Journal of Geometry and Physics 8 (1992) 283–316
North-Holland

JOURNAL OF
GEOMETRY AND
PHYSICS

On the connection between harmonic maps and the self-dual Yang–Mills and the sine-Gordon equations

Karen Uhlenbeck
The University of Texas at Austin, Austin, TX 78712, USA

Received 4 May 1991

This is an algebraic paper which demonstrates some algebraic connections between the Yang–Mills and the sine-Gordon equation.

Keywords: *harmonic maps, self-dual Yang–Mills equation, sine-Gordon equation*
1991 MSC: 43A99, 81T13

The algebraic study of harmonic maps as pursued in this paper had its first appearance in papers by Calabi in 1967 [1] and Chern in 1970 [2] on minimal surfaces in spheres. Those papers were widely read, but were developed more fully only after the appearance of work by theoretical physicists, who became interested in harmonic maps from two-dimensional domains as toy models for gauge theories. Work of these physicists [3,4] on classification problems greatly stimulated the mathematics community [5], and by now the literature is too large to detail in an introduction. However, the extensive work of Zakrzewski and his students is fundamental to newer developments [16,17].

My own interest was also inspired by work of theoretical physicists on Kac–Moody Lie algebras [6–8]. The concise development of their ideas led to my first paper, "Harmonic maps into Lie groups" [9], of which the present paper is a follow-up. As we said before, originally some of this work with chiral models, or minimal surfaces in groups, was meant as a toy model for gauge theories. At one time it was believed that there might be a more direct relationship with conformal field theories, but I believe that seems doubtful at the moment.

This is an algebraic paper, which simply demonstrates some algebraic connections among the various equations. The relationship between the harmonic map equation and the Yang–Mills equation has been known for many years. It was, in fact, Roger Penrose who pointed out to me what is in the end the correct interpretation of the reality conditions. To obtain the sigma model naturally from the Yang–Mills equations, one needs to use the somewhat unusual signature $++--$ in four-dimensional space.

Section 1 derives the harmonic maps from E^2 and $E^{1,1}$ into a Lie group G as dimensionally reduced self-dual Yang–Mills on $E^{2,2}$ with structure group G. Section 2 describes the gauge-invariant harmonic map equations and the construction of the physicists' Wess–Zumino terms from this gauge-invariant equation. Section 3 describes the equation which must appear as we dimensionally reduce self-dual Yang–Mills equations in $E^{2,2}$ along planes which rotate from definite E^2 to indefinite $E^{1,1}$ signatures. Section 4 contrasts the global algebraic structures of harmonic maps from E^2 (or S^2) and from $E^{1,1}$. Section 5 gives what I consider an elegant gauge-theoretic derivation of the relationship between harmonic maps from $E^{1,1}$ into S^2 and the sine-Gordon equation. Section 6 translates our loop group action from harmonic maps to the sine-Gordon equation and shows how the Bäcklund transformations are really the same in each case.

My original goal in pursuing these ideas is to find a geometric approach to a quantum theory. An additional paper on symplectic and Poisson structure will follow. However, the original goal is elusive.

The reader will find this paper much easier to read if portions of the preceding paper, "Harmonic maps into Lie groups" [9], are read first.

1. Dimensional reduction from Yang–Mills

The self-dual Yang–Mills equations have become a fundamental tool in smooth four-manifold topology. The three-dimensional monopole equation, obtained by dimensional reduction from the four-dimensional self-dual equations, has generated nearly as much interest [10,11]. Further reduction to two dimensions yields a set of equations in a Riemann surface which have been extensively studied by Hitchin and are interesting in their own right [12].

A further bit of background information is that the Minkowski (signature $-+++$ or $+---$) version of the self-dual Yang–Mills equations is incompatible with the usual reality conditions required by a compact gauge group. However, a third possibility exists: that of using the signature $++--$ for space–time. Dimensional reduction to three dimensions yields what should be a fascinating set of hyperbolic first-order equations. In this paper we discuss the further reduction to two dimensions. Most of the reductions lead either to an elliptic or to a hyperbolic equation which is equivalent to the harmonic map equation from domains in E^2 or $E^{1,1}$ into the gauge group G.

We first discuss the elliptic reduction. Both the self-dual formalism and the elliptic harmonic map equation are more familiar to differential geometers. To describe the harmonic map equation, we recall that if $s:M\to G$ is a map, then s is harmonic when

$$d(s^{-1}*ds)=0\,. \tag{1}$$

In the case M is two dimensional, $*: T^*M \to T^*M$ depends only on the conformal structure, and the equation for harmonic maps is a conformal invariant. We lose very little, especially in treating the local theory, by assuming $M = \Omega \subseteq \mathrm{E}^2$ (or in fact $\Omega \subseteq \mathrm{E}^{1,1}$ in the indefinite case).

We always want to work in characteristic coordinate patches. For $\Omega \subseteq \mathrm{E}^2 \simeq \mathbb{C}$, we use the complex coordinates $(z, \bar{z})$, in which the harmonic map equation (1) becomes

$$\partial(s^{-1}\bar{\partial}s) + \bar{\partial}(s^{-1}\partial s) = 0 . \tag{2}$$

Here $\bar{\partial} = \partial/\partial\bar{z}$ and $\partial = \partial/\partial z$. Our formulation fundamentally involves the introduction of the Lie algebra $\mathfrak{g}$-valued one-form

$$A = \tfrac{1}{2}s^{-1}\,\mathrm{d}s = A_{\bar{z}}\,\mathrm{d}\bar{z} + A_z\,\mathrm{d}z = \tfrac{1}{2}(s^{-1}\bar{\partial}s\,\mathrm{d}\bar{z} + s^{-1}\partial s\,\mathrm{d}z) . \tag{3}$$

Note $\{A_z(p), A_{\bar{z}}(p)\} \subseteq \mathfrak{g} \otimes \mathbb{C} = \mathfrak{g}_{\mathbb{C}}$.

We get two equations for A when s is harmonic. Equation (4a) is the harmonic equation and eq. (4b) is the local consistency condition for the existence of $s^{-1}\,\mathrm{d}s = 2A$:

$$\mathrm{d}{*A} = (\partial A_{\bar{z}} + \bar{\partial}A_z)\,\mathrm{d}z \wedge \mathrm{d}\bar{z} = 0 , \tag{4a}$$

$$\mathrm{d}A + (A \wedge A) = (\partial A_{\bar{z}} - \bar{\partial}A_{\bar{z}} + 2[A_z, A_{\bar{z}}])\,\mathrm{d}z \wedge \mathrm{d}\bar{z} = 0 . \tag{4b}$$

In a simply connected domain $\Omega \subseteq \mathbb{C}$, a solution to (4) is equivalent to the existence of a harmonic map $s: \Omega \to G$ satisfying (2) and (3). However, the form in which our harmonic map equation is to be recognized is through the parameterized family of flat connections $D(\lambda)$ for $\lambda \in \mathbb{C}^*$,

$$D_\lambda = [\bar{\partial} + (1-\lambda)A_{\bar{z}}]\,\mathrm{d}\bar{z} + [\partial + (1-\lambda^{-1})A_z]\,\mathrm{d}z . \tag{5}$$

The connection one-form of D_λ has values in the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ for $\lambda \in \mathbb{C}^*$ but in the Lie algebra for $\lambda = (\bar{\lambda})^{-1} \in \mathrm{S}^1$.

Theorem 1.1. *The Lie algebra-valued one-form A represents a harmonic map $s: \Omega \to G$ if and only if the associated connections D_λ are flat $\lambda \in \mathrm{S}^1$, i.e.,*

$$[\bar{\partial} + (1-\lambda)A_{\bar{z}}, \partial + (1-\lambda^{-1})A_z] = 0 . \quad \square \tag{6}$$

Now we review the self-dual Yang–Mills equation in a similar formalism. First we remind our readers of the usual definite self-dual equations on $\mathcal{O} \subseteq \mathrm{E}^4$. The independent variable is the Lie algebra-valued one-form

$$A = \sum_{i=1}^{4} A_i\,\mathrm{d}x^i .$$

The one-form A is self-dual when the curvatures

$$F_{ij}=[\partial/\partial x^i+A_i,\ \partial/\partial x^j+A_j]$$

satisfy

$$F=\sum F_{ij}\,\mathrm{d}x^i\wedge\mathrm{d}x^j=*F\,.$$

In the signature of E^4, this becomes in coordinates just

$$F_{12}=F_{34}\,,\qquad F_{13}=-F_{24}\,,\qquad F_{14}=F_{23}\,. \tag{7}$$

However, it is more useful to introduce characteristic coordinates into $\mathrm{E}^4\simeq\mathbb{C}^2$ by setting $z=x^1+\mathrm{i}x^2$, $w=x^3+\mathrm{i}x^4$, $\bar z=x^1-\mathrm{i}x^2$, $\bar w=x^3-\mathrm{i}x^4$. Then (7) becomes

$$F_{z,\bar w}=F_{w,\bar z}=0\,,\qquad F_{z,\bar z}+F_{\bar w,w}=0\,. \tag{8a,b}$$

The curvature or Lax pair formulation is that A is self-dual if and only if the curvature in certain families of complex two-planes vanishes. This can be written

$$\left[\frac{\partial}{\partial\bar z}+A_{\bar z}+\lambda\left(\frac{\partial}{\partial\bar w}+A_{\bar w}\right),\frac{\partial}{\partial z}+A_z-\lambda^{-1}\left(\frac{\partial}{\partial w}+A_w\right)\right]=0\,, \tag{9}$$

where $\lambda\in\mathbb{C}\mathrm{P}^1$ under the appropriate projectivized formula. Note that we have given the self-dual formulation, which differs from the more natural and usual anti-self-dual formulation by the change of orientation which exchanges $w\leftrightarrow\bar w$ $(x^4\to -x^4)$.

Now the change in signature from E^4 to $\mathrm{E}^{2,2}$ makes a very slight but very important change in these equations. Equation (7) becomes

$$F_{12}=-F_{34}\,,\qquad F_{13}=-F_{24}\,,\qquad F_{14}=F_{23}\,.$$

This leaves (8a) unchanged but replaces (8b) by

$$F_{z,\bar z}+F_{w,\bar w}=0\,. \tag{8c}$$

The Lax pair or twistor formulation for this case follows that for the case of E^4 and definite signature. The proof is just a calculation.

Theorem 1.2. *The Lie algebra-valued one-form A solves the indefinite self-dual Yang–Mills equations in $\mathrm{E}^{2,2}$ if and only if the curvatures in the families of planes parameterized by $\lambda\in\mathbb{C}\mathrm{P}^1$ vanish:*

$$0=\left[\frac{\partial}{\partial\bar z}+A_{\bar z}+\lambda\left(\frac{\partial}{\partial\bar w}+A_{\bar w}\right),\frac{\partial}{\partial z}+A_z+\lambda^{-1}\left(\frac{\partial}{\partial w}+A_w\right)\right].\quad\square \tag{10}$$

The essential difference between the two cases lies in the fact that in the definite case there are no real characteristic planes. The complex planes in the twistor formulation solve

$$\bar{z}=\lambda^{-1}\bar{w}+\alpha\ , \qquad z=-\lambda w+\beta\ .$$

However, in the indefinite $E^{2,2}$ case, the planes become

$$\bar{z}=\lambda^{-1}\bar{w}+\alpha\ , \qquad z=\lambda w+\beta\ ,$$

which for $\lambda=\bar{\lambda}^{-1}$ and $\alpha=\bar{\beta}$ are in a real plane $\mathbb{R}^2\subseteq\mathbb{R}^4$. This leads to the following proposition.

Theorem 1.3. *Given the* $\mathfrak{g}$*-valued one-form A solving (10) in $\mathcal{O}$, a simply connected domain, and $\rho\in S^1$, there exists a gauge transformation $U_\rho\colon\mathcal{O}\to G$ such that $\tilde{A}=u_\rho^{-1}(Au_\rho+\mathrm{d}u_\rho)$ satisfies*

$$\tilde{A}+\rho\tilde{A}_w=0\ , \qquad A_z+\rho A_w=0\ . \tag{11}$$

Proof. Since the curvatures of the real planes parameterized by $\rho\in S^1$ vanish, there exists a gauge transformation which trivializes the connections on these planes.

We do not actually want to carry out this gauge change at this point. One of our goals is to obtain a gauge-invariant formulation of the harmonic map equation.

