(\gamma) E_\lambda}{(\gamma - 1)((\lambda - \gamma)} \, d\gamma,$$

where the contours enclose 1 *and* -1. *If we deform to contours enclosing* 0 *and* ∞ *we obtain a simpler formula which holds for any v holomorphic in the neighborhood of* 0 *and* ∞:

$$\delta E_\lambda = \frac{\lambda - 1}{2\pi i} \oint \frac{E_\lambda E_\gamma^{-1} v(\gamma) E_\gamma}{(\gamma - 1)(\lambda - \gamma)} \, d\gamma.$$

5. The representation of $\mathscr{A}(S^2, G)$ on holomorphic maps $\mathbb{C}^* \to G$

The formulas we have derived show the existence of a large number of vector fields tangent to the space of harmonic maps. Moreover, these same formulas show the dependence is actually on the twistor parameter λ. In fact, this map from the Kac-Moody algebra to vector fields is actually a representation [8], [9]. However, the computations are quite difficult. We have found it easier to construct the group representation. In this section we omit the q dependence.

The group action we find is a modification of the standard Birkhoff factorization. In the case of Birkhoff factorization, the 2-sphere $\mathbb{C}^* \cup \{0\} \cup \{\infty\} = S^2 = \mathbb{C}P^1$ is divided into two overlapping regions:

$$S_\varepsilon^+ = \{\lambda \colon |\lambda| \geq (1 + \varepsilon)^{-1}\}, \qquad S_\varepsilon^- = \{\lambda \colon |\lambda| \leq (1 + \varepsilon)\}.$$

Let

$$X^+ = \{e \colon S_\varepsilon^+ \to G \text{ is holomorphic on } S_\varepsilon^+ \text{ for some } \varepsilon > 0\},$$
$$X^- = \{f \colon S_\varepsilon^- \to G \text{ is holomorphic on } S_\varepsilon^- \text{ for some } \varepsilon > 0\}.$$

The Birkhoff factorization we want to use is the factorization of an analytic map $g: S^1 \to G$ on $\lambda = 1$ extended to a holomorphic function on $S_\varepsilon^+ \cap S_\varepsilon^-$ for some ε into

$$g = f_L^+ f_R^- = f_L^- f_R^+,$$

where $f_L^+, f_R^+ \in X^+$ and $f_L^-, f_R^- \in X^-$. However, the Birkhoff factorization generally requires a central term Q,

$$g = f_L^+ Q f_R^- = f_R^- Q' f_R^+.$$

The terms Q and Q' identify the holomorphic structure of the bundles on $S^2 = \mathbb{C}P^1$ with g and g^{-1} as clutching functions. They can be chosen to be I only if these bundles are holomorphically trivial. The reader is at this point referred to Pressley and Segal [17] or Helton [13] for extensive information on this general type of problem.

We shall be carrying out the construction of an action of the group X^- on the space X^+ via this recipe. For $e \in X^+$ and $f \in X^-$ write

$$f^\#(e)(\lambda) = f(\lambda)e(\lambda)R(\lambda)$$

for $e \in X^+$, $f^\#(e) \in X^+$, $f \in X^-$ and $R = R(f, e) \in X^-$. This involves factoring $g = fe$ in the opposite order:

$$g = fe = f^\#(e)R^{-1}.$$

Clearly, when we can define such a factorization uniquely, it does construct a representation $\rho: X^- \to \mathrm{Diff}(X^+)$. However, there are definite obstructions to carrying out the factorization. Crane shows this problem arises in Yang-Mills [6].

The group action for the sigma model is based on a different division of the sphere into two regions. It is not this different division which insures that the factorization can be carried out. The key lies in the reality condition.

The two natural domains for our problem are

$$S^+ = C^* = S^2 - (\{0\} \cup \{\infty\}), \qquad S_\varepsilon^- = \{\lambda : |\lambda| < \varepsilon \text{ or } |\lambda| > \varepsilon^{-1}\}.$$

In the standard Riemann-Hilbert problem or Birkhoff decomposition, the contour is $|\lambda| = 1$. Our contour, as can be seen in the description in Theorem 4.3, is a pair of small circles about 0 and ∞, in the intersection of S^+ and S_ε^-. We replace X^+ by

$$X^k = \{e : \mathbb{C}^* \to G \text{ such that } e \text{ and } e^{-1} \text{ have Laurent expansions of order } k \text{ and } e(1) = I\},$$

$$e(\lambda) = \sum_{|\alpha| \le k} e_\alpha \lambda^\alpha, \qquad e^{-1}(\lambda) = \sum_{|\alpha| \le k} e_\alpha^{-1} \lambda^\alpha.$$

We allow $1 \leq k \leq \infty$. Our reality conditions restrict to

$$X^k_{\mathbf{R}} = \{e \in X^k : e(\lambda)^{-1} = e(\sigma(\lambda))^*\}.$$

We replace X^- by the group of meromorphic maps

$$\mathscr{A}^2(S^2, G) = \{f : S^2 - \{p_1, \cdots, p_l\} \to G \text{ meromorphic with no zeros or poles at } (0, \infty) \text{ and } f(1) = I\}.$$

The smaller group $\mathscr{A}^2(S^2, G)$ is somewhat easier to understand than the maps which are meromorphic in neighborhoods of $(0, \infty)$, which is the group corresponding to X^-. As before our reality condition is

$$\mathscr{A}^2_{\mathbf{R}}(S^2, G) = \{f \in \mathscr{A}(S^2, G) \text{ such that } f(\lambda)^{-1} = f(\sigma(\lambda))^*\}.$$

Now the description of our group action is similar to Birkhoff factorization. We write $f^\#(e) = f \cdot e \cdot R$ for $f, R \in \mathscr{A}^2(S^2, G)$ and $e, f^\#(e) \in X^k$. In words, e is holomorphic in $\mathbb{C}^*$, f puts in zero and poles on the left, and R_f takes them off on the right. In the abelian case, $R = f^{-1}$. However, when G is abelian, this factorization is certainly nontrivial.

Lemma 5.1. *If $f^\#(e)$ can be defined, there exists a unique $f^\#(e)$ taking the value I at $1 \in \mathbb{C}^*$. Moreover, if $f^\#(e)$ and $g^\#(f^\#(e))$ are defined and normalized to be I at $\lambda = 1$, then*

$$(gf)^\#(e) = g^\#(f^\#(e)).$$

Proof. Given the existence of some factorization, we can always replace $f^\#(e)$ and R_f by $f(e)^\#(f^\#(e)(1))^{-1}$ and $R(f^\#(e)(1))^{-1}$. These have the desired normalization. Supposing we have two such normalized factorizations

$$\tilde{f}^\#(e)R^{-1} = f^\#(e)R^{-1} = fe.$$

Solve algebraically to get

$$Q = f^\#(e)^{-1}\tilde{f}^\#(e) - R^{-1}\tilde{R}.$$

The function Q is holomorphic in $\mathbb{C}^*$ from the first term and in a neighborhood of $(0, \infty)$ from the second. It is also I at 1. Therefore Q is holomorphic on S^2 and the constant I by Liouville's theorem. The fact that $g^\# f^\# = (gf)^\#$ follows from this uniqueness.

The success of the group action depends on two separate facts. The first is the existence of the factorization for all "simplest type" $f \in \mathscr{A}^2_{\mathbf{R}}(S^2, G)$. These will induce the Backlund transformations of the next section. The second is the factorization of an arbitrary f into factors of "simplest type".

Definition 5.2. *We shall say $f \in \mathscr{A}(S^2, G)$ is of simplest type if $f(\lambda) = \pi + \xi(\lambda)\pi^\perp$, where π is Hermitian projection on a complex subspace, $\pi^\perp = (I - \pi)$ is projection on the orthogonal subspace and $\xi(\lambda)$ is a rational complex function of degree one which is 1 at $\lambda = 1$.*

If we in addition require that f satisfies the reality condition, then we can check that

$$\xi(\lambda) = \left(\frac{\lambda - \alpha}{\bar{\alpha}\lambda - 1}\right)\left(\frac{\bar{\alpha} - 1}{1 - \alpha}\right) = \xi_\alpha(\lambda) \tag{18}$$

is completely identified by its zero at α. Since $f \in \mathscr{A}(S^2, G)$, $\alpha \neq (0, \infty)$. Also if $|\alpha| = 1$, $\xi_\alpha(\lambda) \equiv 1$ is of no interest.

Theorem 5.3. *If $f(\lambda) = \pi + \xi_\alpha(\lambda)\pi^\perp$ is of simplest type, then $e^\# = f^\#(e) = feR_f$ is always defined.*

Proof. After a little thought, one realizes that R_f should also be of simplest type in order to cancel with f:

$$R = \tilde{\pi} + \xi_{\sigma(\alpha)}(\lambda)\tilde{\pi}^\perp = \tilde{\pi} + \xi_\alpha(\lambda)^{-1}\tilde{\pi}^\perp.$$

At $\lambda = \alpha$, we should have

$$\pi e(\alpha)\tilde{\pi}^\perp = 0, \tag{19}$$

and at $\lambda = 1/\bar{\alpha} = \sigma(\alpha)$ we need

$$\pi^\perp e(1/\bar{\alpha})\tilde{\pi} = 0. \tag{20}$$

However, these two equations are compatible, since $e(1/\bar{\alpha})^* = (e(\alpha))^{-1}$, or (20) is the same as the adjoint equation

$$\tilde{\pi}(e(\alpha))^{-1}\pi^\perp = 0. \tag{20'}$$

We use (19) and (20)′ to define $\tilde{\pi}$ to be projection on the subspace $\tilde{\mathsf{V}}$. Here $\tilde{\mathsf{V}}$ is the subspace $\tilde{\mathsf{V}} = e(\alpha)^*\mathsf{V}$, where V is the subspace π projects onto. Then $\tilde{\mathsf{V}}^\perp$ is perpendicular to $e(\alpha)^{-1}\mathsf{V}$ and $e(\alpha)\tilde{v}^\perp =$ span of $e(\alpha)e(\alpha)^{-1}\mathsf{V}^\perp = \mathsf{V}^\perp$.

The reader may amuse himself by constructing an example where the factorization is clearly impossible. Without the reality condition, the two conditions corresponding to (19) and (20) can be incompatible.

Theorem 5.4. *Every $f \in \mathscr{A}_{\mathbf{R}}(S^2, G)$ factors into a product of factors of simplest type and trivial factors in the center. (This factorization is clearly neither unique nor continuously defined.)*

Proof. Since f is meromorphic, $f(\lambda)$ has a nonzero kernel at a finite number of points $\lambda = (\alpha_1, \alpha_2, \cdots, \alpha_m)$, where $|\alpha_i| \neq 1$. We prove the theorem by induction on the total degree of f. For a factor of "simplest type", $\det(\pi + \xi_\alpha(\lambda)\pi)^\perp = \xi_\alpha(\lambda)^m$ where $m = \operatorname{rank}\pi^\perp$. Recall (18): $\xi_\alpha = ((\lambda - \alpha)/(\bar{\alpha}\lambda - 1))((\bar{\alpha} - 1)/(\alpha - 1))$. This determinant has a zero of order m

at α and a pole of order m at $(\bar{\alpha})^{-1}$ and total degree m. With our reality conditions, zeros and poles always come with this pairing.

To define the total degree of f, we need to pair the zeros at $1/\bar{\alpha}$ and α, since the combination of a zero at α and a pole at α cancels out in the determinant. To compute the total order of the zeros at $(\alpha, 1/\bar{\alpha})$, choose one of them, say α, for which $f(\lambda)$ has a zero at α or equivalently a pole at $1/\bar{\alpha}$. Let

$$\hat{f}(\lambda) = \xi_\alpha(\lambda)^n f(\lambda)$$

for $n \geq 0$ have no pole at $\lambda = a$. The requirement $\hat{f}(\alpha) \neq 0$ if $n \geq 1$ determines n. Define the order $m > n \geq 0$ of the zero of $\det \hat{f}(\lambda)$ at $\lambda = \alpha$ to be the total order of the zeros at $(\alpha, 1/\bar{\alpha})$. It is reasonably easy to see that if both α and $1/\bar{\alpha}$ are zeros (i.e., $n > 0$) this definition is independent of the choice of α or $1/\bar{\alpha}$. The total degree of f is the sum of the orders of conjugate pairs.

We prove our theorem by induction on the total degree of f. If its degree is zero, f must be constant, and in fact $f(\lambda) = I$ since $f(1) = I$. Suppose we have demonstrated the factorization for all of f of total degree less than m. Now suppose g satisfies the reality condition and is of degree m. Select any zero of g and write

$$\hat{g}(\lambda) = (\xi_\alpha(\lambda))^n g(\lambda)$$

for $n \geq 0$, so that $\hat{g}(\alpha)$ is well defined and zero only if $n = 0$. Let $\vee$ be the kernel of $\hat{g}(\alpha)$. If $n = 0$, then $\vee$ may be $\mathbb{C}^N$. Then in this trivial case $g(\alpha) = 0$ and

$$g(\lambda) = (\xi_\alpha(\lambda))^{-1} f(\lambda),$$

where $f(\lambda)$ has degree $M - N$. Note $\xi_\alpha(\lambda)^{-1}$ is ineffective in the loop group.

In the case $\vee$ is not all of $\mathbb{C}^N$, define π to be the hermitian projection on the orthogonal complement of $\vee$, so $\pi^\perp$ is the projection on $\vee$. Let

$$f(\lambda) = \hat{g}(\lambda)(\pi + \xi_\alpha(\lambda)^{-1}\pi^\perp).$$

Then

$$f(\alpha) = \hat{g}(\alpha)\pi + \frac{(\bar{\alpha}\alpha - 1)(\bar{\alpha} - 1)}{\alpha + 1} g'(\alpha)\pi^\perp$$

is finite. Thus $\det f(\lambda) = \det \hat{g}(\lambda)(\xi_\alpha(\lambda))^{-s}$, where the power $s > 0$ is the rank of the kernel of $\hat{g}(\alpha)$. The total degree of

$$f(\lambda) = g(\lambda)(\pi + \xi_\alpha(\lambda)^{-1}\pi^\perp)\xi_\alpha\varepsilon(\lambda)^n$$

is $m - s \geq 0$. This completes the proof of the factorization lemma.

This does not quite complete the proof of Theorem 5.4. Due to the lack of uniqueness in the order of factors, we need a separate proof of the smoothness of the factorization.

Theorem 5.5. *There exists a smooth action of $\mathscr{A}_{\mathbf{R}}(S^2, G)$ on $X_{\mathbf{R}}^R$.*

Proof. The action exists for the factors of simplest type by Theorem 5.3. By Theorem 5.4, every element of $\mathscr{A}_{\mathbf{R}}(S^2, G)$ can be decomposed into factors of simplest type. The existence and uniqueness of the action follows from Lemma 5.1. We have only to demonstrate smoothness.

