J. DIFFERENTIAL GEOMETRY
31 (1990) 627-710

HARMONIC MAPS FROM A 2-TORUS TO THE 3-SPHERE

N. J. HITCHIN

0. Introduction

There have been many advances in recent years in the theory of harmonic maps of Riemann surfaces to spheres or symmetric spaces. These give constructions which produce harmonic maps from algebraic curves in associated complex manifolds (e.g. [7], [10]) and which provide many new examples, but unfortunately none of these methods says anything about the simplest and most basic situation of a harmonic map into the 2-sphere or the 3-sphere. Indeed, Bryant has shown that, however many derivatives one takes, there is no way of obtaining a construction like this for a minimal surface in the 3-sphere S^3, a particular case of a harmonic map.

In this paper we shall tackle this same simple basic situation, in the case where the Riemann surface is a torus with any conformal structure. We shall show that the equations for a harmonic map from a torus to S^3 reduce in a different way to algebraic geometry, in fact the geometry of a hyperelliptic curve, which we call the *spectral curve* Σ. This curve has finite genus and is constrained by integrality conditions on the periods of certain differentials of the second and third kind. These constraints are difficult to handle in general, but we shall show the existence of new examples of harmonic maps, and in particular minimal tori in S^3, by finding suitable curves. Furthermore, the method of solution shows that harmonic maps to S^3 are by no means rigid in general. They admit deformations which are parametrized by a real torus, of dimension p, where p is the genus of Σ.

This method of solution has its origins both in the theory of integrable systems like the KdV equation or sinh-Gordon equation and in the study of magnetic monopoles via the Bogomolny equations, where in both cases an algebraic curve lies at the heart of the solution. It is the analogy with

Received March 22, 1988.

the Bogomolny equations, static solutions of the self-dual Yang-Mills equations, which we follow. We begin in §1 by introducing a system of equations which is globally more general than the harmonic map equations, but lies in a natural way within the realm of gauge theories on a 2-dimensional manifold.

We consider a connection A on a principal SU(2)-bundle P over the torus M, and an auxiliary *Higgs field* $\Phi \in \Omega^{1,0}(M; \operatorname{ad} P \otimes \mathbb{C})$. The basic equations we treat then

$$(*) \qquad d''_A \Phi = 0, \qquad F_A = [\Phi, \Phi^*],$$

where F_A is the curvature of A. They differ by a sign from the self-duality equations on M which are dealt with in detail in [15].

The link between solutions to $(*)$ and harmonic maps is provided by considering the two connections $\nabla_A + \Phi - \Phi^*$ and $\nabla_A - \Phi + \Phi^*$ which, as a consequence of the equations, are flat. If they have trivial holonomy, then the two corresponding covariant constant sections of P are related by a *harmonic map* from M to SU(2). If the holonomy is nontrivial, then the equations describe a harmonic section of a flat 3-sphere bundle. Many of the natural geometric conditions one might impose upon the map are naturally interpreted in this formalism. In particular a conformal harmonic map (whose image would be a minimal surface in S^3) is distinguished by the property $\det \Phi = 0$, and a map to a totally geodesic S^2 is distinguished by the property that A is reducible.

The gauge-theoretic equations $(*)$ are equivalent to the statement that the curvature of the $SL(2, \mathbb{C})$ connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ vanishes for all $\zeta \in \mathbb{C}^*$. The holonomy of this flat connection, and its dependence on ζ, forms the subject matter of §§2 and 3. We consider the holonomy around two generators of the fundamental group $\pi_1(M)$ and show in §2 that the eigenvalues $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ of these two holonomy matrices, which are 2-valued functions on $\mathbb{C}^*$, actually have a finite number of branch-points. This argument involves the full 2-dimensional compactness of M and the study of the family of elliptic operators $d''_A - \zeta\Phi^*$. Furthermore, the branch points of $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ are shown to coincide, and this depends on the fact that $\pi_1(M)$ is abelian. In §3 the limiting behavior of the holonomy as $\zeta \to 0$ and $\zeta \to \infty$ is investigated. This again uses the regularity of families of elliptic operators—standard singular perturbation theory for the holonomy around one generator would not be sufficient to obtain enough information. From the results of these two sections, we show in §4 that $\theta = d\mu/\mu$ and $\tilde{\theta} = d\tilde{\mu}/\tilde{\mu}$ are well defined differentials of the second kind on a ramified double covering of $\mathbb{C}P^1$ (the compactification of $\mathbb{C}^*$)

with double poles at 0 and ∞. From their definition their periods lie in $2\pi i\mathbb{Z}$.

The nonsingular hyperelliptic curve obtained this way is not quite the spectral curve, but is instead its normalization—the spectral curve Σ is allowed to have singularities. In §5 we define the spectral curve by introducing the appropriate singularities. The differentials are well defined on Σ as is the *eigenspace bundle* E_x for $x \in M$. This is the eigenspace of the holonomy of the flat connection around closed curves through x. It is a line bundle over the whole of Σ of degree $-(p+1)$ where p is the arithmetic genus of Σ.

To illustrate the development thus far, we compute in §6 the spectral curve of a simple example—the Clifford torus. This is the minimal surface in S^3 defined by $|z_1| = |z_2| = 1/\sqrt{2}$. The spectral curve is here rational—the double covering of $\mathbb{C}P^1$ branched over 0 and ∞.

As $x \in M$ varies, the eigenspace bundle varies in the Picard variety of Σ, the space of equivalence classes of holomorphic line bundles of degree $-(p+1)$. We show in §7 that, considering this variety as an orbit of the abelian group $H^1(\Sigma\colon \mathscr{O})/H^1(\Sigma;\mathbb{Z})$, the variation is *linear* in the real and imaginary parts of a uniformizing parameter of M, which is effectively a linearization of the equation $(*)$. Using the fact that E_x is quaternionic with respect to a real structure on Σ, we deduce a number of properties of the bundle, including the fact that it is nonspecial.

At this stage we have shown that any solution of the equation $(*)$ determines a hyperelliptic curve Σ, differentials θ and $\tilde{\theta}$ with periods in $2\pi i\mathbb{Z}$ and a line bundle E_0 (corresponding to an origin in M), satisfying certain reality properties. In §8 we reverse the whole process, and show how to construct a solution of the equations from such data. This construction is framed largely in geometric form. Its analytic counterpart, which would be necessary to derive explicit solutions, would be couched in terms of Baker-Akhiezer functions, θ-functions or τ-functions as in [26], for example. Special classes of solutions correspond to special properties of the data. For example $\det\Phi = 0$ (the case of a conformal map) requires 0 and ∞ to be branch points of Σ, and A reducible (the case of a map to S^2) requires an extra involution τ on Σ such that the linear variation of the eigenspace bundle takes place in the Prym variety of τ. The condition for obtaining an actual harmonic map is

$$\mu(1) = \mu(-1) = \tilde{\mu}(1) = \tilde{\mu}(-1) = 1.$$

It can also be expressed in terms of constraints on the periods of integrals of the third kind on Σ having simple poles over $+1$ and -1 in $\mathbb{C}P^1$.

There is only one class of harmonic maps which cannot be constructed this way, and those are the ones for which μ and $\tilde{\mu}$ are *constant*. There are then no branch points, no differentials, no spectral curve. These maps are, however, the conformal maps to S^2, and they can be found straightforwardly in terms of systems of divisors on the Riemann surface M.

In producing the construction, we find that the constraints to produce a harmonic map are actually independent of the choice of line bundle E_0. So, given one harmonic map we obtain others by varying E_0 in the Picard variety of quaternionic bundles of degree $-(p+1)$. When Σ is nonsingular, this is a p-dimensional real torus, so gives a p-dimensional deformation space for the map.

The construction would be worth little if there were no examples at all, and so §§9–12 are devoted to finding curves satisfying the required properties for genus 0–3. In the case of a rational curve, it is easy to find the constraints—they are of an algebraic rather than transcendental nature. One may also easily write down the harmonic maps. We find that the only conformal maps are coverings of the Clifford torus. The elliptic solutions in §10 include two well-known examples. The first is the Gauss map of the Delaunay surface. This is the surface of revolution in $\mathbb{R}^3$ generated by rolling an ellipse along the x-axis, and rotating the curve traced out by its focus. This is a classical example of a surface of constant mean curvature, and by the theorem of Ruh and Vilms [24] the Gauss map is harmonic. This map is doubly periodic and so gives a harmonic map of the torus. Elliptic functions are clearly involved here. The second case concerns minimal tori in S^3 which are invariant under a circle subgroup of $\mathrm{SO}(4)$. By interpreting these as closed geodesics in a surface of revolution, Hsiang and Lawson [16] characterized these in terms of rational values of an elliptic integral. For us these two uses of elliptic functions are simply a manifestation of the fact that the spectral curve is elliptic, i.e., of genus 1.

For the case of genus 2 in §11 we consider a curve invariant by a cyclic group of order 4 and derive the condition for a harmonic map as a certain hyperelliptic integral taking values in $\mathbb{Q}(i)$. Since 0 and ∞ are branch points in this case, we find minimal tori in S^3.

In these cases, where $p \leq 2$, the deformations of the harmonic map obtained by varying E_0 in the Picard variety are all accounted for by conformal transformations of the torus within itself. Only in the case of $p \geq 3$ will we obtain a true deformation of the image of the map. Thus the genus 3 example produced in §12 is important for possessing this property. It turns out that it is in fact already known and consists of the Gauss map of Wente's surface [28], an immersed torus of constant mean curvature in $\mathbb{R}^3$. This is a harmonic map to S^2, but if E_0 moves out of the Prym variety, the image of the torus moves off the 2-sphere. We make contact here also with Abresch's analytic formula for Wente's surface in terms of elliptic functions [1]—the Prym variety in question is covered by the product of two elliptic curves. The deformations of Wente's Gauss map will however require fully hyperelliptic functions of genus 3 to describe.

From the point of view presented here, all the information about a harmonic map is stored in the geometry of the spectral curve. It is not in general easy to extract information of a geometric nature about the map from the curve. We do however succeed in §13 in calculating the *energy* in terms of coefficients of expansions of the differentials θ and $\tilde{\theta}$ about their poles. Another question to ask is which minimal tori are actually embedded? Some rough information is provided by the theorem of Choi and Schoen [8] which shows that the space of embedded minimal surfaces is *compact*. Knowing that the genus p of the spectral curve depends continuously on the parameters, this tells us that only finitely many values of p give embedded minimal tori. It is conjectured that the Clifford torus is the only embedded minimal torus, which is precisely the statement $p = 0$.

There are some results, however, which are naturally suggested by our methods. For example, a spectral curve of genus p giving an immersed minimal torus has branch points $0, \infty, \alpha_1, \cdots, \alpha_p, \overline{\alpha}_1^{-1}, \cdots, \overline{\alpha}_p^{-1}$ giving $2p$ real degrees of freedom. The principal parts of θ and $\tilde{\theta}$ give four more, but the constraints on their periods give $2p$ conditions. There are four further constraints for an actual harmonic map, and so assuming transversality, for each p we have a discrete family of spectral curves and so in particular a countable number of possible moduli for the torus. There appears this way to be a severe constraint on the modulus of a torus in order for it to be conformally minimally immersed in S^3. This contrasts with the case of S^4 in which any Riemann surface can be conformally minimally immersed [7].

A proper treatment of this phenomenon requires an investigation into the geometry of the moduli space of solutions to the full gauge-theoretic equations $(*)$, and this we postpone for a later paper.

1. Gauge-theory formalism

We shall begin by considering a reformulation of the equations for a harmonic map of a Riemann surface M to a compact Lie group G in terms of a connection on a principal G-bundle over M and an auxiliary field. This will allow us greater freedom and generality and to make use of approaches suggested by parallel gauge-theoretic problems.

Suppose first that $f: M \to G$ is a smooth map from a compact Riemann surface to a Lie group endowed with a bi-invariant metric.

The *derivative* $df \in \Omega^1(M; f^*(TG))$ is a one-form with values in the pull-back of the tangent bundle of G. The Levi-Civita connection on TG is pulled back to a connection A on f^*TG, with structure group G. We denote by d_A the usual covariant exterior derivative $d_A: \Omega^p(M; f^*(TG)) \to \Omega^{p+1}(M; f^*(TG))$. The derivative df automatically satisfies the equation

$$d_A(df) = 0. \tag{1.1}$$

The map f is *harmonic* if df satisfies the equation $d_A^*(df) = 0$. If we use the Hodge star operator $*: \Omega^1(M) \to \Omega^1(M)$, this is equivalent to

$$d_A(*df) = 0. \tag{1.2}$$

Since the star operator in the middle dimension is conformally invariant, these equations are themselves dependent only on the *conformal* structure of the Riemann surface M.

Now the Levi-Civita connection of a compact Lie group may be expressed in terms of two other connections. We have the flat connection obtained by trivializing the tangent bundle by *left* translation, and similarly another flat connection obtained by *right* translation. Denoting their covariant derivatives by ∇_L and ∇_R, then the Levi-Civita connection ∇ is the average of the two:

$$\nabla = \tfrac{1}{2}(\nabla_R + \nabla_L). \tag{1.3}$$

(This relation is usually expressed by means of the covariant derivatives of left-invariant vector fields, giving the formula $\nabla_X Y = \frac{1}{2}[X, Y]$.)

The two trivial connections ∇_L and ∇_R are gauge-equivalent, the gauge transformation being just the adjoint representation on the Lie algebra—the space of left-invariant vector fields. Thus

$$g^{-1}\nabla_L g = \nabla_R,$$

or equivalently

$$g^{-1}dg = \nabla_R - \nabla_L. \tag{1.4}$$

Given a map $f: M \to G$, we pull back the connections and obtain from (1.3) and (1.4)

$$\nabla_A = \tfrac{1}{2}(\nabla_R + \nabla_L), \qquad f^{-1}df = \nabla_R - \nabla_L.$$

The expression $f^{-1}df$ is here a 1-form with values in the Lie algebra of left-invariant vector fields, but left-translating it back, it is the derivative df of the map, a 1-form with values in $f^*(TG)$. This allows us to rewrite the harmonic map equations in gauge-theoretic terms on M. We have a principal G-bundle P with connection ∇_A and trivial connections ∇_L and ∇_R, such that $\nabla_A = \frac{1}{2}(\nabla_L + \nabla_R)$. The *difference* $\nabla_R - \nabla_L$ is a 1-form 2ϕ with values in the vector bundle $\operatorname{ad} P$ associated to P by the adjoint representation. (1.1) and (1.2) now become

$$d_A\phi = 0, \qquad d_A * \phi = 0. \tag{1.5}$$

These simplify if we write $\phi = \Phi - \Phi^*$ where $\Phi \in \Omega^{1,0}(M; \operatorname{ad} P \otimes \mathbb{C})$ is the component of ϕ of type $(1, 0)$. Since

$$*(\Phi - \Phi^*) = i(\Phi + \Phi^*),$$

equations (1.5) give

$$\begin{aligned} d_A(\Phi - \Phi^*) &= d''_A\Phi - d'_A\Phi^* = 0, \\ d_A(\Phi + \Phi^*) &= d''_A\Phi + d'_A\Phi^* = 0, \end{aligned}$$

and hence

$$d''_A\Phi = 0. \tag{1.6}$$

Here $d''_A \colon \Omega^{1,0}(M; \operatorname{ad} P\otimes\mathbb{C}) \to \Omega^{1,1}(M; \operatorname{ad} P\otimes\mathbb{C})$ is the Cauchy-Riemann operator defined by the connection A—the (0,1) part of the covariant derivative. The d''_A gives $\operatorname{ad} P \otimes \mathbb{C}$ the structure of a holomorphic vector bundle (see [5]), sometimes called the Koszul-Malgrange structure. With this interpretation, (1.6) states that Φ is a holomorphic section of $\operatorname{ad} P \otimes_{\mathbb{C}} K$, where K is the canonical bundle of holomorphic 1-forms on M.

There is a further equation arising from the fact that $\nabla_L = \nabla_A - \phi$ and $\nabla_R = \nabla_A + \phi$ are flat connections. Knowing the curvature of ∇_L to be zero gives

$$\begin{aligned} 0 = d_L^2 &= (d_A - \phi)^2 = d_A^2 + \phi^2 \quad (\text{since } d_A\phi = 0) \\ &= F_A + (\Phi - \Phi^*)^2 = F_A - \Phi\Phi^* - \Phi^*\Phi. \end{aligned}$$

Using the usual extension of the Lie bracket to Lie-algebra valued forms, we may write this equation together with (1.6) as

$$\boxed{\begin{aligned} d''_A\Phi &= 0, \\ F_A &= [\Phi, \Phi^*]. \end{aligned}} \tag{1.7}$$

We shall call A the *connection* associated to the harmonic map, and Φ the *Higgs field.* This is a notation arising from gauge theories in mathematical physics. In fact, locally, (1.7) may be interpreted as the self-dual Yang-Mills equations in $\mathbb{R}^4$ with signature $(++--)$ which are invariant under translation in the last two variables. The analogous reduction of the self-dual Yang-Mills equations in Euclidean $\mathbb{R}^4$ leads to the equations

$$d''_A\Phi = 0, \qquad F_A = -[\Phi, \Phi^*],$$

which were studied in [15]. There it was shown that there are no non-abelian solutions to these equations on a two-dimensional torus. By contrast we shall see that there is a nontrivial theory for equations (1.7).

Let us restrict ourselves now to the case $G = \mathrm{SU}(2)$, with its bi-invariant metric given by the Killing form. This is of course a metric of constant curvature on $\mathrm{SU}(2) \cong S^3$.

There are three special types of harmonic map from a Riemann surface to $\mathrm{SU}(2)$ which not only have their own individual geometrical interest, but also, as we shall see, lead to distinct approaches in solving the equations.

The first type is a *conformal* harmonic map: the pull-back of the metric on $\mathrm{SU}(2)$ is a metric on M in the conformal class of the Riemann surface. Geometrically, the image of M is then an immersed *minimal surface* in $\mathrm{SU}(2)$ [18]. A slight extension of this is the notion of *branched conformal* harmonic map, where the pulled-back metric becomes degenerate at a finite number of points—the branch points of the map—but elsewhere is in the conformal class.

The second type of map is a harmonic map to a *totally geodesic 2-sphere* in $S^3 \cong \mathrm{SU}(2)$. The Gauss map of a surface of constant mean curvature in $\mathbb{R}^3$ is a harmonic map to the 2-sphere [24].

The third type is a combination of the two: a *branched conformal map* to the 2-sphere. Such a map is by definition holomorphic or antiholomorphic —the equations reduce entirely to the Cauchy-Riemann equations—and presents no problems in finding all solutions. One simply considers linear systems of divisors on the Riemann surface M.

We shall give now the gauge theoretic interpretation, in terms of the connection A and Higgs field Φ, of the above special cases. Since the connection is an $\mathrm{SU}(2)$-connection we may consider it as being defined on a rank 2 complex vector bundle V with a symplectic form $\langle\ ,\ \rangle$. The Higgs field Φ is then a holomorphic section of $\operatorname{End} V \otimes K$ whose trace is zero.

Proposition (1.8). *Let $f\colon M \to \mathrm{SU}(2)$ be a harmonic map of a compact Riemann surface, and let (A, Φ) be the corresponding connection and Higgs field. Then f is a branched conformal map if and only if $\det\Phi = 0$. The branch points of f are the points of M where Φ vanishes.* (*Since Φ is a holomorphic section of $\operatorname{End} V \otimes K$ and V is of rank 2, $\det\Phi$ is a holomorphic section of K^2—a quadratic differential on M.*)

Proof. The pull-back of the bi-invariant metric on $\mathrm{SU}(2)$ under the map f is

$$\begin{aligned} g &= -\operatorname{tr}(f^{-1}df)^2 = -4\operatorname{tr}\phi^2 \\ &= -4(\operatorname{tr}\Phi^2 - 2\operatorname{tr}\Phi\Phi^* + \operatorname{tr}\Phi^{*2}). \end{aligned}$$

A metric is in the conformal class of M if it is locally of the form $h\,dz\,d\bar{z}$ with respect to a holomorphic coordinate z, so g is in this class if the dz^2 and $d\bar{z}^2$ components vanish, i.e. if and only if $\operatorname{tr}\Phi^2 = 0$. Since Φ is locally a 2×2 matrix with trace 0, $\det\Phi = -\frac{1}{2}\operatorname{tr}\Phi^2 = 0$ if and only if M is branched conformal.

The pulled-back metric is now $8\operatorname{tr}\Phi\Phi^*$ and this is degenerate only when $\Phi = 0$. Thus the branch points are the zeros of Φ. q.e.d.

In the case where $\det\Phi = 0$, we may define further holomorphic invariants. Since $\det\Phi = 0$ and $\operatorname{tr}\Phi = 0$, we have $\Phi^2 \equiv 0$. For all points of M, Φ therefore has a nonzero kernel and if Φ is not identically zero there is a well-defined holomorphic line bundle $L \subset V$ with $L \subseteq \ker\Phi$. Since $\Phi^2 = 0$, $\operatorname{im}\Phi \subseteq \ker\Phi$, so Φ defines a homomorphism of line bundles

$$\Phi\colon V/L \to L \otimes K$$

which vanishes at the zeros of Φ.

Since V has a symplectic form, V/L is isomorphic to L^*, thus the homomorphism above defines a holomorphic section $a \in H^0(M; L^2K)$ which vanishes at the zeros of Φ.

Consider next the holomorphic subbundle $L \subset V$. This is preserved by d''_A and we may ask if the covariant derivative d_A preserves it. The obstruction to this is given by the section $b \in \Omega^{1,0}(M; L^{-2})$ defined by

$$bs^2 = \langle d'_A s, s\rangle$$

for some local holomorphic section s of L. Now,

$$(d''_A b)s^2 = d''_A(bs^2) = \langle d''_A d'_A s, s\rangle$$

since s is holomorphic. But

$$F_A = d'_A d''_A + d''_A d'_A \quad \text{and} \quad d''_A s = 0,$$

so

$$(d''_A b)s^2 = \langle F_A s, s\rangle = \langle \Phi\Phi^* s - \Phi^*\Phi s, s\rangle \quad \text{from (1.7)}$$
$$= \langle \Phi^* s, \Phi s\rangle - \langle \Phi^*\Phi s, s\rangle = 0 \quad \text{since } \Phi s = 0.$$

Thus b is a *holomorphic section* of $L^{-2}K$. The product $ab \in H^0(M; K^2)$ is a quadratic differential naturally associated to the solution. In terms of the harmonic map it is essentially the *second fundamental form* of the image of M in S^3 [18]. The section b plays another role—that of determining the holomorphic structure on V. We already know that V is given by an *extension*

$$0 \to L \to V \xrightarrow{\pi} L^* \to 0$$

and hence by a class in $H^1(M; L^2)$.

To find a representative in Dolbeault cohomology for this class, we take a C^∞ line bundle L'' complementary to L and identify it with L^* by π. If t is a local section of L', then we define $b' \in \Omega^{0,1}(M; L^2)$ by

$$\langle d''_A t, t\rangle = b't^2.$$

It is easy to see that the cohomology class of b' in $H^1(M; L^2)$ measures the obstruction to finding a *holomorphic* complementary subbundle in V and hence is the extension class.

The SU(2) structure on the vector bundle V can be described in terms of the symplectic form $\langle\ ,\ \rangle$ and a quaternionic structure $j: V \to V$ which is antilinear and satisfies $j^2 = -1$. The hermitian form $(\ ,\ )$ is given by $(v, w) = \langle v, jw\rangle$.

Now let s be a local holomorphic section of L and consider the section $t = js$ of $jL = L^\perp$ in V. We have

$$b't^2 = \langle d''_A t, t\rangle = \langle d''_A js, js\rangle$$
$$= \langle jd'_A s, js\rangle = -\overline{\langle d'_A s, s\rangle} = -\overline{bs^2}.$$

Hence using the unitary structure on L to identify $\overline{L}^*$ and L, we have

$$b' = -\overline{b} \in \Omega^{0,1}(M; L^2)$$

representing the extension class.