The first dimensional reduction we accomplish merely by assuming that we have a solution to (11) in $\mathcal{O}=\Omega\times\mathbb{C}\subseteq\mathbb{C}\times\mathbb{C}\simeq E^{2,2}$ which is independent of the variables $(x^3,x^4)\simeq(w,\bar{w})\in\mathbb{C}$. There are fancier ways to say this (we do so later). For some mysterious reason, we obtain an identification of $(\mathrm{d}w\sim\mathrm{d}z,\ \mathrm{d}\bar{w}\sim\mathrm{d}\bar{z})$ if we are to obtain conformal equations. Let $B=B_{\bar{z}}\,\mathrm{d}\bar{z}+B_z\,\mathrm{d}z$, where $A_{\bar{w}}=B_{\bar{z}}$, $A_w=B_z$. □

Theorem 1.4. *Let $A=A_z\,\mathrm{d}z+A_{\bar{z}}\,\mathrm{d}\bar{z}$ and $B=A_{\bar{w}}\,\mathrm{d}\bar{z}+A_w\,\mathrm{d}z$ represent a dimensionally reduced solution to the indefinite self-dual Yang–Mills equations in $\mathcal{O}=\Omega\times\mathbb{C}\subseteq E^{2,2}$ to $\Omega\subseteq E^2$. Then*

$$[\partial/\partial\bar{z}+A_{\bar{z}}+\lambda B_{\bar{z}},\ \partial/\partial z+A_z+\lambda^{-1}B_z]=0\ . \tag{12}$$

Furthermore, if Ω is simply connected, then there exists an (essentially) unique gauge transformation $u\colon\Omega\to G$ such that if

$$\tilde{A}=u^{-1}(Au+\mathrm{d}u)\ , \qquad \tilde{B}=u^{-1}Bu\ ,$$

then $\tilde{A}+\tilde{B}=0$ and

$$[\partial/\partial\bar{z}+(1-\lambda)\tilde{A}_{\bar{z}},\ \partial/\partial z+(1-\lambda^{-1})A_z]=0\ .$$

Furthermore $\tilde{A}=\tfrac{1}{2}s^{-1}\,\mathrm{d}s$, where s is harmonic.

Proof. We obtain (12) from (8a) and (8c) by assuming that the one-form A satisfies $(\partial/\partial\bar{w})A=(\partial/\partial w)A=0$. The gauge change u simply trivializes the flat connection in Ω,

$$D_\lambda = ((\bar{\partial}+A_{\bar{z}}+\lambda B_{\bar{z}})\,d\bar{z},\ (\partial+A_z+\lambda^{-1}B_z)\,dz)\ ,$$

at $\lambda=1$. If you like, $u=u_1$ trivializes the family of flat connections at $\lambda=1$, u_{-1} trivializes the same family at $\lambda=-1$, and $s=u_1^{-1}u_{-1}$ is harmonic. Because the trivialization of a flat connection is unique up to an element of G, s is unique up to multiplication by constants in G on both the left and the right. □

Note that the converse of this theorem is naively true. However, one elegant corollary gives a more geometric explanation of the S^1 action described in section 8 of our first paper [9].

Corollary 1.5. *The rotation $\lambda\to\rho\lambda$, $A\to A$ and $B\to\rho^{-1}B$ results in an* S^1 *action on harmonic maps.*

Proof. This is an obvious description of a transformation of the self-dual Yang–Mills equations which preserves the real structure of G when $\rho=\bar{\rho}^{-1}$ and commutes with dimensional reduction. Therefore, it leads to an action on harmonic maps which is unique up to dividing on the left and right by constants. Let u_ρ trivialize the self-dual connection D_λ. Then $\rho^*s=u_\rho^{-1}u_{-\rho}$. □

In dealing with the Minkowski reduction, we need only remark on the differences between the Euclidean and Minkowski versions. For harmonic maps, we simply replace the characteristic coordinates in E^2, $(z,\bar{z})=(x+iy,\ x-iy)$, by light cone coordinates in $E^{1,1}$, $(\xi,\eta)=(x+t,\ x-t)$. It will serve us conceptually to replace λ by σ also. Then, if we prime the numbers for the Minkowski versions of our Euclidean equations, we have

$$\frac{\partial}{\partial\eta}\left(s^{-1}\frac{\partial s}{\partial\xi}\right)+\frac{\partial}{\partial\xi}\left(s^{-1}\frac{\partial s}{\partial\eta}\right)=0, \tag{2'}$$

$$A=\tfrac{1}{2}s^{-1}\,ds=A_\eta\,d\eta+A_\xi\,d\xi=s^{-1}\frac{\partial s}{\partial\eta}d\eta+s^{-1}\frac{\partial s}{\partial\xi}d\xi\,, \tag{3'}$$

$$\frac{\partial}{\partial\eta}A_\xi+\frac{\partial}{\partial\xi}A_\eta=0\,,\qquad \frac{\partial}{\partial\eta}A_\xi-\frac{\partial}{\partial\xi}A_\eta+2[A_\eta,A_\xi]=0\,, \tag{4a',b'}$$

$$D_\sigma=(\partial/\partial\eta+(1-\sigma)A_\eta\,d\eta\,,\ \partial/\partial\xi+(1-\sigma^{-1})A_\xi\,d\xi)\ ; \tag{5'}$$

note that the flat connection D_σ is defined to lie in $\mathfrak{g}_{\mathbb{C}}$ for $\sigma\in\mathbb{C}^*$, but in $\mathfrak{g}$ for $\sigma\in\mathbb{R}$; clearly $\lambda\neq\sigma$, as D_λ was real for $\lambda\in S^1$;

$$[\partial/\partial\eta+(1-\sigma)A_\xi\,,\ \partial/\partial\xi+(1-\sigma^{-1})A_\xi]=0\,. \tag{6'}$$

The equivalent of theorem 1.1 says that in a simply connected domain $\Omega\subseteq E^{1,1}$, (6′) is equivalent to the existence of the harmonic map $s:\Omega\to G$. This completes a condensed description of the Minkowski version of harmonic maps.

Of course, we need not repeat our introductory remarks on self-dual Yang–Mills equations. However, it will be easier to understand the Minkowski reduction if we introduce variables which are characteristic for this reduction. Let $\xi=x^1+x^3$, $\eta=x^1-x^3$, $\xi'=x^2+x^4$ and $\eta'=x^2-x^4$.

Lemma 1.6. *Equation (10) written in null-cone variables becomes*

$$0=\left[\frac{\partial}{\partial\eta}+A_\eta+\sigma\left(\frac{\partial}{\partial\xi'}+A_{\xi'}\right),\frac{\partial}{\partial\xi}+A_\xi+\sigma^{-1}\left(\frac{\partial}{\partial\eta}+A_\eta\right)\right], \tag{13}$$

where $\sigma=-\mathrm{i}[(\lambda+1)/(\lambda-1)]$.

Proof. Naively we consider (10) to be the vanishing of the commutator of two operators $[L_1, L_2]=0$,

$$L_1=\frac{\partial}{\partial x^1}+A_1+\mathrm{i}\left(\frac{\partial}{\partial x^2}+A_2\right)+\lambda\left[\frac{\partial}{\partial x^3}+A_3+\mathrm{i}\left(\frac{\partial}{\partial x^3}+A_3\right)\right],$$

$$L_2=\lambda\left[\frac{\partial}{\partial x^1}+A_1-\mathrm{i}\left(\frac{\partial}{\partial x^2}+A_2\right)\right]+\frac{\partial}{\partial x^3}+A_3-\mathrm{i}\left(\frac{\partial}{\partial x^4}+A_4\right).$$

Add and subtract the two to get the two operators

$$L_1+L_2=(\lambda+1)\left(\frac{\partial}{\partial x^1}+A_1+\frac{\partial}{\partial x^3}+A_3\right)-\mathrm{i}(\lambda-1)\left(\frac{\partial}{\partial x^2}+A_2-\frac{\partial}{\partial x^4}-A_4\right),$$

$$L_2-L_1=(+\lambda-1)\left(\frac{\partial}{\partial x^1}+A_1-\frac{\partial}{\partial x^3}-A_3\right)-\mathrm{i}(\lambda+1)\left(\frac{\partial}{\partial x^2}+A_2+\frac{\partial}{\partial x^4}+A_4\right).$$

Now introduce the characteristic coordinates (ξ, η) and (ξ', η') as described above. At the same time, divide the first by $2(\lambda+1)$ and the second by $2(\lambda-1)$. This gives the two operators

$$W_1=\frac{L_1+L_2}{2(\lambda+1)}=\frac{\partial}{\partial\xi}+A_\xi-\mathrm{i}\frac{\lambda-1}{\lambda+1}\left(\frac{\partial}{\partial\eta'}+A_{\eta'}\right),$$

$$W_2=\frac{L_2-L_1}{2(\lambda-1)}=\frac{\partial}{\partial\eta}+A_\eta-\mathrm{i}\frac{\lambda+1}{\lambda-1}\left(\frac{\partial}{\partial\xi'}+A_{\xi'}\right).$$

Substitute $\sigma=-\mathrm{i}(\lambda+1)/(\lambda-1)$, and the vanishing of the commutator of L_1 and L_2 is equivalent to the vanishing of the commutator of W_1 and W_2 except at $\lambda=\pm 1$. The substitution of σ for λ replaces $\lambda\in S^1$ by $\sigma\in\mathbb{R}$, which is consistent with our comments following (5′). □

We now follow the same reasoning as for the Euclidean case. Dimensional reduction follows from assuming invariance with respect to $\partial/\partial\eta'$ and $\partial/\partial\xi'$.

Theorem 1.7. *The self-dual Yang–Mills equations on $\Omega\times E^{1,1}\subseteq E^{1,1}\times E^{1,1}=E^{2,2}$ when dimensionally reduced to Ω become*

$$[\partial/\partial\eta+A_\eta+\sigma B_\eta, \partial/\partial\xi+A_\xi+\sigma^{-1}B_\xi]=0\ , \tag{14}$$

where $\sigma\in\mathbb{C}^*$, $A=A_\eta\,\mathrm{d}\eta+A_\xi\,\mathrm{d}\xi$ and $B=A_{\xi'}\,\mathrm{d}\eta+A_{\eta'}\,\mathrm{d}\xi$. If this equation is satisfied and Ω is simply connected then there is an essentially unique gauge transformation $u:\Omega\to G$ such that if $u^{-1}Au+u^{-1}\,\mathrm{d}u=\tilde{A}$ and $u^{-1}Bu=\tilde{B}$, then $\tilde{A}+\tilde{B}=0$ and $\tilde{A}=\frac{1}{2}s^{-1}\,\mathrm{d}s$ is the one-form associated to a harmonic map $s:\Omega\to G$. □

The proof is the same as for the Euclidean version.

2. The indefinite self-dual equations and Wess–Zumino terms

Some of the results in this chapter can be found in ref. [13]. We begin with a more global description of the reduced elliptic equations. Let M be any piece of surface, simply connected or not, and consider M as a complex one-manifold. Let P be a principal bundle with structure group G over M. The Lie algebra-valued one-forms $A=A_{\bar z}\,\mathrm{d}\bar z+A_z\,\mathrm{d}z$ of our previous chapter become invariantly described as the connection D on P. The unknowns $(A_{\bar w}, A_w)$ have already been identified with a Lie algebra-valued one-form $B=A_{\bar w}\,\mathrm{d}\bar z+A_w\,\mathrm{d}z$, so $B_{\bar z}=A_{\bar w}$, $B_z=A_w$. Globally B is to be a section of the bundle $\mathrm{Ad}\,P\otimes T^*_{\mathbb{C}}M$, where $\mathrm{Ad}\,P$ is the associated bundle $\mathrm{Ad}\,P=P\otimes_{\mathrm{Ad}}\mathfrak{g}$. In Hitchin's and Simpson's description of the related problem obtained from the self-dual equations on E^4, B plays the role of a Higgs field. Let $'$ and $''$ indicate the (1,0) and (0,1) parts of complex one-forms. The case we are considering has been treated in detail by Valli [14].

Definition. The indefinite self-dual equations on P consist of the equations

$$D'B''=D''B''=0\ ,\qquad [D',D'']+[B',B'']=0\ . \tag{15}$$

The self-dual equations studied by Hitchin are $[D',D'']-[B',B'']=0$. It is necessary to complexify the group to handle them.