Recall the general solution to the linear Riemann-Hilbert problem. Let Γ_ε be the contour $|\lambda| = \varepsilon, \varepsilon^{-1}$ oriented in the usual way about $(0, \infty)$. Suppose also that T is holomorphic in punctured neighborhoods of the points $(0, \infty)$. We wish also to normalize our solution to be zero at 1. We define

$$T_\infty(\lambda) = -\frac{\lambda - 1}{2\pi i} \oint_{\Gamma(\varepsilon)} \frac{T(\xi)}{(\xi - 1)(\xi - \lambda)}\, d\xi$$

for $\lambda \in \mathbb{C} - \{0\}$, $\varepsilon < |\lambda| < \varepsilon^{-1}$. For $|\lambda| < \delta$ or $|\lambda| > \delta^{-1}$ we have

$$T_R(\lambda) = \frac{\lambda - 1}{2\pi i} \oint_{\Gamma(\delta)} \frac{T(\xi)}{(\xi - 1)(\xi - \lambda)}\, d\xi.$$

On the punctured neighborhood of 0 and ∞

$$T(\lambda) = T_\infty(\lambda) + T_R(\lambda).$$

Also $T_\infty(1) = 0$, and $T_R(\lambda)$ is holomorphic in neighborhoods of 0 and ∞. Now assume we have factored

$$g(\lambda) = f(\lambda)e(\lambda) = e^\#(\lambda)R^{-1}(\lambda)$$

and look at the differential of this factorization:

$$(e^\#)^{-1}\delta e^\# - \delta R R^{-1} = (e^\#)^{-1}\delta g R^{-1}.$$

If we let $T(\lambda) = e^\#(\lambda)^{-1}\delta g(\lambda)R(\lambda)$, then the formula for $e^\#(\lambda)^{-1}\delta e^\#(\lambda) = T_\infty(\lambda)$ yields the correct result and normalization. Note that if we write $\delta g(\lambda) = v(\lambda)f(\lambda)e(\lambda)$ as in our earlier calculations, we obtain

$$\delta e^\#(\lambda) = -\left(\frac{\lambda - 1}{2\pi i}\right) e^\#(\lambda) \oint_{\Gamma(\varepsilon)} \frac{e^{\#-1}(\xi)v(\xi)e^\#(\xi)}{(\xi - 1)(\xi - \lambda)}\, d\xi. \tag{21}$$

It is clear that these formulas lead to a correct factorization in a small neighborhood, which must agree with our factorization via simplest factors by the uniqueness criteria. This also agrees with Theorem 4.3.

6. The representation of $\mathscr{A}_{\mathbf{R}}(S^2, G)$ on extended harmonic maps and Backlund transformations

Once we understand the representation of $\mathscr{A}_{\mathbf{R}}(S^2, G)$ on $X^k_{\mathbf{R}}$, our calculations become extremely easy. We let $E: \mathbb{C}^* \times \Omega \to G$ be an extended solution of the harmonic map equation. We define $f(^{\#}E)(q) = (f^{\#}E(q))$ for all $q \in \Omega$. Elementary properties follow immediately from §5. First we define

$$\mathscr{M}^k(G) = \{E: \mathbb{C}^* \times \Omega \to G \text{ satisfying (a)–(d)}\}: \tag{22}$$
$$\begin{aligned}
&\text{(a)}\quad E^{-1}\bar{\partial}E = (1-\lambda)A_{\bar{z}},\ E^{-1}\,\partial E = (1-\lambda^{-1})A_z,\\
&\text{(b)}\quad E_1(q) = I,\\
&\text{(c)}\quad E^*_{\bar{\lambda}} = (E_{1/\lambda})^{-1},\\
&\text{(d)}\quad E_\lambda(q) = \textstyle\sum_{|\alpha|\le k} E_\alpha(q)\lambda^\alpha.
\end{aligned}$$

Presumably the correct description of $\mathscr{M}^k(G)$ is as an algebraic variety.

Theorem 6.1. *Given $f \in \mathscr{A}_{\mathbf{R}}(S^2, G)$, $f^{\#}: \mathscr{M}^k(G) \to \mathscr{M}^k(G)$. Moreover, $f \to f^{\#}$ is a representation.*

Proof. The proof of the theorem, modulo the harmonic map property, is already contained in (5.1) and (5.2) and the inherent assumption (easily proved) of analyticity in all variables. Note we prefer to ignore the topology on $\mathscr{A}_{\mathbf{R}}(S^2, G)$ and $\mathscr{M}^k(G)$ as all the appropriate ones will be equivalent, and interest in them is unknown. Appropriate complex structures would be another matter! Thus we will only verify the harmonic map property. This we reduce to Theorem 2.3, and we make the computation of $\partial/\partial\bar{z} = \bar{\partial}$. The ∂ term can be computed algebraically from the $\bar{\partial}$ term, or as an exercise following our computation. Our first observation is that the expression

$$(1-\lambda)^{-1}(f^{\#}E)^{-1}_{\lambda}\bar{\partial}(f^{\#}E)_\lambda \tag{23}$$

is holomorphic by construction except at possibly 0, ∞ and 1. The condition that $(f^{\#}E)_1 \equiv I$ means that the numerator vanishes at $\lambda = 1$, and that no pole is introduced by the denominator at $\lambda = 1$. However,

$$(f^{\#}E)^{-1}_{\lambda}\bar{\partial}(f^{\#}E)_\lambda = (f(\lambda)E_\lambda R_\lambda){-}1\bar{\partial}(f(\lambda)E_\lambda r_\lambda).$$

Since f is independent of q, we get this expression equal to

$$R^{-1}_{\lambda}(E^{!}_{\lambda}\bar{\partial}E_\lambda)R_\lambda + R^{-1}_{\lambda}\bar{\partial}R_\lambda. \tag{24}$$

Now divide (24) by $(1-\lambda)$ to get (23). Because E_{-1} is harmonic, the first term is $R^{-1}_{\lambda}A_{\bar{z}}R_\lambda$, which is holomorphic near 0 and ∞. Likewise, by construction the second term is also holomorphic near 0 and ∞. By Liouville's theorem

$$(1-\lambda)^{-1}(f^{\#}E)^{-1}_{\lambda}\bar{\partial}(f^{\#}E)_\lambda = \tilde{A}_{\bar{z}}$$

has no λ dependence. It follows that $(f^{\#}E)_{-1}$ is harmonic from the equivalent computation for ∂ and Theorem 2.3. Clearly $\tilde{A}_{\bar{z}}\,d\bar{z} + \tilde{A}_z\,dz = \tilde{A}$ is the one-form belonging to the new harmonic map. Finally, since $E_\lambda(p) \equiv I$ for all λ, it follows that $(f^{\#}E)_\lambda(p) = f(\lambda)If^{-1}(\lambda) = I$, the normalization required.

An appropriate retranslation of all these calculations into the infinitesimal framework shows that the map into vector fields described in §3 is really a representation (i.e., brackets go to Lie brackets). This is an elementary exercise. However, note that we have also shown that the elements in the image of the representation are in the infinitesimals of integral curves in $\mathscr{M}^k(G)$.

It is interesting to see what kind of new solutions are produced from known solutions by "factors of simplest type". The argument here is quite simple. We notice that the action of factors of simplest type on the left of extended solutions is cancelled by factors of simplest type on the right. That is, the following result is a corollary of Theorem 6.1 (using Theorem 5.3). Our notation will be simplified if we let

$$\mathscr{P} = \{\pi \in L(\mathbb{C}^N, \mathbb{C}^N) \colon \pi^2 = \pi \text{ and } \pi^* = \pi\}.$$

Recall

$$\xi_\alpha(\lambda) = \frac{(\lambda - \alpha)}{(\bar{\alpha}\lambda - 1)} \frac{(\bar{\alpha} - 1)}{(1 - \alpha)}.$$

Corollary 6.2. *Let $f = \pi + \xi_\alpha(\lambda)\pi^\perp$, where $\pi \in \mathscr{P}$. Then*

$$f^{\#}(E_\lambda) = (\pi + \xi_\alpha(\lambda)\pi^\perp)E_\lambda(\tilde{\pi} + \xi_\alpha(\lambda)^{-1}\tilde{\pi}^\perp),$$

where $\tilde{\pi} \colon \Omega \to \mathscr{P}$. Moreover, $\tilde{\pi}(q)$ is the Hermitian projection on the subspace $(E_\alpha(q))^* V$, where $\pi \colon \mathbb{C}^N \to V$.

The mathematical description of a Backlund transformation is as a method of obtaining new solutions of a system of partial differential equations from old solutions via solving ordinary differential equations. We see that this description applies to our new solutions produced from factors of simplest type. We need to choose the subspace V and $\alpha \in \mathbb{C}^* - \{S^1\}$. The extended E_λ is produced by integrating a pair of consistent ordinary differential equations in $\partial/\partial x$ and $\partial/\partial y$ which are compatible. Then the new solution is obtained algebraically from E_λ. Note however that the algebra is horrendous.

It is true that when applied to harmonic maps $E^{1,1} \to S^2 \subseteq \mathrm{SU}(2)$ with a somewhat different normalization convention on E_λ, these factors of simplest type produce the ordinary Backlund transformations of Sine Gordon. The description of this relationship is beyond the scope of this paper (which is already overly long).

On the other hand, we find it useful to give a direct description of the Backlund transformation, bypassing the extended solution.

Theorem 6.3. *Let* $s\colon \Omega \to \mathrm{U}(N)$ *be a harmonic map, and* $A = \frac{1}{2}s^{-1}\,ds$ *as usual. Then a family of new solutions parametrized by* $\alpha \in \mathbb{C}^*$ *and* $\pi \in \mathscr{P}$ *can be found by solving the consistent pair of ordinary differential equations for* $\hat{\pi}\colon \Omega \to \mathscr{P}$ *with* $\hat{\pi}(p) = \pi$;

$$
\begin{aligned}
\bar{\partial}\hat{\pi} &= (1-\alpha)\hat{\pi}A_{\bar{z}} - (1-\bar{\alpha}^{-1})A_{\bar{z}}\hat{\pi} + (\alpha - \bar{\alpha}^{-1})\hat{\pi}A_{\bar{z}}\hat{\pi}, \\
\partial\hat{\pi} &= (1-\alpha^{-1})\hat{\pi}A_{z} - (1-\bar{\alpha})A_{z}\hat{\pi} + (\alpha^{-1} - \bar{\alpha})\hat{\pi}A_{z}\hat{\pi}.
\end{aligned} \tag{25}
$$

The new solution can be written

$$\hat{s} = (\pi - \gamma\pi^{\perp})s(\hat{\pi} - \bar{\gamma}\hat{\pi}^{\perp}), \qquad \textit{where } \gamma = \frac{1-\bar{\alpha}}{1+\bar{\alpha}}\frac{1+\alpha}{1-\alpha} \in S^1.$$

Proof. Note that $\pi - \pi^{\perp} \in \mathrm{U}(N)$ is ineffective and can be left off the left. To prove this theorem we note that we already know our new solution is of this form from the information that our new extended solution is of the form

$$\hat{E}_\lambda = (\pi + \xi_\alpha\pi^{\perp})E_\lambda(\hat{\pi} + \xi_\alpha^{-1}\hat{\pi}^{\perp}).$$

Now, rather than following through the Birkhoff factorization, we go back to Theorem 2.3. From this theorem $\hat{E}_\lambda$ is an extended solution if and only if

$$
\begin{aligned}
\hat{A}_{\bar{z}} &= (\hat{\pi} + \xi_\alpha\hat{\pi}^{\perp})[A_{\bar{z}}(\hat{\pi} + \xi_\alpha^{-1}\hat{\pi}^{\perp}) + (1-\lambda)^{-1}\bar{\partial}(\hat{\pi} + \xi_\alpha^{-1}\hat{\pi}^{\perp})], \\
\hat{A}_{z} &= (\hat{\pi} + \xi_\alpha\hat{\pi}^{\perp})[A_{z}(\hat{\pi} + \xi_\alpha^{-1}\hat{\pi}^{\perp}) + (1-\lambda^{-1})\partial(\hat{\pi} + \xi_\alpha^{-1}\hat{\pi}^{\perp})]
\end{aligned}
$$

are constant in λ. We check that the coefficients of the poles at $\lambda = (\alpha, 1/\bar{\alpha})$ vanish. This gives four equations:

$$
\begin{aligned}
\hat{\pi}[A_{\bar{z}}\hat{\pi}^{\perp} + (1-\alpha)^{-1}\bar{\partial}\hat{\pi}^{\perp} &= 0, \\
\hat{\pi}^{\perp}[A_{\bar{z}}\hat{\pi} + (1-\bar{\alpha}^{-1})^{-1}\bar{\partial}\hat{\pi}] &= 0, \\
\hat{\pi}[A_{z}\hat{\pi}^{\perp} + (1-\alpha^{-1})^{-1}\partial\hat{\pi}^{\perp}] &= 0, \\
\hat{\pi}^{\perp}[A_{z}\hat{\pi} + (1-\bar{\alpha})^{-1}\partial\hat{\pi}] &= 0.
\end{aligned}
$$

Remember $\hat{\pi}^{\perp} = I - \hat{\pi}$. Algebraic manipulation yields the equations given in the theorem.

It is certainly true that the Backlund transformation can be written in a simpler form. However, it seems appropriate to leave this as the subject of future papers.

7. The additional S^1 action

The construction of this S^1 action follows from the construction of Chuu-Lian Terng of this action on harmonic maps from $E^{1,1}$ into Grassmannians. We construct the group action, and observe how the additional infinitesimal symmetry θ interacts with the Dolan representation of $\sum(\mathfrak{g})$.

Of course, we construct a $\mathbb{C}^*$ action and show it preserves the reality conditions when $|\gamma| = 1$, $\gamma \in \mathbb{C}^*$. This action also is based on the extended solutions rather than the solutions themselves. Recall $\mathscr{M}^k(G)$ is defined in (22).

Theorem 7.1. *If $\gamma \in \mathbb{C}^*$ and $E \in \mathscr{M}^k(G)$, then*

$$\gamma^\# E_\lambda = E_{\lambda\gamma} E_\gamma^{-1} \in \mathscr{M}^k(G) \quad \textit{for } |\gamma| = 1.$$

Proof. All conditions are quite simple to verify except the harmonic map condition. The restriction $|\gamma| = 1$ is needed only to verify the reality condition. The harmonic map condition again reduces to the combination of Theorems 2.2 and 2.3. Also, we do the necessary computation for $\bar\partial$ and leave the other in ∂ as an exercise in algebra or the chain rule,

$$\begin{aligned}(\gamma^\# E_\lambda)^{-1}\bar\partial\gamma^\# E_\lambda &= E_\gamma E_{\lambda\gamma}^{-1}\bar\partial(E_{\lambda\gamma}E_\lambda^{-1}) \\ &= E_\gamma(1-\lambda)A_{\bar z}E_\gamma^{-1} - E_\gamma(1-\gamma)A_{\bar z}E_\gamma^{-1} \\ &= \gamma(1-\lambda)E_\gamma A_{\bar z}E_\gamma^{-1}.\end{aligned}$$

Since $(\gamma^\# E_\lambda^{-1}\bar\partial\gamma^\# E_\lambda)(1-\lambda)^{-1}$, (and the similar expression $(\gamma^\# E_\lambda^{-1}\partial\gamma^\# E_\lambda) \times (1-\lambda^{-1})^{-1}$) are independent of λ, we have the appropriate harmonic condition verified.

Now what about the relationship of the action of $\mathscr{A}(S^2, G)$ and $\mathbb{C}^*$? Let $\gamma^\# f(\lambda) = f(\gamma\lambda)$ if $f \in \mathscr{A}(S^2, G)$, $\gamma \in \mathbb{C}^*$.