Suppose this class is trivial, so that $b' = d''a$ for some $a \in \Omega^0(M; L^2)$. Then

$$b = \overline{-d''_L a} = d'_L \overline{a},$$

where d'_L is the $(1, 0)$ part of the unitary connection on L^{-2}. But $d''_L b = 0$ as shown above, so $d''_L d'_L \overline{a} = 0$. It follows that

$$\int_M (d'_L \overline{a},\ d'_L \overline{a}) = \int_M (d''_L d'_L \overline{a}, \overline{a}) = 0,$$

so $d'_L \overline{a} = 0$, i.e., $b = 0$.

Thus the extension defining V is *nontrivial if* $b \neq 0$.

We turn next to the second type of harmonic map, a map to a totally geodesic 2-sphere.

Proposition (1.9). *Let $f: M \to \mathrm{SU}(2)$ be a harmonic map of a compact Riemann surface, and let (A, Φ) be the corresponding connection and Higgs field. Then,*

(i) *f maps to a totally geodesic 2-sphere if and only if A is reducible to a $\mathrm{U}(1)$ connection.*

(ii) *f maps to a totally geodesic 2-sphere, if and only if there exists a gauge transformation g such that $g^2 = -1$, $g^{-1}\Phi g = -\Phi$ and g leaves A invariant.*

(*Of the two criteria given here, the first is a simple gauge-theoretic notion, but the second has more relevance when we actually solve the equations.*)

Proof. Any totally geodesic 2-sphere in S^3 is equivalent under an isometry of S^3 to the fixed-point set of the involution $g \to -g^{-1}$. The action on the tangent bundle of S^3 restricted to S^2 reduces the structure group from $\mathrm{SO}(3)$ to $\mathrm{SO}(2)$ and hence reduces the $\mathrm{SU}(2)$ connection (the connection on the spinor bundle) to $\mathrm{U}(1)$. Thus, pulling back to M gives a reducible connection.

Conversely suppose A is reducible to $\mathrm{U}(1)$. We may write $V = U \oplus U^*$ as an orthogonal direct sum for a complex line bundle U with a unitary connection. With respect to this decomposition the Higgs field Φ may be written as

$$\Phi = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix},$$

where the entries are holomorphic sections of the line bundles: $\alpha \in H^0(M; K)$, $\beta \in H^0(M; U^2 K)$, and $\gamma \in H^0(M; U^{-2} K)$. Now

$$[\Phi, \Phi^*] = \begin{pmatrix} \beta\overline{\beta} - \gamma\overline{\gamma} & 2\alpha\overline{\gamma} - 2\beta\overline{\alpha} \\ 2\gamma\overline{\alpha} - 2\alpha\overline{\beta} & -(\beta\overline{\beta} - \gamma\overline{\gamma}) \end{pmatrix}, \tag{1.10}$$

where we have used the unitary structure on U to identify $\overline{U}$ with U^{-1}. However A is reducible so

$$F_A = \begin{pmatrix} F_U & 0 \\ 0 & -F_U \end{pmatrix}.$$

Consequently from the equation $F_A = [\Phi, \Phi^*]$, we obtain

$$\alpha\overline{\gamma} = \beta\overline{\alpha}. \tag{1.11}$$

Assume now that the holomorphic 1-form α is not identically zero. Consider the quadratic differential

$$\det\Phi = -\alpha^2 - \beta\gamma = \frac{\alpha}{\overline{\alpha}}(\alpha\overline{\alpha} + \gamma\overline{\gamma}) \quad \text{from (1.11).}$$

If $\det\Phi$ vanishes at a point to order k, then it follows that α and γ must both vanish to order $\frac{1}{2}k$, and since they are both holomorphic this means in particular that k is even. Thus the zeros of the quadratic differential $\det\Phi$ have even multiplicities $2m_i$, and at each of these zeros the differential α has a zero of multiplicity m_i. Therefore for some constant λ we have

$$\det\Phi = \lambda\alpha^2.$$

However γ and similarly β have at least the same zeros as α and are sections of $U^{-2}K$ and U^2K respectively. It follows that U^2 must be trivial, and β and γ are themselves constant multiples of α. Thus, holomorphically, $V \cong U \oplus U$, and Φ is the product of a constant matrix C and the 1-form α.

If α is not identically zero, then (1.11) also leads to $\beta\overline{\beta} = \gamma\overline{\gamma}$ and consequently from (1.10), $[\Phi, \Phi^*] = 0$, so the constant matrix is normal and therefore diagonalizable. An eigenspace of Φ defines a holomorphic line subbundle of V, but since it is also an eigenspace for Φ^* by normality it is also antiholomorphic, i.e., preserved by d'_A as well as d''_A. This means the connection A *and* the Higgs field Φ reduce to the group $\mathrm{U}(1)$, or in other words we have a harmonic map to a circle subgroup of $\mathrm{SU}(2)$—a geodesic in S^2.

There only remains the case where α vanishes identically. We then have

$$\Phi = \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix}.$$

Consider the covariant constant gauge transformation $g = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$. We have $g^{-1}\Phi g = -\Phi$. Now consider the left-invariant connection $\nabla_L = \nabla_A - \Phi$ and the right-invariant connection $\nabla_R = \nabla_A + \Phi$. Since $g^{-1}\Phi g = -\Phi$ and g preserves ∇_A, we have

$$g^{-1}\nabla_L g = g^{-1}(\nabla_A - \Phi)g = \nabla_A + \Phi = \nabla_R.$$

Thus g represents the harmonic map f. Since $g^2 = -1$, it maps to a 2-sphere in $\mathrm{SU}(2)$.

Hence if A is reducible, f maps to a circle or a 2-sphere. Considering the second part of the proposition, if g leaves A invariant, then it is a covariant constant gauge transformation and thus reduces the structure group from SU(2) to U(1), leading to the first part.

Remark (1.12). Locally, the equations for a harmonic map to the 2-sphere take a more familiar form. Suppose we choose a local coordinate in a neighborhood where $\det\Phi$ is nonzero, such that $\det\Phi = -dz^2$. Now write $\beta = s^2 dz$ and $\gamma = s^{-2} dz$ for a local nonvanishing section s of U. The hermitian inner product U gives $\|s\|^2 = h$, and the isomorphism $\overline{U} \cong U^{-1}$ can be written as $\overline{s} \to hs^{-1}$. The equation $F_A = [\Phi, \Phi^*]$ now becomes

$$F_U = \beta\overline{\beta} - \gamma\overline{\gamma} = (h^2 - h^{-2})\, dz\, d\overline{z}.$$

Moreover, the curvature F_U of the line bundle is given by

$$F_U = d'' d' \log h.$$

Putting $h^2 = e^u$ we obtain the equation

$$-\frac{1}{2}\frac{\partial^2 u}{\partial z \partial \overline{z}} = e^u - e^{-u},$$

which is essentially the Euclidean version of the *sinh-Gordon* equation. Its relationship with surfaces of constant mean curvature in $\mathbb{R}^3$ is classical [11].

The final case of a branched conformal map to S^2 is obtained by considering both conditions: $\det\Phi = 0$ and A reducible.

Now we restrict attention to the case where M is of genus 1—*a torus.* In this case, the canonical bundle K is holomorphically trivial, so choosing a trivialization the Higgs field Φ is simply a trace zero endomorphism of the holomorphic vector bundle V. There are three nontrivial cases to consider according to the description above.

(a) *The general case:* $\det\Phi \neq 0$, *A is irreducible.* Here, since $\det\Phi$ is a holomorphic function and hence constant, Φ has distinct eigenvalues $\pm\lambda$ which are constants. The eigenspaces split the vector bundle V holomorphically as a direct sum $V \cong L \oplus L^*$.

(b) $\det\Phi \neq 0$, *A reducible.* In this case, as we saw in Proposition (1.9), either we have a U(1) solution to the equations, corresponding to a harmonic map to the circle, or the Higgs field has the form

$$\Phi = \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix}$$

relative to the decomposition $V = U \oplus U^*$ defined by the U(1) reduction of A. Since $\beta \in H^0(M; U^2)$, $\gamma \in H^0(M; U^{-2})$ and neither vanish because $\det \Phi \neq 0$, the line bundle U^2 is holomorphically trivial. Thus $V \cong U \oplus U$ and Φ is constant under this isomorphism. The eigenvalues $\pm\lambda$ of Φ are basic invariants of the solution.

(c) $\det \Phi = 0$, *A irreducible*. Here, as in Proposition (1.8), we have the kernel of Φ defining a subbundle $L \subset V$, a section $a \in H^0(M; L^2)$ defined by Φ: $V/L \to L$, and a section $b \in H^0(M; L^{-2})$ measuring the obstruction to A preserving the subbundle L. Since A is irreducible, b is nonzero, and therefore $\deg L^2 \leq 0$. If $\deg L^2 < 0$ then a vanishes, so Φ vanishes. But then $F_A = [\Phi, \Phi^*]$ vanishes and any flat SU(2) connection on a torus reduces to U(1). Thus $\deg L^2 = 0$ and the section b trivializes L^2.

The product $c = ab$ is a nonzero constant which is a basic invariant of the solution. The vector bundle V is a nontrivial extension of a line bundle L of order two by itself, as we showed above. Also, since a is nonvanishing everywhere, Φ does not vanish anywhere, and therefore from (1.8) any harmonic conformal map is an immersion.

(d) $\det \Phi = 0$, *A reducible*. In this case, as in (c), if $b \neq 0$ then V is a nontrivial extension of a line bundle of order 2 by itself. But if A is reducible, V is a direct sum of line bundles which contradicts this. Hence $b = 0$, so $V = L \oplus L^*$ where A reduces to a U(1) connection on L, the kernel of Φ. Thus, with respect to this direct sum decomposition,

$$\Phi = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$$

with $a \in H^0(M; L^2)$.

Finally, having given a gauge-theoretic description of a harmonic map to a Lie group G, and a geometrical interpretation in terms of A and Φ of natural constraints on the map, we may ask if *every* solution to the equations

$$d''_A\Phi = 0, \qquad F_A = [\Phi, \Phi^*]$$

on a compact Riemann surface arises from a harmonic map. Given such a solution, if we put $\phi = \Phi - \Phi^*$, then from the first equation $d_A\phi = d_A * \phi = 0$. Thus

$$(d_A - \phi)^2 = F_A + \phi^2 = F_A - [\Phi, \Phi^*] = 0,$$
$$(d_A + \phi)^2 = F_A + \phi^2 = F_A - [\Phi, \Phi^*] = 0.$$

Hence $\nabla_A + \phi$ and $\nabla_A - \phi$ are *flat* connections. In the case where (A, Φ) arises from a harmonic map these are not only flat but *trivial*—the holonomy is trivial. Conversely suppose the holonomy of these two flat connections is trivial, then we may find a covariant constant trivialization of the principal bundle P with respect to $\nabla_A - \phi$. The covariant constant trivialization with respect to $\nabla_A + \phi$ gives, relative to the first trivialization, a map to the group G. It is easy to check now that the equations $d_A\phi = d_A * \phi = 0$ make this map harmonic. Thus the *triviality* of the two flat connections is the condition to obtain a global harmonic map to the group. Locally, of course, in a simply connected neighborhood, this triviality is satisfied.

The global geometric analogue of the harmonic map which corresponds to a general solution of (1.7) consists of patching together these local maps. It may be described by taking the flat $G \times G$ connection on M by putting $\nabla_A - \phi$ on one factor and $\nabla_A + \phi$ on the other and forming the flat G-bundle associated to the right and left action of $G \times G$ on G (i.e., considering G as a symmetric space). A solution to the gauge-theoretic equations corresponds to a *harmonic section* of this flat bundle. Most of our work will deal with general solutions to (1.7), and hence harmonic sections, and only finally shall we consider the special conditions which give a true harmonic map.

2. Holonomy

The relationship between solutions to (1.7) and flat connections is much deeper than the identification of the flat connections $\nabla_A - \phi$ and $\nabla_A + \phi$ in terms of a harmonic map. It forms the basis of a method of solution of the equations themselves.

We introduce an indeterminate $\zeta \in \mathbb{C}^*$ and consider the complex connection

$$\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*. \tag{2.1}$$

The curvature of this connection is

$$\begin{aligned} F = (d_A + \zeta^{-1}\Phi - \zeta\Phi^*)^2 &= F_A - [\Phi, \Phi^*] \quad \text{since } d_A\Phi = d_A\Phi^* = 0 \\ &= 0 \quad \text{since } F_A = [\Phi, \Phi^*]. \end{aligned}$$

Thus, when $G = \mathrm{SU}(2)$, we have a family of flat $\mathrm{SL}(2, \mathbb{C})$ connections parametrized by $\zeta \in \mathbb{C}^*$. When $|\zeta| = 1$, the connections are unitary, and when $\zeta = +1$ or -1, they are the connections arising from the harmonic map, corresponding to the left and right invariant connections on $\mathrm{SU}(2)$.

This interpretation of the equations arises from many sources. It may be regarded as part of the Zakharov-Shabat formalism, used by Uhlenbeck [27] in analyzing the harmonic maps from S^2 to a Lie group, but may also be considered as a vestige of the Atiyah-Ward twistor space method of solving the self-dual Yang-Mills equations [5]. The use of a 1-parameter family of linear differential equations in solving a nonlinear problem which we shall employ here has its closest analogue in the solution of the periodic KdV equation [20].

Given a flat $SL(2, \mathbb{C})$ connection on M and a base point, it' holonomy describes a representation of the fundamental group

$$\pi_1(M) \to SL(2, \mathbb{C}).$$

We consider the situation where M is a torus, in which case $\pi_1(M) \cong \mathbb{Z} \oplus \mathbb{Z}$. Choosing generators, a representation is then a pair of commuting matrices in $SL(2, \mathbb{C})$. Since the family (2.1) of flat connections depends holomorphically on $\zeta \in \mathbb{C}^*$, we have holomorphically varying holonomy matrices $H(\zeta)$ and $\tilde{H}(\zeta)$ corresponding to the holonomy around the two generators of $\pi_1(M)$; they satisfy

$$\det H(\zeta) = \det \tilde{H}(\zeta) = 1, \qquad H(\zeta)\tilde{H}(\zeta) = \tilde{H}(\zeta)H(\zeta). \tag{2.2}$$

The holonomy matrix $H(\zeta)$ depends on the choice of base point in M, but its conjugacy class and hence eigenvalues do not. These will be our primary concern in this section.

Since $\det H(\zeta) = 1$, the eigenvalues μ and μ^{-1} of $H(\zeta)$ satisfy

$$\mu^2 - h(\zeta)\mu + 1 = 0,$$

where $h(\zeta) = \operatorname{tr} H(\zeta)$, and are thus given by

$$\mu = \tfrac{1}{2}[h(\zeta) \pm \sqrt{h(\zeta)^2 - 4}].$$

This is a 2-valued function on $\mathbb{C}^*$ with branch points at the odd-order zeros of $h(\zeta)^2 - 4$. We shall show next that there are only a finite number of these. This finiteness theorem, which is essential for reducing the harmonic map equations to algebraic geometry, is not simply a consequence of the compactness of the closed curve around which we are solving the ordinary differential equation given by the covariant derivative (2.1). It involves the full two-dimensional compactness of the torus. Indeed the analogous result for the periodic KdV equation does not hold—one needs in general hyperelliptic curves of infinite genus [20] to find all periodic solutions.

Proposition (2.3). *Let (A, Φ) be a solution to (1.7) on a torus M, and let $H(\zeta)$ and $\tilde{H}(\zeta)$ be the holonomy matrices of the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ around two generators of $\pi_1(M)$. Then the function $h(\zeta)^2 - 4$, where $h(\zeta) = \operatorname{tr} H(\zeta)$, has a finite number of odd order zeros in $\mathbb{C}^*$.*

Proof. Let $\zeta = \zeta_0$ be an odd zero of $h(\zeta)^2 - 4$ and put $t = \zeta - \zeta_0$. The holonomy matrices $H(\zeta)$ and $\tilde{H}(\zeta)$ have power series expansions around $\zeta = \zeta_0$:

$$H(t) = \sum_{i=0}^{\infty} A_i t^i, \qquad \tilde{H}(t) = \sum_{i=0}^{\infty} B_i t^i.$$

Now $h(\zeta_0)^2 = 4$ so $\operatorname{tr} A_0 = \pm 2$. Assume for the moment that $\operatorname{tr} A_0 = +2$.

If $h(\zeta)^2 - 4$ has a zero at $\zeta = \zeta_0$ of order $(2m+1)$, then

$$(2 + t(\operatorname{tr} A_1) + t^2(\operatorname{tr} A_2) + \cdots)^2 = 4 + t^{2m+1} g(t), \tag{2.4}$$

where $g(0) \neq 0$. Thus, equating coefficients,

$$\operatorname{tr} A_i = 0, \quad 0 < i < 2m+1, \qquad \text{and} \qquad \operatorname{tr} A_{2m+1} \neq 0. \tag{2.5}$$

Suppose $A_0 = I$. Then since

$$\det(A_0 + tA_1 + t^2 A_2 + \cdots) = 1,$$

we have

$$1 + \sum_{i=1}^{\infty} t^i (\operatorname{tr} A_i) + \det\left(\sum_{i=1}^{\infty} t^i A_i\right) = 1,$$

and using (2.5)

$$t^{2m+1}(\operatorname{tr} A_{2m+1}) + \cdots + \det(tA_1 + t^2 A_2 + \cdots) = 0. \tag{2.6}$$

If $A_i = 0$ for $0 < i < k$, but $A_k \neq 0$ then from (2.6)

$$t^{2m+1}(\operatorname{tr} A_{2m+1}) + \cdots + t^{2k}(\det A_k) + \cdots = 0.$$

Since $\operatorname{tr} A_{2m+1} \neq 0$, there must be a nonzero term in $\det(t^k A_k + \cdots)$ to cancel this coefficient, hence $2m+1 > 2k$. This means that the coefficient of t^{2k}, i.e., $\det A_k$, must vanish, since it is now the lowest order term. On the other hand, since $k < 2k < 2m+1$, we also have from (2.5) that $\operatorname{tr} A_k = 0$.

Thus A_k is nonzero but $\det A_k = \operatorname{tr} A_k = 0$, so A_k is conjugate to a matrix $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. If A_0 is *not* the identity, then since it has equal eigenvalues $+1$, it must be conjugate to a matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Thus in either case the first nonscalar coefficient in $A_0 + tA_1 + t^2 A_2 + \cdots$ always has a unique eigenspace.

Now the holonomy matrix $\tilde{H}(t) = \sum_{i=0}^{\infty} B_i t^i$ *commutes* with $H(t)$, so B_0 commutes with the first nonscalar coefficient of $H(t)$. Thus B_0 must be of the form

$$B_0 = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$$

or, since $\det B_0 = 1$,

$$B_0 = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}. \tag{2.7}$$

Hence A_0 and B_0 have a common eigenspace with eigenvalue 1 for A_0 and ± 1 for B_0. Taking into account the omitted value $\operatorname{tr} A = -2$, we have the possible eigenvalues ± 1 for A_0 and B_0.

We pass now to the interpretation in terms of the connection $\nabla_A + \zeta_0^{-1}\Phi - \zeta_0\Phi^*$ on the vector bundle V. If we tensor V with a suitably chosen flat line bundle of order 2 to remove the sign ambiguities in A_0 and B_0, then we have a flat $\mathrm{SL}(2, \mathbb{C})$ connection such that the holonomy matrices have a common eigenspace with eigenvalue $+1$. This gives a globally defined covariant constant section s of V, and hence a solution to the two equations

$$\begin{aligned} d'_A s + \zeta_0^{-1}\Phi s &= 0, \\ d''_A s - \zeta_0 \Phi^* s &= 0. \end{aligned} \tag{2.8}$$

Now suppose that $d'_A s + \zeta^{-1}\Phi s = 0$ has a nonzero solution for *all* $\zeta \in \mathbb{C}^*$. It does in particular then for $\zeta = e^{i\theta}$. We consider the flat *unitary* connection

$$\nabla_B = \nabla_A + e^{-i\theta}\Phi - e^{i\theta}\Phi^*,$$

so that $d'_B s = 0$. Since $F_B = d'_B d''_B + d''_B d'_B = 0$, we have $d'_B d''_B s = 0$ and hence using the hermitian inner products,

$$(d''_B s, d''_B s) = d'(d''_B s, s). \tag{2.9}$$

Integrating over M and using Stokes' theorem we deduce (this is the usual Weitzenböck vanishing argument) that $d''_B s = 0$ and so s is covariant constant.

Now since s is covariant constant, the $\mathrm{SU}(2)$ holonomy leaves fixed a vector and is then trivial. We thus see that if $d'_A + \zeta^{-1}\Phi$ has a nontrivial kernel for *all* $|\zeta| = 1$, then the holonomy of $\nabla_A + e^{-i\theta}\Phi - e^{i\theta}\Phi^*$ is trivial for all θ, so $H(\zeta) = \tilde{H}(\zeta) = 1$ on the unit circle. By holomorphicity, $H(\zeta) = \tilde{H}(\zeta) = 1$ for all $\zeta \in \mathbb{C}^*$. In this case $h(\zeta) \equiv 2$ and there are no odd order zeros. We assume, then, that the elliptic operator $d'_A + \zeta^{-1}\Phi$ is generically invertible. It depends holomorphically on ζ for $\zeta \in \mathbb{C}^* \cup \infty$, so there are only a *finite* number of values of ζ with $|\zeta| \geq 1$ for which $(d'_A + \zeta^{-1}\Phi)s = 0$ has a solution.

Similarly, applying an analogous vanishing theorem, the elliptic operator $d''_A - \zeta\Phi^*$ which is holomorphic for $|\zeta| \leq 1$ has only a finite number of values for $|\zeta| \leq 1$ for which it is noninvertible.

Thus if (2.8) is to hold, it can do so for only a finite number of values ζ_0, proving the proposition.

Proposition (2.3) concerned the holonomy around *one* generator of the fundamental group. This gave a finite number of points in $\mathbb{C}^*$ which were branch points of $\sqrt{h(\zeta)^2 - 4}$. In fact, consideration of the other generator gives the same branch points:

Proposition (2.10). *Let (A, Φ) be a solution to (1.7) on a torus M, and let $h(\zeta)$, $\tilde{h}(\zeta)$ be the trace of the holonomy of the connection (2.1) around the two generators of $\pi_1(M) = \mathbb{Z} \oplus \mathbb{Z}$. Then the odd zeros of $h^2(\zeta) - 4$ are also odd zeros of $\tilde{h}^2(\zeta) - 4$ unless $\tilde{H}(\zeta) \equiv \pm 1$.*

Proof. From the proof of Proposition (2.3), formula (2.7) shows that, at an odd zero of $h^2(\zeta) - 4$, $\tilde{h}^2(\zeta) - 4$ also vanishes. If $\tilde{H}(\zeta) \not\equiv \pm 1$ it vanishes with finite order at $\zeta = \zeta_0$. Then putting $t = \zeta - \zeta_0$, suppose the order is $2n$, so that

$$\tilde{h}^2 - 4 = t^{2n}\tilde{g}(t), \quad \text{where } \tilde{g}(0) \neq 0,$$

and the eigenvalue $\tilde{\mu}$ satisfies $\tilde{\mu} = \tilde{h} \pm t^n\sqrt{\tilde{g}(t)}/2$. This has a power series expansion in t.

Let K denote the field of fractions of convergent power series. Then $\tilde{H}(t)$ is a 2×2 matrix with both entries and distinct eigenvalues in K, and therefore has a basis of eigenvectors in K.

Now $H(t)$ has entries in the same field and commutes with $\tilde{H}(t)$. It therefore preserves the eigenvectors of $\tilde{H}(t)$, and they are consequently eigenvectors of $H(t)$. Thus $H(t)$ has eigenvalues in K.