Theorem 2.1. *Let (D, B) be a solution of the indefinite self-dual equations on P. In any simply connected region of M, P is trivial and there exists a trivialization $P\simeq M\times G$ in which*

$$D'+B'=\partial\ ,\qquad D''+B''=\bar\partial\ .$$

Furthermore

$$-B'=\tfrac{1}{2}s^{-1}\,\partial s\ ,\qquad -B''=\tfrac{1}{2}s^{-1}\,\bar\partial s\ ,$$

where $s: M \to G$ is a harmonic map.

Proof. This is an invariant formulation of the contents of theorem 1.4. □

The converse is clear.

Corollary 2.2. *If $s: M \to G$ is a harmonic map, then*

$$D = d + \tfrac{1}{2}s^{-1}\,\mathrm{d}s\,, \qquad B = -\tfrac{1}{2}s^{-1}\,\mathrm{d}s$$

gives a solution to the infinite self-dual equations on the trivial principal bundle $P = M \times G$. □

However, much more is true. The next theorem is an easy calculation, but the result is very striking. It is hard not to hope this may be helpful in quantization of harmonic maps.

First, a word of introduction. This is well known at least in the physics literature and is meant to be expository. Compact Lie groups with a biinvariant metric (and we are indeed assuming G has a biinvariant metric) all have a canonical closed three-form

$$\gamma = \tfrac{1}{3}([s^{-1}\,\mathrm{d}s, s^{-1}\,\mathrm{d}s] \cdot s^{-1}\,\mathrm{d}s)\,. \tag{16}$$

Fix a base map $s_0: M \to G$. Then we consider homotopic s, with a homotopy $s_t: M \to G$ for $t \in [0,1]$. (If M has a boundary, we will assume for now that $s_t | \partial M \equiv s_0 | \partial M$.) The variational problem for harmonic maps has the Lagrangian

$$E(s) = \frac{1}{2}\int_M (s^{-1}\,\mathrm{d}s, s^{-1}\,\mathrm{d}s)\,\mathrm{d}\mu\,.$$

The Lagrangian for the Wess–Zumino term is

$$\int_{M \times [0,1]} \gamma^* s = \frac{1}{3}\int_{M \times [0,1]} ([s^{-1}\,\tilde{\mathrm{d}}s, s^{-1}\,\tilde{\mathrm{d}}s], s^{-1}\,\tilde{\mathrm{d}}s)\,.$$

Be careful, as $\tilde{\mathrm{d}} = \mathrm{d} + (\partial/\partial t)\mathrm{d}t$ is a three-dimensional object in this integrand.

If $G = \mathrm{SU}(2) = \mathrm{S}^3$, $\int_{M \times [0,1]} \gamma^* s$ is precisely the volume in S^3 enclosed between the surfaces $s(M)$ and $s_0(M)$. The Wess–Zumino term is exactly that which arises in the consideration of the classical isoperimetric problems.

Ordinarily in geometry one prescribes $\int_{M \times [0,1]} \gamma^* s = Q$ as a constraint, but it could be considered as part of the variational problem.

Lemma 2.3. *Fix s_0. Then the variation of the Wess–Zumino term along $\delta s = s\varphi$ is just*

$$\int_M ([s^{-1}\,\mathrm{d}s, s^{-1}\,\mathrm{d}s]\cdot\varphi)\,,$$

which contributes $*[s^{-1}\,\mathrm{d}s, s^{-1}\,\mathrm{d}s]$ *to the Euler–Lagrange equation.*

Proof. Extend the variation along the homotopy $\delta s_t = s_t\varphi_t$, where $\varphi_0=0$, $\varphi_1=\varphi$. Naively the variation can be seen to be

$$\int_{M_3} ([s^{-1}\,\hat{\mathrm{d}}s, s^{-1}\,\hat{\mathrm{d}}s], \hat{\mathrm{d}}\varphi + [s^{-1}\,\hat{\mathrm{d}}s, \varphi])\,.$$

However, integration by parts leaves only the boundary term

$$\int_{M_3} \hat{\mathrm{d}}([s^{-1}\,\mathrm{d}s, s^{-1}\,\hat{\mathrm{d}}s], \varphi) = \int_{M\times 1} ([s^{-1}\,\mathrm{d}s, s^{-1}\,\mathrm{d}s], \varphi)\,.$$

This had to happen, as γ is closed. □

Theorem 2.4. *Let* (D, B) *be a solution of the indefinite self-dual equations on a simply connected surface* M. *Then for* $\lambda\in S^1$, *there exists a G-invariant bundle map* $P_\lambda: P\to M\times G$ *varying smoothly with* λ *such that*

$$P_\lambda\circ D' + \lambda^{-1}B'\circ P_\lambda^{-1} = \partial\,, \qquad P_\lambda\circ D'' + \lambda B''\circ P_\lambda^{-1} = \bar\partial\,.$$

Moreover, the G-invariant composition $P_\rho\circ P_\lambda^{-1}: M\times G\to M\times G$ *can be represented as a map* $W_{\rho,\lambda}: M\to G$. *Then* $W = W_{\rho,\lambda}$ *satisfies the harmonic equation with Wess–Zumino term*

$$\bar\partial(w^{-1}\,\partial w) + \partial(w^{-1}\,\bar\partial w) + \tau[w^{-1}\,\bar\partial w, w^{-1}\,\partial w] = 0$$

for $\tau = (\rho+\lambda)/(\rho-\lambda)$.

Proof. We can work in any gauge, so we choose the one we have already used frequently, where $P = M\times G$, $D = \mathrm{d}+A$ and $B-A=0$. The self-dual equations tell us that for $\lambda\in S^1$,

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

is a G-connection which is flat. Let E_λ be the trivialization, i.e., $E_\lambda^{-1}\,\mathrm{d}E_\lambda = D_\lambda$. There are a number of possible normalizations of E_λ. In the absence of other information we can pick a base point p and require $E_\lambda(p)\equiv I$.

Now in this coordinate description, $W_{\lambda,\rho} = w = E_\lambda E_\rho^{-1}$ (we drop the subscripts in the computation). Just compute to get

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

Here we use

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

In the next step we use

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

Then

$$\partial(w^{-1}\bar\partial w) = (\rho-\lambda)\partial(E_\rho A_{\bar z}E_{\bar\rho}^{-1}) = \rho^{-1}(\rho-\lambda)E_\rho[A_{\bar z}, A_z]E_{\bar\rho}^{-1}. \tag{17}$$

By reversing $(\partial, \rho, \lambda)$ and $(\bar\partial, \rho^{-1}, \lambda^{-1})$ we get

$$\bar\partial(w^{-1}\partial w) = \rho(\rho^{-1}-\lambda^{-1})E_\rho[A_z, A_{\bar z}]E_\rho = (\rho/\lambda)\partial(w^{-1}\bar\partial w). \tag{18}$$

Furthermore

$$\bar\partial(w^{-1}\partial w) - \partial(w^{-1}\partial w) = [w^{-1}\partial w, w^{-1}\bar\partial w]. \tag{19}$$

It is now straightforward to compute

$$\begin{aligned} 0 &= \bar\partial(w^{-1}\partial w) - (\rho/\lambda)\partial(w^{-1}\bar\partial w) \\ &= \tfrac{1}{2}(1-\rho/\lambda)[\bar\partial(w^{-1}\partial w) + \partial(w^{-1}\bar\partial w)] \\ &\quad + \tfrac{1}{2}(1+\rho/\lambda)[\bar\partial(w^{-1}\partial w) - \partial(w^{-1}\bar\partial w)] \\ &= \frac{1}{2}\frac{\lambda-\rho}{\lambda}\left(\bar\partial(w^{-1}\partial w) + \partial(w^{-1}\bar\partial w) + \frac{\lambda+\rho}{\rho-\lambda}[w^{-1}\bar\partial w, w^{-1}\partial w]\right). \end{aligned}$$

□

Corollary 2.5. *For real solutions $W: M \to G$, in complex coordinates, τ will be purely imaginary.*

Proof. Actually, the geometric term is (naturally) real. In the theorem above

$$\tau = \frac{\lambda+\rho}{\rho-\lambda} = \frac{\lambda/\rho+1}{(\lambda/\rho)^{-1}-1}$$

is purely imaginary as λ/ρ has modulus 1. These two facts are reconciled by the fact that $dz \wedge d\bar z$ (the Kähler form) is purely imaginary (skew). □

We conclude by some remarks on the Minkowski version. Here it makes less sense to think of M as a piece of a compact surface with a principal bundle P because we are dealing with wave equations. However, the substitution of $\bar z \to \eta$, $z \to \xi$, $\lambda \to \sigma$ and $\rho \to \nu$ yields precisely the same theory. Of course, if one starts with periodic solutions, the difficulty of determining which of these solutions are periodic is hard. However, energy conditions on planes $(\mathbb{R}, \tau) \subseteq E^{1,1}$ would be preserved. Here the Wess–Zumino τ is real. According to Witten, the desired parameter for conformal field theories is $\tau = \pm 1$. This corresponds to σ (or λ) $= (0, \infty)$, which is exactly the singularity from our point of view.

3. The null equations

In reducing the indefinite self-dual equations from four to two dimensions, we need to choose a plane of vector fields $(\alpha\, \partial/\partial p + \beta\, \partial/\partial q)$ which leave the problem invariant. This leaves a dual plane of one-forms representing dynamical variables. If this dual plane is definite, it inherits a definite metric and the discussion leading to theorem 1.4 is canonical up to Euclidean motions of this plane. If the dual plane is indefinite but non-degenerate, the same goes for the application of theorem 1.7, up to Lorentz transformations of this plane. Of course, fixing variables in $E^{2,2}$, rotating the planes and watching the equations vary as the inner product changes is a little more difficult.

The above discussion misses some cases: when the plane of dynamical variables has one degenerate null direction and when it is totally null. The first is clearly somewhat more relevant, since we would have to pass through this case as we rotate from Euclidean to Minkowski space. The second case is included for completeness.

We find the real formulation of the self-dual equations (13) more useful,

$$\left[\left(\frac{\partial}{\partial\eta} + A_\eta\right) + \sigma\left(\frac{\partial}{\partial\xi'} + A_{\xi'}\right), \left(\frac{\partial}{\partial\xi} + A_\xi\right) + \sigma^{-1}\left(\frac{\partial}{\partial\eta'} + A_{\eta'}\right)\right] = 0 .$$

Now we need to choose one null direction (say η) and one perpendicular definite direction (say x^3). This means the invariant plane of vector fields was $\alpha\, \partial/\partial\xi + \beta\, \partial/\partial x^4$. This leaves the equation

$$\left[\left(\frac{\partial}{\partial\eta} + A_\eta\right) + \sigma\left(\frac{1}{2}\frac{\partial}{\partial x^3} + A_{\xi'}\right), A_\xi + \sigma^{-1}\left(\frac{1}{2}\frac{\partial}{\partial x^3} + A_{\eta'}\right)\right]. \tag{20}$$

We shall lose nothing by fixing gauge with $A_\eta = -A_{\xi'}$, $A_\xi = -A_{\eta'}$ and replacing x^3 by $2x^3 = x$. So (20) becomes equivalent to the vanishing of the commutator of the two operators $\partial/\partial\eta + \partial/\partial x + (1-\sigma)A_\eta$ and $\partial/\partial x + (\sigma - 1)A_\xi$. Subtract σ times the second from the first to get the equivalent equation

$$[\partial/\partial\eta + (1-\sigma)A_\eta + \sigma(1-\sigma)A_\xi\,, \partial/\partial x + (\sigma-1)A_\xi] = 0 . \tag{21}$$

Let

$$E_\sigma^{-1} \frac{\partial}{\partial\eta} E_\sigma = (1-\sigma)(A_\eta + \sigma A_\xi)\,, \qquad E_\sigma^{-1} \frac{\partial}{\partial x} E_\sigma = (\sigma - 1)A_\xi \tag{22}$$

locally be the gauge which trivializes the connections for $\sigma \neq 1$.