Proposition 7.2. $\gamma^\#(f^\# E) = (\gamma^\# f)(\gamma^\# E)$.

Proof. This again follows from the uniqueness. We have

$$\gamma^\#(f^\# E) = f(\gamma\lambda)E_{\gamma\lambda}(q)S_{\gamma\lambda}(E_{\gamma\cdot}(q), f(\gamma\cdot)) = (f^\# E(q))(\gamma\lambda).$$

This is all straightforward. However, the interaction with the representation as presented by Dolan is somewhat more complicated. We must figure out not the infinitesimal action with respect to the variable λ, but that with respect to the variable $\tau = -i(\lambda - 1)/(\lambda + 1)$ for which we have a grading. The energetic can compute the full action. Note that for $|\gamma| = 1$, $\lambda \to \gamma\lambda$ does not preserve the imaginary axis, so that the action of γ does not preserve the unit circle, $|\tau| = 1$, which was the basis for the chosen grading in the Kac-Moody Lie algebra. So it comes as no surprise that the infinitesimal action of θ does not preserve the grading.

Proposition 7.3. *The infinitesimal generators of $\mathbf{C}^*$ on the variable τ is $\theta(\tau) = \frac{1}{2}(\tau^2 + 1)$.*

Proof.
$$\left.\frac{d}{d\theta}\right|_{\theta=0} - i\left(\frac{e^{\tau\theta}\lambda - 1}{e^{\tau\theta}\lambda + 1}\right) = \frac{2\lambda}{(\lambda+1)^2} = \frac{1}{2}(\tau^2+1).$$

Theorem 7.4. *If we adjoin the infinitesimal generator θ of the S^1 action to the affine Kac-Moody Lie algebra with the relation*
$$[\theta, \alpha_k t^k] = \frac{k}{2} a_k (t^{k+1} + t^{k-1}),$$
we obtain a larger Lie algebra, which has a representation on the vector fields $\mathscr{M}^k(G)$ extending the Dolan representation.

Proof. This is just a translation of the group action which we just discussed into infinitesimals. It is not hard to show, independently of our derivation, that the enlarged object is a Lie algebra by verifying the Jacobi identity.

The existence of this larger representation is what convinces us that our representation has very little to do with the complicated algebraic representations usually constructed. To change the grading so that θ preserves the grading uncouples the two halves of the Kac-Moody algebra, leaving it without much algebraic interest. Naturally this decoupling will correspond to the decoupling of the germs of the maps at $\pm i$.

8. Harmonic maps into Grassmannians

The mathematics literature does not deal with harmonic maps into Lie groups, but concerns itself primarily with harmonic maps into S^k, $\mathbf{C}P^{N-1}$ or various real or complex Grassmannians. In our theory we can treat the case of harmonic maps into the complex Grassmannians $G_{k,N}$ of k-planes in N space. We consider
$$G_{k,N} \subset G_{\mathbf{R}} = \mathrm{U}(N),$$
$G_{k,N} = \{\phi \in G_{\mathbf{R}} \colon \phi^2 = I$ and the eigenspace corresponding to $+1$ is k-dimensional$\}$.

We identify ϕ with the k-dimensional subspace corresponding to the eigenspace of $+1$. This embedding of $G_{k,N} \subset \mathrm{U}(N)$ does put a multiple of the usual (Kähler) metric on $G_{k,N}$. Moreover, $G_{k,N} \subset \mathrm{U}(N)$ is totally geodesic.

Proposition 8.1. $\Omega \xrightarrow{s} G_{k,N} \xrightarrow{\varphi} \mathrm{U}(N)$. *Then s is harmonic if and only if φs is harmonic.*

Proof. Such a theorem is true for any $\varphi: X \to Y$, φ totally geodesic.

We now wish to look at harmonic maps $s: \Omega \to U(N)$ satisfying $s^2 = I$ or $s = s^{-1}$. Our first observation is that we shall have to change our normalization $s(p) = I$. For the Grassmannian $G_{k,N}$, we introduce a new normalization $E_\lambda(p) = Q_k(\lambda)$. In fact, $Q_k(\lambda)$ can be chosen in any way which satisfies

$$(26) \qquad \begin{aligned} &\text{(a)} \quad Q_k(1) = I, \\ &\text{(b)} \quad Q_k(-1) = \begin{pmatrix} +I & 0 \\ 0 & -I \end{pmatrix}, \end{aligned}$$

where $Q_k(-1)$ has k eigenvalues $+1$ and $N-k$ eigenvalues -1,

$$\text{(c)} \quad Q_k(-\lambda)Q_k(-1) = Q_k(\lambda).$$

We can always conjugate a harmonic map $s: \Omega \to U(N): s^2 = I$ into one satisfying $Q_k(-1) = s(p)$. With this new normalization, the choice of E_λ is uniquely determined.

Our first observation is that the $\mathbb{C}^*$ action acts to carry harmonic maps satisfying $s^2 = I$ into those satisfying $(\gamma^* s)^2 = 1$ for $\gamma \in \mathbb{C}^*$.

Proposition 8.2. *If $E_\lambda(p) = Q_k(\lambda)$, we have for $|\gamma| = 1$*

$$\gamma^* s = E_{-\gamma} E_\gamma^{-1} : \Omega \to G_{k,N}$$

harmonic if $s: \Omega \to G_{k,N}$.

Proof. By our previous discussions, $\gamma^* s: \Omega \to U(N)$ is harmonic. The change is normalization does not affect them. We just need to check that $(\gamma^* s)^2 = I$.

An extended solution for $\gamma^* s$ is given by

$$(\gamma^* E)_\lambda = E_{\lambda\gamma} E_\gamma^{-1}.$$

Since

$$(-1^* E)_\lambda(p) = E_{-\lambda}(p) E_{-1}^{-1}(p) = Q_k(\lambda),$$

this is the unique solution satisfying the normalization condition for $\gamma = 1$. Since

$$((-1)^* E)_{-1} = E_1 E_{-1}^{-1} = s^{-1} = s,$$

the extended solutions for $\gamma = -1$ and $\gamma = 1$ are the same. This means that

$$E_\lambda = E_{-\lambda} E_{-1}^{-1} = E_{-\lambda} s.$$

Therefore $(\gamma^* s)^2 = (\gamma^* E_{-1})^2 = (E_{-\gamma} s^{-1} E_{-\gamma}^{-1})^2 = E_{-\gamma} s^{-2} E_{-\gamma}^{-1} = I$. This verifies $\gamma^* s: \Omega \to G_{k,N}$.

We wish to deliver an appropriate subgroup of $\mathscr{A}_{\mathbf{R}}(S^2, G)$ which will preserve the condition $s^2 = I$ or $s = s^{-1}$. Rather than simply make a definition, we do the computation first as motivation.

In order to preserve the condition $s = s^{-1}$, we need to relate the action of $f \in \mathscr{A}(S^2, G)$ on s to an action of some element $g \in \mathscr{A}(S^2, G)$ on s^{-1}. Once we do this, we will not only understand how to preserve $s = s^{-1}$, but we will understand completely why the action of $\mathscr{A}(S^2, G)$ operating on left pull-backs of s corresponds to an action of $\mathscr{A}(S^2, G)$ on left pull-backs of s^{-1}, or right pull-backs of s. This explains fully our comment in the very first section: it does not matter whether we use right or left pull-backs.

Corresponding to the harmonic map $s = E_{-1}$ we have

$$(f^{\#}E)_\lambda = f(\lambda)E_\lambda S_\lambda.$$

For the time being, we assume the normalization $E_\lambda(p) = Q(\lambda)$ is chosen. Then the extended solution corresponding to $s^{-1} = E_1 E_{-1}^{-1}$ is $\tilde{E}_\lambda = E_{-\lambda}E_{-1}^{-1}$. Then

$$(g^{\#}\tilde{E})_\lambda = g(\lambda)\tilde{E}_\lambda\tilde{S}_\lambda = g(\lambda)E_{-\lambda}E_{-1}^{-1}\tilde{S}_\lambda.$$

Our goal is that $(f\#E)_\lambda$ should be the extended solution for $\tilde{s} = (f\#E)_{-1}$ and $(g^{\#}\tilde{E})_\lambda$ should be the extended solution for $\tilde{s}^{-1} = (g^{\#}\tilde{E})_{-1}$. This means

$$(g^{\#}\tilde{E})_\lambda = (f^{\#}E)_{-\lambda}(f^{\#}E)_{-1}^{-1}$$

or

$$g(\lambda)E_{-\lambda}E_{-1}^{-1}\tilde{S}_\lambda = f(-\lambda)E_{-\lambda}S_\lambda\tilde{s}^{-1}.$$

Let $g(\lambda) = f(-\lambda)$. Then

$$(g^{\#}\tilde{E})_\lambda = f(-\lambda)E_{-\lambda}E_{-1}^{-1}\tilde{S}_\lambda = (f^{\#}E)_{-\lambda}Q,$$

where Q is chosen so $(f^{\#}E)_{-1}Q = I$, or $Q = (f^{\#}E)_{-1}^{-1} = \tilde{s}^{-1}$, as required.

Now the situation for Grassmannians is not quite so simple, since the normalization required, $E_\lambda(p) = Q_k(\lambda)$, is not preserved by the action of $\mathscr{A}_{\mathbf{R}}(S^2, G)$. We simply allow the group to act on extended solutions, carrying extended solutions to extended solutions.

Theorem 8.3. *Suppose E_λ is an extended harmonic map satisfying*

(a) $(E_{\bar{\lambda}}^* = (E_{\lambda^{-1}})^{-1}$,

(b) $E_1 = I$,

(c) $E_\lambda = E_{-\lambda}E_{-1}^{-1}$.

Then the action of $\{f \in \mathscr{A}_{\mathbf{R}}(S^2, G) \colon f(-\lambda) = f(\lambda)\}$ on E preserves these conditions.

Proof. Conditions (a) and (b) are verified (the normalization was necessary only to have $\mathscr{A}_{\mathbf{R}}(S^2, G)$ act in a unique fashion on solutions). To verify that (c) is preserved, note that

$$\begin{aligned}(f^{\#}E)_\lambda &= f(\lambda)E_\lambda S_\lambda = f(-\lambda)E_{-\lambda}E_{-1}^{-1}S_\lambda \\ &= f(-\lambda)E_{-\lambda}S_{-\lambda}(f(-1)E_{-1}S_{-1})^{-1}.\end{aligned}$$

It follows that

$$(f^\# E)_\lambda = (f^\# E)_{-\lambda}((f^\# E)_{-1})^{-1}.$$

A straightforward computation on the change of variables shows that $\lambda \to -\lambda$ corresponds to $\tau \to -1/\tau$. The condition $f(-\lambda) = f(\lambda)$ corresponds to $f(\tau) = f(-1/\tau)$, or in the Fourier series expansion $\sum_\alpha a_\alpha \tau^\alpha$, we must have $a_\alpha = (-1)^\alpha a_\alpha$. We may conclude that half of the affine Kac-Moody Lie algebra $\sum(\mathfrak{g}_{\mathbf{R}})$ has a representation in the vector fields tangent to harmonic maps into Grassmannians.

PART II

9. The single uniton

The question arises of identifying the "simplest" harmonic maps $\Omega \to \mathrm{U}(N)$. We choose to define the level of complexity of a harmonic map by the minimal number of terms needed in the expansion of an extended solution.

Definition 9.1. An *n-uniton* is a harmonic map $s\colon \Omega \to \mathrm{U}(N)$ which has an extended solution

$$E_\lambda\colon \mathbb{C}^* \times \Omega \to G = \mathrm{GL}(N, \mathbb{C})$$

with

(a) $E_\lambda = \sum_{\alpha=0}^n T_\alpha \lambda^\alpha$ for $T_\alpha\colon \Omega \to \mathfrak{g}$,
(b) $E_1 = I$,
(c) $E_{-1} = Qs^{-1}$ for $Q \in \mathrm{U}(N)$ constant,
(d) $(E_{\bar\lambda})^* = (E_{\lambda^{-1}})^{-1}$.

There are a few observations which we can make about these conditions. For $n = 0$, $E_\lambda \equiv I$ is the only extended solution, which represents $s \equiv Q^{-1}$, or the constant harmonic maps. For any harmonic map s and constant $Q \in \mathrm{U}(N)$, both $Q^{-1}s$ and sQ are also harmonic maps. In condition (c), we allow the same extended solution to represent s and $Q^{-1}s$. For right multiplication, the extended solution E_λ which represents s can be replaced by $Q^{-1}E_\lambda Q$ to represent sQ. In either case, s, $Q^{-1}s$ and sQ have the same uniton number.

The important condition (d) is the reality condition. Finally, note that the uniton number n can always be enlarged in a fake way by multiplying an extended solution E_λ by a $Q\colon \mathbb{C} \to G$ (constant in p). This difficulty is resolved in §13.

The term "uniton" is meant to be analogous to the term "instanton". The harmonic map equations are partial differential equations in two real variables with certain remarkable similarities to the self-dual Yang-Mills equations, which are equations in four real variables whose solutions are called instantons. Neither unitons nor instantons are much like "solitons", which are special solutions of equations with a time parameter. Presumably, the $E^{1,1} \to G_{\mathbf{R}}$ equations for harmonic maps have solitons. However, they can have no unitons. In the $E^{1,1}$ case, the appropriate reality condition on the extended solution is $(E_{\bar{\lambda}})^* = (E_\lambda)^{-1}$, which admits no finite power series solutions in λ whatsoever. If we are willing to abandon the reality condition, there may be more similarities between the Minkowski and Euclidean problems.

Our goal in this section is to understand the single or one uniton. As a preliminary step to our first proposition, recall that if $\eta \subset \Omega \times \mathbf{C}^N$ is a $k < N$ dimensional complex sub-bundle, then we can construct $\pi\colon \Omega \to \mathscr{P} \subset L(\mathbf{C}^N, \mathbf{C}^N)$, where $\pi(q)$ is the Hermitian projection on $\eta_q \subset \mathbf{C}^N$. Algebraically we can identify $\pi \in \mathscr{P}$ by noting that

$$(27)\qquad \begin{aligned} &\text{(a)}\quad \pi^*(q) = \pi(q) \text{ for all } q \in \Omega,\\ &\text{(b)}\quad \pi^2(q) = \pi(q) \text{ for all } q \in \Omega,\\ &\text{(c)}\quad \pi(q) \text{ has rank } k \text{ at every point } q \in \Omega. \end{aligned}$$

Moreover, every time we find a π satisfying (27)(a)–(c), we can find the sub-bundle η (just the image). We identify sub-bundles and the Hermitian projections on them without any further comment. We also suppress the dependence on $q \in \Omega$, which we take for granted. Moreover, note that $\pi^\perp = (I - \pi)$ is the Hermitian projection on the bundle $\eta^\perp$ which is orthogonal to η. In discussing harmonic maps into Grassmannians in §8, we could have mentioned that a harmonic map $s\colon \Omega \to G_{k,N}$ is represented by

$$s = (\pi - \pi^\perp) = (2\pi - I),$$

where $\pi(q)$ is the Hermitian projection into the image subspace $s(q)$ at each $q \in \Omega$.