However ζ_0 is an odd zero of $h^2 - 4$, so $h^2 - 4 = t^{2m+1}g(t)$, and $\mu = h \pm t^{m+1/2}\sqrt{g(t)}/2$, which does not lie in K.

Thus $\tilde{h}^2 - 4$ must have an odd zero too. q.e.d.

We must next examine the behavior of $h(\zeta)$ as ζ tends to zero or infinity: to invoke algebraic geometry by compactifying $\mathbb{C}^*$ to the projective line $\mathbb{C}P^1$.

3. Limiting behavior of the holonomy

We consider in this section the behavior of the holonomy of the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ as $\zeta \to 0$ or ∞. In fact we need only consider one limit, as follows. The $\mathrm{SU}(2)$ structure on the bundle V gives it a *quaternionic* structure $j: V \to V$, an antilinear map such that $j^2 = -1$. Since the connection A preserves this structure we have $j^{-1}\nabla_A j = \nabla_A$, and since Φ takes values in the complexification of the Lie algebra of SU(2) we have $j^{-1}\Phi j = -\Phi^*$. Thus

$$j^{-1}(\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*)j = \nabla_A - \overline{\zeta}^{-1}\Phi^* + \overline{\zeta}\Phi,$$

and consequently the holonomy matrix satisfies

$$j^{-1}H(\zeta)j = H(\overline{\zeta}^{-1}), \tag{3.1}$$

or equivalently $H(\overline{\zeta}^{-1})^* = H(\zeta)^{-1}$. The trace of the holonomy $h(\zeta)$ thus satisfies the reality condition $h(\overline{\zeta}^{-1}) = \overline{h(\zeta)}$ which determines the behavior at ∞ in terms of the behavior at 0. Consider then the situation as $\zeta \to 0$.

To introduce the method, we begin by considering the exceptional case where $H(\zeta) \equiv 1$ and $\tilde{H}(\zeta) \equiv 1$ (after tensoring by a flat $\mathbb{Z}_2$-bundle).

Proposition (3.2). *Suppose that* (A, Φ) *is a solution to* (1.7). *Then the holonomy of the flat connection* $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ *is trivial for all* $\zeta \in \mathbb{C}^*$ *if and only if* (A, Φ) *defines a conformal map of* M *to* S^2.

Proof. We make use of the $(0, 1)$ part of the connection $d''_A - \zeta\Phi^*$ which depends holomorphically on ζ and at $\zeta = 0$ gives the d''_A-operator. It defines a holomorphic structure V_ζ on V, depending holomorphically on $\zeta \in \mathbb{C}$.

The flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ is trivial so we have for each $\zeta \in \mathbb{C}^*$ two linearly independent covariant constant sections v_1 and v_2 of V. These are in particular holomorphic with respect to $d''_A - \zeta\Phi^*$ and holomorphically trivialize V_ζ. Thus any holomorphic section v of V_ζ is a constant linear combination of v_1 and v_2 and so is also *covariant constant.*

The holomorphic structures V_ζ define a holomorphic bundle $\tilde{V}$ on $M \times \mathbb{C}$, and the direct image sheaf $p_*\tilde{V}$ by projection onto $\mathbb{C}$ is a locally free sheaf—a holomorphic vector bundle W. The fiber W for $\zeta \neq 0$ is simply $H^0(M; V_\zeta) \cong \mathbb{C}^2$ and at $\zeta = 0$ is a 2-dimensional subspace of $H^0(M; V_0)$.

Over some neighborhood of 0 in $\mathbb{C}$ we take a holomorphic section s of W, with $s(0) = s_0 \neq 0$. For each ζ, $s(\zeta)$ is a section of V, holomorphic with respect to $d''_A - \zeta\Phi^*$, but also as noted above, *covariant constant* with respect to $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ for $\zeta \neq 0$.

Writing $s(z, \zeta) = s_0(z) + \zeta s_1(z) + \zeta^2 s_2(z, \zeta)$, with $s_2(z, \zeta)$ holomorphic in ζ, we therefore have

$$(d'_A + \zeta^{-1}\Phi)(s_0 + \zeta s_1 + \zeta^2 s_2) = 0 \quad \text{for } \zeta \in \mathbb{C}^*,$$

and hence

$$\Phi s_0 = 0, \tag{3.3}$$

$$d'_A s_0 + \Phi s_1 = 0. \tag{3.4}$$

From (3.3) $\det\Phi = 0$ and since $\Phi^2 = 0$, Φs_1 lies in the kernel of Φ. Thus from (3.4) d'_A and d''_A preserve the kernel of Φ, so the connection A is reducible. This is case (d) of §1, and thus defines a holomorphic section of a flat S^2-bundle. Since however the holonomy at $\zeta = \pm 1$ is trivial, it is an actual map to S^2. Conversely the connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ on M arising from a conformal map to S^2 is the pull-back of the corresponding connection on S^2 which clearly has trivial holonomy since S^2 is simply connected. q.e.d.

We turn then instead to the case of nontrivial holonomy which must be of type (a), (b), (c) and (d) which we treat now in turn.

Proposition (3.5). *Let* (A, Φ) *be a solution to* (1.7) *with* $\det\Phi = -\lambda^2 dz^2 \neq 0$ *and* A *irreducible, and let* $\mu(\zeta)$ *and* $\tilde\mu(\zeta)$ *be eigenvalues of the holonomy of the flat connection* $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ *around the generators of* $\pi_1(M)$. *Then there is a punctured neighborhood of* $0 \in \mathbb{C}$ *in which*

$$\pm\log\mu(\zeta) = \lambda\zeta^{-1} + a + \zeta b(\zeta),$$
$$\pm\log\tilde\mu(\zeta) = \lambda\tau\zeta^{-1} + \tilde a + \zeta\tilde b(\zeta),$$

where $b(\zeta)$ *and* $b(\tilde\zeta)$ *are holomorphic.*

Proof. We consider as above the holomorphic bundle $\tilde V$ on $M \times \mathbb{C}$ given by the holomorphic structure $d''_A - \zeta\Phi^*$. We now know also that the holonomy is not identically trivial. From Proposition (2.3) we can deduce that there are no odd order zeros of $h(\zeta)^2 - 4$ in some punctured neighborhood $\dot U$ of 0 in $\mathbb{C}$. This means (as in the proof of Proposition (2.10)) that we can locally define holomorphic eigenvalues and eigenspaces of the holonomy matrices, and therefore write

$$V_\zeta = L_\zeta \oplus L^*_\zeta, \tag{3.6}$$

where L_ζ and L^*_ζ are the common eigenspaces of the two holonomy matrices. (Note that this decomposition may not hold globally on $\dot U$ because the two eigenspaces may interchange as ζ passes around 0.)

We may also assume, from the proof of Proposition (2.3), that the operator $d''_A - \zeta\Phi^*$ on V_ζ or $V_\zeta \otimes \tilde L$, where $\tilde L$ is a line bundle of order 2, is invertible for $\zeta \in \dot U$. Since the bundle L_ζ has a flat connection defined on it, it is necessarily of degree zero, hence the invertibility of this operator implies that L_ζ^2 is not holomorphically trivial.

Now consider the bundle $\operatorname{End}_0 V_\zeta$ of trace free endomorphisms. From (3.6) this has a covariant constant decomposition

$$\operatorname{End}_0 V_\zeta \cong L_\zeta^2 \oplus 1 \oplus L_\zeta^{-2}$$

with respect to the connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$.

The covariant constant section of $\operatorname{End}_0 V_\zeta$ which splits V_ζ as a direct sum is holomorphic, and since L_ζ^2 is holomorphically nontrivial it generates the 1-dimensional space $H^0(M; \operatorname{End}_0 V_\zeta)$.

We now have the information that the eigenspaces of the holonomy are determined by the *holomorphic structure* of V_ζ alone, and so we proceed as before using the regularity of the holomorphic structure on V_ζ as $\zeta \to 0$.

Take the bundle $\tilde{V}$ on $M \times U$ and consider the direct image sheaf of $\operatorname{End}_0 \tilde{V}$. Since for $\zeta \neq 0$, $\dim H^0(M; \operatorname{End}_0 V_\zeta) = 1$, this is a holomorphic line bundle on U and we take a local holomorphic section

$$\psi(z, \zeta) = \psi_0(z) + \zeta\psi_1(z) + \zeta^2\psi_2(z, \zeta)$$

with ψ_2 holomorphic.

For $\zeta \neq 0$, this is a covariant constant, hence

$$(d'_A + \zeta^{-1}\Phi)(\psi_0 + \zeta\psi_1 + \zeta^2\psi_2) = 0,$$

which gives

$$[\Phi, \psi_0] = 0, \tag{3.7}$$

$$d'_A\psi_0 + [\Phi, \psi_1] = 0. \tag{3.8}$$

Now let z be a uniformizing parameter on the torus M, so that dz is a holomorphic 1-form. We may write $\Phi = \phi\, dz$ where, since $\det\Phi \neq 0$, the holomorphic endomorphism ϕ has eigenvalues $\pm\lambda$, and $V = L \oplus L^*$, the eigenspaces of ϕ.

From (3.7), ψ_0 is a constant multiple of ϕ which by scaling the section we can take as equal to ϕ. Now

$$\det\psi = \det(\phi + \zeta\psi_1 + \zeta^2\psi_2) = -\lambda^2 - \zeta \operatorname{tr}\phi\psi_1 + \cdots,$$

and since $\lambda \neq 0$, if ζ is in some neighborhood of 0, the eigenvalues of ψ are holomorphic in ζ. The corresponding eigenspaces of ψ are L_ζ and L_ζ^*, and so these extend holomorphically as $\zeta \to 0$ to the eigenspaces L and L^* of ϕ. The line bundle L_ζ therefore extends over $\zeta = 0$.

Let $s(z, \zeta) = s_0(z) + \zeta s_1(z) + \zeta^2 s_2(z, \zeta)$ be any local nonvanishing holomorphic section of this line bundle. Then since L_ζ is preserved by the connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$,

$$(d'_A + \zeta^{-1}\Phi)(s_0 + \zeta s_1 + \zeta^2 s_2) = \theta(s_0 + \zeta s_1 + \zeta^2 s_2)$$

for some $(1, 0)$ form θ holomorphic in z—the connection form of the flat connection on L_ζ.

Setting $\theta = \zeta^{-1}\lambda\, dz + \alpha$, we obtain

$$\alpha(s_0 + \zeta s_1 + \zeta^2 s_2) = d'_A s_0 + \Phi s_1 - \lambda s_1\, dz + \zeta d'_A s_1 + \cdots .$$

Thus α is *holomorphic* in $\zeta \in \mathbb{C}$.

At $\zeta = 0$, we have $\alpha = \alpha_0$, where

$$\alpha_0 s_0 = d'_A s_0 + (\phi - \lambda) s_1\, dz.$$

Note that $(\phi-\lambda)s_1$ lies in the $-\lambda$ eigenspace of ϕ, thus α_0 is a connection form for the connection on L (where $V = L \oplus L^*$) obtained by projecting (with respect to this eigenspace decomposition) the connection A into L. Since α_0 is holomorphic, the connection is flat.

We have succeeded, then, in finding a connection form θ for L_ζ. We represent now the torus M by $\mathbb{C}/\Gamma$ where Γ is the lattice generated by $z = 1$ and $z = \tau$ (τ in the upper half-plane), and use the line segments $[0, 1]$ and $[0, \tau]$ to give generators of the fundamental group. Integrating the form

$$\theta = \zeta^{-1}\lambda\, dz + \alpha_0 + \zeta\alpha_1 + \cdots$$

over these segments we obtain the proposition.

Proposition (3.9). *Let* (A, Φ) *be a solution to* (1.7) *with* $\det\Phi = -\lambda^2 dz^2 \neq 0$ *and* A *reducible. Then there is a punctured neighborhood of* $0 \in \mathbb{C}$ *in which the eigenvalues* $\mu(\zeta)$, *and* $\tilde{\mu}(\zeta)$ *of the holonomy satisfy*

$$\pm\log\mu(\zeta) = \lambda\zeta^{-1} + ik\pi + \zeta b(\zeta),$$
$$\pm\log\tilde{\mu}(\zeta) = \lambda\tau\zeta^{-1} + i\tilde{k}\pi + \zeta\tilde{b}(\zeta) \qquad (k, \tilde{k} \in \mathbb{Z}),$$

where $b(\zeta)$ *and* $\tilde{b}(\zeta)$ *are even holomorphic functions.*

Proof. This is a special case of the previous argument, and so we only need to investigate the extra information which we have from the reducibility of A. By (1.9) we see that there exists a gauge transformation g with $g^2 = -1$ such that g preserves A, and $g^{-1}\Phi g = -\Phi$.

Applying g to the flat connection we obtain

$$g^{-1}(\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*)g = \nabla_A - \zeta^{-1}\Phi + \zeta\Phi^*.$$

Thus the eigenvalues of the holonomy at ζ are the same as at $-\zeta$.

From (3.5) we have

$$\log \mu(\zeta) = \lambda\zeta^{-1} + a + \zeta b(\zeta),$$
$$\log \mu(-\zeta) = -\lambda\zeta^{-1} + a - \zeta b(-\zeta).$$

We must therefore have $\mu(\zeta)^{-1} = \mu(-\zeta)$, and so $a = ik\pi$ for $k \in \mathbb{Z}$ and $b(\zeta)$ is even.

Thus in this case the flat projected connection on L has holonomy group $\mathbb{Z}_2$.

Proposition (3.10). *Let (A, Φ) be a solution to (1.7) with $\det\Phi = 0$ and A irreducible, and let $c = -\kappa^2 dz^2$ be the quadratic differential invariant. Then there is a punctured neighborhood of $0 \in \mathbb{C}$ in which the eigenvalues $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ of the holonomy satisfy*

$$\pm\log \mu(\zeta) = \kappa\zeta^{-1/2} + ik\pi + \zeta^{1/2} b(\zeta^{1/2}),$$
$$\pm\log \tilde{\mu}(\zeta) = \kappa\tau\zeta^{-1/2} + i\tilde{k}\pi + \zeta^{1/2}\tilde{b}(\zeta^{1/2}) \qquad (k, \tilde{k} \in \mathbb{Z}).$$

Proof. Here we follow the argument of Proposition (3.5) up to (3.7) and (3.8), for until then we made no use of the determinant of Φ. We have a section ψ of $\operatorname{End}_0 V_\zeta$ over $M \times U$ where

$$\psi = \psi_0 + \zeta\psi_1 + \zeta^2\psi_2.$$

From (3.7), since $\Phi = \phi\, dz$ is nonzero, ψ_0 can again be taken to be ϕ. Now, however

$$\det\psi = \det\phi - \zeta\operatorname{tr}(\phi\psi_1) + \cdots = -\zeta\operatorname{tr}(\phi\psi_1) + \cdots,$$

and we cannot find eigenvalues of ψ holomorphic in ζ if $\operatorname{tr}(\phi\psi_1)$ is nonzero. In fact this coefficient is always nonzero in this case, as follows.

If we choose a local holomorphic trivialization of V so that $\phi = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, then the connection A has connection matrix

$$\alpha\, dz = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} dz.$$

From (3.8),

$$d'_A\phi = -[\alpha, \phi]\, dz = -[\phi, \psi_1]\, dz.$$

Thus $\alpha = -\psi_1 + k\phi$ and $-\operatorname{tr}(\phi\psi_1) = \operatorname{tr}(\phi\alpha) = c$, the lower off-diagonal term in the above connection matrix. This measures the failure of the connection A to preserve the line bundle $L = \ker\phi$. We have here tacitly trivialized L, and invariantly speaking, $\operatorname{tr}(\phi\alpha)$ is a quadratic differential: it is the *second fundamental form* c mentioned in §1. In particular it is nonzero if A is irreducible, so we have $\det\psi = c\zeta + \cdots$ for a nonzero constant $c = -\kappa^2$.

Now set $\eta^2 = \zeta$, and we have

$$\det\psi = -\kappa^2\eta^2(1 + a_1\eta^2 + \cdots). \tag{3.11}$$

Thus ψ has well defined holomorphically varying eigenvalues $\pm\kappa\eta(1 + \frac{1}{2}a_1\eta^2 + \cdots)$ on an open set $\hat{U} \subset \mathbb{C}$ which double covers $U \subset \mathbb{C}$. Consequently there is a well-defined holomorphic line bundle L_η over $M \times U$ which for $\eta \neq 0$ is an eigenspace for the holonomy of the flat connection $\nabla_A + \eta^{-2}\Phi - \eta^2\Phi^*$.

Let $s(z, \eta) = s_0(z) + \eta s_1(z) + \eta^2 s_2(z, \eta)$ be a local nonvanishing holomorphic section. Then since s is an eigenvector of ψ,

$$(\phi + \eta^2\psi_1 + \eta^4\psi_2)(s_0 + \eta s_1 + \cdots) = \kappa\eta(1 + \tfrac{1}{2}a_1\eta^2 \cdots)(s_0 + \eta s_1 + \cdots).$$

Thus

$$\phi s_1 = \kappa s_0. \tag{3.12}$$

Since for $\eta \neq 0$, the line bundle is preserved by the connection and we also have

$$(d'_A + \eta^{-2}\Phi)(s_0 + \eta s_1 + \cdots) = \theta(s_0 + \eta s_1 + \cdot) \tag{3.13}$$

for some $(1, 0)$ form θ. Writing

$$\theta = \frac{a_{-2}}{\eta^2} + \frac{a_{-1}}{\eta} + a_0 + \cdots ,$$

we have $a_{-2}s_0 = \Phi s_0$. But $\det\Phi = 0$, so the only eigenvalue of Φ is zero, hence $a_{-2} = 0$.

We have, further, $a_{-1}s_0 + \Phi s_1 = 0$. But then from (3.12) $a_{-1} = -\kappa\, dz$ and so

$$\theta = -\frac{\kappa\, dz}{\eta} + a_0 + \cdots .$$

Integrating over the segments giving the generators of $\pi_1(M)$, we have

$$\pm\log\mu(\eta) = -\frac{\kappa}{\eta} + a + \cdots , \qquad \pm\log\tilde{\mu}(\eta) = -\frac{\kappa\tau}{\eta} + \tilde{a} + \cdots .$$

Now since the flat connection $\nabla_A + \eta^{-2}\Phi - \eta^2\Phi^*$ depends only on η^2, we must have

$$\log\mu(-\eta) = \pm\log\mu(\eta) + 2k\pi i,$$

i.e., $a = -a \mod 2\pi i\mathbb{Z}$, so a, $\tilde{a} \in \pi i\mathbb{Z}$.

Proposition (3.14). *Let* (A, Φ) *be a solution to* (1.7) *with* $\det \Phi = 0$ *and* A *reducible. Then the eigenvalues* $\mu(\zeta)$ *and* $\tilde{\mu}(\zeta)$ *of the holonomy are constant.*

Proof. We follow the proof of (3.10) but now, from case (d) of §1, the connection matrix A has the form

$$\alpha\, dz = \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} dz,$$

and hence $\det \psi = c\zeta^n + \cdots$ for some nonzero constant and $n > 1$. Repeating the argument, we have

$$(\phi + \eta^2 \psi_1 + \eta^4 \psi_2)(s_0 + \eta s_1 + \cdots) = \kappa \eta^n (1 + \tfrac{1}{2} a_1 \eta^2 \cdots)(s_0 + \eta s_1 + \cdots),$$

and so

$$\Phi s_0 = 0 \quad \text{and} \quad \Phi s_1 = 0. \tag{3.15}$$

Considering (3.13), we have

$$(d'_A + \eta^{-2}\Phi)(s_0 + \eta s_1 + \cdots) = \theta(s_0 + \eta s_1 + \cdots),$$

but then from (3.15), θ is holomorphic in η, and consequently so are $\mu(\eta)$ and $\tilde{\mu}(\eta)$. Hence $h(\zeta) = \mu + \mu^{-1}$ is a holomorphic function on $\mathbb{C}P^1$ and so constant. Therefore $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ are constant.

Remark (3.16). Note that, as a consequence of these propositions, if $\mu(\zeta) \equiv \pm 1$, then so is $\tilde{\mu}(\zeta)$ and we are reduced to the situation of $H(\zeta) \equiv \pm 1$ and $\tilde{H}(\zeta) \equiv \pm 1$, the trivial case of a conformal map to the 2-sphere. In particular, in Proposition (2.10) either $h^2(\zeta) - 4$ and $\tilde{h}^2(\zeta) - 4$ are *both* identically zero or *both* have exactly the same odd zeros.

4. The hyperelliptic curve

We now have enough information to relate a solution of (1.7), and in particular a harmonic map, to an algebraic geometric object—a *hyperelliptic curve.* In §2 we studied the eigenvalues $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ of the holonomy matrices of the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ and showed they had a finite number of branch points in $\mathbb{C}^*$. In §3 we saw that if $\det \Phi \neq 0$, then $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ are single-valued in a punctured neighborhood of 0 (and ∞), and if $\det \Phi = 0$, then they are single-valued in a double covering of such a neighborhood. We begin by defining a smooth hyperelliptic curve associated to this data.

Definition (4.1). If (A, Φ) is a solution to (1.7) on a torus M, and $\alpha_1, \cdots, \alpha_n, \overline{\alpha}_1^{-1}, \cdots, \overline{\alpha}_n^{-1}$ are the odd order zeros of $h(\zeta)^2 - 4$ as in §2, then we define the associated *hyperelliptic curve* $\hat{\Sigma}$ to be:

(i) if $\det\Phi \neq 0$, the double covering of $\mathbb{C}P^1$ branched over $\alpha_1, \cdots, \alpha_n, \overline{\alpha}_1^{-1}, \cdots, \overline{\alpha}_n^{-1}$,

(ii) if $\det\Phi = 0$, the double covering of $\mathbb{C}P^1$ branched over $\alpha_1, \cdots, \alpha_n, \overline{\alpha}_1^{-1}, \cdots, \overline{\alpha}_n^{-1}$, 0 and ∞.

We denote the covering map by $\pi: \hat{\Sigma} \to \mathbb{C}P^1$.

By the very definition of the curve, the functions $\mu(\zeta)$ and $\tilde{\mu}(\zeta)$ are single-valued on $\hat{\Sigma} \setminus \pi^{-1}\{0, \infty\}$. These are the *eigenvalues* of the holonomy matrices $H(\zeta)$ and $\tilde{H}(\zeta)$. We consider next the *eigenspaces* of the holonomy.

Fix a point $x \in M$, and consider the holonomy matrix $H_x(\zeta)$ obtained by parallel translation of the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ around closed curves through x in the homotopy class of the corresponding generator of $\pi_1(M)$. On $\hat{\Sigma} \setminus \pi^{-1}\{0, \infty\}$ we define a holomorphic vector bundle, the *eigenspace bundle*, by the property

$$E_x(\xi) \subseteq \ker(H_x(\xi) - \mu(\xi)) \quad \text{for } \xi \in \hat{\Sigma}.$$

If the holonomy varies nontrivially (i.e., excluding the case $\det\Phi = 0$, A reducible), this is a line bundle on $\hat{\Sigma} \setminus \pi^{-1}\{0, \infty\}$. Moreover, by the analysis of the limiting behavior as $\zeta \to 0$ in §3, the bundle actually extends over $\zeta = 0$ to become an eigenspace L_x of $\Phi_x: V_x \to V_x \otimes K_x$. We therefore obtain a holomorphic line bundle E_x defined on the whole curve $\hat{\Sigma}$. Since $H_x(\zeta)$ and $\tilde{H}_x(\zeta)$ commute, the bundle E_x could equally be defined by considering holonomy around the other generator.

The eigenvalue $\mu(\zeta)$ has an essential singularity at $\pi^{-1}\{0, \infty\}$ according to Propositions (3.5) and (3.10), but from those results the differential form $\theta = d\log\mu = d\mu/\mu$ is a meromorphic form with a double pole at $\pi^{-1}(0)$ and $\pi^{-1}(\infty)$ and zero residue—a *differential of the second kind*—and hence an algebraic object.