Theorem 3.1. *The dimensionally reduced indefinite self-dual Yang–Mills equations on a simply connected domain $\Omega \subseteq \mathbb{R}^2$ are equivalent to any of the equations ($\sigma \neq 1$, $\sigma \in \mathbb{R}$)*

$$\frac{\partial}{\partial x}\left(E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma\right)+(1-\sigma)^{-1}\left[E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma, E_\sigma^{-1}\frac{\partial}{\partial \eta}E_\sigma^{-1}\right]=0 .$$

Proof. Proceed as above to get the definition of $E_\sigma:\Omega \to G$, $\sigma \neq 1$ real. From eq. (21), we have that the coefficient of σ^2 in the commutator is zero. This coefficient is

$$\begin{aligned} &-[\partial/\partial x - A_\xi, A_\xi] + [A_\xi, -A_\eta] \\ &= \frac{1}{1-\sigma}\left(\frac{\partial}{\partial x}\left(E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma\right)+\frac{1}{1-\sigma}\left[E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma, E_\sigma^{-1}\frac{\partial}{\partial \eta}E_\sigma\right]\right).\end{aligned}$$

This yields the equation as a consequence. Now suppose we have a solution of any of these equations. We define A_η and A_ξ using eq. (22),

$$A_\xi = (\sigma-1)^{-1}E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma ,$$

$$A_\eta = (1-\sigma)^{-1}\left(E_\sigma^{-1}\frac{\partial}{\partial \eta}E_\sigma + \sigma E_\sigma^{-1}\frac{\partial}{\partial x}E_\sigma\right).$$

The null self-dual equations (21) encode three equations: the coefficients of σ^0, σ^1, σ^2. The coefficient of σ^2 has been checked to be the expression in the theorem. Equation (22) has as consistency equation the vanishing of the equation at $\sigma \neq 1$. The coefficients have been chosen to vanish at $\sigma = 1$. These three independent conditions verify that (21) is identically satisfied. □

Corollary 3.2. *The null harmonic map equation can be regarded as the limit of either a Minkowski or a Euclidean harmonic map equation with inner product* (1, ϵ^2) *and Wess–Zumino term* $\epsilon^{-1}/(1-\sigma)$.

Proof. We first write the harmonic map equation in orthogonal laboratory coordinates,

$$\frac{\partial}{\partial x}\left(s^{-1}\frac{\partial s}{\partial x}\right) \pm \frac{\partial}{\partial t}\left(s^{-1}\frac{\partial s}{\partial t}\right) + t'\left[s^{-1}\frac{\partial s}{\partial x}, s^{-1}\frac{\partial s}{\partial t}\right]=0 ,$$

where t' is now real in both cases. Now let the signatures decay by allowing the metric on forms to decay in the t variable. In other words

$$\frac{\partial}{\partial x}\left(s^{-1}\frac{\partial s}{\partial x}\right) + \epsilon^2\frac{\partial}{\partial t}\left(s^{-1}\frac{\partial s}{\partial t}\right) + t\epsilon\left[s^{-1}\frac{\partial s}{\partial x}, s^{-1}\frac{\partial s}{\partial t}\right]=0 .$$

Set $t=1/\epsilon(1-\sigma)$ and let $\epsilon\to 0$ to get the null harmonic map equation for s at σ. □

It came as a surprise to me to discover that there are two cases of totally null planes possible. In eq. (13), we can either let η and ξ' be the dynamic variables, or η and η'. We take the second case first. Here (13) becomes

$$[\partial/\partial\eta+A_\eta+\sigma A_\xi,\ \partial/\partial\eta'+A_{\eta'}+\sigma A_{\xi'}]=0\ .$$

Proposition 3.3. *One case of the totally null self-dual reduction is equivalent to the harmonic map equation with only a Wess–Zumino term*

$$[s^{-1}\,\partial s/\partial\eta,\ s^{-1}\,\partial s/\partial\eta']=0\ .$$

Proof. Gauge fix with $A_\eta+A_\xi=0$, $A_{\eta'}+A_{\xi'}=0$. Let s trivialize the connection D_σ,

$$(\sigma-1)A_\eta=s^{-1}\,\partial s/\partial\eta\ ,\qquad (\sigma-1)A_{\eta'}=s^{-1}\,\partial s/\partial\eta\ ,$$

for any $\sigma\neq 1$. The coefficient of σ^2 is exactly the Wess–Zumino term. The converse is just as easy to check. □

Proposition 3.4. *The second case of a totally null reduction consists of a connection with a covariant constant section.*

Proof. We now have the commutator

$$[\partial/\partial\eta+A_\eta+\sigma\,\partial/\partial\xi'+A_{\xi'},\ A_\xi+\sigma^{-1}A_{\eta'}]=0\ .$$

Clearly the connection $(\partial/\partial\eta+A_\eta,\ \partial/\partial\xi'+A_{\xi'})$ has no condition on it, although in a standard way we can expect to be able to fix the gauge so that $A_\eta+\sigma A_{\xi'}=0$. The three independent conditions are

$$[\partial/\partial\eta+A_\eta,A_{\eta'}]=0\ ,\qquad [\partial/\partial\xi'+A_{\xi'},A_\xi]=0\ ,$$
$$[\partial/\partial\xi'+A_{\xi'},A_{\eta'}]+[\partial/\partial\eta+A_\eta,A_\xi]=0\ .$$

No further gauge fixing is a priori possible. However, if we derived this as the limit of one of the other cases, we might think the gauge choice $A_{\eta'}=\sigma A_\xi$ is a possible reasonable assumption (for some σ). Then we have an equation for a covariant constant section of our bundle $(P\times_{\mathrm{ad}}\mathfrak{g})\otimes T^*M$. The converse clearly does lead to a solution with no restriction. □

4. The group action for the elliptic and hyperbolic problems

The traditional theory of integrable systems is fundamentally tied up with time evolution. In studying harmonic maps into groups, we have at hand an elliptic version and a "Wick rotated" hyperbolic version. The hyperbolic type has very close connections with the sine-Gordon equation and may indeed be integrable.

In pursuing the connection between the two, we hope to shed some light on the meaning of "integrable" for elliptic, non-time evolutionary problems. This is certainly in the spirit of Noether's theorem, which connects symmetries of the integral with divergence-free vector fields on the domain and is blind to the signature. However, the symmetries we construct for harmonic maps are non-local and do not fit into the framework of Noether's theorem. They are, in fact, symmetries constructed by factorization of a loop group.

We review the algebraic structure of the "loop group" action on the solution space of harmonic maps from simply connected $\Omega\subset S^2$ into $G=\mathrm{U}(N)$, the unitary group or $G=\mathrm{SU}(N)\subset\mathrm{U}(N)$ as given in our earlier paper [9]. Then we contrast these formulas with the Wick-rotated formulas. Although the formulas are the same, differences appear in the global structure of the solution space.

We already started this project of comparison in the first section, with the description of the two Lax pairs in characteristic coordinates $(z,\bar z)\in\mathbb{C}$ and $(\xi,\eta)\in\mathrm{E}^{1,1}$. We restate our fundamental theorem, which is theorem 2.4 and its converse of ref. [9]. For the remainder of this section $G=\mathrm{U}(N)$, $G_{\mathbb{C}}=\mathrm{GL}(N,\mathbb{C})$.

Theorem 4.1. *The map $s:\Omega\to G$ is harmonic if and only if there is an extended harmonic map*

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

such that the expressions

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

are constant (in λ) and satisfy in addition

(a) $E_1=1$,

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

(c) $E^*_{\tau(\lambda)}=E_\lambda^{-1}$ *for* $\tau(\lambda)=(\bar\lambda)^{-1}$.

The proof is not difficult [9]. The constant expressions will turn out to be $A_{\bar z}$ and A_z and E_λ trivializes the flat connection $(\bar\partial+(1-\lambda)A_{\bar z},\partial+(1-\lambda^{-1})A_z)$. The involution $\tau(\lambda)=\bar\lambda^{-1}$ has $|\lambda|=1$ as its fixed point set and (c) is equivalent to $E_\lambda\in G$ being in the real group for $|\lambda|=1$.

The Minkowski or hyperbolic version is only slightly different.

Theorem 4.1′. *The map $s:\Omega\to G$ for simply connected $\Omega\subseteq\mathrm{E}^{1,1}$ is harmonic if and only if there exists an extended harmonic map*

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

with the expressions

$$(1-\sigma)^{-1}E_\sigma^{-1}\,\partial E_\sigma/\partial\eta\,,\qquad (1-\sigma^{-1})^{-1}E_\sigma^{-1}\,\partial E_\sigma/\partial\xi$$

independent of $\sigma\in\mathbb{C}^*$. *In addition we require*

(a) $E_1=I$,
(b) $E_{-1}=s$,
(c) $E^*_{\tau(\sigma)}=E_\sigma^{-1}$ *for* $\tau(\sigma)=\bar\sigma$.

The proof is no different from the proof for the elliptic case. Here E_σ trivializes the flat connection

$$(\partial/\partial\eta+(1-\sigma)A_\eta,\,\partial/\partial\xi+(1-\sigma^{-1})A_\xi)\;.$$

The involution $\sigma\to\bar\sigma$ has the real line $\mathbb{R}$ as fixed point, which is the set $\sigma\in\mathbb{C}^*$ where E_σ is unitary.

The fundamental algebraic structure we will be discussing later is the action of a loop type Lie algebra (or half of it, if you like). However, in discussing the group actions we find it simpler to work with a smaller group. The group we work with is

$$A_{\mathbb{R}}=A_{\mathbb{R}}(S^2,\mathrm{GL}(N,\mathbb{C}))\subseteq\text{Meromorphic maps}(S^2,\mathrm{GL}(N,\mathbb{C}))\;.$$

Here $f\in A_{\mathbb{R}}$ is meromorphic from S^2 to $\mathrm{GL}(N,\mathbb{C})$, with no poles or zeros at $(0,\infty)$. In addition we require $f(\tau(\lambda))^*=f(\lambda)^{-1}$, our standard reality condition, and a convenient normalization $f(1)=I$. We also defined $X^\infty_{\mathbb{R}}$ as the object in which $E(z,\bar z)$ sits. It consists of the holomorphic maps from $\mathbb{C}^*\to\mathrm{GL}(N,\mathbb{C})$ satisfying our by now usual reality condition. So $f\in A_{\mathbb{R}}$ has poles and zeros away from $(0,\infty)$ and $e\in X^\infty_{\mathbb{R}}$ is holomorphic everywhere *except* at $(0,\infty)$. We define the map $f^\#:X^\infty_{\mathbb{R}}\to X^\infty_{\mathbb{R}}$ by the Birkhoff factorization

$$f(\lambda)e(\lambda)=f^\#(e)(\lambda)r^{-1}(\lambda)\;,\quad f,r\in A_{\mathbb{R}},\,e,f^\#(e)\in X^\infty_{\mathbb{R}}\;.\tag{23}$$

Theorem 4.2 (theorem 6.3 of ref. [9]). *The map*

$$f^\#:E_\cdot(z,\bar z)\to f^\#E_\cdot(z,\bar z)$$

for fixed $(z,\bar z)\in\mathbb{C}$ *determines a representation of* $A_{\mathbb{R}}$ *on the space of extended harmonic maps.*

Now the Minkowski version of theorem 4.2 is identical.

Corollary 4.3. *Theorem 4.2 is valid for Minkowski as well as Euclidean domains, if we replace the Euclidean involution* $\tau(\lambda)=\bar\lambda^{-1}$ *by its Minkowski counterpart* $\tau(\sigma)=\bar\sigma$.

The proofs of the theorem and its hyperbolic counterpart are identical: One factors elements of $A_{\mathbb{R}}$ [or $A'_{\mathbb{R}}$ with $\tau(\sigma)=\bar{\sigma}$] into "simplest factors". For the Euclidean case, these are

$$f(\lambda)=\pi+\xi_\alpha(\lambda)\pi^{\perp}\,,\quad \xi_\alpha(\lambda)=\frac{\lambda-\alpha}{\bar{\alpha}\lambda-1}\,\frac{\bar{\alpha}-1}{\alpha-1}\,, \tag{24}$$

π is the Hermitian projection on a subspace of $\mathbb{C}^N$. We showed in two ways that these could always be made to act, first by direct construction in the factorization, and then by showing that the same solutions could be arrived at by a pair of consistent ordinary differential equations (in $(z, \bar{z})$ or equivalently (x, y)).

For the Minkowski case, the simplest factors are different due to the reality condition. We have

$$f(\sigma)=\pi+\mu_\alpha(\sigma)\pi^{\perp}\,,\quad \mu_\alpha(\sigma)=\left(\frac{\sigma-\alpha}{\sigma-\bar{\alpha}}\right)\left(\frac{1-\bar{\alpha}}{1-\alpha}\right). \tag{25}$$

At first glance, there is very little difference. However, we pass now to the global theory.