Proposition 9.2. *$s\colon \Omega \to \mathrm{U}(N)$ is a one-uniton if and only if $s = Q(\pi - \pi^\perp)$ for $Q \in \mathrm{U}(N)$, where π satisfies* (27) *and $\pi^\perp \bar{\partial}\pi = 0$.*

Proof. We have $E_\lambda = T_0 + \lambda T_1$ and $E_1 = T_0 + T_1 = I$, so $E_\lambda = T_0 + \lambda(I - T_0)$. The reality condition (d) is equivalent to

$$(I - T_0)^*(T_0 = 0, \qquad T_0^*(I - T_0) = 0,$$
$$I = (I - T_0)^*(I - T_0) + T_0^* T_0 = I - T_0^* - T_0 + 2T_0^* T_0.$$

The difference between the first two equations gives $T_0 - T_0^* = 0$. Once we know $T_0 = T_0^*$, both equations read $T_0^2 = T_0$. It follows that $T_0 = \pi$ and

$T_1 = (I - T_0) = \pi^\perp$ are projections on orthogonal sub-bundles:
$(1-\lambda)^{-1}(E_\lambda)^{-1}\bar\partial E_\lambda = (1-\lambda)^{-1}(\pi+\lambda^{-1}\pi^\perp)\bar\partial(\pi+\lambda)(I-\pi) = \pi\bar\partial\pi+\lambda^{-1}\pi^\perp\bar\partial\pi.$
The dual condition is

$$(1-\lambda^{-1})^{-1}E_\lambda^{-1}\partial E_\lambda = -(1-\lambda^{-1})^{-1}\partial(E_\lambda)^{-1}E_\lambda = -\partial\pi(\pi+\lambda\pi^\perp).$$

The harmonic condition that both expressions be independent of λ reduces to $\pi^\perp(\bar\partial\pi) = ((\partial\pi)\pi^\perp)^* = 0$.

Checking back to §8, we note that $s = \pi - \pi^\perp$ is a harmonic map into $G_{k,N}$. It is in fact a special kind of harmonic map. Recall that $G_{k,N}$ is a Kähler manifold with its usual metric. It is well known that every holomorphic map $\Omega \to X$, where X is Kähler, is automatically harmonic. These basic unitons correspond to holomorphic maps.

Choose $u_j\colon \Omega \to \mathbb{C}^N$ so that $(u_{-1}(q)), \cdots, u_k(q))$ form an orthogonal basis for $\pi(q)$. The u's are not unique and it is not necessary to choose them globally. In terms of k-planes, the image k-plane is (projectively) at q

$$u_1(q) \wedge \cdots \wedge u_k(q) \in \bigwedge_k \mathbb{C}^N,$$

and the formula for our projection is (exactly)

$$\pi(q) = \sum_{j=1}^k u_j(q) \times u_j(q)^*. \tag{28}$$

The map into $G_{k,N}$ is holomorphic if

$$\bar\partial(u_1 \wedge \cdots \wedge u_k) = \lambda(u_1 \wedge \cdots \wedge u_k)$$

for $\lambda\colon \Omega \to \mathbb{C}$, i.e., if

$$\bar\partial u_j = \sum_l c_{l,j} u_l. \tag{29}$$

This is equivalent to the statement that the complex sub-bundle $\eta \subset \Omega \times \mathbb{C}^N$, which is the image of π, is holomorphic. In particular, this means that if $f\colon \Omega \to \eta$ is a section, then $\bar\partial f\colon \Omega \to \eta$ is also a section.

Theorem 9.3. *The map $s\colon \Omega \to \mathrm{U}(N)$ is a single uniton if and only if $s = Q(\pi - \pi^\perp)$, where $Q \in \mathrm{U}(N)$, and $\pi\colon \Omega \times \mathbb{C}^N$ is the orthogonal projection into a holomorphic sub-bundle $\eta \subset \Omega \times \mathbb{C}^N$.*

Proof. Use the formula for π given by (28), where $(u_1, \cdots, u_k, v_{k+1}, \cdots, v_N)$ are locally an orthogonal basis for $\mathbb{C}^N$, and η is spanned by the u's. Then

$$(I-\pi)\bar\partial\pi = \sum_{l=k+1}^N \sum_{j=1}^k (\bar\partial u_j, v_l) v_l \otimes u_j^*. \tag{30}$$

If the bundle η is holomorphic, $(\bar\partial u_j, v_l) = 0$ by (29), and $(I - \pi)\bar\partial\pi = 0$. On the other hand, if $(I-\pi)\bar\partial\pi = 0$, since $u_l \otimes v_j$ are linearly independent in $L(\mathbb{C}^N, \mathbb{C}^N)$, it follows that $(\bar\partial u_j, v_l) = 0$, $j = 1, 2, \cdots, k$ and $l = k + 1, \cdots, N$, or (29) holds. This shows η is holomorphic. The theorem follows as a consequence of Proposition 9.2.

We hope it might be of some interest to show how the Kac-Moody Lie algebra acts in this simplest case. Return to the formula in Theorem 4.3. We indicate infinitesimals by δf to distinguish from the group action. Because E_γ is so simple, and δf is holomorphic in a neighborhood of 0 and ∞, the expression

$$\frac{E_\gamma^{-1}\delta f(\gamma)E_\gamma}{(1-\gamma)(\lambda-\gamma)} = \frac{(\pi + \gamma^{-1}\pi^\perp)\delta f(\gamma)(\pi + \gamma\pi^\perp)}{(1-\gamma)(\lambda-\gamma)}$$

has only a single pole at zero and a single pole at infinity, the contour integral in (4.3) is very easy to compute by residues.

Theorem 9.4. *The action of the group $\mathscr{A}_{\mathbb{R}}(S^2, G)$ on the single unitons reduces to an effective action of $G = \{(f(0), f(\infty)\colon f(0)^* = f(\infty)^{-1})\} \subset G \times G$. The infinitesimal action is*

$$\delta E_\lambda = (\lambda - 1)(\pi^\perp \delta f)(0)\pi - \pi\delta f(\infty)\pi^\perp),$$

where $\delta f(0)^* + \delta f(\infty) = 0$.

Proof. By residue calculations we obtain at zero

$$\frac{1}{2\pi i}\oint_{|\gamma|=\varepsilon} \frac{E_\gamma^{-1}\delta f(-\gamma)E_\lambda}{(1-\gamma)(\lambda-\gamma)}\,d\gamma = \frac{1}{\lambda}\pi^\perp\delta f(0)\pi.$$

At infinity

$$\frac{1}{2\pi i}\oint_{|\gamma|=1/\varepsilon} \frac{E_\gamma^{-1}\delta f(-\gamma)E_\gamma}{(1-\gamma)(\lambda-\gamma)} = \pi\delta f(\infty)\pi^\perp.$$

Applying this to Theorem 4.3 we get

$$\begin{aligned}\delta E_\lambda &= (\lambda - 1)E_\lambda(\lambda^{-1}\pi^\perp\delta f(0)\pi - \pi\delta f(\infty)\pi^\perp) \\ &= (\lambda - 1)(\pi^\perp\delta f(0)\pi - \pi\delta f(\infty)\pi^\perp).\end{aligned}$$

Since only $f(0)$ and $f(\infty)$ affect the formulas, the reality condition $(f(\bar\lambda))^* = (f(\lambda^{-1}))^{-1}$ becomes $f(0)^* = f(\infty)^{-1}$, or infinitesimally $\delta f(0)^* + \delta f(\infty) = 0$.

Note that $G_{\mathbb{R}}$ acts naturally by conjugation. For $a \in \mathfrak{g}_{\mathbb{R}}$, $a = \delta f(0) = \delta f(\infty)$,

$$\delta E_\lambda = E_\lambda a - aE_\lambda = (\lambda - 1)(\pi^\perp a - a\pi^\perp) = (\lambda - 1)(\pi^\perp a\pi - \pi a\pi^\perp).$$

This helps check the signs. Note that the group of holomorphic transformations of $G_{k,N}$ acts on holomorphic maps into $G_{k,N}$ by composition. In the case of $G_{1,2} = \mathbb{C}P^1$, we have checked that our action of $\mathrm{SL}(2,\mathbb{C})$ (the center of $\mathrm{GL}(N,\mathbb{C})$ acts trivially) is exactly that of composition with the conformal group of $\mathbb{C}P^1$. We conjecture that the constructed action of $\mathrm{SL}(N,\mathbb{C})$ on $G_{k,N}$ is similar, but have not checked this.

10. The fixed points of the S^1 action

Before we proceed further, we give some examples of harmonic maps, which all happen to have images in the Grassmannians, and are based on the holomorphic (single uniton) map of the previous sections. Harmonic maps from S^2 into $G_{1,N} = \mathbb{C}P^{N-1}$ have been classified ([10], [11], [7]) and we sketch how they fit into the theory as 2-unitons. In particular, we show how the Grassmann subalgebra of the Kac-Moody representations acts on these 2-unitons.

Throughout this section, we ignore the fact that there are point singularities in our constructions. It is well known that the singularities of these complex constructions have complex codimension two, and the apparent point singularities in our one complex variable constructions are removable. Also, we regard $S^2 = \mathbb{R}^2 \cup \{\infty\}$ or $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$ without comment, although some details of this relationship are given in the next section.

First, let us redescribe the construction of the single unitons. Choose k meromorphic maps $u_j \colon \Omega \to \mathbb{C}^N$ which are independent except at isolated points. Use the Gramm-Schmidt process to orthogonalize, giving $(s_1, \cdots, s_k)$. Then

$$s = \left(2\sum_{j=1}^{k} s_j \otimes s_j^* - I\right)$$

is harmonic (apparent point singularities are removable). If $\pi = \sum_{j=1}^{k} s_j \otimes s_j^*$, then $s = (\pi - \pi^\perp) = (2\pi - I)$ and an extended solution is of the form

$$E_\lambda = (\pi + \lambda \pi^\perp).$$

Here $\pi(q)$ is the Hermitian projection on the subspace $\eta(q) \subset \mathbb{C}^N$ spanned by the $u(q)$'s. The bundle

$$\eta = \{(q, \eta(q)), q \in \Omega\} \subset \Omega \times \mathbb{C}^N$$

is a holomorphic sub-bundle.

For a more complicated construction, we repeat this process n times, obtaining n holomorphic sub-bundles $\eta_\alpha \subset \Omega \times \mathbb{C}^N$ and n Hermitian projections on them. There are specific constraints however:

$$\text{(31)} \qquad \begin{array}{ll} \text{(a)} & \eta_\alpha \subset \eta_{\alpha+1} \text{ is proper,} \\ \text{(b)} & f\colon \text{if } f\colon \Omega \to \eta_\alpha, \text{ then } \partial f\colon \Omega \to \eta_{\alpha+1}. \end{array}$$

The holomorphic condition requires that

$$\text{(c)} \quad \text{if } f\colon \Omega \to \eta_\alpha, \text{ then } \bar\partial f\colon \Omega \to \eta_\alpha.$$

Another way to explain the conditions is that the meromorphic maps used to construct η_α and their ∂ derivatives should be contained in the span of those used to construct $\eta_{\alpha+1}$.

If we assume $\eta_\alpha \subset \eta_{\alpha+1}$ is proper, a counting argument shows that $\eta_\alpha = \Omega \times \mathbb{C}^N$ for $\alpha \geq \alpha_0 \geq N$. Assume $\alpha_0 = n$ and $\eta_n = \Omega \times \mathbb{C}^N$.

Theorem 10.1. *Let π_α be the projection on η_α, $\alpha = 1, 2, \cdots, n$, where η_α satisfies* (31). *Then*

$$E_\lambda = \pi_0 + \sum_{\alpha=1}^{n} \lambda^\alpha (\pi_\alpha - \pi_{\alpha-1})$$

is an extended solution to a harmonic map $s = E_{-1}$.

Proof. Note $E_1 = \pi_n = I$, which verifies one normalization. Since $\eta_{\alpha-1} \subset \eta_\alpha$, we have $\tilde\pi_\alpha = \pi_\alpha - \pi_{\alpha-1}$ is the projection on $\eta_\alpha \cap \eta_{\alpha-1}^\perp = \{(q, v)\colon (v) \in \eta_\alpha(q) \text{ and } (v, w) = 0 \text{ for } (w) \in \eta_{\alpha-1}(a)\} \subseteq \eta_\alpha$. If $\pi_0 = \tilde\pi_0$, then

$$E_\lambda = \sum_{\alpha=0}^{n} \lambda^\alpha \tilde\pi_\alpha, \tag{32}$$

where the $\tilde\pi_\alpha$ are all projections into mutually orthogonal subspaces. The harmonic map is

$$E_{-1} = \sum_{\alpha=0}^{n} (-1)^\alpha \tilde\pi_\alpha.$$

For $\alpha > \beta$, we automatically get from $\tilde\pi_\alpha \pi_\beta = 0$

$$\tilde\pi_\alpha \bar\partial \pi_\beta = \tilde\pi_\alpha \pi_\beta \bar\partial \pi_\beta = 0.$$

This uses the condition that $\eta_\beta \subset \Omega \times \mathbb{C}^N$ is holomorphic so $\pi_\beta^\perp \bar\partial \pi_\beta = 0$ and $\bar\partial \pi_\beta = \pi_\beta \bar\partial \pi_\beta$. For $\alpha < \beta$, $\pi_\alpha \pi_\beta^\perp = 0$ and $\bar\partial \pi_\beta^\perp = -\bar\partial \pi_\beta$. We conclude

$$\pi_\alpha \bar\partial \pi_\beta = -\pi_\alpha \bar\partial \pi_\beta^\perp = \bar\partial \pi_\alpha \pi_\beta^\perp = (\pi_\beta^\perp \partial \pi_\alpha)^*. \tag{33}$$

We show (33) is zero in two steps. Locally we can write

$$\pi_\beta^\perp \partial \pi_\alpha = \pi_\beta^\perp \left(\sum_j \partial s_j \otimes s_j^* \right) + \pi_\beta^\perp \sum_j s_j \otimes (\bar\partial s_j)^*.$$

Because $s_j(q) \in \eta_\alpha(q)$ and $\pi_\beta(q)\eta_\alpha(q) = 0$, the second term vanishes. However $\partial s_j(q) \in \eta_{\alpha+1}(q)$ from (31)(b), and $\eta_{\alpha+1}(q) \subset \eta_\beta(q)$, so the first term also vanishes for the same reasons.

$$(1-\lambda)^{-1}E_\lambda^{-1}\bar{\partial}E_\lambda = E_\lambda^{-1}\sum_{\beta=0}^{n-1}\lambda^\beta\bar{\partial}\pi_\beta = \sum_{\beta=0}^{n-1}\lambda^{\beta-\alpha}\sum_{\alpha=0}^{n}(\pi_\alpha-\pi_{\alpha-1})\bar{\partial}\pi_\beta.$$

We have just shown that the only nonzero term is

$$\sum_{\beta=0}^{N-1}\lambda^0\pi_\beta\bar{\partial}\pi_\beta,$$

which is independent of λ. A similar calculation in ∂ verifies the hypothesis of Theorem 2.3, and E_{-1} is harmonic.