We now prove some basic properties of the hyperelliptic curve:

Proposition (4.2). *Let $\hat{\Sigma}$ be the hyperelliptic curve associated to a solution (A, Φ) of (1.7) on a torus M. Then:*

(i) *$\hat{\Sigma}$ has a real structure (an antiholomorphic involution) $\rho: \hat{\Sigma} \to \hat{\Sigma}$ which commutes with π and induces the real structure $\zeta \to \overline{\zeta}^{-1}$ on $\mathbb{C}P^1$,*

(ii) *the hyperelliptic involution $\sigma: \hat{\Sigma} \to \hat{\Sigma}$ which interchanges the branches of the covering commutes with π and has no real fixed points,*

(iii) *the differentials θ and $\tilde{\theta}$ satisfy*

$$\sigma^*\theta = -\theta; \quad \sigma^*\tilde{\theta} = -\tilde{\theta}; \qquad \rho^*\theta = -\overline{\theta}; \quad \rho^*\tilde{\theta} = -\overline{\tilde{\theta}},$$

(iv) *there is an isomorphism* j *between the bundles* $(\sigma\rho)^*\overline{E}_x$ *and* E_x, *such that* $j^2 = \mathrm{id}$,

(v) *each period of* θ *and* $\tilde{\theta}$ *is of the form* $2in\pi$ *for some* $n \in \mathbb{Z}$.

Proof. (i) The abstract definition of a branched double covering may be given a more concrete realization which helps to define the curve and its associated structures. If we put

$$p(\zeta) = (-1)^n \prod_{i=1}^{n} (\overline{\alpha}_i\zeta^2 - (\alpha_i\overline{\alpha}_i + 1)\zeta + \alpha_i),$$

if $\det\Phi \neq 0$, $l = n$, and

$$p(\zeta) = (-1)^n \zeta \prod_{i=1}^{n} (\overline{\alpha}_i\zeta^2 - (\alpha_i\overline{\alpha}_i + 1)\zeta + \alpha_i),$$

if $\det\Phi = 0$ and $l = n + 1$, then

$$p(\overline{\zeta}^{-1}) = \overline{\zeta}^{-2l}\overline{p(\zeta)}, \tag{4.3}$$

so p may be considered as a section of the line bundle $\mathcal{O}(2l)$ over $\mathbb{C}P^1$, real with respect to the real structure $\zeta \to \overline{\zeta}^{-1}$.

Consider the total space of $\mathcal{O}(l)$ and the tautological section η of $\mathcal{O}(l)$ over it. The hyperelliptic curve $\hat{\Sigma}$ is then the zero set of the section $\eta^2 - p$ of $\mathcal{O}(2l)$ over $\mathcal{O}(l)$. The hyperelliptic involution σ is thus just fiberwise multiplication by -1 in $\mathcal{O}(l)$, and the projection π simply the restriction of the projection of the line bundle $\mathcal{O}(l)$. In terms of the affine coordinate ζ on $\mathbb{C}P^1$, the equation is $\eta^2 = p(\zeta)$ and $\sigma(\eta, \zeta) = (-\eta, \zeta)$. The real structure on $\mathbb{C}P^1$ extends to $\mathcal{O}(l)$ for any l by

$$\rho(\eta, \zeta) = (\overline{\eta\zeta^{-l}}, \overline{\zeta}^{-1});$$

this clearly commutes with σ and π, and from (4.3) takes the curve into itself.

(ii) The fixed points of σ are given by $\eta = 0$, i.e., the zeros of $p(\zeta)$, the branch points of the covering. From (2.7) there are no odd order zeros of $h^2 - 4$ on the circle, so $\alpha_i \neq \overline{\alpha}_i^{-1}$, hence there are no real fixed points of σ.

(iii) From (3.1), $\mu(\overline{\zeta}^{-1}) = \overline{\mu}(\zeta)^{\pm 1}$. However, the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ is unitary for $\zeta = e^{i\theta}$, so

$$\mu(e^{i\theta}) = \overline{\mu}^{-1}(e^{i\theta}),$$

hence $\mu(\overline{\zeta}^{-1}) = \overline{\mu}(\zeta)^{-1}$ for all ζ, so that $\rho^*\mu = \overline{\mu}^{-1}$.

Thus $\rho^*\theta = \rho^* d\mu/\mu = -d\overline{\mu}/\overline{\mu} = -\overline{\theta}$ and similarly for $\tilde{\theta}$.

By the definition of the curve, $\sigma^*\mu = \mu^{-1}$, hence $\sigma^*\theta = \sigma^* d\mu/\mu = -d\mu/\mu = -\theta$ and similarly for $\tilde{\theta}$.

(iv) Consider the quaternionic structure $j\colon V_x \to V_x$ and let $E_x(\xi) \subset V_x$ be the eigenspace bundle. Thus if $v \in E_x(\xi)$, then $H(\zeta)v = \mu(\zeta)v$ and

$$j^{-1}H(\zeta)j\cdot(j^{-1}v) = j^{-1}\mu(\zeta)v = \overline{\mu(\zeta)}j^{-1}v,$$

so

$$H(\overline{\zeta}^{-1})j^{-1}v = \overline{\mu(\zeta)}j^{-1}v \quad \text{from (3.1)}$$

or

$$H(\overline{\zeta}^{-1})j^{-1}v = \mu(\overline{\zeta}^{-1})^{-1}j^{-1}v \quad \text{from (iii) above.}$$

This implies that if $v \in E_x(\xi)$, then $j^{-1}v \in E_x(\sigma\rho\xi)$, hence $j\colon (\sigma\rho)^*E_x \to E_x$ is an antilinear isomorphism satisfying $j^2 = -1$.

(v) From (3.7) and (3.10), θ and $\tilde{\theta}$ have zero residues at their poles; thus they have well-defined integrals over a basis for $H_1(\hat{\Sigma}; \mathbb{Z})$, their *periods*. On the other hand $\theta = d\log\mu$ and μ is a well-defined function on $\hat{\Sigma}\backslash\pi^{-1}\{0, \infty\}$. Thus the periods are differences of values of the logarithm and take values in $2\pi i\mathbb{Z}$. Similarly for $\tilde{\theta}$.

This final property imposes a severe constraint on the hyperelliptic curve.

5. The spectral curve

The hyperelliptic curve $\hat{\Sigma}$ of §4 will define the hyperelliptic functions by means of which we shall solve the harmonic map equations, but it will be more convenient to deal with another, possibly singular, curve Σ of which $\hat{\Sigma}$ is the normalization. There are two phenomena which demand the introduction of this curve. The first is the consideration of a continuous family of hyperelliptic curves corresponding to solutions of (1.7) on the torus. If within such a family two branch points coalesce, then we would expect a *singular* hyperelliptic curve to be the limit. Secondly, we have not determined the *degree* of the eigenspace bundle E_x. The introduction of the spectral curve Σ provides a natural answer to both problems.

Consider the eigenspace line bundle E_x on $\hat{\Sigma}$. Both E_x and σ^*E_x are subbundles of the trivial bundle $\hat{\Sigma}\times V_x$, and evaluation of the symplectic form $\langle\ ,\ \rangle$ on V defines a holomorphic section ω of the line bundle $E_x^*\otimes\sigma^*E_x^*$.

Now ω vanishes at those points of $\hat{\Sigma}$ at which $E_x \subset V_x$ and $\sigma^*E_x \subset V_x$ coincide. This certainly happens at the branch points, but we must also consider the other cases of coincidence, together with questions of multiplicity.

Suppose, then, that ω vanishes at a point ξ_i which is not a branch point. Let $\pi(\xi_i) = \beta_i \in \mathbb{C}P^1$ and suppose $\beta_i \neq 0, \infty$. Choose a local parameter t on $\mathbb{C}P^1$ such that $t = 0$ at β_i. Then the holonomy matrix $H(t)$ has a power series expansion in t. If ω vanishes to order m_i, then we may find local holomorphic trivializations of the two eigenspace line bundles in V_x of the form

$$v_1(t) = v_0(t) + t^{m_i}u(t), \qquad v_2(t) = v_0(t) - t^{m_i}u(t), \tag{5.1}$$

where $\omega(v_0(0), u(0)) \neq 0$.

Now $v_1(t)$ and $v_2(t)$ are local trivializations of the eigenspace bundle E_x on the two branches of $\hat{\Sigma}$ over this coordinate patch, but if we consider the singular curve $s^2 = t^{2m_i}$ in $\mathbb{C}^2$, then $v_0(t) + su(t) \in V_x$ is a nonvanishing function on the singular curve which agrees with $v_1(t)$ and $v_2(t)$ on the two branches $s = t^{m_i}$ and $s = -t^{m_i}$.

If we regard the hyperelliptic curve (as in the proof of (5.2)) as given by the equation $\eta^2 = p(\zeta)$ in the total space of the line bundle $\mathcal{O}(l)$, then we may also consider as above the eigenspace bundle as defined on the singular curve $\eta^2 = (\zeta - \beta_i)^{2m_i}p(\zeta)$ in the total space of $\mathcal{O}(l + m_i)$. Continuing with the zeros $\xi_1, \cdots, \xi_k$ of ω which are not branch points, we have the singular curve

$$\eta^2 = \prod_{i=1}^{k}(\zeta - \beta_i)^{2m_i}p(\zeta) \tag{5.2}$$

in the total space of $\mathcal{O}(l + \sum_{i=1}^{k} m_i)$.

The section ω may vanish also at the branch points with multiplicity greater than one, but it must necessarily be odd. Suppose over each branch point α_i (and then also at its conjugate $\overline{\alpha}_i^{-1}$) ω has multiplicity $2n_i + 1$. Then a similar argument shows that E_x is defined on the singular curve

$$\eta^2 = \prod_{i=1}^{l}(\zeta - \alpha_i)^{2n_i}(\zeta - \overline{\alpha}_i^{-1})^{2n_i}p(\zeta) \tag{5.3}$$

in the total space of $\mathcal{O}(l + \sum_{i=1}^{l} 2n_i)$.

Note that the analysis of §3 shows that $\zeta = 0, \infty$ does not contribute multiplicities to ω, and hence these points never produce singularities in this manner.

Putting (5.2) and (5.3) together we have:

Definition (5.4). If $\eta^2 = p(\zeta)$ is the hyperelliptic curve associated to a solution (A, Φ) of (1.7), and the section ω of the line bundle $E_x^* \otimes \sigma^* E_x^*$ vanishes with multiplicity $2n_i + 1$ at the branch points α_i and $\overline{\alpha}_i^{-1}$ and with multiplicity m_i over the remaining points $\beta_1, \cdots, \beta_k$, then the *spectral curve* Σ is the curve in the total space of $\mathcal{O}(l+2\sum_{i=1}^{l} n_i+\sum_{i=1}^{k} m_i)$ defined by the equation

$$\eta^2 = p(\zeta)q(\zeta)r(\zeta), \tag{5.5}$$

where $q(\zeta)$ is the section of $\mathcal{O}(4\sum_{i=1}^{l} n_i)$ which vanishes with multiplicity $2n_i$ at the branch points $\alpha_i, \overline{\alpha}_i^{-1}$ $(1 \le i \le l)$, and $r(\zeta)$ is the section of $\mathcal{O}(2\sum_{i=1}^{k} m_i)$ which vanishes with multiplicity $2m_i$ at $\pi(\beta_i)$.

Remarks. 1. Clearly we would expect generically that the only points where the eigenspace bundles coincide are the branch points, and there with multiplicity one. In this case the hyperelliptic curve and the spectral curve coincide.

2. Since the eigenspaces of the holonomy at different points $x, y \in M$ are related by parallel translation of the connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ for $\zeta \neq 0, \infty$, the spectral curve is in fact independent of the point x used in its definition.

The spectral curve Σ is defined by an equation of the form $\eta^2 = P(\zeta)$ in the total space of $\mathcal{O}(n)$ where P is a section of $\mathcal{O}(2n)$ on $\mathbb{C}P^1$ and η is the tautological section of $\mathcal{O}(n)$. Its *arithmetic genus* p is given by the adjunction formula in the surface $\mathcal{O}(n)$:

$$2p - 2 = KC + C^2 = -4 - 2n + 4n = -4 + 2n,$$

and hence $p = n - 1$.

Its *geometric genus* g is the genus of its normalisation, and applying the adjunction formula for $\hat{\Sigma}$ in $\mathcal{O}(l)$ we have $2g - 2 = -4 + 2l$ and hence $g = l - 1$.

In (5.4) we had

$$2n = l + \sum_{i=1}^{k} m_i + \sum_{i=1}^{l} 2n_i = \sum_{i=1}^{l}(2n_i + 1) + \sum_{i=1}^{k} m_i,$$

which is the sum of the multiplicities of the zeros of ω. Since ω is a section of the line bundle $E_x^* \otimes \sigma^* E_x^*$, we deduce that

$$\deg E_x^* = n = p + 1. \tag{5.6}$$

Moreover, as we have seen, E_x may be considered as a line bundle defined on the possibly singular curve Σ. Its degree therefore depends only on the arithmetic genus p of Σ.

Applying the holonomy matrix $H(t)$ to (5.1) we see that the eigenvalues $\mu_1(t)$ and $\mu_2(t) = \mu_1(t)^{-1}$ may be written in the form:

$$\mu_1(t) = \mu_0(t) + t^{m_i}\nu(t), \qquad \mu_2(t) = \mu_0(t) - t^{m_i}\nu(t), \tag{5.7}$$

and again $\mu(t)+s\nu(t)$ is a well-defined function on the singular curve $s^2 = t^{2m_i}$. It follows that μ is a well-defined function not only on $\hat{\Sigma}\setminus\pi^{-1}\{0, \infty\}$ but also on the spectral curve $\Sigma\setminus\pi^{-1}\{0, \infty\}$, and that the differentials θ and $\tilde{\theta}$ may also be considered to exist on Σ.

All the data of Proposition (4.2) can then be transferred to the spectral curve Σ together with the additional fact that $\deg E_x^* = p+1$ where p is the arithmetic genus of Σ.

Note that in a neighborhood of a point of the hyperelliptic curve $\hat{\Sigma}$ which maps to a singular point of the spectral curve Σ with multiplicity m,

$$\mu(t) = \pm(1 + t^m f(t))$$

and so $\theta = d\mu/\mu$ has a zero of order $(m-1)$. So also does $\tilde{\theta} = d\tilde{\mu}/\mu$, derived from the holonomy around the other generator. The differentials θ and $\tilde{\theta}$ are however linearly independent; for otherwise from the reality property of Proposition (4.2) there would be a real linear relationship between them, but considering the principal parts at $\zeta = 0$ as in Propositions (3.5) and (3.10) this would give a real linear relationship between 1 and τ. Since τ is in the upper half-plane, this is impossible.

This linear independence imposes a constraint on the singularities of the spectral curve. For example, suppose the geometric genus g of the spectral curve is zero. The hyperelliptic curve $\hat{\Sigma}$ is then given by $\eta^2 = p(\zeta)$ where $p(\zeta)$ is a section of $\mathcal{O}(2)$ (a quadratic polynomial) on $\mathbb{C}P^1$. If $\det\Phi \neq 0$, then 0 and ∞ are not zeros of $p(\zeta)$, and a differential with zero residue at 0 and ∞ and double poles there and satisfying $\sigma^*\alpha = -\alpha$ is of the form

$$\alpha = \frac{(a+b\zeta^3)}{\zeta^2}\frac{d\zeta}{\eta},$$

using (η, ζ) as affine coordinates.

Putting in the reality condition $\rho^*\alpha = -\overline{\alpha}$, we obtain $b = \overline{a}$ and hence, by using the principal parts of θ and $\tilde{\theta}$,

$$\theta = \frac{(\eta_0\lambda + \overline{\eta}_0\overline{\lambda}\zeta^3)}{\zeta^2}\frac{d\zeta}{\eta}, \qquad \tilde{\theta} = \frac{(\eta_0\lambda\tau + \overline{\eta}_0\overline{\lambda\tau}\zeta^3)}{\zeta^2}\frac{d\zeta}{\eta}.$$

These have common zeros only if $\tau = \overline{\tau}$, which is impossible for τ in the upper half-plane.

If $\det\Phi = 0$, then $\hat{\Sigma}$ is given by $\eta^2 = \zeta$, and the differentials with zero residue, double poles and satisfying $\sigma^*\alpha = -\alpha$ are of the form

$$\alpha = \frac{(a + b\zeta)}{\zeta}\frac{d\zeta}{\eta},$$

and clearly any two differentials of this form with common zeros are linearly dependent.

Thus if $g = 0$, θ and $\tilde{\theta}$ have no common zeros and so the spectral curve has no singularities. The hyperelliptic curves which yield soliton solutions of the KdV equations are all rational curves with double points, so we may say by contrast that there are no *pure soliton* solutions for the harmonic map equations of the torus to S^3.

6. An example—the Clifford torus

The simplest example of a harmonic map from a torus to S^3 is provided by the *Clifford torus*. This is an embedded minimal surface in S^3, in fact the only known embedded minimal torus [18]. It is an orbit of the maximal torus of SO(4) acting on S^3, of maximal area. Analytically, the harmonic map is defined by $g\colon S^1 \times S^1 \to \mathrm{SU}(2)$, where

$$g(e^{i\theta}, e^{i\phi}) = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{i\theta} & e^{i\phi} \\ -e^{-i\phi} & e^{-i\theta} \end{pmatrix}. \tag{6.1}$$

This is clearly a square torus—its conformal structure is defined by $\tau = i$ in the upper half-plane. From (6.1), we obtain

$$g^{-1}dg = \frac{i}{2}\begin{pmatrix} d\theta - d\phi & e^{-i(\theta-\phi)}(d\theta + d\phi) \\ e^{i(\theta-\phi)}(d\theta + d\phi) & -d\theta + d\phi \end{pmatrix}. \tag{6.2}$$

We are working here in the gauge provided by the left-invariant trivialization, so that the connection matrix of ∇_L is zero and from (1.4) of ∇_R is $g^{-1}dg$. Thus the connection matrix in this gauge for $\nabla_A = \frac{1}{2}(\nabla_L + \nabla_R)$ is $\frac{1}{2}g^{-1}dg$. The difference $\nabla_R - \nabla_L = g^{-1}dg = 2\phi$ according to the notation of §1. Thus

$$\Phi - \Phi^* = \phi = \tfrac{1}{2}g^{-1}dg, \tag{6.3}$$

$$i(\Phi + \Phi^*) = +\tfrac{1}{2}g^{-1}dg. \tag{6.4}$$

Now $z = \theta + i\phi$ is a uniformizing parameter on the torus and ${}^*dz = i\,dz$, so $*(d\theta + i\,d\phi) = i(d\theta + i\,d\phi)$ and $*d\theta = -d\phi$, $*d\phi = d\theta$. Hence, from (6.2) and (6.4),

$$i(\Phi + \Phi^*) = \frac{i}{4}\begin{pmatrix} -d\phi - d\theta & e^{-i(\theta-\phi)}(-d\phi + d\theta) \\ e^{i(\theta-\phi)}(-d\phi + d\theta) & d\phi + d\theta \end{pmatrix}, \tag{6.5}$$

and thus from (6.2), (6.3) and (6.4)

$$2\Phi = \frac{1}{4}\begin{pmatrix} -(1-i)d\theta-(1+i)d\phi & e^{-i(\theta-\phi)}((1+i)d\theta-(1-i)d\phi) \\ e^{i(\theta-\phi)}((1+i)d\theta-(1-i)d\phi) & (1-i)d\theta+(1+i)d\phi \end{pmatrix},$$

and the flat connection $\nabla_A+\zeta^{-1}\Phi-\zeta\Phi^*$ has connection matrix

$$\begin{aligned} A &= \tfrac{1}{2}g^{-1}dg+\zeta^{-1}\Phi-\zeta\Phi^* \\ &= \Phi-\Phi^*+\zeta^{-1}\Phi-\zeta\Phi^* \quad \text{from (6.3)} \\ &= (1+\zeta^{-1})\Phi-(1+\zeta)\Phi^*. \end{aligned} \tag{6.6}$$

Conjugating by the $\mathrm{SU}(2)$-valued function

$$h = \begin{pmatrix} e^{-i(\theta-\phi)/2} & 0 \\ 0 & e^{i(\theta-\phi)/2} \end{pmatrix}$$

gives the connection matrix $h^{-1}dh+h^{-1}Ah$, which is the constant matrix

$$\begin{aligned} &\begin{pmatrix} -\frac{i}{2}(d\theta-d\phi) & 0 \\ 0 & \frac{i}{2}(d\theta-d\phi) \end{pmatrix} \\ &\quad+\frac{1}{8}(1+\zeta^{-1})\begin{pmatrix} -(1-i)d\theta-(1+i)d\phi & (1+i)d\theta-(1-i)d\phi \\ (1+i)d\theta-(1-i)d\phi & (1-i)d\theta+(1+i)d\phi \end{pmatrix} \\ &\quad-\frac{1}{8}(1+\zeta)\begin{pmatrix} -(1+i)d\theta-(1-i)d\phi & (1-i)d\theta-(1+i)d\phi \\ (1-i)d\theta-(1+i)d\phi & (1+i)d\theta+(1-i)d\phi \end{pmatrix}. \end{aligned} \tag{6.7}$$

Note that $h(\theta, \phi)$ is a *two-valued* gauge transformation of the trivial bundle:

$$h(\theta+2\pi m, \phi+2\pi n) = (-1)^{m+n}. \tag{6.8}$$

Geometrically, this means that (6.7) is the connection matrix for a connection on the bundle obtained after tensoring with the flat $\mathbb{Z}_2$-bundle on M corresponding to the element of $H^1(M; \mathbb{Z}_2)$ given by (6.8).

Consider now the flat connection restricted to the first factor in $M = S^1\times S^1$—the circle with parameter θ. We have

$$\begin{aligned} \nabla &\equiv \frac{d}{d\theta}+\begin{pmatrix} -i/2 & 0 \\ 0 & i/2 \end{pmatrix}+\frac{1}{8}(1+\zeta^{-1})\begin{pmatrix} -(1-i) & (1+i) \\ (1+i) & (1-i) \end{pmatrix} \\ &\quad-\frac{1}{8}(1+\zeta)\begin{pmatrix} -(1+i) & (1-i) \\ (1-i) & (1+i) \end{pmatrix} = \frac{d}{d\theta}-B(\zeta), \end{aligned} \tag{6.9}$$

and since B has constant coefficients, the holonomy matrix is given by

$$H(\zeta) = \exp 2\pi B(\zeta), \tag{6.10}$$

and the eigenvalues by $\mu^{\pm 1}(\zeta) = \exp 2\pi\nu(\zeta)$, where $\pm\nu(\zeta)$ are the eigenvalues of $B(\zeta)$, so

$$\nu^2(\zeta) = -\det B(\zeta).$$

From (6.9)

$$\begin{aligned}\det B(\zeta) = & -(-\tfrac{i}{2} - \tfrac{1}{8}(1+\zeta^{-1})(1-i) + \tfrac{1}{8}(1+\zeta)(1+i))^2 \\ & -(\tfrac{1}{8}(1+\zeta^{-1})(1+i) - \tfrac{1}{8}(1+\zeta)(1-i))^2 \\ = & -(-\tfrac{i}{2} - \tfrac{1}{4}(1+\zeta^{-1})(1-i))(-\tfrac{i}{2} + \tfrac{1}{4}(1+\zeta)(1+i)),\end{aligned}$$

so

$$\nu^2(\zeta) = -\tfrac{i}{8}\zeta^{-1}(\zeta - i)^2. \tag{6.11}$$

Similarly, considering the holonomy around the S^1 factor with parameter ϕ, we obtain eigenvalues $\tilde{\mu}^{\pm 1}(\zeta) = \exp \pm\tilde{\nu}(\zeta)$, where

$$\tilde{\nu}^2(\zeta) = \tfrac{i}{8}\zeta^{-1}(\zeta + i)^2. \tag{6.12}$$

From (6.11) and (6.12) we see that the only branch points of ν and $\tilde{\nu}$ are at $\zeta = 0$ and $\zeta = \infty$, and hence $h(\zeta)^2 - 4 = (e^{\nu} - e^{-\nu})^2$ has no odd order zeros for $\zeta \in \mathbb{C}^*$. The Clifford torus is however conformally (actually *isometrically*) embedded in S^3, so $\det \Phi = 0$, and the nonsingular model $\hat{\Sigma}$ of the spectral curve is defined from §4 by taking the double covering of $\mathbb{C}P^1$ branched over 0 and ∞. This is another copy of the projective line with affine parameter $\eta^2 = \zeta$.