Theorem 4.4. *Harmonic maps $s\colon S^2\to G=U(N)$ have finite uniton number n, which is the smallest degree of $E_\lambda(z, \bar{z})$ represented as a polynomial in λ. The action of $A_{\mathbb{R}}$ preserves this uniton number. Changes in uniton number can be achieved by looking at singular Bäcklund transformations corresponding to $\alpha\to(0, \infty)$. These singular transformations consist of a consistent Cauchy–Riemann equation and a linear constraint.* □

Theorem 14.6 of ref. [9] shows that the entire solution space can be constructed canonically by applying singular Bäcklund transformations. The group $A_{\mathbb{R}}$ does not act transitively. In the very simplest case its effective action is like $SL(2, \mathbb{C})$ acting on a Grassmannian submanifold $\mathbb{C}P^1\subseteq SU(2)$ containing the image $s(S^2)$. For $s\colon S^2\to\mathbb{C}P^1\subseteq SU(2)$ holomorphic of degree k, $SL(2, \mathbb{C})$ is transitive only for $k=1$.

Without going into the details of the action in the hyperbolic (or Minkowski) case, its becomes immediately clear that the *reality condition* forbids any finite polynomial solutions at all in $X^n_{\mathbb{R}}$ for $n<\infty$. The equation $f(\lambda)f(\bar{\lambda})^*=I$ has no meromorphic solutions with all the zeros and poles at $(0, \infty)$. A similar algebraic condition shows us that there are no singular Bäcklund transformations. The limits of Bäcklund transformations as $\alpha\to(0, \infty)$ seem to approach a rather boring identity.

On the other hand, in the next section we will obtain concrete information about the hyperbolic action by examining the relationship of harmonic maps into $\mathbb{C}P^1\subseteq SU(2)$ with the classical sine-Gordon equation. Here it is known that repeated applications of the classical Bäcklund transformation, which is identical in this case to our Bäcklund transformation, obtain large portions of the solution space from a trivial solution. The algebraic connection of these two global constructions as Wick rotations of each other is not at all clear.

5. Sine-Gordon equation

It is a well-known theorem that the sine-Gordon equation is equivalent to the theory of harmonic maps into $S^2=\mathbb{C}P^1$. This can be found in Pohlmeyer [15]. This section derives this relationship using the "gauge-free" discussion of harmonic maps which comes from section 1.

We first need a lemma about harmonic maps into Grassmannians, which will identify those harmonic maps into $U(N)$ which lie in some $G(r, N)$. Recall that for the purposes of our theory

$$\bigcup_r G(r,N)=\{s\in U(N): s^2=I\}\,.$$

Lemma 5.1. *Suppose* $s: E^{1,1}\to G(r,N)$. *Then*

$$(\partial/\partial\eta)s+[A_\eta, s]=0\,, \tag{a}$$

$$sA_\eta+A_\eta s=\{s, A_\eta\}=0\,, \tag{b}$$

$$(\partial/\partial\xi)s+[A_\xi, s]=0\,, \tag{c}$$

$$sA_\xi+A_\xi s=\{s, A_\xi\}=0\,. \tag{d}$$

Proof. We need only prove (a) and (b), since the two light cones behave exactly alike. We just compute, using $s=s^{-1}$,

$$[A_\eta, s]=\frac{1}{2}\left(s^{-1}\frac{\partial s}{\partial\eta}s-ss^{-1}\frac{\partial s}{\partial\eta}\right)=\frac{1}{2}\left(-\frac{\partial s^{-1}}{\partial\eta}s^2-\frac{\partial s}{\partial\eta}\right)=\frac{\partial s}{\partial\eta}\,.$$

Also

$$sA_\eta+A_\eta s=ss^{-1}\frac{\partial s}{\partial\eta}+s^{-1}\frac{\partial s}{\partial\eta}s=\frac{\partial s}{\partial\eta}-\frac{\partial s^{-1}}{\partial\eta}s^2=0\,. \qquad\square$$

We now introduce $\varphi=is: E^{1,1}\to\mathfrak{g}\cap G$ as an element lying in both $U(N)$ and $u(N)$. Regard it as a map into the Lie algebra. As gauge transformations carry the connection around, φ becomes a special extra element. Recall from the gauge-invariant formalism:

$$d+A\to d+u^{-1}Au+u^{-1}\,du\,,\qquad A\to B=u^{-1}Au\,,\qquad \varphi\to u^{-1}\varphi u\,. \tag{26}$$

Lemma 5.2. *In the gauge-invariant formulation of harmonic maps into the Grassmannian $G(r, N)$, there exists an extra map $\varphi: \mathrm{E}^{1,1} \to \mathfrak{g} \cap G$ with r eigenvalues* i, *$N-r$ eigenvalues* $-\mathrm{i}$ *satisfying*

(a) $\mathrm{d}\varphi + [A, \varphi] = 0$,

(b) $\{\varphi, A\} = 0$.

Proof. In the usual gauge, φ satisfies these equations because s does. These equations are gauge invariant. □

There certainly is a theory for more general Lagrangians. But for $\mathbb{C}\mathrm{P}^1 = \mathrm{S}^2 \subseteq \mathrm{U}(2)$, φ, A_ξ, A_η all map into su(2), which is only three-dimensional. Recall also the invariant Lax pair, which we rewrite here for convenience:

$$[\partial/\partial\eta + A_\eta + \lambda B_\eta, \partial/\partial\xi + A_\xi + \lambda^{-1} B_\xi] = 0 \,. \tag{27}$$

Lemma 5.3. $(\partial/\partial\eta)\,|B_\xi| = 0$ *and* $(\partial/\partial\xi)\,|A_\eta| = 0$.

Proof. As a consequence of the harmonic map equation, in the invariant formalism

$$\frac{\partial}{\partial\eta} B_\xi + [A_\eta, B_\xi] = 0 \,.$$

By the Leibnitz rule

$$\frac{\partial}{\partial\eta} B_\xi^2 + [A_\eta, B_\xi^2] = 0 \,.$$

Taking the trace results in

$$\frac{\partial}{\partial\eta} \operatorname{tr}(B_\xi^2) = -\frac{\partial}{\partial\eta} |B_\xi|^2 = 0 \,.$$

This was, of course, a gauge invariant statement. □

Lemma 5.4. *Suppose* $|B_\xi|(\xi, \eta) = b(\xi)$ *is non-zero for all* ξ. *Then there exists a smooth gauge change so that*

$$\varphi = \mathrm{i}s = 2\boldsymbol{j} \,, \qquad B_\xi = \sqrt{2}\, b(\xi)\boldsymbol{i},$$

where $\boldsymbol{i}, \boldsymbol{j}$ *and* $\boldsymbol{k}$ *form the usual quaternion basis for* SU(2),

$$\boldsymbol{i} = \frac{1}{2}\begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, \qquad \boldsymbol{j} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \boldsymbol{k} = \frac{1}{2}\begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}.$$

Moreover, this gauge transformation is unique in SU(2) *up to*

$$\pm I = \pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Proof. The fact that $\frac{1}{2}\varphi$ and $[1/\sqrt{2}\, b(\xi)]B_\xi$ serve as standard generators for Q in su(2) follows from the anti-commutation relations, and it is easy to show that the Q they generate can be conjugated uniquely (up to $\pm I$) to the standard one. □

Proposition 5.5. *Let* $s: \mathrm{E}^{1,1} \to \mathbb{C}\mathrm{P}^1 \subset \mathrm{SU}(2)$ *be harmonic. Suppose* $|A_\eta(\xi, \eta)| = a(\eta) \neq 0$, $|A_\xi(\xi, \eta)| = b(\xi) \neq 0$. *Then there is a gauge transformation, which is unique up to multiplication by* $\pm I$, *in which*

$$A_\eta = 0\,, \qquad B_\eta = \sqrt{2}\, a(\eta) \exp(-u\boldsymbol{j})\, \boldsymbol{i} \exp(u\boldsymbol{j})\,,$$

$$A_\xi = \frac{\partial u}{\partial \xi}\boldsymbol{j}\,, \qquad B_\xi = \sqrt{2}\, b(\xi)\boldsymbol{i}\,,$$

$$\varphi = 2\boldsymbol{j}\,.$$

Proof. First we note that we can choose the gauge of lemma 5.4 as the gauge of this theorem. Before we proceed further, note that there is a *similar* but different choice of gauge in which

$$\hat{\varphi} = 2\boldsymbol{j}\,, \qquad \hat{B}_\eta = \sqrt{2}\, a(\eta)\boldsymbol{i}\,,$$

as there is no difference between the two light cones. Let $U: \mathrm{E}^{1,1} \to \mathrm{SU}(2)$ be the change of gauge between the two, i.e.,

$$\varphi = 2\boldsymbol{j} = U^{-1}\, 2\boldsymbol{j}\, U = U^{-1}\hat{\varphi} U\,,$$

$$B_\eta = U^{-1}\sqrt{2}\, a(\eta)\boldsymbol{i}U = U^{-1}\hat{B}_\eta U\,.$$

It is immediately clear that U commutes with $\boldsymbol{j}$, and hence $U = \exp u\boldsymbol{j}$ for $u: \mathrm{E}^{1,1} \to \mathbb{R}$, since we are using a simply connected domain.

It remains only to determine A_ξ and A_η. Looking back to lemma 5.2, we see

$$\mathrm{d}\varphi + [A, \varphi] = \mathrm{d}\boldsymbol{j} + [A, \boldsymbol{j}] = [A, \boldsymbol{j}] = 0\,.$$

So immediately we know $A = a\boldsymbol{j}$ for some real one-form a. However, we see further from the validity of the Lax pair that

$$\frac{\partial}{\partial \eta} B_\xi + a_\eta[\boldsymbol{j}, B_\xi] = \sqrt{2}\, b(\xi)\frac{\partial}{\partial \eta}\boldsymbol{i} + \mathrm{d}\eta[\boldsymbol{j}, \mathrm{i}] = 0\,.$$

Hence $a_\eta = 0$. In the same vein

$$\frac{\partial}{\partial \xi} B_\eta = \sqrt{2}\, \frac{\partial}{\partial \eta} [\exp(-uj)\, \boldsymbol{i} \exp uj] = -\frac{\partial u}{\partial \xi} [j, B_\eta] = -a_\xi [j, B_\eta] \,.$$

It follows that $A_\xi = a_{\xi j} = (\partial u / \partial \xi) j$. □

Theorem 5.6. *Let* $s: \mathrm{E}^{1,1} \to \mathbb{C}\mathrm{P}^1 \subseteq \mathrm{SU}(2)$ *be a harmonic map with*

$$|A_\xi| = \tfrac{1}{2} |s^{-1}\, \partial s / \partial \xi| = a(\xi) \neq 0 \,, \qquad |B_\eta| = \tfrac{1}{2} |s^{-1}\, \partial s / \partial \eta| = b(\eta) \neq 0 \,.$$

Then there exists a choice of gauge, which is unique up to multiplication by $\pm I$, *in which the Lax pair for this harmonic map is of the form*

$$0 = \left[\frac{\partial}{\partial \eta} + \lambda a(\eta) \exp(-uj)\, \boldsymbol{i} \exp(uj) \,, \frac{\partial}{\partial \xi} + \frac{\partial u}{\partial \xi} j + \lambda^{-1} \sqrt{2}\, b(\xi)\, \boldsymbol{i} \right].$$

The function u satisfies the equation

$$\frac{\partial}{\partial \eta} \frac{\partial}{\partial \xi} u + a(\eta) b(\xi)\, 4 \sin 2u = 0 \,.$$

Proof. This, in fact, is just the culmination of all our work except for the derivation of the equation for u. This comes from the λ^0 term of the commutator (Lax pair)

$$[\partial/\partial \eta + A_\eta, \partial/\partial \xi + A_\xi] + [B_\eta, B_\xi] = 0 \,.$$

This becomes, for us

$$0 = \frac{\partial}{\partial \eta} \left(\frac{\partial u}{\partial \xi} j \right) + 2a(\eta) b(\xi) [\exp(-uj)\, \boldsymbol{i} \exp uj, \boldsymbol{i}] \,.$$

Since $\boldsymbol{i}$ and j anticommute

$$\exp(uj)\, \boldsymbol{i} = \boldsymbol{i} \exp(-uj) \,.$$

So we see that

$$\begin{aligned} [\exp(-uj)\, \boldsymbol{i} \exp uj, \boldsymbol{i}] &= \exp(-2uj)\ (\boldsymbol{i})^2 - (\boldsymbol{i})^2 \exp(2uj) \\ &= \exp(2uj) + \exp(-2uj) = 2 \sin 2uj \,. \end{aligned} \tag{28}$$

This completes the derivation. □

For those who still have not recognized the equation, we note the following corollary.