The details of the Eells-Wood analysis of harmonic maps into $\mathbb{C}P^{N-1}$ explicitly show that every such harmonic map is a 2-uniton. We refer the reader to their paper, but briefly translate their discussion into ours. The subspace η_0 is generated as the span of $\{s, \partial s, \cdots, \partial^{k-1}s\}$, where $s\colon S^2 \to \mathbb{C}^N$ is meromorphic ($s\colon S^2 \to \mathbb{C}^N/\mathbb{C}^* = \mathbb{C}P^{N-1}$ is holomorphic (with removable singularities)). The sub-bundle η_1 is spanned by $\{s, \partial s, \cdots, \partial^k s\}$ and $\eta_\alpha = S^2 \times \mathbb{C}^N$. Then

$$E_\lambda = \pi_0 + \lambda(\pi_1 - \pi_0) + \lambda^2(I - \pi_1),$$
$$E_{-1} = (I + \pi_0 - \pi_1) - (\pi_1 - \pi_0).$$

The projection $\pi = \pi_1 - \pi_0$ is a projection onto a one-dimensional sub-bundle

$$\eta = \left\{ (q,v)\colon v = \sum_{j=0}^{k} a_j\partial^j s(q) \text{ and } (v, \partial^j s(q)) = 0 \text{ for } j < k \right\}.$$

Technically, $E_{-1} = \pi^\perp - \pi$, so the harmonic map to $\mathbb{C}P^{N-1}$ associated with this extended solution is $s = -E_{-1}$. This is due only to the convention of associating the k-plane in $G_{k,N}$ with the $+1$ eigenspace of s, rather than the -1 eigenspace.

Proposition 10.2. *The simple multi-uniton construction yields an extended solution of the form*

$$E_\lambda = (\pi_0 + \lambda\pi_0^\perp)(\pi_1 + \lambda\pi_1^\perp)\cdots(\pi_{n-1} + \lambda\pi_{n-1}^\perp)$$
$$= (\pi_{n-1} + \lambda\pi_{n-1}^\perp)\cdots(\pi_1 + \lambda\pi_1^\perp)(\pi_0 + \lambda\pi_0^\perp),$$

where π_α is the Hermitian projection on $\eta_\alpha \subset \Omega\times\mathbb{C}^N$. Rank $\pi_\alpha >$ Rank $\pi_{\alpha+1}$. *Moreover, E_λ is an extended solution when*

(a) $\eta_\alpha \subset \eta_{\alpha+1}$,
(b) *if* $f\colon \Omega \to \eta_\alpha$, *then* $\bar\partial f\colon \Omega \to \eta_\alpha$,
(c) *if* $f\colon \Omega \to \eta_\alpha$, *then* $\partial f\colon \Omega \to \eta_{\alpha+1}$.

In §12 we generalize this construction. It is highly nonunique, due to the lack of uniqueness in the factorization. In our case, the factors commute. In the general case, they do not.

A simple calculation shows that the infinitesimal action of the representation of the Kac-Moody Lie algebra does not preserve this construction unless $n = 1$. If we add the constraint $f(-\lambda)$ $(= f(\lambda))$, which is needed to preserve the Grassmannian image as we say in Theorem 8.3, we find the Laurent series expansion of f about zero and infinity must be even. Since for 2-unitons only the one-jet at 0 and ∞ act effectively, in this case the action reduces to an action of $\mathrm{SL}(N, \mathbb{C})$ as described for the single unitons. From contour integration we get:

Proposition 10.3. *If $n = 2$, then $\mathfrak{g}$ acts infinitesimally on the Grassmannian 2-unitons by*

$$\delta E_\lambda = (\lambda - 1)[\pi_0^\perp \delta f(0)\pi_0 - \pi_0 \delta f(\infty)\pi_0^\perp + \lambda(\pi_1^\perp \delta f(0)\pi_1 - \pi_1 \delta f(\infty)\pi_1^\perp)],$$

where $\delta f(0)^* + \delta f(\infty) = 0$ *and* $\delta f'(0) = \delta f'(\infty) = 0$,

$$E_\lambda = \pi_0 + \lambda(\pi_1 - \pi_0) + \lambda^2 \pi_1^\perp.$$

Then

$$\begin{aligned}\delta\pi_0 &= -\pi_0^\perp \delta f(0)\pi_0 + \pi_0 \delta f(\infty)\pi_0^\perp,\\ \delta\pi_0^\perp &= \pi_1^\perp \delta f(0)\pi_1 - \pi_1 \delta f(\infty)\pi_1^\perp.\end{aligned}$$

The actual significance of the solutions constructed in this section is that they are the only solutions which are fixed by the S^1 action of §7.

Proposition 10.4. *An extended harmonic map E_λ is fixed by the S^1 action if and only if it satisfies the hypotheses of Theorem* 10.1.

Proof. In Theorem 7.1, we have $\gamma^\# E_\lambda = E_{\lambda\gamma}E_\gamma^{-1}$. In our case, $E_\lambda = \sum_{\alpha=0}^n \lambda^\alpha \tilde\pi_\alpha$, where the $\tilde\pi_\alpha$ are Hermitian projections on mutually orthogonal sub-bundles. Certainly

$$\gamma^\# E_\lambda = \left(\sum_{\alpha=0}^n (\lambda\gamma)^\alpha \tilde\pi_\alpha\right)\left(\sum_{\beta=0}^n \gamma^{-\beta}\tilde\pi_\beta^*\right) = \sum_{\alpha=0}^n (\lambda\gamma)^\alpha \gamma^{-\alpha}\tilde\pi_\alpha = E_\gamma.$$

Therefore, these extended solutions are fixed points for the S^1 action.

Now suppose conversely that $\gamma^\# E_\lambda = E_\lambda$. Then

$$\gamma^\# E_\lambda = \left(\sum_{\alpha=0}^n (\lambda\gamma)^\alpha T_\alpha\right)\left(\sum_{\beta=0}^n \gamma^{-\beta} T_\beta^*\right) = \sum_{\alpha=0}^n \lambda^\alpha \sum_{\beta=0}^n \gamma^{\alpha-\beta} T_\alpha T_\beta^* = \sum_{\alpha=0}^n \lambda^\alpha T_\alpha.$$

It follows that

$$\sum_{\beta=0}^{n} \gamma^{\alpha-\beta} T_\alpha T_\beta^* = T_\alpha,$$

or $T_\alpha T_\beta^* = 0$ for $\alpha \neq \beta$ and $T_\alpha T_\alpha^* = T_\alpha$. Since $T_\alpha^* = (T_\alpha T_\alpha^*)^* = T_\alpha T_\alpha^* = T_\alpha$, we can conclude that $T_\alpha = \tilde{\pi}_\alpha$, where $\tilde{\pi}_\alpha$ is the Hermitian projection on a sub-bundle $\tilde{\eta}_\alpha \subset \Omega \times \mathbb{C}^N$. Because $\tilde{\pi}_\alpha \tilde{\pi}_\beta = 0$ for $\alpha \neq \beta$, the sub-bundles are all mutually orthogonal. From the harmonic map equations $\bar{\partial} E_\lambda = (1-\lambda) E_\lambda A_{\bar{z}}$ and $\partial E_\lambda = (1-\lambda^{-1}) E_\lambda A_z$, we obtain

$$\bar{\partial} \tilde{\pi}_\alpha = (\tilde{\pi}_\alpha - \tilde{\pi}_{\alpha-1} A_{\bar{z}}, \qquad \partial \tilde{\pi}_\alpha = (\tilde{\pi}_\alpha - \tilde{\pi}_{\alpha+1}) A_z,$$

when $\tilde{\pi}_{-1} = \tilde{\pi}_{n+1} = 0$. From these two equations, we verify Theorem 10.2(b), (c) for the sub-bundles

$$\eta_\alpha = \bigoplus_{\beta \leq \alpha} \tilde{\pi}_\beta.$$

11. Global conservation laws and finiteness

In our constructions up to this point, the only topological property we have used is simple connectivity of $\Omega \subset \mathbb{R}^2$. However, our global classification theorems apply only to harmonic maps from $S^2 = \mathbb{R}^2 \cup \{\infty\} \to G_{\mathbb{R}}$. Recall that if $\Omega = \mathbb{R}^2 = S^2 - \{\infty\}$, a harmonic map $s\colon \Omega \to X$ can be regarded as going from the domain $S^2 - \{\infty\} \to X$ due to the conformal invariance in the domain of the harmonic map equations. The following theorem is a special case of a more general theorem proved in Sacks-Uhlenbeck [18].

Theorem 11.1. *If $s\colon \mathbb{R}^2 \to G_{\mathbb{R}}$ is harmonic and $\int_{\mathbb{R}^2} |ds|^2\, dx < \infty$, then $s\colon \mathbb{R}^2 = S^2 - \{\infty\} \to G_{\mathbb{R}}$ extends to a smooth harmonic map $\tilde{s}\colon S^2 \to G_{\mathbb{R}}$.*

Corollary 11.2. *If $s\colon \mathbb{R}^2 \to G_{\mathbb{R}}$ is harmonic and $\int_{\mathbb{R}^2} |ds|^2\, dx < \infty$, then the entire §§1–8 apply to $\tilde{s}\colon \Omega \to G_{\mathbb{R}}$, where $\Omega = S^2$.*

Proof. The only properties we used in §§1–8 were simple connectivity and connectivity of Ω. In particular, $E_\lambda\colon S^2 \to G$ plays the same role for $\Omega \subset S^2$ as $\Omega \subset \mathbb{R}^2$.

From now on in this section we assume the harmonic map $s\colon S^2 \to G_{\mathbb{R}}$. Note that for any n-uniton, if we write the equation $\bar{\partial} E_\lambda = (1-\lambda) E_\lambda A_{\bar{z}}$ and expand in power series, the first term is $E_0 A_{\bar{z}} = 0$. It follows that $A_{\bar{z}}$ must have a kernel. In fact, much more is true on S^2.

Theorem 11.3. *If $s\colon S^2 \to G_{\mathbb{R}}$ is harmonic, then $A_{\bar{z}} = -(A_z)^* = \frac{1}{2} s^{-1} \bar{\partial} s$ is nilpotent, that is, $A_{\bar{z}}^k = 0$ for $k \leq N$.*

Proof. Return to equation (11). On S^2, $A_{\bar z}\, d\bar z$ is a section of $\mathfrak{g} \otimes \Sigma^+$, where Σ^+ is the appropriate line bundle generated by $d\bar z$. The power $(A_{\bar z})^p(d\bar z)^p$ is a section of a similar product $\mathfrak{g} \otimes (\Sigma^+)^p$ over S^2. For all p

$$\partial((A_{\bar z})^p(d\bar z)^p + [A_z, (A_{\bar z})^p(d\bar z)^p])\, dz = 0.$$

It follows that $\partial\, \mathrm{tr}(A_{\bar z})^p(d\bar z)^p = 0$. But $\mathrm{tr}(A_{\bar z})^p(d\bar z)^p$ is a section of $(\Sigma^+)^p$, which has no nonzero anti-holomorphic sections. It follows that $\mathrm{tr}(A_{\bar z})^p = 0$, so all the eigenvalues of $A_{\bar z}$ vanish.

Incidentally, the minimal surface condition on the harmonic map $s\colon \Omega \to G_{\mathbf{R}}$ is $\mathrm{tr}(A_{\bar z})^2 = 0$.

Even more important to our theory is the fact that our solutions are n-unitons for $n < \infty$. The following theorem is related to Theorem 3.3.

Theorem 11.4. *If $E_\lambda\colon \Omega \to G$ is an extended harmonic map, then for all $\lambda \in \mathbb{C}^*$, $E_\lambda = X$ solves*

$$LX = \bar\partial\partial X - \partial X A_z - \bar\partial X A_z = 0. \tag{34}$$

Proof. First rewrite (34) slightly as

$$2LX = \partial(\bar\partial X - 2XA_{\bar z}) + \bar\partial(\partial X - 2XA_z).$$

Here we used (9)(a). From (2)

$$\begin{aligned} LE_\lambda &= -\partial((1+\lambda)E_\lambda A_{\bar z}) - \bar\partial((1+\lambda^{-1})E_\lambda A_{\bar z}) \\ &= -E_\lambda((1+\lambda)(1-\lambda^{-1})A_z A_{\bar z} + (1+\lambda^{-1})(1-\lambda)A_{\bar z}A_z \\ &\quad + (1+\lambda)\partial A_{\bar z} + (1+\lambda^{-1})\bar\partial A_z) \\ &= E_\lambda((\lambda - \lambda^{-1})[A_{\bar z}, A_z] + \lambda\partial A_{\bar z} + \lambda^{-1}\bar\partial A_{\bar z} = 0. \end{aligned}$$

Now L is clearly an elliptic operator, and on S^2 has a finite-dimensional kernel. This is the basis of our finiteness theorem.

Theorem 11.5. *Let $E\colon \mathbb{C} \times S^2 \to G$ be the unique extended solution for a harmonic map $S^2 \to G_{\mathbf{R}}$ with $E_\lambda(p) = I$ for $p \in S^2$. Then E_λ has a finite power series expansion*

$$E_\lambda = \sum_{\alpha=-m}^{n} \lambda^\alpha T_\alpha.$$

Proof. We let $E_\lambda = \sum_{\alpha=-\infty}^{\infty} \lambda^\alpha T_\alpha$, where the Laurent expansion of E in λ is unique, but possibly infinite. Since $L(E_\lambda) = 0$, $L(T_\alpha) = 0$ for all α. Note $T_\alpha(0) = 0$ for $\alpha \neq 0$ for $\alpha \neq 0$, and that

$$T_\alpha = \frac{1}{2\pi i} \oint_{|\lambda|=R} E_\lambda \lambda^{-\alpha-1}\, d\lambda.$$

The terms T_α depend on $q \in S^2$; $T_\alpha \colon S^2 \to \mathfrak{g}$. Suppose $n = \infty$. Then, due to the finiteness of the kernel of L, for some $q \geq 1$

$$T_{q+1} + \sum_{j=1}^{q} a_j T_{q-j+1} = 0.$$

From the uniqueness of E_λ, we then have for $l \geq 1$

$$T_{q+l} + \sum_{j=1}^{q} a_j T_{q-j+l} = 0.$$

Thus

$$\left(1 + \sum_{j=1}^{q} a_j \lambda^j\right) E_\lambda = \tilde{E}_\lambda,$$

where $\tilde{E}_\lambda$ has a finite pole of order q at $\lambda = \infty$. But

$$\begin{aligned}
|T_\alpha| &= \frac{1}{2\pi i}\left|\lim_{R\to\infty} \oint_{|\lambda|=R} E_\lambda \lambda^{-\alpha-1}\, d\lambda\right| \\
&= \frac{1}{2\pi}\left|\lim_{R\to\infty} \oint_{|\lambda|=R} \tilde{E}_\lambda \left(1 + \sum_{j=1}^{q} a_j \lambda^j\right)^{-1} \lambda^{-\alpha-1}\, d\lambda\right| \\
&\leq \frac{1}{2\pi} \lim_{R\to\infty} \oint_{|\lambda|\to\infty} |\tilde{E}_\lambda| \left|\left(\sum_{j=1}^{q} a_j \lambda^j\right)^{-1}\right| R^{-\alpha-1}\, d\lambda \\
&= \frac{1}{2\pi} \lim_{R\to\infty} (|A_a|\, |a_l|^{-1} R^{-\alpha-l+q}).
\end{aligned}$$

Here $|a_l| \neq 0$ and A_q is the top term in the Laurent series for $\tilde{E}_\lambda$ at $\lambda = \infty$. Clearly $T_\alpha = 0$ for $\alpha > q - l$.