On $\hat{\Sigma}$ we have the meromorphic differential θ defined by $\theta = d\mu/\mu = 2\pi d\nu$. Putting $\zeta = \eta^2$, (6.11) gives

$$\theta = 2\pi\sqrt{\frac{-i}{8}}\, d\left(\eta - \frac{i}{\eta}\right) = 2\pi\sqrt{\frac{-i}{8}}\left(1 + \frac{i}{\eta^2}\right) d\eta, \tag{6.13}$$

and similarly

$$\tilde{\theta} = 2\pi\sqrt{\frac{i}{8}}\left(1 - \frac{i}{\eta^2}\right) d\eta. \tag{6.14}$$

7. The eigenspace bundle

We have seen in §5 that the eigenspace bundle E_x corresponding to a point $x \in M$ is a holomorphic line bundle over the spectral curve Σ of degree $-(p+1)$ where p is the arithmetic genus of Σ. We consider now the variation of this line bundle in the Picard group $H^1(\Sigma; \mathcal{O}^*)$ of the equivalence classes of line bundles on Σ as x varies on the torus M. This variation will turn out to be *linear*, and will provide effectively a linearization of (1.7).

Recall that $E_x(\xi) \subset V_x$ is defined by the property

$$E_x(\xi) \subseteq \ker(H_x(\zeta) - \mu(\xi)), \qquad \xi \in \Sigma,$$

where $\pi(\xi) = \zeta$ and $H_x(\zeta)$ is the holonomy of the flat connection $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$ around closed curves through x in a given homotopy class.

Let $y \in M$ be another point and choose a path joining x to y. Let $P_{yx}(\zeta)\colon V_x \to V_y$ denote parallel translation of the flat connection from x to y along the path. Then

$$H_y(\zeta)P_{yx}(\zeta) = P_{yx}(\zeta)H_x(\zeta),$$

and so

$$P_{yx}(E_x) = E_y. \tag{7.1}$$

Parallel translation thus gives an *isomorphism between* E_x *and* E_y for $\pi(\xi) \neq 0, \infty$. We therefore have a nonvanishing section of the line bundle $E_x^* \otimes E_y$ on the open set $\Sigma \setminus \pi^{-1}\{0, \infty\}$. To see what happens as $\zeta \to 0$ or ∞, let us consider separately the two cases $\det\Phi \neq 0$ and $\det\Phi = 0$.

In the first case $\zeta = 0$ is not a branch point, and from (3.6) we can find a neighborhood U of 0 on which the vector bundle V on M globally splits as a direct sum $V = L_\zeta \oplus L_\zeta^*$ of eigenspaces of the holonomy of the flat connection and which converge moreover as $\zeta \to 0$ to the eigenspaces of $\Phi\colon V \to V \otimes K$. The eigenspace bundle $E_x(\xi)$ is $(L_\zeta)_x$ or $(L_\zeta^*)_x$ on the two components U_1, U_2 of $\pi^{-1}(U) \subset \Sigma$.

From Proposition (3.5), however, the connection matrix of the flat connection on L_ζ has the form

$$\theta = \zeta^{-1}\lambda\, dz + \alpha_0 + \zeta\alpha_1 + \cdots,$$

so parallel translation in the line bundle L_ζ from the point x, given by $z = x_1 + ix_2 \in \mathbb{C}$, to the point y, given by $z = y_1 + iy_2$, is of the form

$$P_{yx} = e^{\zeta^{-1}\lambda[(y_1 - x_1) + i(y_2 - x_2)]} h(\zeta, x, y), \tag{7.2}$$

where $h(\zeta, x, y)$ is holomorphic in ζ and nonvanishing. The corresponding parallel translation in L^* is

$$P_{yx} = e^{-\zeta^{-1}\lambda[(y_1 - x_1) + i(y_2 - x_2)]} h^{-1}(\zeta, x, y). \tag{7.3}$$

The trivialization of $E_x^* \otimes E_y$ by parallel translation does not therefore extend over $\zeta = 0$, but does if we multiply it by $e^{-\zeta^{-1}\lambda[(y_1-x_1)+i(y_2-x_2)]}$ on U_1 and $e^{\zeta^{-1}\lambda[(y_1-x_1)+i(y_2-x_2)]}$ on U_2.

We may argue similarly at $\zeta = \infty$, or use the real structure ρ to transfer everything from 0 to ∞. What we have found is a description of the line bundle $E_x^* \otimes E_y$ in terms of *transition functions.* We cover the spectral curve Σ with five open sets:

$$U_1, U_2, \rho(U_1), \rho(U_2), \Sigma \setminus \pi^{-1}\{0, \infty\} = U_0,$$

and define (taking x to be $z = 0$ for simplicity) the transition functions

$$(7.4) \qquad \begin{array}{ll} e^{\zeta^{-1}\lambda(y_1+iy_2)} & \text{on } U_0 \cap U_1, \\ e^{-\zeta^{-1}\lambda(y_1+iy_2)} & \text{on } U_0 \cap U_2, \\ e^{-\zeta\bar{\lambda}(y_1-iy_2)} & \text{on } U_0 \cap \rho(U_1), \\ e^{\zeta\bar{\lambda}(y_1-iy_2)} & \text{on } U_0 \cap \rho(U_2). \end{array}$$

Since U_1, U_2, $\rho(U_1)$, and $\rho(U_2)$ may be chosen to be pairwise disjoint there are no further cocycle conditions.

Now exponentiation identifies the Picard group or Jacobian of line bundles of degree zero on Σ with the additive abelian group $H^1(\Sigma; \mathscr{O})/H^1(\Sigma; \mathbb{Z})$, and (7.4) shows that $E_x^* \otimes E_y$ is the element in this group represented by the Čech cocycle

$$(7.5) \qquad \begin{array}{ll} \zeta^{-1}\lambda(y_1 + iy_2) & \text{on } U_0 \cap U_1, \\ -\zeta^{-1}\lambda(y_1 + iy_2) & \text{on } U_0 \cap U_2, \\ -\zeta\bar{\lambda}(y_1 - iy_2) & \text{on } U_0 \cap \rho(U_1), \\ \zeta\bar{\lambda}(y_1 - iy_2) & \text{on } U_0 \cap \rho(U_2). \end{array}$$

In particular, this depends *linearly* on y_1 and y_2.

The second case of $\det \Phi = 0$ is dealt with similarly. Here since $\zeta = 0$ is a branch point, we take a neighborhood of 0 whose inverse image in Σ is a connected set U_1 with parameter η given by $\eta^2 = \zeta$. From §3 there is a holomorphic line bundle L_η over M, which is an eigenspace for the holonomy of the flat connection and converges to $\ker \Phi$ as $\eta \to 0$. The connection matrix as in Proposition (3.10) is of the form

$$\theta = -\eta^{-1}\kappa\, dz + a_0 + \cdots,$$

so consideration of parallel translation leads to a description of the line bundle $E_x^* \otimes E_y$ by exponentiation of a Čech cocycle defined by *three* open sets: U_1, $\rho(U_1)$, U_0. It is

$$\begin{aligned} &-\eta^{-1}\kappa(y_1+iy_2) \quad &&\text{on } U_0\cap U_1, \\ &\eta\bar{\kappa}(y_1-iy_2) &&\text{on } U_0\cap\rho(U_1). \end{aligned} \tag{7.6}$$

In either case, the eigenspace bundle E_y is of the form $E_x\otimes L_y$ where L_y is a line bundle of degree zero varying *linearly* in y.

Remark (7.7) From the geometrical origin of the line bundle in terms of equations on the torus $M=\mathbb{C}/\Gamma$ it is clear that L_y must be trivial if $y_1+iy_2=m+n\tau$, so that $y\to L_y$ actually maps the torus M linearly to a real torus in the Picard group $H^1(\Sigma;\mathscr{O})/H^1(\Sigma;\mathbb{Z})$. This observation is equivalent to the fact that each period of the differentials θ and $\tilde{\theta}$ on Σ is of the form $2in\pi$ as in Proposition (4.2).

To see this, suppose we consider the case $\det\Phi\neq 0$, and let $\mathscr{M}$ denote the sheaf over Σ of local meromorphic functions on Σ whose only poles are simple ones at $\pi^{-1}\{0,\infty\}$. Since multiplication by the section ζ of $\mathscr{O}(2)$ on $\mathbb{C}P^1$ which vanishes at 0 and ∞ makes a local section of $\mathscr{M}$ regular, $\mathscr{M}=\pi^*\mathscr{O}(2)$ and we have the exact sequence of sheaves

$$0\to\mathscr{O}\xrightarrow{\zeta}\mathscr{M}\to\mathscr{O}(2)|_D\to 0, \tag{7.8}$$

where D is the divisor $\pi^{-1}\{0,\infty\}$.

The sheaf $d\mathscr{M}$ of local differentials with double poles at $\pi^{-1}\{0,\infty\}$ and zero residues appears naturally in the exact sequence

$$0\to\mathbb{C}\to\mathscr{M}\to d\mathscr{M}\to 0, \tag{7.9}$$

and putting together (7.8) and (7.9) we have a commutative diagram:

$$\begin{array}{ccccccccc} & & & & 0 & & 0 & & \\ & & & & \uparrow & & \uparrow & & \\ & & & & \mathscr{O}(2)_D & \cong & \mathscr{O}(2)_D & & \\ & & & & \uparrow & & \uparrow & & \\ 0 & \to & \mathbb{C} & \to & \mathscr{M} & \to & d\mathscr{M} & \to & 0 \\ & & & & \uparrow & & \uparrow & & \\ 0 & \to & \mathbb{C} & \to & \mathscr{O} & \to & d\mathscr{O} & \to & 0 \\ & & & & \uparrow & & \uparrow & & \\ & & & & 0 & & 0 & & \end{array}$$

Now $\theta\in H^0(\Sigma;d\mathscr{M})$ and its image in $H^1(\Sigma;\mathbb{C})$ in the long exact sequence of the middle row consists of the class of the periods of θ. This maps by exactness to zero in $H^1(\Sigma;\mathscr{M})$, therefore its period class in the long exact sequence of the bottom row maps to an element in $H^1(\Sigma;\mathscr{O})$ coming from $H^0(\Sigma;\mathscr{O}(2)_D)$ in the exact sequence for the middle column.

Since the periods of θ are of the form $2in\pi$, $n \in \mathbb{Z}$, the corresponding element in $H^1(\Sigma; \mathscr{O})$ actually lies in $H^1(\Sigma; \mathbb{Z})$. Finding this element just involves spelling out the coboundary map

$$H^0(\Sigma; \mathscr{O}(2)_D) \xrightarrow{\delta} H^1(\Sigma; \mathscr{O}).$$

From Proposition (3.5), $\pm \log \mu = \lambda\zeta^{-1} + a + \zeta b(\zeta)$, so $\pm\theta = d(\lambda\zeta^{-1}) + \cdots$ and gives the value $\pm\lambda$ on the two points of D lying over $\zeta = 0$, with $\pm\bar{\lambda}$ over the corresponding points at ∞. The coboundary map δ is defined by extending this section s of $\mathscr{O}(2)$ on D to a neighborhood U of D and taking the Čech cocycle $\zeta^{-1}s$ of $\mathscr{O}$ on $\Sigma \setminus D \cap U$ to represent an element of $H^1(\Sigma; \mathscr{O})$. This, however, is precisely the description (7.5) for $y_1 = 1$, $y_2 = 0$, showing that L_y is trivial at this point. Using $\tilde{\theta}$ one similarly sees that $y_1 + iy_2 = \tau$ gives a trivial bundle and by linearity $y = m + n\tau$. The case $\det \Phi = 0$ may be treated using (7.6) in an exactly parallel manner.

We obtain therefore a map $l \colon M \to \mathrm{Pic}^{p+1}(\Sigma)$ to the Picard variety of line bundles of degree $(p+1)$ by defining $l(y) = E_y^*$, and we have seen that, with respect to a uniformizing parameter of M, the map l is real linear and maps to a real torus. This image may not be a 2-dimensional torus. Indeed if $p = 0$ or 1, the Picard variety is itself only 0 or 1-dimensional. However, we shall see that this is the only case where l fails to be an immersion. We need the following lemma:

Lemma (7.10). *If Σ is the spectral curve and $\pi \colon \Sigma \to \mathbb{C}P^1$ the projection, then*

$$\pi^* \colon H^0(\mathbb{C}P^1; \mathscr{O}(k)) \to H^0(\Sigma; \mathscr{O}(k))$$

is an isomorphism for $k < p + 1$.

Proof. The spectral curve Σ by its definition lies in the total space of $\mathscr{O}(p+1)$ and is given by an equation $\eta^2 = P(\zeta)$, where P is a section of $\mathscr{O}(2p+2)$ on $\mathbb{C}P^1$ and η is the tautological section of $\mathscr{O}(p+1)$. We compactify $\mathscr{O}(p+1)$ to the projective bundle $X = P(\mathscr{O}(p+1) \oplus \mathscr{O})$ over $\mathbb{C}P^1$. The divisor classes F of a fiber and Z of the zero section generate $H^2(X; \mathbb{Q})$ and satisfy

$$F^2 = 0, \qquad F \cdot Z = 1, \qquad Z^2 = p + 1. \tag{7.11}$$

Since the spectral curve satisfies

$$\Sigma \cdot F = 2 \quad \text{and} \quad \Sigma \cdot Z = 2p + 2,$$

its divisor class is $2Z$.

Let L be the line bundle on X given by the divisor class Z. Then there is an exact sequence of sheaves

$$0 \to \mathcal{O}(k) \otimes L^{-2} \to \mathcal{O}(k) \to \mathcal{O}(k)_\Sigma \to 0. \tag{7.12}$$

By covering $\mathbb{C}P^1$ with the standard patches it is easy to see that for $k < p+1$ all sections of $\mathcal{O}(k)$ on $\mathcal{O}(p+1)$ are pulled back from $\mathbb{C}P^1$ and hence also on X.

From the long exact sequence of (7.12) the lemma will follow if $H^0(X; \mathcal{O}(k) \otimes L^{-2}) = 0$ and $H^1(X; \mathcal{O}(k) \otimes L^{-2}) = 0$. The first is clearly true by restriction to fibers, since $\mathcal{O}(k) \otimes L^{-2}$ is of negative degree on a fiber.

Now since $\mathcal{O}(k) \otimes L^{-1}$ is of degree -1 on each fiber by (7.11) and $H^0(\mathbb{C}P^1; \mathcal{O}(-1)) = H^1(\mathbb{C}P^1; \mathcal{O}(-1)) = 0$, then by the Leray spectral sequence $H^i(X; \mathcal{O}(k) \otimes L^{-1}) = 0$ for $i = 0$ and 1.

Consider then the exact sequence

$$0 \to \mathcal{O}(k) \otimes L^{-2} \to \mathcal{O}(k) \otimes L^{-1} \to \mathcal{O}(k) \otimes L_Z^{-1} \to 0$$

restricting to the zero section for the last term. From the exact cohomology sequence we have

$$H^0(Z; \mathcal{O}(k) \otimes L^{-1}) \cong H^1(X; \mathcal{O}(k) \otimes L^{-2}),$$

but $Z \cong \mathbb{C}P^1$ and $\mathcal{O}(k) \otimes L^{-1} \cong \mathcal{O}(k-p-1)$ from (7.11). Thus if $k < p+1$, the degree of $\mathcal{O}(k) \otimes L^{-1}$ is negative on Z, and so $H^0(Z; \mathcal{O}(k) \otimes L^{-1})$ vanishes and hence also the required cohomology group $H^1(X; \mathcal{O}(k) \otimes L^{-2})$. q.e.d.

From the lemma, we see in particular that $H^0(\Sigma; \mathcal{O}(2))$ is 3-dimensional if $p+1 > 2$. Thus, considering the exact sequence of sheaves on Σ

$$0 \to \mathcal{O} \to \mathcal{O}(2) \to \mathcal{O}(2)_D \to 0,$$

where D is the divisor of the section ζ of $\mathcal{O}(2)$, the long exact cohomology sequence shows that for $p > 1$ the dimension of the image of $H^0(D; \mathcal{O}(2))$ in $H^1(\Sigma; \mathcal{O})$ via the coboundary map δ is 2-dimensional. Considering the real elements of $H^0(D; \mathcal{O}(2))$ which come from the y_1 and y_2 coefficients in (7.5), we see that this image has two real dimensions, and from the Remark above this is the image of the map l, translated by E_x^*.

So, apart from the cases $p = 0$ and 1, which we shall consider in detail in §§9 and 10, the image of $\mathcal{M}$ in the Picard variety is a 2-dimensional torus.

This torus is real in a specific sense given by Proposition (4.2) and its interpretation on the spectral curve Σ. This is that the holomorphic line bundle E_x is quaternionic with respect to the real structure $\sigma\rho: \Sigma \to \Sigma$.

Recall that if $\tau: X \to X$ is a real structure on a complex manifold, i.e., an antiholomorphic involution, and L is a holomorphic line bundle on X, then so is $\overline{\tau^* L}$. An isomorphism $L \cong \overline{\tau^* L}$ is of two types depending on whether its square gives multiplication by a positive or negative scalar on L. In the first case L is called *real*, in the second *quaternionic*. The space of holomorphic sections $H^0(X; L)$ is a real or quaternionic vector space depending on the type.

If X is the spectral curve Σ, then $\tau = \sigma\rho$ makes E_x quaternionic. With respect to the real structure $\zeta \to \bar{\zeta}^{-1}$ on $\mathbb{C}P^1$ the line bundles $\mathscr{O}(k)$ are *real* and hence from (4.2) real on Σ. The tensor product of a real bundle and a quaternionic one is quaternionic, and the tensor product of two quaternionic ones is real. Thus the bundle $L_y = E_x^* \otimes E_y$ is a *real* element of $H^1(\Sigma; \mathscr{O})/H^1(\Sigma; \mathbb{Z})$ with respect to $\sigma\rho$. We shall prove next some important properties of these quaternionic bundles.

The group of real classes in the Picard group, or Jacobian, $H^1(\Sigma; \mathscr{O})/H^1(\Sigma; \mathbb{Z})$ may in general have several components. In our situation we have:

Proposition (7.13). *The real part of the Jacobian of Σ has one component if p is even and two components if p is odd.*

Proof. First note that $\sigma\rho$ has no fixed points, for such a point would project to a point on the unit circle in $\mathbb{C}P^1$ and be fixed by σ. In other words the two eigenspaces of the holonomy coincide at this point. However, over the unit circle, the flat connection is unitary, so the eigenspaces are orthogonal.

Now the Jacobian of the possibly singular curve Σ fibers over the Jacobian of its normalization $\hat{\Sigma}$ with fiber of the form $\mathbb{C}^{*m} \times \mathbb{C}^n$. The $\mathbb{C}^*$ factors are given by the isomorphism α_ξ of the corresponding line bundle over $\hat{\Sigma}$ over a point ξ with the fiber over $\sigma\xi$. Since $\sigma\rho$ has no fixed points, a real isomorphism is determined by transforming α_ξ by $\sigma\rho$ to the conjugate points giving the connected group $\mathbb{C}^*$ as the real group, embedded in $\mathbb{C}^* \times \mathbb{C}^*$ by $z \to (z, \bar{z})$. The real form of $\mathbb{C}$ is always a connected group isomorphic to $\mathbb{R}$, so the fibers of the projection to $H^1(\hat{\Sigma}; \mathscr{O})/H^1(\hat{\Sigma}; \mathbb{Z})$ are connected.

Consider now the smooth curve $\hat{\Sigma}$. In the definition of Σ in (5.4) we see that since there are no real points of $\mathbb{C}P^1$ over which the eigenspaces of the holonomy coincide, the points $\pi(\beta_1), \cdots, \pi(\beta_k)$ occur in conjugate pairs, hence

$$p+1 = g+1+2\sum_{i=1}^{l} n_i + 2\sum_{i=1}^{k/2} m_i ,$$

and p and g have the same parity.

The hyperelliptic curve $\hat{\Sigma}$ is defined by the equation $\eta^2 = p(\zeta)$ where p is a real polynomial of degree $2g+2$ with no roots on the unit circle (from (4.2)), thus there are $g+1$ roots inside the circle and the winding number of $p(\zeta)$ is $(g+1)$. Consequently if g is even, the double covering of the unit circle in $\hat{\Sigma}$ has one component and if g is odd, two components.

Now decompose $\mathbb{C}P^1$ into the two hemispheres whose intersection is the unit circle and correspondingly decompose $\hat{\Sigma}$ into two submanifolds D_1 and D_2 with common boundary the inverse image of the circle. If g is even, this is a single circle and the Mayer-Vietoris sequence gives

$$\mathbb{Z}_2^{2g} = H^1(\hat{\Sigma}; \mathbb{Z}_2) \cong H^1(D_1; \mathbb{Z}_2) \oplus H^1(D_2; \mathbb{Z}_2).$$

The real structure interchanges D_1 and D_2, so the invariant elements are of the form $(x, (\sigma\rho)^* x)$ and there are $\# H^1(D_1; \mathbb{Z}_2) = 2^g$ of these.

Now $H^1(\hat{\Sigma}; \mathscr{O})/H^1(\hat{\Sigma}; \mathbb{Z})$ is a g-dimensional complex torus and the real points under $\sigma\rho$ will be a disjoint union of n real g-dimensional tori. In particular each torus will have 2^g real half-periods, i.e., 2^g invariant elements of $H^1(D_1; \mathbb{Z}_2)$. Since for even g there are only 2^g of these, we deduce that the real part of the Jacobian of $\hat{\Sigma}$ is connected, and so by our previous argument is that of Σ.

When g is odd, we have two circles on the common boundary of D_1 and D_2, and these contribute a 2-dimensional invariant subspace of $H^1(\hat{\Sigma}; \mathbb{Z}_2)$. There is a complementary $(2g-2)$-dimensional space generated by cycles in D_1 and their conjugates in D_2 which are interchanged by $\sigma\rho$. Thus there is a $(g-1)+2 = (g+1)$-dimensional invariant subspace, i.e., 2^{g+1} invariant half-periods (see also [12]). The real Jacobian must therefore have 2 components. So then does that of Σ. q.e.d.

We see then that the real part of $\mathrm{Pic}^{p+1}(\Sigma)$ has one component if p is even and two if p is odd. The eigenspace bundle is *quaternionic* and so the single component for p even must consist of the quaternionic line bundles. In fact, without knowing the existence of the eigenspace bundle, one can see that this real torus consists of quaternionic bundles, for, as shown in [3], every square root $K^{1/2}$ of the canonical bundle of a real curve with no real points must be quaternionic if the genus is even. Tensoring with the real bundle $\mathscr{O}(2)$ gives us a quaternionic bundle of degree $(g+1)$. Similarly, from [3] a curve of odd genus has 2^g real square roots and the rest quaternionic. Thus the two components in the odd genus case consist precisely of real and quaternionic bundles. In either case we have the following.

Proposition (7.14). *The space of quaternionic line bundles in* $\mathrm{Pic}^{p+1}(\Sigma)$ *is connected.*

The quaternionic line bundles on Σ have the property that they are, in the language of algebraic geometry, *nonspecial*:

Proposition (7.15). *Let* L *be a line bundle of degree* $(p-1)$ *on* Σ *which is quaternionic with respect to the real structure* $\sigma\rho$. *Then* $H^0(\Sigma; L) = 0$.