Corollary 5.7. *If* $a(\xi) \neq 0$, $b(\xi) \neq 0$, *there exists a conformal change of coordinates which locally makes* $\sqrt{2}\, a(\xi) = 1$, $\sqrt{2}\, b(\xi) = 1$. *In this coordinate system*

$$\frac{\partial}{\partial\eta}\frac{\partial}{\partial\xi}u+2\sin 2u=0\,. \qquad \square$$

Here is an interesting fact.

Corollary 5.8. *Suppose we are given the Lax pair for the sine-Gordon equation as in theorem 5.6 with* $a(\eta)\sqrt{2}=b(\xi)\sqrt{2}=1$. *Suppose, also, that* M_λ *is a trivialization for the flat connection*

$$\begin{aligned} M_\lambda^{-1}\frac{\partial}{\partial\eta}M_\lambda &= +\lambda\exp(-uj)\,i\exp(uj)\,, \\ M_\lambda^{-1}\frac{\partial}{\partial\xi}M_\lambda &= +\frac{\partial u}{\partial\xi}j+\lambda^{-1}i\,. \end{aligned} \tag{29}$$

Then the original harmonic map is of two forms

$$s=BM_{-1}M_1^{-1}A^{-1}\,, \qquad s=iAM_1jM_1^{-1}A^{-1}$$

for some constant matrices A and B.

Proof. The twist AM_1 is just the gauge change from our original harmonic map formulation to the sine-Gordon form. The unknown B appears because Bs trivializes the flat connection at $\lambda=1$ just as well as s. Our recipe for trivializing was good only up to constants, of course. $\square$

6. A loop group action on solutions to the sine-Gordon equation

We reviewed in section 4 how a loop type group acts on solutions of the extended harmonic map equation. There is, of course, an ambiguity due to choice of normalization of M_λ or E_λ. Modulo this difficulty, the group acts on harmonic maps. Moreover, the subgroup satisfying $f(\lambda)=f(-\lambda)$ acts on harmonic maps into Grassmannians (modulo the same difficulty). However, we wish to derive a formulation of this action which is compatible with the more traditional approach to the sine-Gordon equation.

Our chosen normalization, which we refer to as the sine-Gordon gauge, is regular in some sense at $\lambda=0$, where the singularity in the Lax pair is the coefficient of λ^{-1}, the constant $B_\xi=i$.

In the harmonic map gauge, where we have already constructed the loop group action, our factorization had been selected to normalize E_λ at $\lambda=0$ so $E_1=I$, and we fixed the choice of E_λ by fixing $E_\lambda(p)$ for some $p\in \mathrm{E}^{1,1}$. In the sine-Gordon gauge, customarily we fix the singularity at $\lambda=0$ to agree with the essential singularity of the vacuum solution given by $u\equiv 1$. The eigenfunctions which appear in M_λ would usually be normalized to agree with the vacuum eigenfunctions as $x \to -\infty$ (or $x \to +\infty$) if the growth conditions of the solution admit this possibility. Note for precision the following fact.

Definition 6.1. *The vacuum solution* $u \equiv 0$ *to sine-Gordon is the solution with the Lax pair*

$$A_\xi = A_\eta = 0 , \qquad B_\xi = B_\eta = \boldsymbol{i} . \qquad \square$$

It is an easy calculation to see that for the vacuum solution

$$M_\lambda(\xi, \eta) = \exp(\lambda\eta + \lambda^{-1}\xi)\, \boldsymbol{i} . \tag{30}$$

In our choices of normalization, it is essential to note that multiplication of M_λ by elements of the group or by meromorphic maps into the center of U(2) has no effect on the computations. We are in effect interested in the group PSU(2) = SO(3) and its complexification PSL(2, $\mathbb{C}$). However, since we have been using the matrix groups U(2) and GL(2, $\mathbb{C}$), we shall continue to do so.

We employ the canonical normalization $M_\lambda(p) = I$ for $p = (0, 0)$ rather than that involving limits as $x \to \pm\infty$. This allows us to avoid energy decay questions. However, our computations are therefore not compatible with the symplectic structure or the standard computations. They can, however, be adjusted to agree with these.

Lemma 6.2. *Let* M_λ *be the trivialization of the Lax pair for the solution of the sine-Gordon equation with* $M_\lambda(p) = I$. *Then*

$$M_\lambda \boldsymbol{j} = \boldsymbol{j} M_{-\lambda} .$$

Proof. If $M_\lambda(0) = I$, then the solutions of

$$M_\lambda^{-1} \frac{\partial M}{\partial \eta} = \lambda \exp(-u\boldsymbol{j})\, \boldsymbol{i} \exp(u\boldsymbol{j}) ,$$

$$M_\lambda^{-1} \frac{\partial M}{\partial \xi} = + \frac{\partial u}{\partial \xi} \boldsymbol{j} + \lambda^{-1} \boldsymbol{i}$$

are unique. Let

$$D_\lambda = (D + A + \lambda^{-1} B_\xi \, d\xi + \lambda B_\eta \, d\eta) ,$$

where

$$A = \frac{\partial u}{\partial \xi} \boldsymbol{j} \, d\xi , \qquad B = \boldsymbol{i} \, d\xi + \exp(-u\boldsymbol{j})\, \boldsymbol{i} \exp(u\boldsymbol{j}) \, d\eta .$$

Note that $\{i, j\} = 0$, so

$$(D_\lambda)j = j(D_{-\lambda}) .$$

It follows that $-jM_{-\lambda}j$ and M_λ both trivialize D_λ and agree at $(\xi, \eta) = (0, 0)$ and hence agree everywhere. □

We now define the group which is to act on trivializations of the Lax pair for the sine-Gordon equation [recall that elements which map to the center will be ineffective, which is why we can allow $r(\lambda) \neq 1$ in (d)]:

$$\mathscr{A}_{\mathrm{sG}} = \{h: \mathbb{C}\mathrm{P}^1 \to \mathrm{U}(2) \text{ meromorphic and satisfying conditions (a)–(d)}\} , \quad (31)$$

(a) $h(0) = I$,
(b) $\lim_{\lambda\to\infty} \lambda^{-m} h(\lambda) = a \in \mathrm{U}(2)$,
(c) $h(\lambda)j = jh(-\lambda)$,
(d) $h(\lambda)h(\bar{\lambda})^* = r(\lambda)I$ for r a meromorphic function on $\mathbb{C}\mathrm{P}^1$.

It is not true that elements can be factored into products of Bäcklund transformations in $\mathscr{A}_{\mathrm{sG}}$. So we will take the route of deriving the existence of the group action from the theorem we have already obtained for the action of $\mathscr{A}_{\mathbb{R}}$ on extended harmonic maps in the harmonic map gauge.

Lemma 6.3. *Suppose $h \in \mathscr{A}_{\mathrm{sG}}$. Then h can be written as*

$$h(\lambda) = q(\lambda) A f(\lambda) ,$$

where q is meromorphic in λ with $q(0) = 1$, $f \in \mathscr{A}_{\mathbb{R}}$ and A is constant.

Proof. We claim that

$$r(\lambda) = a \frac{(\lambda-\alpha_1)(\lambda-\bar{\alpha}_1)\cdots(\lambda-\alpha_m)(\lambda-\bar{\alpha}_m)}{(\lambda-\beta_1)(\lambda-\bar{\beta}_1)\cdots(\lambda-\beta_n)(\lambda-\bar{\beta}_n)} .$$

Namely we claim that the zeros and poles of r occur in complex conjugate pairs. This is clear for those which are not real.

For those that are real, it is sufficient to prove that r cannot have a simple zero at α and by dividing h by $(\lambda-\alpha)^p$ to cancel the even orders. So we want to assume that r has a simple pole at $\alpha \in \mathbb{R}$ and obtain a contradiction. If r has a simple zero at α real, since

$$\det h(\lambda) \det \bar{h}(\bar{\lambda}) = r(\lambda)^2 ,$$

both h and $h(\bar{\lambda})^*$ have simple poles at α with one-dimensional eigenspace. Let $h(\alpha)v = 0$. Now let w be orthogonal to v. Since $h(\alpha) \neq 0$, $h(\alpha)w \neq 0$. Also

$$(h(\alpha)w)^*(h(\alpha)w) = |h(\alpha)w|^2 \neq 0 .$$

But

$$
\begin{aligned}
(h(\alpha)w)^*h(\alpha)w &= w^*h(\alpha)^*h(\alpha)w \\
&= w^*h(\bar{\alpha})^*h(\alpha)w = r(\alpha)w^*w = 0 .
\end{aligned}
$$

This gives us our contradiction.

Now, define

$$
\tilde{h}(\lambda) = \frac{(\lambda-\beta_1)\cdots(\lambda-\beta_n)}{(\lambda-\alpha_1)\cdots(\lambda-\alpha_m)}\, h(\lambda) ,
$$

$$
f(\lambda) = \tilde{h}^{-1}(1)\tilde{h}(\lambda) ,
$$

$$
q(\lambda) = (-1)^{n+m} \frac{(\lambda-\alpha_1)\cdots(\lambda-\alpha_m)(\beta_1)\cdots(\beta_n)}{(\lambda-\beta_1)\cdots(\lambda-\beta_n)(\alpha_1)\cdots(\alpha_m)} ,
$$

$$
A = \tilde{h}(-1)^{n+m} \frac{(\alpha_1)\cdots(\alpha_m)}{(\beta_1)\cdots(\beta_n)} .
$$

With these definitions

$$
h(\lambda) = q(\lambda)Af(\lambda) . \qquad \square
$$

To give some substance to the computations, we introduce the two-parameter family of generators for Bäcklund transformations,

$$
h_{\gamma,\beta}(\lambda) = I + \lambda\beta \exp(-\beta \boldsymbol{j})\, \boldsymbol{i} \exp(\beta \boldsymbol{j}) . \tag{32}
$$

Lemma 6.4. *For γ and β real, $\beta \neq 0$, $h_{\gamma,\beta} \in \mathscr{A}_{\mathrm{sG}}$.*

Proof. Certainly $h_{\gamma,\beta}(0) = I$. Likewise $\lim_{\lambda\to\infty} \lambda^{-1} h_{\gamma,\beta}(\lambda) = \gamma \exp(-\beta \boldsymbol{j})\, \boldsymbol{i} \exp(\beta \boldsymbol{j})$. Further, since $\{\boldsymbol{i}, \boldsymbol{j}\} = 0$, we have $h_{\gamma,\beta}(\lambda)\boldsymbol{j} = \boldsymbol{j} h_{\gamma,\beta}(-\lambda)$. Moreover, (d) is verified by computation,

$$
\begin{aligned}
h_{\gamma,\beta}(\lambda)(h_{\gamma,\beta}(\bar{\lambda}))^* &= [I + \lambda\gamma \exp(-\beta \boldsymbol{j})\, \boldsymbol{i} \exp(\beta \boldsymbol{j})][I - \lambda\gamma \exp(-\beta \boldsymbol{j})\, \boldsymbol{i} \exp \beta \boldsymbol{j}] \\
&= (1 + (\lambda\gamma)^2) I . \qquad \square
\end{aligned}
$$

To write in terms of the Bäcklund transformations in the harmonic map gauge, let $\phi_\beta = \exp(-\beta \boldsymbol{j})\, \boldsymbol{i} \exp(\beta \boldsymbol{j})$, so $\phi_\beta^2 = -I$. Then $\pi_\beta = \frac{1}{2}(I + \mathrm{i}\phi_\beta)$ is a Hermitian projection onto a one-dimensional subspace which rotates through a circle as β moves from 0 to π,

$$
\begin{aligned}
I+\lambda\gamma\phi_\beta &= I+\mathrm{i}\lambda\gamma(I-2\pi_\beta) \\
&= (1-\mathrm{i}\lambda\gamma)\pi_\beta+(1+\mathrm{i}\lambda\gamma)\pi_\beta^\perp \\
&= (1-\mathrm{i}\gamma\lambda)\left(\pi_\beta-\frac{\lambda-\mathrm{i}\gamma^{-1}}{\lambda+\mathrm{i}\gamma^{-1}}\pi_\beta^\perp\right) \\
&= q(\lambda)A(\pi_\beta+\mu_{\alpha^{-1}}(\lambda)\pi_\beta^\perp)\,.
\end{aligned}
$$

Here $\alpha=\mathrm{i}\gamma^{-1}$, $\mu_\alpha(\lambda)$ is defined in (25), $q(\lambda)=(1-\mathrm{i}\gamma\lambda)$ and

$$A=\pi_\beta-\frac{\gamma+\mathrm{i}}{\gamma-\mathrm{i}}\pi_\beta^\perp\,.$$

This illustrates the relationship between the Bäcklund transformations in the sine-Gordon gauge and those in the harmonic map gauge. It is not at all clear in the harmonic map gauge that these Bäcklund transformations preserve the Grassmann condition, so the sine-Gordon gauge is an improvement from this point of view.