A similar argument at 0 shows that the order end of the Laurent series is also finite.

12. Adding a uniton by singular Backlund transformation

In this section we give a general procedure for generating new, extended solutions from a given solution. At least for the Euclidean (elliptic) case, this procedure differs from the continuous group actions described in §§6 and 7. In fact, the Backlund transformations described in §6 depend on a parameter $\alpha \in C^* - \{S^1\}$. As $|\alpha| \to 1$, the Backlund transformations approximate I. However, as $\alpha \to (0, \infty)$, the pair of ordinary differential equations degenerates into a single Cauchy-Riemann equation and a functional constraint. This degenerate Backlund transformation is the equation for adding or subtracting a uniton.

Care must be taken. Factorization is not unique and depends on many choices. For example, all harmonic maps $S^2 \to \mathrm{SU}(2)$ are known to be single unitons, so we shall have to explain exactly when new solutions can be obtained by this method.

In this section, Ω is any two-dimensional simply connected domain. We abandon the normalization $E_\lambda(p) = I$ and work with arbitrary extended solutions. Also $G = \mathrm{GL}(N, \mathbb{C})$ as before and $G_{\mathbf{R}} = \mathrm{U}(N)$.

Theorem 12.1. *Let $E_\lambda\colon \mathbb{C}^* \times \Omega \to G$ be an extended harmonic map, $E_\lambda\colon \Omega \to G_{\mathbf{R}}$ for $|\lambda| = 1$. Then $\tilde{E}_\lambda = E_\lambda(\pi + \lambda\pi^\perp)$ is an extended map for $\pi\colon \Omega \times \mathbb{C}^N \to \eta$ a Hermitian projection if and only if*

(a) $\pi^\perp A_z \pi = 0$,

(b) $\pi^\perp(\bar\partial\pi + A_{\bar z}\pi) = 0$,

where $A_{\bar z} = \frac{1}{2}s^{-1}\bar\partial s = -(A_z)^*$.

Proof. First check the Hermitian condition. Since $\tilde{E}_\lambda^{-1} = (\pi + \lambda^{-1}\pi^\perp)E_\lambda^{-1}$, we have

$$\tilde{E}_{\bar\lambda^{-1}} = (\pi + \lambda\pi^\perp)(E_{\bar\lambda^{-1}}^{-1}) = (\pi + \bar\lambda\pi^\perp)^* E_{\bar\lambda}^* = (E_{\bar\lambda}(\pi + \bar\lambda\pi^\perp))^* = \tilde{E}_{\bar\lambda}^*.$$

Also, if we recall the fact that $\pi^\perp = I - \pi$, we note

$$\begin{aligned}\tilde{E}_\lambda^{-1}\partial\tilde{E}_\lambda &= (\pi + \lambda^{-1}\pi^\perp)E_\lambda^{-1}\bar\partial(E(\pi + \lambda\pi^\perp)) \\ &= (\pi + \lambda^{-1}\pi^\perp)(E_\lambda^{-1}\bar\partial E_\lambda(\pi + \lambda\pi^\perp) + (1 - \lambda)\bar\partial\pi).\end{aligned}$$

Now from the definition of extended solution, $(1 - \lambda)E_\lambda^{-1}\bar\partial E_\lambda = A_{\bar z}$, so

$$\tilde{A}_{\bar z} = (1 - \lambda)^{-1}\tilde{E}_\lambda^{-1}\bar\partial\tilde{E}_\lambda = (\pi + \lambda\pi^\perp)(A_{\bar z}(\pi + \lambda\pi^\perp) + \bar\partial\pi). \tag{35}$$

From Theorem 2.3, $\tilde{E}_\lambda$ is an extended solution if the λ^{-1} term (a) and the λ term (b) vanish. One checks that the lack of dependence of $(1 - \lambda^{-1})\tilde{E}_\lambda^{-1}\partial\tilde{E}_\lambda$ on λ yields the transpose of (a) and (b).

Conditions (a) and (b) can be interpreted as conditions on the sub-bundle η, the sub-bundle on which π is the Hermitian projection. Condition (a) says $A_z\colon \eta \to \eta$ and condition (b) says that η is holomorphic in the complex structure $\bar\partial + A_{\bar z}$. These two conditions are compatible, since by (11) $[\bar\partial + A_{\bar z}, A_z] = 0$. For further reference, we restate Theorem 12.1 in this language.

Corollary 12.2. *Let $E_\lambda\colon \mathbb{C}^* \times \Omega \to G$ be an extended harmonic map satisfying $E_\lambda\colon S^1 \times \Omega \to G_{\mathbf{R}} = \mathrm{U}(N)$. Then*

$$\tilde{E}_\lambda = E_\lambda(\pi + \lambda\pi^\perp)$$

is an extended harmonic map for $\pi\colon \Omega\times\mathbb{C}^N \to \eta$, *the Hermitian projection onto a sub-bundle* $\eta \subset \Omega\times\mathbb{C}^N$ *of rank between* 1 *and* $N-1$, *if and only if*

(a) η *is holomorphic in the* $\bar\partial + A_{\bar z}$ *complex structure,*

(b) $A_z\colon \eta \to \eta$.

For future reference, there is a variety of ways to construct such bundles η. Our difficulty will be with the uniqueness of the construction. First we make a general definition and give some general comments on the definition.

Definition 12.3. Let $A\colon \Omega\times\mathbb{C}^N \to \Omega\times\mathbb{C}^N$ be an analytic linear bundle endomorphism. Then A has maximal rank on an open set Ω'. Define

$$\text{kernel bundle of } A|\Omega' = \{(v,q)\colon A(q)v = 0, q\in\Omega'\} \subset \Omega'\times\mathbb{C}^N,$$
$$\text{range bundle of } A|\Omega' = \{(v,q)\colon v = A(q)w,\ \text{all } w\in\mathbb{C}^N, q\in\Omega'\}.$$

In all the cases we deal with, A is holomorphic or anti-holomorphic with respect to complex structures on $\Omega\times\mathbb{C}^N$, and the points where its rank is not maximal are isolated. Likewise, it follows that the range and kernel bundles of A are holomorphic in an appropriate associated complex structure and extend to smooth bundles. If the appropriate complex bundles are used instead, as would be proper in algebraic geometry, the kernel and range bundles are just sheaves, which in dimension 1 have double duals which are bundles. We make the computations to show the appropriate complex structure, but tacitly ignore the problem of singularities.

Let η_K be the kernel bundle of A_z (note that over S^2, $A_z\colon S^2\times\mathbb{C}^N \to S^2\times\mathbb{C}^N\times\Sigma^+$, but the definition of kernel bundle is still valid).

Lemma 12.4. *The bundle* η_K *is holomorphic in the* $\bar\partial + A_{\bar z}$ *complex structure.*

Proof. Suppose $f\colon \Omega \to \eta_K$. Then $A_z f = 0$. But

$$(\bar\partial + A_{\bar z})(A_z f) = A_z(\bar\partial + A_{\bar z} f) + [\bar\partial + A_{\bar z}), A_z]f = 0$$

and

$$[(\bar\partial + A_{\bar z}), A_z] + (\bar\partial A_z + [A_{\bar z}, A_z])f = 0$$

from the harmonic map equations. It follows that $(\bar\partial + A_{\bar z})f\colon \Omega \to \eta_K$. This shows that η_K is holomorphic in the $\bar\partial + A_{\bar z}$ complex structure.

Corollary 12.5. *Let* $\eta \subset \eta_K$ *be any sub-bundle of* A_z *which is holomorphic in the complex structure* $\bar\partial + A_{\bar z}$. *Let* π *be the Hermitian projection on* η. *Then* $\tilde E_\lambda = E_\lambda(\pi + \lambda\pi^\perp)$ *is an extended solution.*

To understand the complicated possibilities, we list some other possible choices of η. We check that the cokernel bundle

$$\eta_c = \{(q,v)\colon A_{\bar z}(q)v = -A_z^*(q)v = 0\} \subset \Omega\times\mathbb{C}^N$$

is anti-holomorphic, namely $\partial + A_z$ maps sections of η_c to sections of η_c. Choose any anti-holomorphic bundle of η_c in this structure $\partial + A_z$. Then call this $\eta^\perp \subset \eta_c$. Check that η is now holomorphic in the structure $\bar\partial + A_{\bar z}$. Clearly $\pi^\perp A_z \pi = 0$ since $\pi^\perp$ is the projection on $\eta^\perp$ and $(A_z)^*$ maps elements of $\eta^\perp$ to zero.

Other choices for η include the kernel bundle for $(A_z)^k$ and a variety of other possibilities. However, we knew from the start that the construction is not unique. We shall show that the type of construction leading to Corollary 12.5 is sufficient to produce every n-uniton solution, $n < \infty$.

Theorem 12.6. *Suppose that E_λ is an extended solution and $\tilde E_\lambda = E_\lambda(\pi + \lambda\pi^\perp)$ is also. Then if $A_{\bar z} = \frac{1}{2}E_\lambda\bar\partial E_\lambda^{-1}$ and $\tilde A_z = \frac{1}{2}\tilde E_\lambda^{-1}\bar\partial \tilde E_\lambda$,*

$$\tilde A_{\bar z} = \pi A_{\bar z}\pi + \pi^\perp A_{\bar z}\pi^\perp + \pi\bar\partial\pi = A_{\bar z} + \bar\partial\pi.$$

Proof. The terms in (35) which do not contain λ are

$$\tilde A_{\bar z} = (1 - \lambda^{-1}\tilde E_\lambda^{-1}\bar\partial \tilde E_\lambda = \pi A_{\bar z}\pi + \pi^\perp A_{\bar z}\pi + \pi\bar\partial\pi$$

which is automatically

$$A_{\bar z} + \bar\partial\pi - \pi^\perp A_{\bar z}\pi - \pi^\perp\bar\partial\pi - \pi A_{\bar z}\pi^\perp.$$

The last three terms vanish by Theorem 12.1.

Note that it is extremely hard to make $\tilde A_{\bar z} = A_{\bar z}$, since then $\bar\partial\pi = 0$. Then the sub-bundle $\eta \subset \Omega \times \mathbb{C}^N$ is just $\eta = \Omega \times V_0 \subset \Omega \times \mathbb{C}^N$. Also, if $E_\lambda \colon \mathbb{C}^* \times \Omega \to GL(2, \mathbb{C})$, there is very little choice for π. One checks that in this case $\tilde A_{\bar z} = A_{\bar z} + \partial\pi = 0$ and the attempt to add a uniton to the one uniton solution can only result in cancelling the solution already there.

13. The minimal uniton number

Up until this stage in the second part of our discussion, we have not concerned ourselves with the lack of uniqueness in choosing E_λ. This section will describe a unique choice for which the uniton number n is minimal. Throughout this section we assume the reality condition $(E_{\lambda^{-1}})^{-1} = (E_{\lambda^*})^*$.

Definition 13.1. Given an extended solution $E_\lambda = \sum_{\alpha=0}^n \lambda^\alpha T_\alpha$, $V_0 = V_0(E) =$ linear closure of $\{v \in \mathbb{C}^N : v = T_0(q)w, q \in \Omega, w \in \mathbb{C}^N\}$.

Since everything is analytic, Ω may be taken as any subset of the domain. As usual, we find the uniqueness easier to prove than the existence.

Theorem 13.2. *A solution $E_\lambda = \sum_{\alpha=0}^n \lambda^\alpha T_\alpha$ where $V_0 = V_0(E) = \mathbb{C}^N$ is unique (if it exists).*

Proof. Suppose there were two extended solutions with this property:

$$E_\lambda = \sum_{\alpha=0}^{n} \lambda^\alpha T_\alpha, \qquad E'_\lambda = \sum_{\alpha=0}^{m} \lambda^\alpha T'_\alpha,$$

where $V_0(E) = V_0(E') = \mathbb{C}^N$. Then $E'_\lambda = Q(\lambda)E_\lambda$ for $Q(\lambda) = \sum_{\alpha=-n'}^{m'} \lambda^\alpha Q_\alpha = E'_\lambda E_\lambda^{-1}$. Here $Q(\lambda) \in G$ satisfies the reality condition, and we assume $Q_{-n'} \neq 0$ for $n' \leq n$, $Q_{n'} \neq 0$ for $m' \leq m$. However, if $n' > 0$

$$Q_{-n'}T_0(q) = T'_{-n'}(q) = 0$$

for all $q \in \Omega$. It follows that $Q_{-n'}|V_0(E) = Q_{-n} = 0$ since $V_0(E) = \mathbb{C}^N$, which contradicts the assumption that $n' > 0$. We conclude $n' \leq 0$.

Similarly, we have

$$E_\lambda = (Q(\lambda))^{-1}E'_\lambda = \sum_{\alpha=-m'}^{n'} \lambda^\alpha (Q^*_{-\alpha})E'_\lambda.$$

Repeat the same argument. It shows $m' \leq 0$. It follows that $m' = n' = 0$ and the normalization $E_1 = E'_1 = I$ implies $Q(\lambda) \equiv I$.

The idea behind the existence is simple, although the proof is a little complicated.

Theorem 13.3. *Let n be the minimal uniton number. Then there exists an E_λ with $V_0(E_\lambda) = \mathbb{C}^N$ which has this minimal uniton number.*

Proof. Suppose not. Then choose a solution $E_\lambda = E_{\lambda,0}$ such that $\dim V_0(E)$ is the maximum possible for solutions with minimal uniton number n. Suppose $\dim V_0(E) = M < N$.

Starting with $E_{\lambda,0} = E_\lambda$, by induction we construct the following sequence of solutions $E_{\lambda,j}$, $j \geq 0$ by iteration.

Let $E_{\lambda,j} = \sum_{\alpha=0}^{n} \lambda^\alpha T_{\alpha,j}$. Let $V_j = V_0(E_j) =$ linear closure of $\{v \in \mathbb{C}^N : v = T_{0,j}(q)w$ for all $q \in \Omega$, $w \in \mathbb{C}^N\}$. Let $P_j : \mathbb{C}^N \to \mathbb{C}^N$ be the Hermitian projection on V_j. Define

$$E_{\lambda,j+1} = (\lambda^{-1}P_j^\perp + P_j)E_{\lambda,j}.$$

Since $P_j^\perp T_{0,j} = 0$, it can readily be seen that $E_{\lambda,j+1}$ has the required form. Also

$$T_{0,j+1} = P_j T_{0,j} + P_j^\perp T_{1,j} = T_{0,j} + P_j^\perp T_{1,j}. \tag{36}$$

Since M is the maximal rank possible for $V_0(E_j)$, $\operatorname{rank} P_{k+1} = \dim V_{j+1} \leq M$. However, from the construction, it can be seen that

$$\dim V_{j+1} \geq \dim V_j.$$

It follows that

$$\dim V_{j=1} = \operatorname{rank} P_{j+1} = M$$

for all j. It is also clear that this construction continues for all $j \geq 0$. We claim this is impossible. According to our construction $E_{\lambda,j} = Q_j(\lambda)E_\lambda$, where $Q_j(\lambda) : \mathbb{C}^n \to \mathbb{C}^N$ has the formula

$$Q_j(\lambda) = (P_{j-1}^\perp \lambda^{-1} + P_{j-1}) \cdots (P_0^\perp \lambda^{-1} + P_0).$$

Since $Q_j(\lambda) = E_{\lambda,j}(q)E_\lambda^{-1}(q) = \sum_{\alpha=-n}^n \lambda^\alpha Q_{\alpha,j}$ for all $q \in \Omega$, the leading term

$$\lambda^{-j} P_{j-1}^\perp P_{j-2}^\perp \cdots P_0^\perp \equiv 0$$

for $j > n$. This is impossible by the following proposition.