Proof. Since L is quaternionic, the vector space $H^0(\Sigma; L)$ is quaternionic and hence even-dimensional as a complex vector space. The same will also be true of $\sigma^* L$. Suppose this dimension is nonzero: it is therefore at least two. Fix a section s_0 of $\sigma^* L$ and consider the sections ss_0 of $L \otimes \sigma^* L$ as $s \in H^0(\Sigma; L)$. Now $L \otimes \sigma^* L$ is σ-invariant and hence pulled back from $\mathbb{C}P^1$, and of degree $(2p-2)$, hence

$$L \otimes \sigma^* L \cong \mathscr{O}(p-1).$$

But from Lemma (7.10) every section of this bundle is pulled back from $\mathbb{C}P^1$, and so every divisor of a section must be σ-invariant. However, fixing s_0 and varying s in $H^0(\Sigma; L)$ we get a linear system of divisors of dimension ≥ 1 (since $\dim H^0(\Sigma; L) \geq 2$) which must be σ-invariant. Thus the movable part of the system $|L|$ is of the form $|\mathscr{O}(k)|$, and we have $L \cong \mathscr{O}(k) \otimes L'$ where L' is the fixed part; in particular $\dim H^0(\Sigma; L') = 1$. But $\mathscr{O}(k)$ is real and L is quaternionic, therefore L' is quaternionic. But then $\dim H^0(\Sigma; L')$ must be even which is a contradiction. q.e.d.

Applying (7.15) to $E_x^*(-1)$ we see that $H^0(\Sigma; E_x^*(-1)) = 0$. From the Riemann-Roch theorem for a possibly singular curve in a surface, we have $H^1(\Sigma; E_x^*(-1)) = 0$. Applied to E_x^* it gives

$$\dim H^0(\Sigma; E_x^*) - \dim H^1(\Sigma; E_x^*) = (p+1) + (1-p) = 2. \tag{7.16}$$

Now consider a point in $\mathbb{C}P^1$ whose inverse image in Σ consists of two points—a divisor D—and the exact sequence of sheaves:

$$0 \to E_x^*(-1) \to E_x^* \to E_x^*|_D \to 0.$$

From the long exact sequence we have

$$0 \to H^0(\Sigma; E_x^*) \to H^0(D; E_x^*) \to \cdots$$

but since $H^0(D; E_x^*)$ is two-dimensional, $\dim H^0(\Sigma; E_x^*) \leq 2$. On the other hand from (7.16) $\dim H^0(\Sigma; E_x^*) \geq 2$, hence we have equality and $H^1(\Sigma; E_x^*) = 0$.

By its very definition E_x is a subbundle of the trivial bundle $\Sigma \times V_x$, and hence there is an exact sequence

$$0 \to E_x \to V_x^* \to E_x^* \to 0.$$

Since E_x is of negative degree, $H^0(\Sigma; E_x) = 0$ and hence we have an injection $0 \to V_x^* \to H^0(\Sigma; E_x^*)$. But we saw that $H^0(\Sigma; E_x^*)$ is 2-dimensional, so there is a natural isomorphism

$$V_x^* \cong H^0(\Sigma; E_x^*). \tag{7.17}$$

To summarize, we have seen here that as x varies over the torus M, the eigenspace bundle varies *linearly* in a subtorus of the connected space of quaternionic line bundles of degree $(p+1)$ on Σ. For each quaternionic line bundle L of degree $(p+1)$, the space $H^0(\Sigma; L)$ is 2-dimensional and restricted to the subtorus; it is naturally isomorphic to the vector bundle V^*.

We shall use this information in the next section in order to reverse the foregoing arguments and show how to construct solutions to (1.7) from an algebraic curve satisfying certain constraints.

8. The construction

The object of this section is to provide a construction of solutions to (1.7), and in particular harmonic maps, from an algebraic curve. We shall prove the following:

Theorem (8.1). *Let Σ be a curve in the total space of $\mathcal{O}(p+1)$ over $\mathbb{C}P^1$ defined by the equation $\eta^2 = P(\zeta)$ where η is the tautological section of $\pi^*\mathcal{O}(p+1)$ over $\mathcal{O}(p+1)$, $P(\zeta)$ is a section of $\mathcal{O}(2p+2)$ over $\mathbb{C}P^1$, and $\pi: \mathcal{O}(p+1) \to \mathbb{C}P^1$ is the projection. Suppose Σ satisfies the following conditions:*

(i) *$P(\zeta)$ is a real section of $\mathcal{O}(2p+2)$ with respect to the real structure $\zeta \to \bar{\zeta}^{-1}$ on $\mathbb{C}P^1$,*

(ii) *$P(\zeta)$ has no real zeros (i.e., zeros on the unit circle $\zeta = \bar{\zeta}^{-1}$),*

(iii) *$P(\zeta)$ has at most simple zeros at $\zeta = 0$ and $\zeta = \infty$,*

(iv) *there exist differentials θ, $\tilde{\theta}$ of the second kind on Σ with periods lying in $2\pi i\mathbb{Z}$,*

(v) *θ and $\tilde{\theta}$ have double poles at $\pi^{-1}(0)$ and $\pi^{-1}(\infty)$ and satisfy $\sigma^*\theta = -\theta$, $\sigma^*\tilde{\theta} = -\tilde{\theta}$, $\rho^*\theta = -\bar{\theta}$, $\rho^*\tilde{\theta} = -\bar{\tilde{\theta}}$ where σ is the involution on Σ induced by multiplication by -1 in the fibers of $\mathcal{O}(p+1)$ and ρ is the real structure on $\mathcal{O}(p+1)$ induced from $\zeta \to \bar{\zeta}^{-1}$, and*

(vi) *the principal parts of* θ *and* $\tilde{\theta}$ *are linearly independent over* $\mathbb{R}$.

Then, for each point in the Picard variety of line bundles of degree $(p+1)$ *on* Σ *which are quaternionic with respect to the real structure* $\rho\sigma$, *there exists a solution of* (1.7) *for a torus, such that* Σ *is the spectral curve of the solution and* θ, $\tilde{\theta}$ *the corresponding differentials. The solution is, moreover, unique modulo gauge transformations and the operation of tensoring* V *by a flat* $\mathbb{Z}_2$*-bundle.* (*Note that* (ii) *implies that* $\rho\sigma$ *has no fixed points, and so from* [3] *quaternionic bundles of degree* $(p+1)$ *exist.*)

Proof. (1) The first task is to define the torus and its conformal structure from the curve Σ.

Consider the case where $P(0) \neq 0$, then the differentials θ and $\tilde{\theta}$ have double poles with zero residue on the divisor $D_0 = \pi^{-1}(0)$ consisting of two points. Using ζ as a local parameter we have

$$\theta = a\frac{d\zeta}{\zeta^2} + \text{holomorphic}, \qquad \tilde{\theta} = \tilde{a}\frac{d\zeta}{\zeta^2} + \text{holomorphic},$$

and the principal parts of θ and $\tilde{\theta}$ may be invariantly considered as sections of $\mathscr{O}(2)$ on the divisor D_0 (cf. (7.8)). We consider then the 2-dimensional complex vector space $H^0(D_0; \mathscr{O}(2))$ and the 1-dimensional subspace U of vectors u for which $\sigma^* u = -u$, σ interchanging the two points. Since from (v) $\sigma^*\theta = -\theta$ and $\sigma^*\tilde{\theta} = -\tilde{\theta}$, the principal parts of θ and $\tilde{\theta}$ lie in U. Since by (vi) they are linearly independent over $\mathbb{R}$, they generate a lattice Γ in U and we define the torus $M = U/\Gamma$. Note from (3.5) that in the case where Σ is known to be a spectral curve, the principal parts are $-\lambda d\zeta/\zeta^2$ and $-\lambda\tau d\zeta/\zeta^2$, so that this definition gives a torus conformally equivalent to one with lattice generated by 1 and τ. If $P(0) = 0$, then from (iii) we have a simple zero so that the single point D_0 lying over 0 is a smooth point of Σ and then the principal parts of θ and $\tilde{\theta}$ lie in the 1-dimensional space $U = H^0(D_0; \mathscr{O}(1))$. We define the torus by $M = U/\Gamma$ where Γ is the lattice generated by the principal parts of θ and $\tilde{\theta}$. From (3.10) this is the definition that is required.

(2) Next we must define a rank 2 vector bundle V over M. For this we consider, for $P(0) \neq 0$, the exact sequence of sheaves

$$0 \to \mathscr{O} \to \mathscr{O}(2) \to \mathscr{O}(2)_D \to 0,$$

where D is the divisor $\pi^{-1}\{0, \infty\}$ on Σ, and the exact cohomology sequence

$$0 \to \mathbb{C} \to H^0(\Sigma; \mathscr{O}(2)) \to H^0(D; \mathscr{O}(2)) \xrightarrow{\delta} H^1(\Sigma; \mathscr{O}).$$

Now the divisor D is real and in fact $D = D_0 + \rho\sigma(D_0)$, so the 2-dimensional subspace of the 4-dimensional complex vector space $H^0(D; \mathscr{O}(2))$ on which σ acts as -1 can be written as $U \oplus \overline{U} = U \otimes \mathbb{C}$ where U is the subspace of $H^0(D_0; \mathscr{O}(2))$ defined above. From (v) the principal parts of θ and $\tilde{\theta}$ lie in this subspace and are real with respect to the natural real structure $\sigma\rho$.

The coboundary map δ maps the 2-dimensional space $U \otimes \mathbb{C}$ to $H^1(\Sigma; \mathscr{O})$ commuting with real structures and this gives a real linear map from U to the real part of $H^1(\Sigma; \mathscr{O})$.

From Remark (7.7), which applies to the curve Σ, the images of the principal parts of θ and $\tilde{\theta}$ under δ are integral classes since from (iv) they have periods in $2\pi i\mathbb{Z}$.

Consequently we obtain a (real linear) map

$$l: M = U/\Gamma \to H^1(\Sigma; \mathscr{O})/H^1(\Sigma; \mathbb{Z}) = \operatorname{Pic}^0(\Sigma),$$

which associates a line bundle on Σ, real with respect to $\rho\sigma$, to each $x \in U/\Gamma$, with $L_0 \cong \mathscr{O}$. To be precise, we have associated an *equivalence class* of line bundles on Σ to x, but we may choose a Poincaré line bundle on $\Sigma \times \operatorname{Pic}^0(\Sigma)$, i.e., a line bundle $\mathscr{L}$ which for each $y \in \operatorname{Pic}^0(\Sigma)$ restricts to a bundle L_y in the equivalence class of y. By the universal property of Poincaré bundles, there is a holomorphic map

$$f: \Sigma \times (U \otimes \mathbb{C})/\Gamma \to \Sigma \times \operatorname{Pic}^0(\Sigma),$$

and we pull back the Poincaré bundle $\mathscr{L}$ to obtain a line bundle on $\Sigma \times (U \otimes \mathbb{C})/\Gamma$. For each $x \in U/\Gamma \subset (U \otimes \mathbb{C})/\Gamma$ the line bundle L_x is real.

There is an ambiguity in the choice of Poincaré bundle—tensoring with a line bundle pulled back from $\operatorname{Pic}^0(\Sigma)$—which we now proceed to remove as far as we can. For this purpose consider the involution σ_1 on $\Sigma \times (U \otimes \mathbb{C})/\Gamma$ given by σ on the first factor. The bundle $\sigma_1^*\mathscr{L} \otimes \mathscr{L}$ is trivial on each fiber of the projection p_2, so the direct image sheaf $p_{2*}(\sigma_1^*\mathscr{L} \otimes \mathscr{L})$ is a line bundle S on $(U \otimes \mathbb{C})/\Gamma$. The first Chern class of $\mathscr{L}$ on $\Sigma \times (U \otimes C)/\Gamma$ may be written

$$c_1(L) = a + p_2^*b,$$

where $a \in H^1(\Sigma; \mathbb{Z}) \otimes H^1(U \otimes \mathbb{C}/\Gamma; \mathbb{Z})$ and $b \in H^2(U \otimes \mathbb{C}/\Gamma; \mathbb{Z})$. Since σ_1 acts as -1 on $H^1(\Sigma; \mathbb{Z})$ and $+1$ on $H^1(U \otimes \mathbb{C}/\Gamma; \mathbb{Z})$ we have

$$c_1(\sigma_1^*\mathscr{L} \otimes \mathscr{L}) = -a + p_2^*b + a + p_2^*b = 2p_2^*b.$$

From the Grothendieck-Riemann-Roch formula we have

$$c_1(S) = p_{2*}(td(\Sigma) \cdot 2p_2^*b),$$

which is an even class in $H^2(U\otimes\mathbb{C}/\Gamma;\mathbb{Z})$, and hence the line bundle S has a holomorphic square root $S^{1/2}$. Under the natural notion of equivalence of square roots of a given line bundle S, any two differ by an element of $H^1(U\otimes\mathbb{C}/\Gamma;\mathbb{Z}_2)\cong H^1(U/\Gamma;\mathbb{Z}_2)$.

Choose such a square root and consider $\mathscr{L}'=\mathscr{L}\otimes p_2^*S^{-1/2}$. This is again a Poincaré bundle on $\Sigma\times(U\times\mathbb{C})/\Gamma$ but satisfies in a canonical way the identity

$$p_{2*}(\sigma_1^*\mathscr{L}'\otimes\mathscr{L}')\cong\mathscr{O},$$

and hence we have, for each $x\in U/\Gamma$, a natural isomorphism

$$\sigma^*L_x\otimes L_x\cong\mathscr{O} \tag{8.2}$$

uniquely defined modulo the action of $H^1(U/\Gamma;\mathbb{Z})$, where $L_x=\mathscr{L}'|_{\Sigma\times\{x\}}$.

Now let E^* denote the given quaternionic line bundle of degree $(p+1)$ on Σ and define the bundle $\mathscr{E}$ on $\Sigma\times(U\otimes\mathbb{C})/\Gamma$ by

$$\mathscr{E}=p_1^*E^*\otimes\mathscr{L}'. \tag{8.3}$$

Then $\mathscr{E}_x=\mathscr{E}|_{\Sigma\times\{x\}}$ is quaternionic for $x\in U/\Gamma$, of degree $(p+1)$ and depends linearly on x.

From (7.15) and the discussion following it, we see from the nonspeciality of quaternionic line bundles that $\dim H^0(\Sigma;E_x^*)=2$ and $H^1(\Sigma;E_x^*)=0$, and so the direct image sheaf $p_{2*}\mathscr{E}$ is a rank-2 bundle on $(U\otimes\mathbb{C})/\Gamma$. Restricting to U/Γ we obtain a rank-2 vector bundle V^*. In other words, we set

$$V_x^*=H^0(\Sigma;E_x^*), \tag{8.4}$$

which is consistent with (7.17) for spectral curves. Since E_x^* is quaternionic, each V_x^* is a quaternionic vector space so the bundle V has a quaternionic structure.

(3) The quaternionic structure on V gives its structure group a reduction to the nonzero quaternions. To reduce to SU(2), the unit quaternions, we need a nondegenerate skew form on V.

Now from (8.2) we have an isomorphism

$$\sigma^*E_x^*\otimes E_x^*\cong\sigma^*E^*\otimes E^*,$$

and since E^* is of degree $(p+1)$, the σ-invariant bundle $\sigma^*E^*\otimes E^*$ is isomorphic to $\mathscr{O}(p+1)$. Choose a *real* isomorphism; then we have a natural isomorphism

$$\sigma^*E_x^*\otimes E_x^*\cong\mathscr{O}(p+1) \tag{8.5}$$

compatible with real structures.

Let v_1, $v_2 \in V_x^*$ be two vectors. From (8.4) they correspond to two sections s_1, s_2 of E_x^*. Using (8.5) we have

$$\sigma^* s_1 \otimes s_2 \in H^0(\Sigma; \mathscr{O}(p+1)). \tag{8.6}$$

To proceed further we need to know the sections of $\mathscr{O}(p+1)$ on Σ. From the exact sequence of sheaves (7.12) with $k = p+1$, it follows that

$$0 \to H^0(X; \mathscr{O}(p+1)) \to H^0(\Sigma; \mathscr{O}(p+1)) \to H^1(X; \mathscr{O}(p+1) \otimes L^{-2}),$$

and from the sequence restricting to the zero section Z of X we have

$$\mathbb{C} \cong H^0(\mathbb{C}P^1; \mathscr{O}) \cong H^0(Z; \mathscr{O}(p+1) \otimes L^{-1}) \cong H^1(X; \mathscr{O}(p+1) \otimes L^{-2}).$$

Hence,

$$\dim H^0(\Sigma; \mathscr{O}(p+1)) \leq 1 + \dim H^0(X; \mathscr{O}(p+1)). \tag{8.7}$$

Restricting to the total space of $\mathscr{O}(p+1)$ in X, the sections of the line bundle $\mathscr{O}(p+1)$ are spanned by η and pull-backs of sections of $\mathscr{O}(p+1)$ on $\mathbb{C}P^1$. But η does not extend as a section of $\mathscr{O}(p+1)$ on X since its divisor is of the form $aZ + bZ_\infty$ (Z_∞ the infinity section of the projective bundle X) and intersection with the class F of a fiber gives $a + b > 0$, contradicting $F^2 = 0$ if $aZ + bZ_\infty = (p+1)F$ (see (7.11)). Thus from (8.7)

$$\dim H^0(\Sigma; \mathscr{O}(p+1)) \leq 1 + \dim H^0(\mathbb{C}P^1; \mathscr{O}(p+1)).$$

However, since η is a nonvanishing section of $\mathscr{O}(p+1)$ on Σ, not pulled back from $\mathbb{C}P^1$ we have equality and every section of $\mathscr{O}(p+1)$ on Σ is of the form $\lambda\eta + r(\zeta)$, where λ is a constant and $r(\zeta)$ a section of $\mathscr{O}(p+1)$ on $\mathbb{C}P^1$.

Returning to (8.6), we have from this that

$$\sigma^* s_1 \otimes s_2 = \omega(s_1, s_2)\eta + r(\zeta), \tag{8.8}$$

where $\omega(s_1, s_2)$ is a constant depending bilinearly on s_1 and s_2. Now

$$\begin{aligned}\sigma^* s_2 \otimes s_1 &= \sigma^*(\sigma^* s_1 \otimes s_2) = \sigma^*(\omega(s_1, s_2)\eta + r(\zeta)) \\ &= -\omega(s_1, s_2)\eta + r(\zeta).\end{aligned}$$

Thus $\omega(s_2, s_1) = -\omega(s_1, s_2)$ and ω is skew. If $\omega(s_1, s)$ vanishes for all s, then the divisor of s_1 must be σ-invariant and E_x^* therefore must be pulled back from $\mathbb{C}P^1$. Since E_x^* is quaternionic and not real this is impossible, so ω is nondegenerate and defines a symplectic form on V.

(4) The next goal is to define for each $\zeta \in \mathbb{C}P^1$ ($\zeta \neq 0, \infty$) a flat connection on V over U/Γ. To do this we pass to the universal covering U and define a parallel translation $\Pi_{yx}: V_x^* \to V_y^*$ satisfying

$$\Pi_{zx} = \Pi_{zy}\Pi_{yx} \qquad (x, y, z \in U).$$

By differentiating along paths, this will define in the standard way a connection on V over U.

We shall define Π_{yx} by returning to the argument where parallel translation was used to describe the eigenspace bundle in (7.1).

Consider $E_x \otimes E_y^* \cong L_y \otimes L_x^*$. By the definition of the line bundles L_y, there exists a nonvanishing section P_{yx} of $L_y \otimes L_x^*$ on $\Sigma \setminus \pi^{-1}\{0, \infty\}$ such that P_{yx} extends to the whole of Σ if we multiply by

$$e^{\pm\zeta^{-1}\lambda[(y_1-x_1)+i(y_2-x_2)]} \quad \text{on } U_1 \text{ and } U_2$$

if 0 is not a branch point, and

$$e^{\eta^{-1}\kappa[(y_1-x_1)+i(y_2-x_2)]} \quad \text{on } U_1$$

if 0 is a branch point, and corresponding multipliers at ∞ given by applying the real structure.

Any two such sections differ by a scalar multiple, but using the isomorphism (8.2), we have a trivialization of $\sigma^*(L_y \otimes L_x^*) \otimes L_y \otimes L_x^*$, so we may choose P_{yx} to satisfy

$$(\sigma^* P_{yx})P_{yx} = 1. \tag{8.9}$$

There is only an ambiguity of ± 1 now, which from the connectedness of U may be removed, so that P_{yx} is uniquely determined. By uniqueness and linearity we have

$$P_{zy}P_{yx} = P_{zx}. \tag{8.10}$$

We also have compatibility with the real structure by uniqueness and the reality property of the multipliers.

Now choose a section of $\mathcal{O}(1)$ on $\mathbb{C}P^1$ which vanishes at $\zeta(\neq 0, \infty)$ with D_ζ the corresponding divisor in Σ. From the exact sequence of sheaves

$$0 \to E_x^*(-1) \to E_x^* \to E_x^*|_{D_\zeta} \to 0$$

we have

$$H^0(\Sigma; E_x^*) \cong H^0(D_\zeta; E_x^*), \tag{8.11}$$

since $H^0(\Sigma; E_x^*(-1)) = H^1(\Sigma; E_x^*(-1)) = 0$ from the nonspeciality of quaternionic bundles. Thus any section is determined by its restriction to the divisor D_ζ. We define $\Pi_{yx}(\zeta)$ by the commutative diagram:

$$\begin{array}{ccc} H^0(\Sigma; E_x^*) & \xrightarrow{\Pi_{yx}} & H^0(\Sigma; E_y^*) \\ \| & & \| \\ H^0(D_\zeta; E_x^*) & \underset{\cong}{\xrightarrow{P_{yx}}} & H^0(D_\zeta; E_y^*) \end{array} \tag{8.12}$$

Then $\Pi_{yx}(\zeta): V_x^* \to V_y^*$ is an invertible linear map which is compatible with both the quaternionic structure and symplectic structure of V_x^* and V_y^*, i.e.,

$$\Pi_{yx}(\zeta)^* \omega_y = \omega_x, \tag{8.13}$$

$$j_y \Pi_{yx}(\zeta) = \Pi_{yx}(\overline{\zeta}^{-1}) j_x. \tag{8.14}$$

Moreover, from (8.10) it satisfies

$$\Pi_{zy}(\zeta)\Pi_{yx}(\zeta) = \Pi_{zx}(\zeta),$$

and therefore defines a connection on V^* over U. Since $\Pi_{yx}(\zeta)$ is independent of any path joining x to y, the connection is *flat.*

(5) Finally, we must show that this family of flat connections is of the form $\nabla + \zeta^{-1}\Phi - \zeta\phi^*$, and to do this we consider the limiting behavior as $\zeta \to 0$.

First consider for $\xi \in \Sigma$, the evaluation map

$$H^0(\Sigma; E_x^*) \xrightarrow{\mathrm{ev}_\xi} E_x^*(\xi). \tag{8.15}$$

If all sections vanish at ξ, then the divisor ξ lies in the fixed part of E_x^*, but in the proof of (7.15) we saw that the fixed part of a quaternionic bundle is real. Since $(\sigma\rho)(\xi)$ is also a fixed divisor, then removing the fixed part we obtain a quaternionic line bundle of degree $\leq p - 1$ with a 2-dimensional space of sections. From (7.15) this is impossible, so the evaluation map (8.15) is *surjective.*

From the definition (8.12) of parallel translation we have a commutative diagram for $\xi \notin \pi^{-1}\{0, \infty\}$:

$$\begin{array}{ccc} H^0(\Sigma; E_x^*) & \xrightarrow{\Pi_{yx}(\zeta)} & H^0(\Sigma; E_y^*) \\ {\scriptstyle \mathrm{ev}_\xi}\downarrow & & \downarrow{\scriptstyle \mathrm{ev}_\xi} \\ E_x^*(\xi) & \underset{P_{yx}(\xi)}{\longrightarrow} & E_y^*(\xi) \end{array}$$

so parallel translation preserves the sub-bundle $E_x(\xi)$ of V^* which is the kernel of the evaluation map, and its effect there is $P_{yx}^{-1}(\xi)$.