Proposition 6.5. *Let $h\in\mathscr{A}_{\mathrm{sG}}$. Then the Birkhoff factorization*

$$h(\lambda)M_\lambda=(h^\# M)_\lambda T_\lambda$$

can always be carried out with $T_0=I$ to take extended canonical trivializations to canonical trivializations. Moreover, if M_λ satisfies $M_\lambda(j)=(j)M_{-\lambda}$, so does $(h^\# M)_\lambda$.

Proof. By lemma 6.3 we have $h(\lambda)=g(\lambda)Af(\lambda)$ for $f\in\mathscr{A}_{\mathbb{R}}$. Also, we have $E_\lambda=M_\lambda M_1^{-1}$, since E_λ is the extended solution in the harmonic map gauge. Moreover, by theorem 6.3 of ref. [9], we have

$$f(\lambda)E_\lambda=(f^\# E)_\lambda R_\lambda\,,$$

where $(f^\# E)_\lambda$ is smoothly defined in f and E. Now let

$$
\begin{aligned}
h(\lambda)M_\lambda &= q(\lambda)Af(\lambda)E_\lambda M_1 = q(\lambda)A(f^\# E)_\lambda R_\lambda M_1 \\
&= (h^\# M)_\lambda T_\lambda\,.
\end{aligned}
$$

Here

$$(h^\# M)_\lambda=A(f^\# E)_\lambda R_0 M_1\,, \tag{33}$$

$$T_\lambda=q(\lambda)(R_0M_1)^{-1}R_\lambda M_1\,. \tag{34}$$

The decomposition $h(\lambda)=q(\lambda)Af(\lambda)$ has an ambiguity, a meromorphic function which acts trivially on E_λ and is absorbed into the product $q(\lambda)R_\lambda$, so the action is well defined. Since $(h^\# M)_\lambda$ is obtained from an extended harmonic map by a gauge change R_0M_1 on the left and a constant A on the right, it represents the trivialization of a harmonic map in a gauge-invariant setting.

By construction, $T_0=q(0)(R_0M_1)^{-1}R_0M_1=I$. To check whether the canonical property $M_\lambda(p)=I$ is preserved, note that $E_\lambda(p)=I$ and $M_1(p)=I$. So $(f^{\#}E)_\lambda(p)=I$,

$$(h^{\#}M)_\lambda(p)=A(f^{\#}E)_\lambda(p)R_0(p)M_1(p)=AR_0(p)\ .$$

However, $R_\lambda(p)=f(\lambda)$, so $R_0(p)=f(0)$. But

$$I=h(0)=q(0)Af(0)=Af(0)=AR_0(p)=(h^{\#}M)_\lambda(p)\ ,$$

since $h\in\mathscr{A}_{\mathrm{sG}}$.

Finally, the uniqueness insures us that, if $M_\lambda j=jM_{-\lambda}$ and $h(\lambda)j=jh(-\lambda)$, then $(h^{\#}M)_\lambda$ also has this property. □

Theorem 6.6. *The group $\mathscr{A}_{\mathrm{sG}}$ acts smoothly on canonical trivializations for the Lax pair for the sine-Gordon equation.*

Proof. From proposition 6.5, we have a group action which carries canonical trivializations into trivializations for the harmonic map problem in a gauge-free context. Moreover, if we let $(h^{\#}M)_\lambda=M^{\#}_\lambda=M_\lambda T_\lambda^{-1}$,

$$(M^{\#}_\lambda)^{-1}\frac{\partial}{\partial\eta}M^{\#}_\lambda=A^{\#}_\eta+\lambda B^{\#}_\eta=\left(T_\lambda(A_\eta+\lambda B_\eta)-\frac{\partial T_\lambda}{\partial\eta}\right)T_\lambda^{-1}\ , \qquad (35)$$

$$(M_\lambda)^{-1}\frac{\partial}{\partial\xi}M^{\#}_\lambda=A^{\#}_\xi+\lambda^{-1}B^{\#}_\xi=\left(T_\lambda(A_\xi+\lambda^{-1}B_\xi)-\frac{\partial T_\lambda}{\partial\xi}\right)T_\lambda^{-1}\ . \qquad (36)$$

Our first observation is that the commutation relation $jM_\lambda=M_{-\lambda}j$ implies $A^{\#}$ commutes with j and $B^{\#}$ anti-commutes with j.

Next we use the fact that $T_\lambda=I+\lambda T_1+\cdots$. So, by taking power series at 0 and (35),

$$A^{\#}_\eta+\lambda B^{\#}_\eta=A_\eta+\lambda(B_\eta+[T_1,A_\eta]-\partial T_1/\partial\eta)\ .$$

Since $A_\eta=0$, $A^{\#}_\eta=0$. Moreover, one can read off again that $B^{\#}_\eta$ anti-commutes with j.

Now, by taking (36) instead, we get

$$\lambda^{-1}B^{\#}_\xi+A^{\#}_\xi=\lambda^{-1}B_\xi+(A_\xi+[T_1,B_\xi])\ .$$

We conclude that $B_\xi=B^{\#}_\xi=i$.

Once we know $A^{\#}_\eta=0$, $B^{\#}_\xi=i$ and the fact that $A^{\#}_\xi$ commutes with j and $B^{\#}_\eta$ anti-commutes, the only fact we need to verify is $|B^{\#}_\eta|^2=|B_\eta|^2=2$. This we do by checking (35) using the Laurent series at ∞. We can assume $\lambda^{-m}T(\lambda)\sim T(\infty)$. Then from

$$B^{\#}_\eta=T(\infty)B_\eta T(\infty)^{-1}$$

it is obvious that

$$\mathrm{tr}(B_\eta^{\#})^2=\mathrm{tr}(B_\eta)^2=2\,. \qquad \square$$

It is appropriate to derive the infinitesimal formulas in two ways, both directly and from our formulas for δE_λ.

Theorem 6.7. *Let $g=\delta f$ be holomorphic in a neighborhood N_ϵ of 0 and ∞ with $g(0)=0$, $g(\lambda)+g(\bar{\lambda})^*=0$ and $g(\lambda)j=jg(-\lambda)$. Let Γ_ϵ be a pair of contours about 0 and ∞ in N_ϵ oriented counterclockwise about 0 and ∞. Then the induced infinitesimal action is given by*

$$M_\gamma^{-1}\,\delta M_\gamma=\frac{\gamma}{2\pi\mathrm{i}}\oint_{\Gamma_\epsilon}\frac{M_\lambda^{-1}g(\lambda)M_\lambda}{(\lambda-\gamma)\lambda}\,\mathrm{d}\lambda$$

for γ in the region between the two contours.

Proof. Suppose we have factored smoothly

$$h_t(\lambda)M_\lambda=(h_t^{\#}M)_\lambda T_{t,\lambda}$$

for a one-parameter family with $\mathrm{d}/\mathrm{d}t|_{t=0}h_t(\lambda)=g(\lambda)$. Here we assume $h_t(\lambda)$ is in a suitable extension of $\mathscr{A}_{\mathrm{sG}}$. Then, differentiating at $t=0$, we get

$$g(\lambda)M_\lambda=\delta M_\lambda+M_\lambda Q_\lambda\,,$$

where $T_{0,\lambda}=I$ and $\mathrm{d}/\mathrm{d}t|_{t=0}T_{t,\lambda}=Q_\lambda$. Then

$$M_\lambda^{-1}g(\lambda)M_\lambda=M_\lambda^{-1}\,\delta M_\lambda+Q_\lambda$$

is a decomposition of $M_\lambda^{-1}g(\lambda)M_\lambda$, which is holomorphic in punctured neighborhoods of 0 and ∞, into $M_\lambda^{-1}\,\delta M_\lambda$, which is holomorphic on $\mathbb{C}^*$, and Q_λ, which is holomorphic in neighborhoods of 0 and ∞. The normalization is $Q_0=0$; and the solution is via contour integrals. For γ inside the neighborhoods

$$Q_\gamma=-\frac{\gamma}{2\pi\mathrm{i}}\oint_{\Gamma_\epsilon}\frac{M_\lambda^{-1}g(\lambda)M_\lambda}{(\lambda-\gamma)\lambda}\,\mathrm{d}\lambda\,.$$

For γ in the complement of these neighborhoods

$$M_\gamma^{-1}\,\delta M_\gamma=\frac{\gamma}{2\pi\mathrm{i}}\oint_{\Gamma_\epsilon}\frac{M_\lambda^{-1}g(\lambda)M_\lambda}{(\lambda-\gamma)\lambda}\,\mathrm{d}\lambda\,.$$

The sum reduces to a small contour about γ, which is counterclockwise,

$$M_\gamma^{-1}g(\lambda)\,\delta M_\gamma=\frac{\gamma}{2\pi\mathrm{i}}\oint_{|\gamma-\lambda|=\epsilon/2}\frac{M_\lambda^{-1}g(\lambda)M_\lambda}{(\lambda-\gamma)\lambda}\,\mathrm{d}\lambda\,,$$

which is just the Cauchy formula.

We have, in an alternative derivation,

$$E_\gamma^{-1}\,\delta E_\gamma = \frac{(\gamma-1)}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \frac{E_\lambda^{-1} g(\lambda) E_\lambda}{(\lambda-\gamma)(\lambda-1)}\,\mathrm{d}\lambda$$

for $E_\gamma = M_\gamma \cdot M_1^{-1}$ in the harmonic map gauge. So

$$M_\gamma^{-1}\,\delta M_\gamma = M_1^{-1} E_\gamma^{-1}\,\delta E_\gamma M_1 + M_1^{-1}\,\delta M_1\,.$$

This gives us the formula

$$M_\gamma^{-1}\,\delta M_\gamma = \frac{\gamma-1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \frac{M_\lambda^{-1} g(\lambda) M_\lambda}{(\lambda-\gamma)(\lambda-1)}\,\mathrm{d}\lambda + M_1^{-1}\,\delta M_1\,.$$

If we prescribe the gauge variation

$$M_1^{-1}\,\delta M_1 = \frac{1}{2\pi\mathrm{i}} \oint \frac{M_\lambda^{-1} g(\lambda) M_\lambda}{(\lambda-1)\lambda}\,\mathrm{d}\lambda\,,$$

then we obtain the correct formula also. □

Corollary 6.8. *We claim*

$$q_\infty \boldsymbol{j} = + \frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \frac{M_\lambda^{-1} q(\lambda) M_\lambda}{\lambda}\,\mathrm{d}\lambda$$

for a function q on $\mathrm{E}^{1,1}$. If u is the solution of the sine-Gordon equation

$$\frac{\partial}{\partial\xi}\frac{\partial}{\partial\eta} u + 2\sin(2u) = 0$$

corresponding to M_λ, then q solves the linearization

$$\frac{\partial}{\partial\xi}\frac{\partial}{\partial\eta} q + 4\cos(2u)\, q = 0\,.$$

Proof. We claim that q is just the variation δu of u associated to δM_λ. The formula for $Q_\infty = \mathrm{d}/\mathrm{d}t|_{t=0} T_{t,\infty}$ is just

$$Q_\infty = -\frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \frac{M_\lambda^{-1} g(\lambda) M_\lambda}{\lambda}\,\mathrm{d}\lambda\,.$$

But $Q_\infty \boldsymbol{j} = \boldsymbol{j} Q_\infty$, $Q_\infty = q\boldsymbol{j}$,

$$B_\eta^t = \exp(-u_t \boldsymbol{j})\,\boldsymbol{i}\exp(u_t \boldsymbol{j}) = T_{t,\infty}(B_\eta) T_{t,\infty}^{-1}\,.$$

It is not difficult to see that

$$\delta u \boldsymbol{j} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} u_t \boldsymbol{j} = -\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} T_{t,\infty} = -Q_\infty .$$

This can be checked directly also, since, if

$$Q = q\boldsymbol{j} = \frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \frac{M_\lambda^{-1} g(\lambda) M_\lambda}{\lambda}\, \mathrm{d}\lambda ,$$

then

$$\frac{\partial Q}{\partial \xi} = \frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} [M_\lambda^{-1} g(\lambda) M_\lambda, A_\eta]\, \mathrm{d}\lambda .$$

Working this out, we get that

$$\begin{aligned}
\frac{\partial^2 q\boldsymbol{j}}{\partial\xi\,\partial\eta} = \frac{\partial^2 Q}{\partial\xi\,\partial\eta} &= \frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_\epsilon} \left[\left[\frac{M_\lambda^{-1} q(\lambda) M_\lambda}{\lambda}, A_\eta\right], \boldsymbol{i}\right] \mathrm{d}\lambda \\
&= q[[\boldsymbol{j}, \exp(-u\boldsymbol{j})\, \boldsymbol{i} \exp(u\boldsymbol{j})], \boldsymbol{i}] \\
&= 2q[\exp(-u\boldsymbol{j})\, \boldsymbol{k} \exp(u\boldsymbol{j}), \boldsymbol{i}] \\
&= -2q(\exp(wu\boldsymbol{j}) + \exp(-2u\boldsymbol{j}))\boldsymbol{j} \\
&= -2q\cos(2u)\boldsymbol{j} ,
\end{aligned}$$

as claimed. □

We conclude this section by asserting that the action of the "simple" factors $h_{\gamma,\beta}$ generates a two-parameter family of solutions which precisely corresponds to the classical two-parameter family of Bäcklund transformations for the sine-Gordon equation. There are many ways to perform the calculations. The choice of α corresponds to an initial condition which may vary with the choice of M_λ.