Proposition 13.4. *If $P_k^\perp P_{k-1}^\perp v = 0$, then $P_{k-1}^\perp v = 0$.*

Proof. Suppose not. Let $w = P_{k-1}^\perp v$. Since $P_k^\perp w = 0$, we have $w = P_k v' = P_{k-1}^\perp v$. However, by (36), we see that this means $\dim V_k > \dim V_{k-1}$, which is impossible.

14. The unique factorization theorem

In §12 we gave a general method for changing any solution with finite uniton by adding a factor to the right of the extended solution. Canonically this should, we suspect, add a uniton, but there are so many ways to do it that the process is highly nonunique. In this section we prescribe conditions which insure uniqueness of the construction, which is equivalent to a unique factorization of $E_\lambda = \sum_{\alpha=0}^n \lambda^\alpha T_\alpha$. Appropriately, we assume the condition of uniqueness of the extended solution discovered in §13 always holds. If $E_\lambda = \sum_{\alpha=0}^n \lambda^\alpha T_\alpha$, then

$$V_0 = \text{linear span of}\{v = T_0(q)w, q \in \Omega, w \in \mathbb{C}^N\} = \mathbb{C}^N.$$

Note the results of this section apply to harmonic maps $S^2 \to U(N)$ primarily. However any harmonic map $\Omega \to \mathrm{U}(N)$ with finite uniton number also can be obtained by our construction.

Lemma 14.1. *Let $P: \mathbb{C}^N \to \mathbb{C}^N$ be a Hermitian projection, and let η_P be the kernel bundle of PT_0. Then η_P extends to a bundle over Ω which is holomorphic in the complex structure $\bar\partial + A_{\bar z}$. Moreover, $A_z: \eta_P \to \eta_P$.*

Proof. Let π_P be the Hermitian projection on η_P, and $\pi_P^\perp = I - \pi_P$ be the Hermitian projection on the orthogonal complement. The first two terms in the power series for $\bar\partial E_\lambda = (1-\lambda)E_\lambda A_{\bar z}$ and $\partial E_\lambda = (1-\lambda^{-1})E_\lambda A_z$ give us, after multiplication by P,

$$\bar\partial(PT_0) - (PT_0)A_{\bar z} = 0, \qquad (PT_0)A_z = 0. \tag{37}$$

From the first equation and $PT_0\pi_P^\perp = PT_0$ we have

$$\begin{aligned} 0 &= \bar\partial(PT_0) - (PT_0)\pi_P^\perp A_{\bar z} - (\bar\partial(PT_0) - (PT_0)A_{\bar z})\pi_P^\perp \\ &= (PT_0)\pi_P^\perp(\bar\partial\pi_P^\perp + [A_{\bar z}, \pi_P^\perp]) = -(PT_0)\pi_P^\perp(\bar\partial\pi_P + [A_{\bar z}, \pi_P]). \end{aligned}$$

Since $(PT_0)\pi_P = 0$, we have

$$\pi_P^\perp(\bar\partial\pi_P + [A_{\bar z}, \pi_P]) = 0.$$

This implies the holomorphicity of π_P in the $\bar\partial + A_{\bar z}$ complex structure. Likewise

$$PT_0 A_z = PT_0 \pi_P^{\perp} Az = 0$$

or $\pi_P^{\perp} A_z = 0$, which implies $\pi_P^{\perp} A_z \pi_P = 0$, or $A_z\colon \eta_P \to \eta_P$.

Lemma 14.2. *Let π_I be the projection on η_I, the kernel bundle of T_0. Then*

$$\tilde E_\lambda = \lambda^{-1} E_\lambda(\pi_I + \lambda \pi_I^{\perp})$$

is a new solution of uniton number $n-1$. Moreover, if $\tilde A_z = (1-\lambda)^{-1} E_\lambda^{-1} \bar\partial \tilde E_\lambda$, then $\tilde A_z \pi_I^{\perp} = 0$.

Proof. By Theorem 12.1 and Lemma 14.1, $\tilde E_\lambda = \sum_{\alpha=-1}^{n} \lambda^\alpha \tilde T_\alpha$ is a new extended solution. It could be that $\tilde T_{-1} \neq 0$, or $\tilde T_n \neq 0$. However $\tilde T_{-1} = T_0 \pi_I = 0$, and $\tilde T_n = T_n \pi_I^{\perp}$. By construction, range $\pi_I^{\perp}$ = range T_0^* = orthogonal subspace to the kernel of T_0. Since $T_n T_0^* = 0$ by the reality condition, it follows that $T_n|\operatorname{range} T_0^* = 0$, or $T_n \pi_I^{\perp} = 0$.

This shows us that we can build up every n-uniton solution from a $n-1$ uniton solution. Moreover, this decomposition is canonical (we just gave the recipe). However, we are interested in the details of this construction.

Proposition 14.3. *If E_λ is an extended solution with $V_0 = \mathbb{C}^N$, then* $\operatorname{rank} \tilde T_0 > \operatorname{rank} T_0$.

Proof. From the multiplication formula

$$\tilde T_0 = T_0 \pi_I^{\perp} + T_1 \pi_I = T_0 + T_1 \pi_I. \tag{38}$$

Clearly, $\operatorname{rank} \tilde T_0 \geq \operatorname{rank} T_0$. We need to show strict inequality. If equality holds, then, since range T_0 = range $T_0 \pi_I^{\perp}$, we must have

$$\operatorname{range}(T_1 \pi_I) \subset \operatorname{range} T_0.$$

Of course, we do these computations on the open set where the rank is maximal. By the usual arguments, the set where the rank drops is a set of isolated points, and the singularities in the bundles are removable there.

Our goal is to show it is impossible for $\operatorname{range}(T_1\pi_I) \subset \operatorname{range} T_0$ almost everywhere. First we list the relevant terms from the power series expansions of the equations describing the extended solutions

$$T_0 A_z = 0, \quad \bar\partial T_0 = T_0 A_{\bar z}, \quad \partial T_0 = T_0 A_z - T_1 A_{\bar z}.$$

Since $T_0 \pi_I^{\perp} = T_0$ and $T_0 \pi_I = 0$, we conclude from the first equation that $\pi_I^{\perp} A_z = 0$, or $A_z = \pi_I A_z$. From the second equation, we conclude that the range bundle of T_0 is holomorphic. If range $T_1 \pi_I \subset$ range T_0, we conclude from the third equation that the range bundle of T_0 is also antiholomorphic. This implies that the range bundle is a constant subspace

$V_0 \subset \mathbb{C}^N$. By assumption, $V_0 = \mathbb{C}^N$. However, there are several arguments which show $\operatorname{rank} T_0 < n$. In particular, $T_0 T_n^* = 0$ from the reality condition.

Corollary 14.4. *If $s\colon \Omega \to \mathrm{U}(N)$ is a harmonic map with finite uniton number, then its minimal uniton number is less than N.*

Proof. We reformulate the theorem, and show instead that if E_λ is an extended solution with $V_0 = \mathbb{C}^N$ of uniton number j, then $\operatorname{rank} T_0 \le N - j$. This is easily demonstrated by induction. For $j = 1$, the 1-uniton solutions are of the form $\pi + \lambda\pi$, and $\operatorname{rank}\pi \le N - 1$. Proposition 14.3 is the induction step.

We now describe how to build uniquely the n-unitons from the $n-1$ unitons. Recall that we assume $\tilde{E}_\lambda = \sum_{\alpha=0}^{n-1} \tilde{T}_\alpha$ is an n-uniton solution, $\tilde{E}_\lambda^{-1}\bar{\partial}\tilde{E}_\lambda = (1-\lambda)\tilde{A}_{\bar{z}}$ and $\tilde{E}_\lambda^{-1}\partial\tilde{E}_\lambda = (1-\lambda^{-1})\tilde{A}_z$. By Lemma 12.4, $\tilde{\eta}_K =$ kernel bundle of $\tilde{A}_z = \{(v,q)\colon \tilde{A}_z(q)v = 0\}$ is holomorphic in the complex structure $\bar{\partial} + \tilde{A}_{\bar{z}}$. Likewise, by Lemma 14.1, $\tilde{\eta}_I =$ kernel bundle of $\tilde{T}_0$ is holomorphic in the same complex structure. Furthermore, by duality $\tilde{\eta}_I' =$ kernel bundle of $\tilde{T}_n$ is anti-holomorphic with respect to the $\partial + A_z$ structure. Finally, suppose $Z \in \mathbb{C}P^{N-1}$ and $P(Z)\colon \mathbb{C}^N \to \mathbb{C}^N$ is the projection on the one-dimensional complex subspace Z. From Lemma 14.1, $\tilde{\eta}_{P(Z)} = \tilde{\eta}_Z =$ kernel bundle of $P(Z)\tilde{T}_0$ is holomorphic in the $\bar{\partial} + \tilde{A}_{\bar{z}}$ complex structure. The bundles $\tilde{\eta}_Z$ for $Z \in \mathbb{C}P^{N-1}$ are a complex $N-1$ dimensional family of bundles of $\Omega \times \mathbb{C}^N$ of dimension $N-1$. This explains the terminology of the next step.

Proposition 14.5. *If $\tilde{E}_\lambda = \lambda^{-1}E_\lambda(\pi_I + \lambda\pi_I^\perp)$ as in Lemma 14.2, $V_0 = \mathbb{C}^N$ (for E_λ), and $\eta = \eta_I^\perp =$ range bundle of $T_0^* =$ Hermitian complement to η_I, the kernel bundle of T_0, then*

(a) *η is holomorphic in the $\bar{\partial} + \tilde{A}_{\bar{z}}$ complex structure,*

(b) *$\eta \subset \eta_{\tilde{K}} =$ kernel bundle of $\tilde{A}_z$,*

(c) *$\eta \cap \tilde{\eta}_I = 0$, where $\tilde{\eta}_I$ is the kernel bundle of $\tilde{T}_0$,*

(d) *$\eta \not\subset \tilde{\eta}_Z$ for $Z \in \mathbb{C}P^N$, where $\tilde{\eta}_Z$ is the kernel bundle of $P(Z)\tilde{T}_0$.*

Proof. Condition (a) follows from the fact that

$$E_\lambda = \tilde{E}_\lambda(\pi_I^\perp + \lambda\pi_I) = \tilde{E}_\lambda(\pi + \lambda\pi^\perp),$$

where π is the Hermitian projection on $\eta = \eta_I^\perp$. Condition (b) follows from $\tilde{A}_z\pi = 0$ as shown in Lemma 14.2. Condition (c) is equivalent to the statement that $\operatorname{rank}(\tilde{T}_0\pi) = \operatorname{rank}(T_0) = \operatorname{rank}\pi$. Since $\pi^\perp$ is the kernel bundle of T_0, this is true. Condition (d) follows from the fact that $P(Z)T_0 = P(Z)\tilde{T}_0\pi$ cannot be identically zero for any $Z \in \mathbb{C}P^N$, since whenever $P(Z)\tilde{T}_0\pi = 0$, $V_0 \subset Z^\perp$. Since $V_0 = \mathbb{C}^N$, $P(Z)\tilde{T}_0 \ne 0$. This is equivalent to $\eta \not\subset \eta_Z$.

Our main theorem is now at hand.

Theorem 14.6. *Every n-uniton solution can be built in a unique way from an $n-1$ uniton solution by means of factorizing E_λ, the unique extended solutions with $V_0 = \mathbb{C}^N$. We have*

$$E_\lambda = \tilde{E}_\lambda(\pi + \lambda^{-1}\pi^\perp),$$

where $\tilde{E}_\lambda = \sum_{\alpha=0}^{n-1} \lambda^\alpha \tilde{T}_\alpha$ and $\bar{\partial} E_\lambda = (1-\lambda)E_\lambda \tilde{A}_{\bar{z}}$. Here π is a Hermitian projection onto a sub-bundle η of $\Omega \times \mathbb{C}^N$ where

(a) *η is holomorphic in $\bar{\partial} + \tilde{A}_{\bar{z}}$,*

(b) *$\eta \subset$ kernel bundle of $\tilde{A}_z$,*

(c) *$\eta \cap$ kernel bundle of $\tilde{T}_0 = 0$,*

(d) *$\eta \not\subset \tilde{\eta}_Z$ for $Z \in \mathbb{C}P^N$, where $\tilde{\eta}_z$ is the kernel bundle of $P(Z)\tilde{T}_0$.*

Proof. By the previous proposition, E_λ can be built up this way from a unique $\tilde{E}_\lambda$. However, if we define E_λ this way, we certainly have

$$\lambda^{-1}E_\lambda(\pi^\perp + \lambda\pi) = \tilde{E}_\lambda.$$

From (a) and (b), E_λ is an extended solution. From (c), $\operatorname{rank} T_0 = \operatorname{rank}(\tilde{T}_0\pi) = \operatorname{rank}\pi$, and $T_0 \neq 0$. Moreover, π is clearly projection onto the perpendicular of the kernel of T_0, as in the construction. Moreover, we know $\tilde{E}_\lambda$ has one less uniton than E_λ, so E_λ is in truth $E_\lambda = \sum_{\alpha=0}^{n} \lambda^\alpha T_\alpha$ with $T_0 \neq 0$, $T_n \neq 0$. Finally $V_0 =$ linear closure of the range of $T_0 = \mathbb{C}^N$. If not, $P(Z)T_0 = 0$ for some $Z \in \mathbb{C}P^{N-1}$, and $P(Z)T_0 = P(Z)\tilde{T}_0\pi = 0$, or $\tilde{\eta}_Z = \text{kernel } P(Z)\tilde{T}_0 \supset \eta$. Condition (d) forbids this.

We leave it to the readers who have made it this far to discover a nice set of parameters for the moduli space of the solutions. In fact, conditions (c) and (d) are in the right dimensions going to be lower dimensional varieties which lead to redundancy in the description of harmonic maps due to lack of uniqueness in the extended solutions.

15. Complex Grassmannian manifolds again

Recall from §8 that we wish to consider a harmonic map $s\colon \Omega \to G_{k,N}$ as a harmonic map $s \to G_{k,N}$ satisfying the extra condition that $s^2 = I$ and $\operatorname{rank}(I + s) = k$ at each point $x \in \Omega$. We prefer our extended solution to satisfy

$$E_\lambda = E_{-\lambda}(E_{-1})^{-1}.$$

It will turn out that our extended solutions which satisfy $V_0 = \mathbb{C}^N$ will almost have this property, although this is not at all obvious. First we discover what property they will have.