Now (8.15) is valid even for $\xi \in \pi^{-1}\{0, \infty\}$, so the sub-bundle $E_x(\xi) \subset H^0(\Sigma; E_x^*)$ which is preserved by parallel translation extends to $\pi^{-1}(0)$. Suppose first that 0 is not a branch point; then $\pi^{-1}(0)$ consists of two points, and the two bundles $E_x(\xi)$ and $E_x(\sigma\xi)$ have zero intersection in some neighborhood. Choosing a local trivialization of V^* given by sections of these two bundles, parallel translation has the form

$$\operatorname{diag}(e^{\pm\zeta^{-1}\lambda[(y_1-x_1)+i(y_2-x_2)]}H^{\pm 1}(\zeta, x, y)),$$

where H is holomorphic in ζ, from the behavior of P_{yx} as $\zeta \to 0$. Thus the connection matrix is of the form

$$A = \zeta^{-1}\begin{pmatrix} \lambda\, dz & 0 \\ 0 & -\lambda\, dz \end{pmatrix} + a(\zeta, x), \tag{8.16}$$

putting $z = x_1 + ix_2$.

If $\zeta = 0$ is a branch point, then the two sub-bundles $E_x(\xi)$ and $E_x(\sigma\xi)$ for $\pi(\xi) \neq 0$ still extend to $\pi(\xi) = 0$, but coincide there.

Choosing a local coordinate $\eta^2 = \zeta$ there is a local gauge such that these two sub-bundles are spanned by $\binom{1}{\eta}$ and $\binom{1}{-\eta}$.

In this gauge the connection matrix of the flat connection has a Laurent expansion

$$A = \sum_{-\infty}^{\infty} A_n\zeta^n = \sum_{-\infty}^{\infty} A_n\eta^{2n},$$

and the two eigenspaces preserved by the connection have holonomy of the form

$$e^{\kappa\eta^{-1}[(y_1-x_1)+i(y_2-x_2)]}H(\eta, x, y).$$

Thus we have

$$\sum_{-\infty}^{\infty} A_n\eta^{2n}\begin{pmatrix} 1 \\ \eta \end{pmatrix} = (\kappa\eta^{-1}\, dz + a(x, \eta))\begin{pmatrix} 1 \\ \eta \end{pmatrix}$$

giving $A_n = 0$ for $n < -1$ and $A_{-1}\binom{1}{0} = 0$, $A_{-1}\binom{0}{1} = \binom{\kappa}{0}$, so

$$A = \zeta^{-1}\begin{pmatrix} 0 & 0 \\ \kappa & 0 \end{pmatrix} dz + a(\zeta, x). \tag{8.17}$$

Hence from (8.16) and (8.17) the connection has the form

$$\nabla_\zeta = \nabla + \zeta^{-1}\Phi + a(\zeta, x),$$

where a is holomorphic in ζ, $\Phi \in \Omega^{1,0}(U; \operatorname{End} V_x)$, and ∇ is some fixed connection. If we absorb the constant coefficient in the power series expansion of $a(\zeta, x)$ into ∇, we have

$$\nabla_\zeta = \nabla + \zeta^{-1}\Phi + \sum_{n=1}^{\infty} \zeta^n \phi_n.$$

Using the real structure (8.14), the connection as $\zeta \to \infty$ is determined by its behavior as $\zeta \to 0$ and we see that $\phi_n = 0$ for $n > 1$, ∇ commutes with the quaternionic structure on V, and $\phi_1 = j^{-1}\Phi j$. Using the compatibility with the symplectic form (8.13) this yields a connection of the form

$$\nabla_\zeta = \nabla + \zeta^{-1}\Phi - \zeta\Phi^*, \tag{8.18}$$

where ∇ is an SU(2) connection.

Now, as in §1, flatness of the connection implies (1.7). Also, from (8.17), if $\zeta = 0$ is a branch point, $\det\Phi = 0$.

We have defined here a solution to (1.7) on the universal covering U of the torus U/Γ. However, the line bundle L_x and the vector bundle V_x defined by it are already defined on the torus. To see that the connection (8.18) is defined on the torus, we need only note that if $x, y \in U/\Gamma$ are sufficiently close, then there is an unambiguous choice of transition functions of the form $e^{\pm\zeta^{-1}\lambda[(y_1-x_1)+i(y_2-x_2)]}$ for the line bundle $L_x^* \otimes L_y$. This makes the connection well-defined on the torus.

In producing the connection we have at each stage effectively reversed the construction of the spectral curve, so Σ is indeed the spectral curve of the connection (8.18). The ambiguity in the definition of the connection was the choice of a square root of the line bundle S. It is easy to see that the $H^1(U/\Gamma; \mathbb{Z}_2)$ ambiguity here changes the line bundle $\mathscr{L}'$ by a flat $\mathbb{Z}_2$-bundle on U/Γ and has the effect of changing the connection (8.18) by tensoring with the same flat $\mathbb{Z}_2$-bundle.

If Σ is a curve satisfying the conditions of Theorem (8.1) with differentials θ, $\tilde{\theta}$, then it is easy to find the eigenvalues of the holonomy of the flat connection. In fact, on the universal covering of Σ we solve $df = \theta$, and since the periods of θ lie in $2\pi i\mathbb{Z}$, then e^f is a well-defined function on $\Sigma \setminus \pi^{-1}\{0, \infty\}$. The additive constant ambiguity in f becomes a multiplicative ambiguity in e^f, but if we insist that $e^f\sigma^*(e^f) = 1$ this is reduced to a ± 1 ambiguity.

Consequently we have functions μ, $\tilde{\mu}$ well defined modulo ± 1, such that

$$\theta = \frac{d\mu}{\mu}, \quad \tilde{\theta} = \frac{d\tilde{\mu}}{\tilde{\mu}}, \qquad \mu\sigma^*\mu = 1, \quad \tilde{\mu}\sigma^*\tilde{\mu} = 1. \tag{8.19}$$

In terms of the proof of Theorem (8.1), if $a \in U$ is the principal part of the differential θ, then the line bundle L_a is trivial and

$$P_{a0} \in L_a \otimes L_0^* \cong \mathcal{O}$$

is a *function* on $\Sigma \times \pi^{-1}\{0, \infty\}$ which extends after multiplying by $e^{\pm\zeta^{-1}\lambda[a_1+ia_2]}$ on the open sets U_1, U_2 and correspondingly at $\rho(U_1)$, $\rho(U_2)$. On the one hand, this means that $\theta = dP_{a0}/P_{a0}$ (and from (8.9) $(\sigma^* P_{a0})P_{a0} = 1$) so $P_{a0} = \mu$. On the other, the definition of the connection in the proof of the theorem means that P_{a0} is the eigenvalue of the holonomy around a generator of $\pi_1(U/\Gamma)$. Thus the functions μ, $\tilde{\mu}$ above are the eigenvalues of the holonomy and θ, $\tilde{\theta}$ the differentials defined in §4 for the corresponding solution to (1.7). q.e.d.

Theorem (8.1) associates to the data $(\Sigma, \theta, \tilde{\theta})$ just an equivalence class of solutions to (1.7), the equivalence being the operation of tensoring by a flat $\mathbb{Z}_2$-bundle (and of course gauge equivalence). To remove this ambiguity we must choose the functions μ and $\tilde{\mu}$ such that $\theta = d\mu/\mu$, $\tilde{\theta} = d\tilde{\mu}/\tilde{\mu}$ and $\mu\sigma^*\mu = 1$, $\tilde{\mu}\sigma^*\tilde{\mu} = 1$. Indeed, since each can be considered as an eigenvalue of the holonomy of the flat connection, tensoring with a flat $\mathbb{Z}_2$-bundle has the effect of multiplying each by the corresponding sign given by an element of $H^1(U/\Gamma; \mathbb{Z}_2)$. Thus the data $(\Sigma, \mu, \tilde{\mu})$ determines the solution of (1.7) up to gauge equivalence.

With this point in mind, we may now give the algebraic geometric equivalent of a harmonic map of the torus to $\mathrm{SU}(2)$, the basic geometric problem which we aimed to solve:

Theorem (8.20). *Let Σ be a curve with equation $\eta^2 = P(\zeta)$ satisfying the conditions of Theorem (8.1), and let μ and $\tilde{\mu}$ be functions on $\Sigma \setminus \pi^{-1}\{0, \infty\}$ satisfying $\theta = d\mu/\mu$, $\tilde{\theta} = d\tilde{\mu}/\tilde{\mu}$, and $\mu\sigma^*\mu = \tilde{\mu}\sigma^*\tilde{\mu} = 1$. Then,*

(i) *$(\Sigma, \mu, \tilde{\mu})$ determines a harmonic map from a torus to S^3 if and only if*

$$\mu(\xi) = \tilde{\mu}(\xi) = 1 \quad \textit{for all } \xi \in \pi^{-1}\{1, -1\},$$

(ii) *the map is conformal if and only if $P(0) = 0$,*

(iii) *the torus maps to a totally geodesic 2-sphere if and only if p is odd, $P(\zeta)$ is an even polynomial, and the point $E \in \mathrm{Pic}^{p+1}(\Sigma)$ and the* *functions* μ *and* $\tilde{\mu}$ *on* $\Sigma\backslash\pi^{-1}\{0, \infty\}$ *are invariant by* $\sigma\tau$, *where* τ *is the involution of* Σ *defined by* $\tau(\eta, \zeta) = (\eta, -\zeta)$,

(iv) *the harmonic map is uniquely determined by* $(\Sigma, \mu, \tilde{\mu})$ *modulo the action of* SO(4) *on* S^3.

Proof. (i) From §1, the solution to (1.7) determines a harmonic map if and only if the flat connections for $\zeta = 1, -1$ are actually *trivial.* Since they are unitary, this occurs if and only if the eigenvalues of the holonomy are $+1$. From the proof of (8.1), this means $\mu(\xi) = \tilde{\mu}(\xi) = 1$ if $\pi(\xi) = \zeta = 1, -1$.

(ii) From (1.8), the map is conformal if and only if $\det\Phi = 0$ and from (8.16) and (8.17) this occurs if and only if 0 is a branch point of the covering $\Sigma \to \mathbb{C}P^1$, i.e., if and only if $P(0) = 0$.

(iii) From (1.9) the torus is mapped to a totally geodesic 2-sphere if and only if there is a gauge transformation g with $g^2 = -1$, leaving A invariant and such that $g^{-1}\Phi g = -\Phi$. As in §3, we then have

$$g^{-1}(\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*)g = \nabla_A - \zeta^{-1}\Phi + \zeta\Phi^*, \tag{8.21}$$

and so the zeros of $P(\zeta)$ are invariant under $\zeta \to -\zeta$. Now a conformal harmonic map to S^2 has trivial spectral curve, so the map cannot be conformal, hence $P(0) \neq 0$. The zeros of $P(\zeta)$ in the disc $|\zeta| < 1$ are therefore paired by the involution, hence the number $(p + 1)$ of these zeros is even, and so the arithmetic genus p is *odd.* Since p is odd, the section $P(\zeta)$ itself is invariant by the natural involution $\zeta \to -\zeta$, so $P(\zeta) = P(-\zeta)$ is an even polynomial.

Now from (8.21)

$$g_x^{-1}H_x(\zeta)g_x = H_x(-\zeta),$$

and so the eigenvalues of the holonomy at ζ and $-\zeta$ are the same. Considering their asymptotic form near $\zeta = 0$ given by Proposition (3.7) and the discussion of case (b) following it, we see that $(\sigma\tau)^*\mu = \mu$ and $(\sigma\tau)^*\tilde{\mu} = \tilde{\mu}$. Note that g gives an isomorphism between the eigenspace bundle E_x and $(\sigma\tau)^*E_x$ whose square is -1.

Conversely suppose Σ satisfies the given conditions. Then $(\sigma\tau)^*\theta = \theta$ and $(\sigma\tau)^*\tilde{\theta} = \tilde{\theta}$, and so the line bundle L_x constructed from the principal parts of θ and $\tilde{\theta}$ satisfies $(\sigma\tau)^*L_x \cong L_x$. Since the point E is also $\sigma\tau$-invariant, from the proof of Theorem (8.1) the bundle E_x^* is $\sigma\tau$-invariant, and we have an isomorphism

$$\sigma\tau\colon H^0(\Sigma; E_x^*) \to H^0(\Sigma; E_x^*), \tag{8.22}$$

which is compatible with the quaternionic structures and symplectic structure, and hence as x varies describes a gauge transformation g of V.

From the construction of the flat connection in Theorem (8.1) we have

$$g^{-1}(\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*)g = \nabla_A - \zeta^{-1}\Phi + \zeta\Phi^*. \tag{8.23}$$

Now the action of τ^2 on E_x^* is ± 1 since τ is an involution. Thus $g^2 = \pm 1$, but if g is an SU(2) gauge transformation with $g^2 = +1$, then g itself is a scalar ± 1; thus in (8.23) $\Phi = 0$ giving a trivial spectral curve. Hence $g^2 = -1$ and from Proposition (1.9), the torus maps to a 2-sphere.

(iv) From Theorem (8.1), the data $(\Sigma, \mu, \tilde{\mu})$ determines the solution to (1.7) modulo gauge transformations. The harmonic map $g(x)$ is defined by choosing covariant constant sections s_1, s_{-1} of the SU(2) principal bundle P associated to V with respect to the trivial connections at $\zeta = 1$ and $\zeta = -1$ and defining $g(x)$ by

$$s_1 = g(x)s_{-1}. \tag{8.24}$$

A gauge transformation may be interpreted as a diffeomorphism of P, covering the identity on M and commuting with the action of SU(2). Because of this last property the function $g(x)$ is independent of gauge equivalence and thus depends only on the choice of sections s_1 and s_{-1}. But any two covariant constant sections are related by an action of SU(2), so a different choice of $\tilde{s}_1$, $\tilde{s}_{-1}$ gives

$$\tilde{s}_1 = hs_1 = hg(x)s_{-1} = hg(x)k^{-1}\tilde{s}_{-1}$$

for $h, k \in \mathrm{SU}(2)$.

The map $g\colon M \to \mathrm{SU}(2)$ is thus well defined modulo right and left actions of SU(2), i.e., the action of $\mathrm{SO}(4) = \mathrm{SU}(2) \times \mathrm{SU}(2)/\pm 1$ on $S^3 = \mathrm{SU}(2)$.

Remarks (8.25). (1) The invariance condition $(\sigma\tau)^* L_x \cong L_x$ can be written as $\tau^* L_x \cong L_x^{-1}$ which means that L_x lies in the *Prym variety* of Σ with respect to the involution τ. Since τ has four fixed points $(\pi^{-1}\{0, \infty\})$, the Riemann-Hurwitz formula gives the genus of Σ/τ as $\frac{1}{2}(p-1)$ and hence the dimension of the Prym variety as $p - \frac{1}{2}(p-1) = \frac{1}{2}(p+1)$. In the guise of the sinh-Gordon equation or the related sine-Gordon equation, the linearization on a Prym variety is well known ([2], [19]).

(2) The constraint (i) in Theorem (8.20) can be more analytically described (at least for a smooth curve Σ) by using the *reciprocity law* for differentials of the second and third kinds (see [13]). If we put

$$\phi = \frac{\eta(1)\, d\zeta}{\eta(\zeta - 1)}, \tag{8.26}$$

where $\eta(1)^2 = (-1)^{p+1} \prod_{i=1}^{p+1}(\overline{\alpha}_i - (\alpha_i\overline{\alpha}_i + 1) + \alpha_i)$, then ϕ is a meromorphic differential with simple poles at the two points $(\xi_1, \sigma\xi_1) \in \pi^{-1}\{1\}$ and with residues $(+1, -1)$. Moreover, using the real structure $\rho(\eta, \zeta) = (\overline{\eta\zeta}^{-(p+1)}, \overline{\zeta}^{-1})$, we have

$$\rho^*\phi = \frac{\eta(1)\overline{\zeta}^p d\overline{\zeta}}{\overline{\eta}(\overline{\zeta} - 1)};$$

so for $p > 0$, ϕ vanishes at ∞.

Let $\delta_1, \cdots, \delta_{2p}$ be cycles on Σ representing a canonical basis for $H_1(\Sigma; \mathbb{Z})$ with common base point ξ_0 and otherwise disjoint, and let N_i denote the periods of ϕ around these cycles. The differential of the second kind θ on Σ has by definition periods around these cycles of the form $2\pi i m_i$ $(m_i \in \mathbb{Z})$.

On the simply connected surface $\Sigma\backslash\bigcup\delta_i$ there is a well-defined function $f(\xi) = \int_{\xi_0}^{\xi}\theta$, and the reciprocity law states that

$$\sum_{i=1}^{p}(m_iN_{p+i} - m_{p+i}N_i) = \sum_{\xi}\mathrm{Res}_\xi(f\phi), \tag{8.27}$$

where the right-hand sum is over the poles of $f\phi$.

In our situation if $\zeta = 0$ is not a branch point, then f has simple poles with residues $\pm\lambda$ (see (3.5)) over $\zeta = 0$, and corresponding simple poles at ∞. If $\zeta = 0$ is a branch point then f has a simple pole with residue κ there (see (3.10)). The differential ϕ has simple poles with residues ± 1 over $\zeta = 1$ and vanishes at ∞. Hence (8.27) gives

$$\sum_{i=1}^{p}(m_iN_{p+i} - m_{p+i}N_i) = -2\lambda\frac{\eta(1)}{\eta(0)} + f(\xi_1) - f(\sigma\xi_1).$$

Now $\theta = d\mu/\mu$, so we can take f to be a single-valued branch of $\log\mu(\xi)$ in $\Sigma\setminus\bigcup\delta_i$ and since $\sigma^*\mu = \mu^{-1}$, this equation gives

$$\sum_{i=1}^{p}(m_iN_{p+i} - m_{p+i}N_i) = -2\lambda\frac{\eta(1)}{\eta(0)} + 2\log\mu(\xi_1). \tag{8.28}$$

If $\zeta = 0$ is a branch point, then near $\zeta = 0$, ϕ has the form

$$\phi = \frac{\eta(1)\,d\zeta}{\eta(\zeta - 1)} = (-1)^p\frac{2\eta(1)\,d\eta}{\prod_{i=1}^{p}\alpha_i} + \cdots$$

and so we obtain:

$$\sum_{i=1}^{p}(m_iN_{p+i} - m_{p+i}N_i) = (-1)^p\frac{2\kappa\eta(1)}{\prod\alpha_i} + 2\log\mu(\xi_1). \tag{8.29}$$

The first part of the constraint (i) now becomes, with $k \in \mathbb{Z}$,

$$(8.30) \qquad \begin{aligned} &\sum_{i=1}^{p}(m_i N_{p+i} - m_{p+i}N_i) + 2\lambda\frac{\eta(1)}{\eta(0)} = 2\pi i k, \text{ or} \\ &\sum_{i=1}^{p}(m_i N_{p+i} - m_{p+i}N_i) + (-1)^{p+1}\frac{2\kappa\eta(1)}{\prod \alpha_i} = 2\pi i k. \end{aligned}$$

To deal with the second part, we use the differential

$$\phi' = \frac{\eta(-1)\,d\zeta}{\eta(\zeta+1)}$$

with periods N_i' and have the constraint

$$(8.31) \qquad \begin{aligned} &\sum_{i=1}^{p}(m_i N'_{p+i} - m_{p+i}N'_i) + 2\lambda\frac{\eta(-1)}{\eta(0)} = 2\pi k', \text{ or} \\ &\sum_{i=1}^{p}(m_i N'_{p+i} - m_{p+i}N'_i) + (-1)^{p+i}\frac{2\kappa\eta(-1)}{\prod \alpha_i} = 2\pi i k. \end{aligned}$$

Note that the *consistency* of (8.30) and (8.31) constrains the equation of the curve Σ. If the equations are consistent then the *principal part* (λ or κ) of the differential is determined by the curve (and a choice of integers). The integrality of the periods of $\theta/2\pi i$ thus imposes possibly further constraints.

Clearly the conditions on $\tilde{\mu}$ in (8.20) can be found by replacing θ by $\tilde{\theta}$ in the above analysis. Note that there is a choice in defining ϕ, depending on which point in $\pi^{-1}\{1\}$ has residue $+1$ and which has -1, thus the consistency condition is between (8.30) and (8.31) or the equation obtained from (8.31) by changing the signs of all the periods N_j' $(1 \le j \le 2p)$.

Theorem (8.20) reduces the question of finding harmonic maps of the 2-torus to the 3-sphere to those of algebraic geometry or just complete hyperelliptic integrals. Many natural questions are not immediately solvable in algebraic geometric terms, however. For example, the question of whether the map is an immersion or an embedding. One rather remarkable result is however clear, concerning deformations of harmonic maps. Note that in Theorem (8.1) we must choose a point E in the quaternionic Picard variety $\mathrm{Pic}^{p+1}(\Sigma)$ to correspond to the origin in $M = U/\Gamma$, the torus. However, condition (i) of Theorem (8.20) makes no reference to this point. Thus, given one harmonic map we have a p-dimensional family of maps, by choosing the point E arbitrarily. Clearly choosing E to be in the image of the map of M to $\mathrm{Pic}^{p+1}(\Sigma)$ given by the eigenspace bundle is the equivalent to a translation of M itself, i.e., a conformal automorphism of M, but if $p > 2$ there is more freedom and we obtain a $(p-2)$-dimensional family of deformations of the harmonic map. If the spectral curve is nonsingular, this family is a $(p-2)$-torus. These deformations preserve μ and $\tilde{\mu}$ and are therefore *isospectral deformations* of the family of flat connections $\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*$. Just what geometrical properties of the map are preserved other than the energy is not clear. In the final few sections we shall consider low genus solutions and show in particular in §12 that there are indeed examples of this phenomenon with $p = 3$.

9. Rational solutions

We consider here the case where Σ is a *rational curve* (i.e., has genus $p = 0$) and look for the corresponding harmonic maps. From §4, the curve is given by the equation

$$\eta^2 = -\overline{\alpha}\zeta^2 + (\alpha\overline{\alpha} + 1)\zeta - \alpha \tag{9.1}$$

for some $\alpha \in \mathbb{C}$, $|\alpha| < 1$, and having the real structure

$$\rho(\eta, \zeta) = (\overline{\eta\zeta}^{-1}, \overline{\zeta}^{-1}). \tag{9.2}$$

Consider first the case $\alpha \neq 0$, in other words let us seek the *nonconformal* harmonic maps.