Lemma 6.9. *Let $h(\lambda) = h_{\gamma,\beta}(\lambda) = I + \lambda\gamma \exp(-\alpha\boldsymbol{j})\, \boldsymbol{i} \exp(\alpha\boldsymbol{j})$. Then the new Lax pair for $(H^{\#}_{\gamma,\beta}M)_\lambda$, where M_λ is a canonical trivialization for the solution u of the sine-Gordon equation, has the form*

$$T_\lambda^{-1}\left(\lambda \mathrm{e}^{-u\boldsymbol{j}} \boldsymbol{i} \mathrm{e}^{u\boldsymbol{j}} T_\lambda + \frac{\partial T}{\partial\eta}\lambda\right) = \lambda \mathrm{e}^{-w\boldsymbol{j}} \boldsymbol{i} \mathrm{e}^{w\boldsymbol{j}} ,$$

$$T_\lambda^{-1}\left[\left(\lambda^{-1}\boldsymbol{i} + \frac{\partial u}{\partial\xi}\boldsymbol{j}\right) T_\lambda + \frac{\partial T}{\partial\xi}\lambda\right] = \lambda^{-1}\boldsymbol{i} + \frac{\partial w}{\partial\xi}\boldsymbol{j} ,$$

where $T_\lambda = I + \gamma\lambda \exp(-w\boldsymbol{j})\, \boldsymbol{i} \exp(v\boldsymbol{j})$. *Here* $w: \mathrm{E}^{1,1} \to R$ *is the new solution,* $v: \mathrm{E}^{1,1} \to R$ *is an auxiliary function.*

Proof. According to our factorization procedure, $(h^{\#}M)\lambda$ is formed from M_λ by

$$(h^{\#}M)_\lambda = h(\lambda) M_\lambda T_\lambda^{-1},$$

where h and T have poles and zeros as maps into $\mathrm{PSL}(2, \mathbb{C})$ at points which must correspond to each other. Since h has poles and zeros only at $\lambda = \pm \mathrm{i}\gamma^{-1}$, the same must be true for T. Moreover, the normalization procedure requires $T_\lambda(0) = I$. The requirement that $T_\lambda \boldsymbol{j} = \boldsymbol{j} T_{-\lambda}$ fixes the form of T_λ as written. This is the same calculation as for Bäcklund transformations for the Euclidean harmonic map problem done in chapter 5 of our previous paper [9]. □

Theorem 6.10. *Let* $h_{\gamma,\beta}$ *be a generator of a Bäcklund transformation. Then the new extended solution has the form*

$$(h^{\#}M)_\lambda = h(\lambda) M_\lambda T_\lambda,$$

where $T_\lambda = I + \lambda\gamma \exp(-v\boldsymbol{j})\, \boldsymbol{i} \exp(v\boldsymbol{j})$. *Moreover* $v: \mathrm{E}^{1,1} \to \mathbb{R}$ *solves the coupled pair of equations*

$$\gamma \frac{\partial v}{\partial \eta} + \sin 2(u - v) = 0, \qquad \frac{\partial}{\partial \xi}(v - u) + \gamma \sin 2v = 0,$$

and $w = 2v - u$ *is the new solution to the sine-Gordon equation.*

Proof. We work first with the equation

$$T_\lambda^{-1}(\lambda \mathrm{e}^{-u\boldsymbol{j}} \boldsymbol{i}\, \mathrm{e}^{u\boldsymbol{j}}) T_\lambda + T_\lambda^{-1} \frac{\partial T_\lambda}{\partial \eta} = \lambda \mathrm{e}^{-w\boldsymbol{j}} \boldsymbol{i}\, \mathrm{e}^{w\boldsymbol{j}}.$$

According to the lemma

$$T_\lambda = I + \gamma\lambda \mathrm{e}^{-v\boldsymbol{j}} \boldsymbol{i}\, \mathrm{e}^{v\boldsymbol{j}}.$$

An easy calculation shows that

$$T_\lambda^{-1} = (1 + \gamma^2\lambda^2)^{-1}(I - \gamma\lambda \mathrm{e}^{-v\boldsymbol{j}} \boldsymbol{i}\, \mathrm{e}^{v\boldsymbol{j}}).$$

Substitute for T_λ and T_λ^{-1} and do the calculations carefully, remembering $\boldsymbol{ij} + \boldsymbol{ji} = 0$ and $\boldsymbol{i} \exp(v\boldsymbol{j}) = \exp(-v\boldsymbol{j})\, \boldsymbol{i}$. We get then

$$\begin{aligned} &\lambda \left(\mathrm{e}^{-2u\boldsymbol{j}} - 2\gamma \mathrm{e}^{-2v\boldsymbol{j}} \frac{\partial v}{\partial \eta} \boldsymbol{j} \right) \boldsymbol{i} \\ &+ \lambda^2 \left(\gamma \mathrm{e}^{2(u-v)\boldsymbol{j}} - \mathrm{e}^{2(v-u)\boldsymbol{j}} + 2\gamma^2 \boldsymbol{j} \frac{\partial v}{\partial \eta} \right) \\ &+ \lambda^3 \gamma^2 \mathrm{e}^{(-4v + 2u)\boldsymbol{j}} \boldsymbol{i} \\ &= (1 + \gamma^2\lambda^2) \lambda \mathrm{e}^{-w\boldsymbol{j}} \boldsymbol{i}\, \mathrm{e}^{w\boldsymbol{j}}. \end{aligned}$$

Clearly the single coefficient of λ^2 must vanish. Since

$$e^{2(u-v)j}-e^{-2(u-v)j}=2\sin 2(u-v)\,j\,,$$

this is equivalent to

$$\gamma\frac{\partial v}{\partial\eta}+\sin 2(u-v)=0\,.$$

Substitute this into the coefficient of λ and it becomes

$$e^{-2uj}+e^{-2vj}(e^{2(u-v)j}-e^{2(v-u)j})\,i=e^{(2u-4v)j}i\,.$$

This agrees with the coefficient of $\lambda^3\gamma^2$, and can be rearranged as $\exp[(-2v+u)j]\,i\exp[(2v-u)j]$. Setting $w=2v-u$ as in the theorem makes the equation correct.

Now follow the same format for the second pair of equations,

$$T_\lambda^{-1}\left[\left(\lambda^{-1}i+\frac{\partial u}{\partial\xi}j\right)T_\lambda+\frac{\partial}{\partial\xi}T_\lambda\right]=\lambda^{-1}i+\frac{\partial w}{\partial\xi}j\,.$$

Multiply the equation as before by $1+\lambda^2\gamma^2$. Then expanding each side in powers of λ,

$$\begin{aligned}&i\lambda^{-1}+\gamma(e^{-2vj}-e^{2vj})+\frac{\partial u}{\partial\xi}j\\&+\lambda\left(2\gamma j\,e^{-2vj}\frac{\partial(u-v)}{\partial\xi}+\gamma^2e^{-4vj}\right)i\\&+\lambda^2\gamma^2\left(-\frac{\partial u}{\partial\xi}+2\frac{\partial v}{\partial\xi}\right)j\\&=\left(\lambda^{-1}i+\frac{\partial w}{\partial\xi}j\right)(1+\lambda^2\gamma^2)\,.\end{aligned}$$

When we equate coefficients of λ^{-1} they agree. Equating coefficients of λ^2 agrees with setting $w=2v-u$. Equate coefficients of λ to get

$$2\gamma^{-1}j\,e^{-2vj}\frac{\partial(u-v)}{\partial\xi}+e^{-4vj}=1\,,$$

$$\gamma^{-1}\frac{\partial(u-v)}{\partial\xi}-\sin 2v=0\,.$$

Substitute in the coefficient of λ^0 to get

$$-2\gamma\sin 2vj+\frac{\partial u}{\partial\xi}j=\frac{\partial}{\partial\xi}(2v-u)\,j=\frac{\partial}{\partial\xi}wj\,.$$

Hence we have a consistent set of equations equivalent to those given in the theorem. □

It should be noted that the equations of theorem 6.10 are more familiar when written in terms of w. Then

$$2v=w+u\,,\qquad 2(v-u)=w-u\,,$$

$$\gamma\frac{\partial(w+u)}{\partial\eta}-2\sin(w-u)=0\,,$$

$$\frac{\partial}{\partial\xi}(w-u)+2\gamma\sin(w+u)=0\,.$$

The unusual factors of 2 appear due to our different coefficient $(\partial\partial/\partial\xi\partial\eta)w+2\sin(2w)=0$. These equations are the equations of the classical Bäcklund transformation, normalized for our coefficients.

References

[1] E. Calabi, Minimal immersions of surfaces in Euclidean spheres, J. Diff. Geom. 1 (1967) 111–125.

[2] S.S. Chern, On the minimal immersions of the two-sphere in a space of constant curvature, in: *Problems in Analysis* (Princeton University Press, Princeton, NJ, 1970) pp. 27–49.

[3] A.M. Din and W.J. Zakrzewski, General classical solutions in the $\mathbb{C}P^{n-1}$ model, Nucl. Phys. B 174 (1980) 397–406.

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

[5] J. Eells and J.C. Wood, Harmonic maps from surfaces to complex projective spaces, Adv. Math. 49 (1983) 217–263.

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

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

[8] V.E. Zakharov and A.B. Shabat, Integration of nonlinear equations of mathematical physics by the method of inverse scattering II, Func. Anal. Appl. 13 (1978) 13–22.

[9] K. Uhlenbeck, Harmonic maps into Lie groups, J. Diff. Geom. 30 (1989) 1–50.

[10] M. Atiyah and N. Hitchin, *The Geometry and Dynamics of Magnetic Monopoles* (Princeton University Press, Princeton, NJ, 1988).

[11] P. Braam, Magnetic monopoles and Kleinian groups, Ph.D. thesis (Oxford, 1988).

[12] N. Hitchin, The self-duality equations on a Riemann surface, Proc. London Math. Soc. 55 (1987) 59–126.

[13] B. Piette, R. Zait and W. Zakrzewski, Solutions of the $U(N)$ σ models with Wess–Zumino–Witten term, Z. Phys. C 39 (1988) 359.

[14] G. Valli, On the energy spectrum of harmonic 2-spheres in unitary groups, Topology 27 (1988) 129–136.

[15] K. Pohlmeyer, Integrable Hamiltonian systems and interactions through quadratic constraints, Commun. Math. Phys. 46 (1976) 207–221.

[16] B. Piette, I. Stokoe and W. Zakrzewski, On stability of solutions of the U(N) chiral model in two dimensions, Z. Phys. C 34 (1988) 449.
[17] W. Zakrzewski, Classical solutions of two-dimensional Grassmannian models, J. Geom. Phys. 1 (1984) 39–63.