Lemma 15.1. *Suppose that E_λ is the extended solution with $V_0 = \mathbb{C}^N$ for a harmonic map into a Grassmannian. Then $E_\lambda = Q_0 E_{-\lambda} E_{-1} Q_0$ for $Q_0^2 = I$ (or $Q_0 \in G_{k,N}$, some k).*

Proof. Since E_λ is an extended harmonic map associated to a map $s\colon \Omega \to G_{k,N}$, we know that there exists a $Q\colon \Omega \times \mathbb{C}^* \to \mathrm{GL}(N, \mathbb{C})$ satisfying the reality condition, so that if $E'_\lambda = Q(\lambda) E_\lambda$

$$E'_\lambda = E'_{-\lambda}(E'_{-1})^{-1}.$$

Rewriting in terms of E_λ gives us

$$Q(-\lambda)^{-1} Q(\lambda) E_\lambda = E_{-\lambda} E_{-1}^{-1} Q(-1)^{-1}.$$

Now $Q(\gamma)^{-1} Q(-\lambda) = \sum_{\alpha=-m'}^{\alpha=m} \lambda^\alpha Q_\alpha$. However, $E_\lambda = \sum_{\alpha=0}^{n} \lambda^\alpha T_\alpha$ where T_0 spans all of $\mathbb{C}^N$. Therefore, if $Q_{-m'} \neq 0$, $Q_{-m'} T_0 \neq 0$. We conclude that $Q(\gamma)^{-1} Q(-\lambda) = \sum_{\alpha=0}^{\alpha=m} \lambda^\alpha Q_\alpha$. By the same reasoning

$$(Q(-\lambda)^{-1} Q(\lambda))^{-1} E_{-\lambda} = E_\lambda Q(-1) E_{-1}.$$

By the reality condition and the same argument $(Q(-\lambda)^{-1} Q(\lambda))^{-1} = \sum_{\alpha=-m}^{0} \lambda^\alpha (Q_{-\alpha})^*$, where $m = 0$. So

$$Q(-\lambda)^{-1} Q(\lambda) = Q_0 = Q(-1) = Q(-1)^{-1}$$

as we require. This completes the proof of the lemma.

We now describe how to add a uniton to a harmonic map into a Grassmannian.

Theorem 15.2. *Suppose $\tilde{s}$ is a harmonic map into a Grassmannian with extended solution $\tilde{E}_\lambda$, $\tilde{s} = Q_0 \tilde{E}_{-1}$, and $\tilde{E}_\lambda = Q_0 \tilde{E}_{-\lambda} \tilde{E}_{-1}^{-1} Q_0$. Then*

$$E_\lambda = \tilde{E}_\lambda(\pi + \lambda \pi^\perp)$$

is the extended solution for a harmonic map $s = Q_0 E_{-1}$ into a Grassmannian if $\pi\colon \Omega \times \mathbb{C}^N \to \eta$ is the Hermitian projection to the bundle $\eta \subset \Omega \times \mathbb{C}^N$ where

(a) *η is a holomorphic in the $\bar\partial + A_{\bar z}$ complex structure,*

(b) *$A_z\colon \eta \to \eta$,*

(c) *$[\tilde{s}, \pi] = [Q_0 \tilde{E}_{-1}, \pi] = 0$.*

Proof. By Corollary 12.2, E_λ is an extended harmonic map. We need to verify the Grassmann condition $E_\lambda = Q_0 E_{-\lambda} E_{-1}^{-1} Q_0$. But

$$E_\lambda = \tilde{E}_\lambda(\pi + \lambda\pi^\perp) = Q_0 \tilde{E}_{-\lambda} \tilde{E}_{-1}^{-1} Q_0 (\pi + \lambda \pi^\perp) = Q_0 \tilde{E}_{-\lambda} \tilde{s}^{-1}(\pi + \lambda\pi^\perp).$$

Since $\tilde{s}^{-1} = (Q_0 \tilde{E}_{-1})^{-1} = s = Q_0 \tilde{E}_{-1}$, we have

$$\begin{aligned} E_\lambda &= Q_0 \tilde{E}_{-\lambda}(\pi + \lambda\pi^\perp)(\tilde{E}_{-1}^{-1} Q_0) \\ &= Q_0 \tilde{E}_{-\lambda}(\pi - \lambda\pi^\perp)(\pi - \lambda\pi^\perp)(\tilde{E}_{-1}^{-1} Q_0) \\ &= Q_0 E_{-\lambda} E_{-1}^{-1} Q_0. \end{aligned}$$

Note for $\lambda = -1$, this condition $E_{-1} = Q_0 E_{-1}^{-1} Q_0$ implies $s = Q_0 E_{-1} = E_{-1}^{-1} Q_0 = s^{-1}$.

Theorem 15.3. *A harmonic map $s: \Omega \to G_{k,N}$ with finite uniton number n is uniquely obtained from a harmonic map $\tilde{s}: \Omega \to G_{k',N}$ of uniton number $n-1$ by setting the unique extended solution E_λ with $V_0 = \mathbb{C}^N$ for s*

$$E_\lambda = \tilde{E}_\lambda(\pi + \lambda\pi^\perp).$$

Here $s = Q_0 E_{-1}$, $\tilde{s} = Q_0 \tilde{E}_{-1}$ and $\tilde{E}_\lambda$ is the unique solution for $\tilde{s}$ with $V_0 = \mathbb{C}^N$. Moreover, π is the Hermitian projection into a bundle η satisfying (a)–(d) *of Theorem* 14.6 *plus the condition*

(e) $[\tilde{s}, \pi] = 0$.

Proof. By Theorem 14.2, E_λ is uniquely obtainable from $\tilde{E}_\lambda$ as described. Moreover, if $\tilde{s}$ is harmonic into a Grassmannian, by Theorem 15.2, s is also.

We need to show that if s lies in a Grassmannian, then $\tilde{s}$ does also and condition (e) is satisfied. We have

$$E_\lambda = \tilde{E}_\lambda(\pi + \lambda\pi^\perp) = \sum_{\alpha=0}^{n} \lambda^\alpha T^n,$$

where $\operatorname{rank} T_0 = \operatorname{rank} \pi$ and kernel T_0 = kernel π. We also have from Lemma 15.1

$$\begin{aligned} E_\lambda &= Q_0 E_{-\lambda} E_{-1}^{-1} Q_0 \\ &= Q_0 \tilde{E}_{-\lambda}(\pi - \lambda\pi^\perp)(\pi - \pi^\perp)\tilde{E}_{-1}^{-1} Q_0 \\ &= Q_0 \tilde{E}_{-\lambda} \tilde{E}_{-1}^{-1} Q_0(\hat{\pi} + \lambda\hat{\pi}^\perp), \end{aligned}$$

where $\hat{\pi} = (Q_0\tilde{E}_{-1})\pi(Q_0E_{-1})^{-1} = \tilde{s}\pi\tilde{s}^{-1}$. But $\operatorname{rank}\hat{\pi} = \operatorname{rank}\pi = \operatorname{rank} T_0$, and kernel $\hat{\pi}$ = kernel π = kernel T_0. It follows that

$$\pi = \hat{\pi} = \tilde{s}\pi\tilde{s}^{-1}, \quad \text{or } [\pi, \tilde{s}] = 0.$$

Again, by uniqueness, $\tilde{E}_\lambda = Q_0 \tilde{E}_{-\lambda} \tilde{E}_{-1}^{-1} Q_0$, which implies that $(\tilde{s})^2 = I$ as required.

16. Additional questions and problems

The original list of questions which accompanied the first version of this paper (written in the fall of 1984) is contained in problems 1–12 below. Since I now understand more physics, I have added four questions which originate in the papers of theoretical physicists (problems 13–16).

Graeme Segal has since given an insightful interpretation of the construction in terms of the loop group and infinite Grassmannians which should answer problem 12 [19]. Extensive work has been done on the classification of harmonic maps into complex Grassmannians [5], [2], which is the subject of problem 9. Several other details of our construction have been elaborated on [16], [24]. However, the work of Hitchin on harmonic maps from $s^1 \times S^1 \to \mathrm{SU}(2)$ and the explosion of results on constant mean curvature surfaces in $\mathbb{R}^3$ [21] indicate that problem 6 is immensely complicated, and certainly not purely algebraic. Similarly, just as the physicists thought, the theory of harmonic maps into flag manifolds (problem 4) is not yielding [12], [15]. I still think the question of the algebraic geometric structures on the moduli space of solutions is important. Work of Verdier is surely important [20]. Problem 7 is restated in problem 15. Much of the rest may be irrelevant.

In reference to the new problems, I am pessimistic about a positive answer to question 13. I believe problem 14 can be carried out rather easily, but as with all supersymmetric theories, the geometry is hard to come by. Problem 15 relates to questions of Wick rotation. Problem 16 is open-ended. String theory relates to everything.

1. All the proofs in this paper relate to $G_{\mathbb{R}} = \mathrm{U}(N)$. The first eight sections should apply directly to any real form of a complex group if phrased in a sufficiently abstract setting. Can anything interesting be said about specific groups?

2. On the other hand, the ideas in Part II are based entirely on $G_{\mathbb{R}} = \mathrm{U}(N)$. What can be said about constructing general harmonic maps of finite uniton numbers into any $G_{\mathbb{R}}$?

3. Is there a decomposition theorem for harmonic maps from 1-connected 2-dimensional Riemannian domains into $\mathrm{GL}(N, \mathbb{R})$? Into any G? The reality condition plays an essential role in our arguments.

4. Can any of the results of this paper be modified to apply to harmonic maps into $M = \mathrm{U}(N)/G$? (We have in mind flag manifolds.)

5. Solutions of the Ernst equations of general relativity are similar to harmonic maps into $\mathrm{SL}(2, \mathbb{R})$ whose images lie in the self-adjoint matrices. Is there any connection between the constructions described in this paper and the Geroch group?

6. The condition that the domain Ω be simply connected is essential in our construction. Can any of the ideas be modified to include the periods which must arise when Ω is not simply connected?

7. When Ω is in $E^{1,1}$, part I goes over directly by replacing $(z, \bar{z})$ by light-light (characteristic) coordinates (η, ξ). However, there are no *real* unitons

with finite uniton number. This is because the reality condition would be $E_\lambda(E_{\bar\lambda})^* = I$. The conjecture is that this system of equations $E^{1,1} \to G_{\mathbb{R}}$ is completely integrable. Since Sine-Gordon is essentially equivalent to $E^{1,1} \to \mathbb{C}P^1$, some details are known.

8. Can a proof of the Glaser-Stora, Din-Zakrzewski and Eells-Wood classification of harmonic maps into $\mathbb{C}P^{N-1}$ be obtained directly from our construction?

9. If $s: \Omega \to G_{k,N}$ is a harmonic map of finite uniton number, is there a bound on the minimal number which is less than $N-1$? (Note if $k=1$, or $N-1$, $n \leq 2$). I believe the answer is $n \leq 2\min(k, N-k)$.

10. Does the minimal uniton number n have any topological interpretations?

11. In principle, our construction does show that the moduli space of solutions is an algebraic variety. However, the description is very awkward. Is there a more elegant description?

12. What is the connection between Parts I and II of the paper? How does the loop group act on the abstract space of solutions?

13. Is the group action of §§5–7 symplectic with respect to a natural Hamiltonian structure?

14. Is there a supersymmetric version of this paper?

15. What is the relationship between the Minkowski and Euclidean harmonic map problems? Are there any structures (classical or quantum) which transform nicely?

16. Does any of the theory in this paper carry over into a quantum theory? If so, does it have anything to do with string theories?

References

[1] A. A. Belavin &V. E. Zakharov, *Yang-Mills equations as inverse scattering problems*, Phys. Lett. **73B** (1978) 53–75.

[2] F. Burstall & J. Wood, *The construction of harmonic maps into complex Grassmannians*, J. Differential Geometry **23** (1986) 255–297.

[3] L. L. Chau, G. M. Lin & W. Y. Shi, *Kac-Moody algebra in the self-dual Yang-Mills equation*, Phys. Rev. D (3) **25** (1982) 1086–1094.

[4] L. L. Chau, M. Prasad & A. Sinha, *Some aspects of the linear system for self-dual Yang-Mills fields*, Phys. Rev. D (3) **24** (1981) 1574–1579.

[5] S. S. Chern & J. Wolfson, *Harmonic maps of S^2 into a complex Grassmann manifold*, Ann. of Math. (2) **125** (1987) 301–335.

[6] L. Crane, *Action of the loop group on the self-dual Yang-Mills equation*, Comm. Math. Phys. **110** (1987), 391–414.

[7] A. M. Din & W. J. Zakrzewski, *General classical solutions in the $\mathbb{C}P^{n-1}$ models*, Nuclear Phys. B **174** (1980) 397–406.

[8] L. Dolan, *Kac-Moody algebra is hidden symmetry of chiral models*, Phys. Rev. Lett. **49** (1981) 1371–1374.

[9] ——, *Kac-Moody algebras and exact solvability in hadronic physics*, Phys. Rep. **109** (1984) 3–94.

[10] J. Eells & J. C. Wood, *Harmonic maps from surfaces to complex projective spaces*, Advances in Math. **49** (1983) 217–263.

[11] V. Glaser & R. Stora, *Regular solutions of the* $\mathbb{C}P^n$ *model and further generalizations*, preprint, CERN, 1980.

[12] M. Guest, *The geometry of maps between generalized flag manifolds*, J. Differential Geometry **25** (1987) 223–248.

[13] J. W. Helton, *Operator theory, analytic functions and electrical engineering*, CBMS Regional Conference, Lincoln, Nebraska, 1985.

[14] N. Hitchin, *Harmonic maps from* T^2 *into* S^3, Harmonic Mappings, Twistors and σ-Models (P. Gauduchon, ed.), World Scientific Press, 1988.

[15] C. Negreiros, *Harmonic maps from compact Riemann surfaces into flag manifolds*, Thesis, University of Chicago, 1987.

[16] B. Piette & W. Zakrzewski, *Properties of classical solutions of the* $U(N)$ *chiral* σ *models in two dimensions*, preprint, Durham, 1987.

[17] A. Pressley & G. Segal, *Loop groups*, Clarendon Press, Oxford, 1986.

[18] J. Sacks & K. Uhlenbeck, *The existence of minimal* 2*-spheres*, Ann. of Math. (2) **113** (1981) 1–24.

[19] G. Segal, *Harmonic maps from a surface into* U_n (*after Uhlenbeck*), hand-written notes, 1986.

[20] J. L. Verdier, *Two dimensional* σ*-models and harmonic maps from* $S62$ *to* S^{2n}, Lecture Notes in Phys., Vol. 180, Springer, Berlin, 1982, 136–141.

[21] H. Wente, *Counter-example to a conjecture of H. Hopf*, Pacific J. Math. **121** (1986) 193–243.

[22] V. E. Zakharov & A. V. Mikhailov, *Relativistically invariant two-dimensional models of field theory which are integrable by means of the inverse scattering problem method*, Soviet Phys. JETP **47** (1978) 1017–1027.

[23] V. E. Zakharov & A. B. Shabat, *Integration of non-linear equations of mathematical physics by the method of inverse scattering*. II, Funktsional Anal. i Prilozhen. **13** (1978) 13–22.

[24] W. J. Zakrzewski, *Some classical solutions of* $U(N)$, preprint.

UNIVERSITY OF CHICAGO
UNIVERSITY OF TEXAS, AUSTIN