Since Σ is simply-connected, the differentials θ and $\tilde{\theta}$ have zero periods and we can find single-valued functions $\log\mu$ and $\log\tilde{\mu}$ such that $\theta = d(\log\mu)$ and $\tilde{\theta} = d(\log\tilde{\mu})$. The function $\log\mu$ has only simple poles at $\zeta = 0$ and $\zeta = \infty$, and hence is of the form

$$\log\mu = \zeta^{-1}(a\eta\zeta + b\eta + c + d\zeta + e\zeta^2).$$

On the other hand, we require $\sigma^*\mu\cdot\mu = 1$, hence

$$\log\mu = \zeta^{-1}(a\eta\zeta + b\eta) + \pi i k \quad \text{for } k \in \mathbb{Z}.$$

We also need the reality condition $\rho^*\mu = \overline{\mu}^{-1}$ which yields

$$\log\mu = a\eta - \overline{a}\zeta^{-1}\eta + \pi i k \tag{9.3}$$

and similarly

$$\log\tilde{\mu} = \tilde{a}\eta - \overline{\tilde{a}}\zeta^{-1}\eta + \pi i\tilde{k}. \tag{9.4}$$

To obtain a harmonic map we need from Theorem (8.20) the constraint $\mu(\xi) = \tilde{\mu}(\xi) = 1$ for all $\xi \in \pi^{-1}\{1, -1\}$. Now when $\zeta = 1$ in (9.1), $\eta^2 = (1+\alpha\overline{\alpha}) - (\alpha + \overline{\alpha})$, so from (9.3)

$$(a - \overline{a})\sqrt{(1+\alpha\overline{\alpha}) - (\alpha+\overline{\alpha})} \in \pi i \mathbb{Z}$$

and similarly for $\zeta = -1$ and for $\tilde{\mu}$. We obtain constraints:

$$\begin{aligned} &(a - \overline{a})\sqrt{(1+\alpha\overline{\alpha}) - (\alpha+\overline{\alpha})} = \pi i m, \\ &i(a + \overline{a})\sqrt{(1+\alpha\overline{\alpha}) + (\alpha+\overline{\alpha})} = \pi i n, \\ &(\tilde{a} - \overline{\tilde{a}})\sqrt{(1+\alpha\overline{\alpha}) - (\alpha+\overline{\alpha})} = \pi i \tilde{m}, \\ &i(\tilde{a} + \overline{\tilde{a}})\sqrt{(1+\alpha\overline{\alpha}) + (\alpha+\overline{\alpha})} = \pi i \tilde{n}. \end{aligned} \tag{9.5}$$

(Since $(1+\alpha\overline{\alpha}) \pm (\alpha+\overline{\alpha})$ is positive we may take the positive square root in these equations.) Consider now the behavior of $\log\mu$ as $\zeta \to 0$. Since $\eta^2 = -\alpha$ at $\zeta = 0$, we have from (9.3) and (9.4),

$$\log\mu = -\frac{\overline{a}\sqrt{-\alpha}}{\zeta} + \cdots, \qquad \log\mu = -\frac{\overline{\tilde{a}}\sqrt{-\alpha}}{\zeta} + \cdots,$$

and so the torus has modulus

$$\tau = \overline{\tilde{a}}/\overline{a} \tag{9.6}$$

by comparison with (3.5).

Putting $r = \sqrt{(1+\alpha\overline{\alpha}) + (\alpha+\overline{\alpha})}$ and $s = \sqrt{(1+\alpha\overline{\alpha}) - (\alpha+\overline{\alpha})}$ from (9.5), we find

$$a = \frac{\pi n}{2r} + \frac{\pi i m}{2s}, \qquad \tilde{a} = \frac{\pi \tilde{n}}{2r} + \frac{\pi i \tilde{m}}{2s},$$

so that from (9.6)

$$\tau = \frac{\tilde{n} + i\tilde{m}x}{n + imx}, \quad \text{where } x = -\frac{r}{s}. \tag{9.7}$$

(9.7) states that the torus is defined by a sublattice of the lattice generated by $\{1, ix\}$ with x real. In other words, it is a finite covering of a *rectangular torus*.

Consider next the case of a *conformal* map. Here from Theorem (8.20) we need $P(0) = 0$, so the spectral curve has equation

$$\eta^2 = \zeta. \tag{9.8}$$

In this case $\log\mu$ has poles at the zeros of η, and putting in the σ-invariance and reality conditions we have

$$\log\mu = a\eta - \overline{a}\eta^{-1} + \pi i k, \qquad \log\mu = \tilde{a}\eta - \overline{\tilde{a}}\eta^{-1} + \pi i \tilde{k}. \tag{9.9}$$

Substituting in the condition $\mu(\xi) = \tilde{\mu}(\xi) = 1$ for $\xi \in \pi^{-1}\{1, -1\}$ gives $(a - \overline{a}) = \pi i m$, $i(a + \overline{a}) = \pi i n$, $(\tilde{a} - \overline{\tilde{a}}) = \pi i \tilde{m}$, and $i(\tilde{a} + \overline{\tilde{a}}) = \pi i \tilde{n}$, and hence

$$(9.10) \qquad a = \frac{\pi}{2}(n + im), \qquad \tilde{a} = \frac{\pi}{2}(\tilde{n} + i\tilde{m}).$$

Here the possible moduli for a harmonic conformal map are the finite coverings of a *square torus.* Note from (6.13) and (6.14) that the Clifford torus corresponds to $n = -1$, $m = 1$, $\tilde{n} = 1$, $\tilde{m} = 1$.

To actually find the harmonic map in the rational case we may remark first of all that since the Picard variety of Σ is trivial, there is a natural identification of the line bundles E_x^* and E_y^* for $x, y \in U/\Gamma = M$. This means that the flat connection constructed in Theorem (8.1) is actually *translation-invariant* on the torus. Trivializing the bundle with the group action, the Higgs field Φ may be written as

$$(9.11) \qquad \Phi = \phi\, dz$$

for a constant trace-free matrix ϕ, and the $(0, 1)$ part of the connection as

$$(9.12) \qquad d_A'' = d'' + \psi\, d\overline{z}$$

for a trace-free complex matrix ψ.

The equations (1.7) then become

$$(9.13) \qquad [\psi, \phi] = 0, \qquad [\psi, \psi^*] = [\phi, \phi^*].$$

The flat connection is

$$d + \psi\, d\overline{z} - \psi^* dz + \zeta^{-1}\phi\, dz - \zeta\phi^*\, d\overline{z}$$

and so, solving the parallel translation equation by exponentiating the constant coefficient matrix, we have

$$(9.14) \qquad \begin{aligned} H(\zeta) &= \exp(-\psi + \psi^* - \zeta^{-1}\phi + \zeta\phi^*), \\ \tilde{H}(\zeta) &= \exp(-\psi\overline{\tau} + \psi^*\tau - \zeta^{-1}\phi\tau + \zeta\phi^*\overline{\tau}), \end{aligned}$$

and hence

$$(9.15) \qquad \begin{aligned} \log\mu &= \sqrt{-\det(\psi - \psi^* + \zeta^{-1}\phi - \zeta\phi^*)}, \\ \log\tilde{\mu} &= \sqrt{-\det(\psi\overline{\tau} - \psi^*\tau + \zeta^{-1}\phi\tau - \zeta\phi^*\overline{\tau})}. \end{aligned}$$

The harmonic map itself is defined by

(9.16)
$$\begin{aligned} g(y_1, y_2) = {} & \exp(-y_1(\psi - \psi^* + \phi - \phi^*) - iy_2(-\psi - \psi^* + \phi + \phi^*)) \\ & \cdot \exp(y_1(\psi - \psi^* - \phi + \phi^*) + iy_2(-\psi - \psi^* - \phi - \phi^*)), \end{aligned}$$

where $z = y_1 + iy_2 \in \mathbb{C}$, the universal covering of M.

The equations (9.13) are easily solved: the first implies that if $\phi \neq 0$ then $\psi = \lambda\phi$, and the second then reads

$$|\lambda|^2[\phi, \phi^*] = [\phi, \phi^*].$$

Now if $|\lambda| \neq 1$, then $[\phi, \phi^*] = 0$ so ϕ is normal and can be diagonalized by an SU(2) transformation. Since ψ commutes with ϕ, ψ will be diagonal too, but this means that both connection and Higgs field are simultaneously reduced to the structure group U(1), in which case there are *no* branch points for the spectral curve, and moreover any harmonic map has the circle as an image.

Thus we have $|\lambda| = 1$, so $\psi = e^{i\theta}\phi$. The general rational harmonic map can therefore be written as

$$\begin{aligned} g(z) = \; & \exp(-(z + e^{i\theta}\bar{z})\phi + (\bar{z} + e^{-i\theta}z)\phi^*) \\ & \cdot \exp(-(z - e^{i\theta}\bar{z})\phi + (\bar{z} - e^{-i\theta}z)\phi^*). \end{aligned}$$

Let us consider from this point of view the construction of the *conformal* map corresponding to the integers $(m, n, \tilde{m}, \tilde{n})$ in (9.10). Since $\det\Phi = 0$, ϕ is nilpotent and from (9.13) there is an SU(2)-basis such that

$$\phi = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 0 & e^{i\theta}\alpha \\ 0 & 0 \end{pmatrix}. \tag{9.17}$$

Now $\det(\psi - \psi^* + \zeta^{-1}\phi - \zeta\phi^*) = e^{-i\theta}|\alpha|^2/\zeta + \cdots$, so from (9.15)

$$\log\mu = ie^{-i\theta/2}|\alpha|/\eta + \cdots,$$

and so from (9.9) and (9.10)

$$i|\alpha|e^{i\theta/2} = \frac{\pi}{2}(n + im), \tag{9.18}$$

i.e.,

$$\alpha = \frac{\pi}{2}(n + im)e^{i\beta} \quad \text{for some } \beta \in \mathbb{R}. \tag{9.19}$$

The $e^{i\beta}$ factor may be removed by an overall conjugation by the matrix

$$\begin{pmatrix} e^{i\beta/2} & 0 \\ 0 & e^{-i\beta/2} \end{pmatrix},$$

and then (9.16) gives

$$g(y_1, y_2) = \exp\left(\pi(my_2 - ny_1)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)\exp\left(-\pi(my_1 + ny_2)\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\right).$$

After an SO(4) rotation this is the map into $S^3 \subset \mathbb{C}^2$ given by

$$z_1 = \frac{1}{\sqrt{2}} e^{i\pi((m+n)y_1-(m-n)y_2)}, \qquad z_2 = \frac{1}{\sqrt{2}} e^{i\pi((m-n)y_1+(m+n)y_2)},$$

which maps onto the Clifford torus $z_1\bar{z}_1 = z_2\bar{z}_2$.

Thus, the only conformal harmonic maps arising from a *rational* spectral curve are finite coverings of the Clifford torus.

10. Elliptic solutions

The spectral curve for genus one is of the form

$$\eta^2 = \prod_{i=1}^{2}(\bar{\alpha}_i\zeta^2 - (\alpha_i\bar{\alpha}_i + 1)\zeta + \alpha_i) \tag{10.1}$$

for $\alpha_i \in \mathbb{C}$, $|\alpha_i| < 1$ and real structure

$$\rho(\eta, \zeta) = (\overline{\eta\zeta}^{-2}, \bar{\zeta}^{-1}). \tag{10.2}$$

Since the normalization of a singular elliptic curve is rational, we know from §5 that Σ must be nonsingular, so that α_1, $\bar{\alpha}_1^{-1}$, α_2, $\bar{\alpha}_2^{-1}$ are distinct.

Consider now the question of finding differentials θ and $\tilde{\theta}$ with periods in $2\pi i\mathbb{Z}$. From the exact sequence of sheaves

$$0 \to \mathscr{O} \to \mathscr{O}(2) \to \mathscr{O}(2)_D \to 0$$

in (7.8) we have the long exact cohomology sequence

$$0 \to \mathbb{C} \to H^0(\Sigma; \mathscr{O}(2)) \to H^0(D; \mathscr{O}(2)) \xrightarrow{\delta} H^1(\Sigma; \mathscr{O}) \to H^1(\Sigma; \mathscr{O}(2)).$$

Here $H^1(\Sigma; \mathscr{O})$ is 1-dimensional (since $p = 1$) and from the arguments following (8.7), $H^0(\Sigma; \mathscr{O}(2))$ is the 4-dimensional space spanned by η and the pull-back of sections of $\mathscr{O}(2)$ on $\mathbb{C}P^1$. By Riemann-Roch, $H^1(\Sigma; \mathscr{O}(2)) = 0$ and so the map δ is surjective.

We may therefore find a class $x \in H^0(D; \mathscr{O}(2))$ satisfying $\sigma^*(x) = -x$ and $\rho^*(x) = -\bar{x}$ such that $\delta(x)$ is a real integral class in $H^1(\Sigma; \mathscr{O})$.

Consider the long exact cohomology sequence of the sequence of sheaves

$$0 \to \mathscr{O}(K) \to d\mathscr{M} \to \mathscr{O}(2)_D \to 0.$$

We have $H^1(\Sigma; \mathscr{O}(K)) \cong \mathbb{C}$ acted on trivially by σ but since $\sigma^*(x) = -x$, it follows that there exists a differential $\beta \in H^0(\Sigma; d\mathscr{M})$ with principal part x.

Decomposing $H^1(\Sigma;\mathbb{C}) \cong H^{1,0} \oplus H^{0,1}$, β has period class $(b, \delta(x))$. Hence adding on a holomorphic differential with period $-b-\overline{\delta(x)}$ we find a differential with the same principal part and period class $(-\overline{\delta(x)}, \delta(x))$ which therefore has periods in $2\pi i\mathbb{Z}$, since $\delta(x)$ is integral.

In fact we may use this argument for a general smooth curve Σ to show that for any $x \in U \subset H^0(D; \mathscr{O}(2))$, there is a differential β with principal part x and *imaginary* periods. Since $\sigma^*\beta + \beta$ and $\rho^*\overline{\beta} + \beta$ are holomorphic differentials with imaginary periods, they must necessarily be zero, so β transforms appropriately under σ and ρ.

Now since $\rho^*\beta = -\overline{\beta}$, the integral over a cycle A which is real (with respect to ρ) is imaginary and over an imaginary cycle B is real. But if all periods are of the form $2\pi i\mathbb{Z}$, then the B-period must be zero. Thus we can find β with periods $2\pi i$ and 0.

If a differential γ has zero periods and double poles at 0 and ∞, then $\gamma = df$ for some meromorphic function of the form

$$f = \frac{1}{\zeta}(a\eta + b + c\zeta + d\zeta^2);$$

if it satisfies $\sigma^*\gamma = -\gamma$ and $\rho^*\gamma = -\overline{\gamma}$, then

$$f = ia\eta/\zeta \quad \text{for } a \in \mathbb{R}.$$

Thus, the lattice generated by differentials $\{\theta, \tilde{\theta}\}$ is a sublattice of that generated by the differentials $\{\beta, iad(\eta/\zeta)\}$ for some $a \in \mathbb{R}$. Conversely, choosing a, we obtain differentials satisfying the conditions of Theorem (8.1) and hence a solution to (1.7) for any curves of the form (10.1).

Let us consider now the question of finding the harmonic maps corresponding to $\theta = n\beta$, $\tilde{\theta} = iad(\eta/\zeta)$. If we choose A, B as the canonical basis for $H_1(\Sigma;\mathbb{Z})$, then in the nonconformal case we have the constraints (8.30) and (8.31) given as

$$nN_2 + 2\lambda\frac{\eta(1)}{\eta(0)} = 2\pi ik, \qquad \pm nN_2' + 2\lambda\frac{\eta(-1)}{\eta(0)} = 2\pi ik', \tag{10.3}$$

and the constraint for the differential $iad(\eta/\zeta)$ is simply

$$ia\eta(1) = 2\pi il, \qquad ia\eta(-1) = 2\pi il'. \tag{10.4}$$

From (10.4), $l'\eta(1) - l\eta(-1) = 0$ and putting this in (10.3) we obtain

$$l'N_2 \pm lN_2' = 2\pi i(kl' - k'l)/n. \tag{10.5}$$

Thus the two constraints are firstly

$$\eta(1)/\eta(-1) = l/l' \in \mathbb{Q}, \tag{10.6}$$

and secondly that the period of the differential

$$\frac{1}{2\pi i}\left(\frac{l'\eta(1)}{(\zeta-1)} \pm \frac{l\eta(-1)}{(\zeta+1)}\right)\frac{d\zeta}{\eta} \tag{10.7}$$

over an imaginary cycle B should be rational.

Conversely, if $\eta(1)/\eta(-1) = l/l'$ with l, l' mutually pime, and if $l'N_2 - lN_2' = 2\pi i m/n \in \mathbb{Q}$, then we may write $m = kl' - k'l$ for k, $k' \in \mathbb{Z}$ and find consistent solutions λ, a to (10.3) and (10.4). Moreover, as above we may find a differential of the second kind satisfying $\sigma^*\theta = -\theta$, $\rho^*\theta = -\overline{\theta}$, having principal part $\pm\lambda\, d\zeta/\zeta^2$ over 0 and zero period over B. Then the reciprocity law (10.3) shows that this differential has period $2\pi i n$, hence (10.6) and (10.7) are the only required constraints. The first is an algebraic constraint on the two complex numbers α_1, α_2 and the second one transcendental.

The case of a *conformal* map is exactly the same, using the second formula of (8.30) and (8.31). In this case, since $\zeta = 0$ and ∞ are branch points, the spectral curve has equation

$$\eta^2 = \zeta((\alpha\overline{\alpha}+1)\zeta - \alpha - \overline{\alpha}\zeta^2).$$

The constraint (10.6) is then

$$\sqrt{\frac{(\alpha\overline{\alpha}+1)-(\alpha+\overline{\alpha})}{(\alpha\overline{\alpha}+1)+(\alpha+\overline{\alpha})}} = \frac{l}{l'} \in \mathbb{Q}, \tag{10.8}$$

and (10.7) may be written as

(10.9)

$$\sqrt{(\alpha\overline{\alpha}+1)-(\alpha+\overline{\alpha})}\,\frac{1}{\pi i}\int_{\alpha}^{\overline{\alpha}^{-1}} \frac{d\zeta}{\sqrt{(\alpha\overline{\alpha}+1)\zeta^2 - \alpha\zeta - \overline{\alpha}\zeta^3}(\zeta^2-1)} \in \mathbb{Q},$$

since the covering of a path joining the two branch points α and $\overline{\alpha}^{-1}$ gives an imaginary cycle B in $H_1(\Sigma;\mathbb{Z})$.

The solutions to (1.7) which have elliptic spectral curves possess a group invariance property just like the rational case. In fact the line bundles L_x for $x \in \ker\delta \subset U$ can all be identified, so the solution is invariant under the 1-dimensional group of translations of $M = U/\Gamma$ parallel to $\ker\delta$. Conversely, since δ is injective for genus $p > 1$ (see §7), any solution with a 1-dimensional invariance group must have genus 0 or 1. As we saw in §9 there is actually a 2-dimensional group for $p = 0$.

Since the bundle V is pulled back from $U/\ker\delta$, there is a gauge in which the invariant connection A has matrix $A_1(x)\,dy$ where y is a linear coordinate such that $\delta(\partial/\partial y) = 0$ and $z = x + iy$. The Higgs field can be written

$$\Phi = \tfrac{1}{2}(A_2(x) + iA_3(x))\,dx,$$

where the $A_i(x)$ take values in the Lie algebra of $\mathrm{SU}(2)$.

The equations (1.7) can then be written

$$\frac{d}{dx}(A_2 + iA_3) = -i[A_1, A_2 + iA_3], \qquad \frac{dA_1}{dx} = \frac{1}{2}i[A_2 + iA_3, A_2 - iA_3],$$

or equivalently

$$\frac{dA_1}{dx} = [A_2, A_3], \quad \frac{dA_2}{dx} = -[A_3, A_1], \quad \frac{dA_3}{dx} = -[A_1, A_2],$$

which we may call Schmid's equations [25], arising in the study of degenerating period matrices. They are closely related to Nahm's equations [22], and the method for solving these in [14] in terms of the geometry of an algebraic curve leads in this case to the same spectral curve Σ. The solution of (10.9) with $A_i(x)$ taking values in the Lie algebra of $\mathrm{U}(k)$ is determined by a curve of genus $(k-1)^2$, and in our context would correspond to harmonic maps into $\mathrm{U}(k)$ with a 1-dimensional isomorphism group.

Given this group invariance, the conformal harmonic maps with elliptic spectral curve map onto minimal tori in S^3 which are invariant under the action of a circle subgroup of $\mathrm{SO}(4)$. These were classified by Hsiang and Lawson [16], whose approach was to show the equivalence of this problem with finding closed geodesics on the quotient space S^3/S^1. This quotient is a surface of revolution, and the quadrature needed to integrate the geodesic flow involves elliptic integrals. The periodicity consists of the rationality of a certain elliptic integral of the third kind, which is essentially the constraint (10.7). It is also shown in [16] that a minimal torus with a 2-dimensional symmetry group is the Clifford torus, coinciding with our result in §9, derived from the spectral curve point of view.

Consider now the question of finding harmonic maps to the 2-sphere. Here, from Theorem (8.20), we require the spectral curve and differentials θ, $\tilde{\theta}$ to be invariant under the involution $\sigma\tau(\eta, \zeta) = (-\eta, -\zeta)$. Thus the curve has equation

$$\eta^2 = (\alpha\bar{\alpha} + 1)^2\zeta^2 - (\alpha + \bar{\alpha}\zeta^2)^2 \quad \text{for some } \alpha \in \mathbb{C},\ |\alpha| < 1. \tag{10.10}$$

It is clear from this equation that $\eta(1)^2 = \eta(-1)^2$, so $\eta(1)/\eta(-1) = \pm 1$ and (10.6) is automatically satisfied. Moreover,

$$(\sigma\tau)^* \frac{d\zeta}{\eta(\zeta - 1)} = \frac{-d\zeta}{\eta(\zeta + 1)}. \tag{10.11}$$

Since τ is a holomorphic involution with no fixed points, it must be a translation by a half-period. In particular, it transforms the imaginary cycle B to a homologous cycle, so

$$\int_B \frac{d\zeta}{\eta(\zeta-1)} = \int_{\tau(B)} \frac{d\zeta}{\eta(\zeta-1)},$$

and so from (10.11)

$$\int_B \frac{d\zeta}{\eta(\zeta-1)} = -\int_B \frac{d\zeta}{\eta(\zeta+1)},$$

i.e.,

$$N_2 + N_2' = 0, \tag{10.12}$$

which is the constraint (10.7).

Thus any spectral curve of the form (10.10) gives a harmonic map to the 2-sphere.

One classical family of such maps arises from the *Delaunay surfaces* [9]. These are surfaces of revolution of constant mean curvature. They are constructed by rotating a plane curve traced out by the focus of an ellipse rolling along the x-axis. The Gauss map is clearly doubly periodic, one period being the length of the circumference of the ellipse, and so defines a harmonic map to the 2-sphere. Explicit formulas involve integrating the arc length of an ellipse and hence *elliptic functions.* On the other hand, we know that the rotational symmetry of the map about the x-axis must lead to an elliptic spectral curve, and hence one of the form (10.10). In fact, a surface of revolution has reflectional symmetry about a meridian. This reflection therefore defines an antiholohorphic involution ν of the surface M and the Gauss map takes this to a standard reflection of the two-sphere into itself, fixing an equator. This equatorial reflection is induced by the isometry $g \to -g^{-1}$ on SU(2) which leaves the Levi-Civita connection fixed but interchanges the left- and right-invariant connections. Pulling back to M, this means that

$$\nu^*(\nabla_A + \zeta^{-1}\Phi - \zeta\Phi^*) = \nabla_A + \zeta^{-1}\Phi^* - \zeta\Phi. \tag{10.13}$$

Thus if α is a branch point, so is $-\alpha^{-1}$. Then since the curve is nonsingular, $-\alpha^{-1} = \pm\bar{\alpha}^{-1}$, so α is real or imaginary.

Actually α is *imaginary*, for (10.13) implies that $\nu^*\Phi = \Phi^*$ and thus if $\Phi = \phi\, dz$, the eigenvalues of ϕ are real on the fixed point set of ν: a meridian. By rotational invariance they are real everywhere, so λ is real. The modulus of the torus M is given by $\tau = ia\eta(0)/\lambda = \pm a\alpha/\lambda$ which, to be in the upper half-plane, must have α imaginary. Thus the spectral curves of the Delaunay surfaces have the form (with $\alpha = ir$):

$$\eta^2 = (1+r^2)^2\zeta^2 + r^2(1-\zeta^2)^2.$$

11. Genus two solutions

Let us now consider spectral curves Σ of arithmetic genus 2. Because the parity of the arithmetic genus p and geometric genus g are the same (see §5), then if Σ is singular $g = 0$. But we already saw in §5 that this is impossible, so Σ is nonsingular. The spectral curve is always therefore of the form

$$\eta^2 = -\prod_{i=1}^{3}(\overline{\alpha}_i\zeta^2 - (\alpha_i\overline{\alpha}_i + 1)\zeta + \alpha_i) \tag{11.1}$$

for α_1, α_2, α_3 distinct with $|\alpha_i| < 1$ and real structure

$$\rho(\eta, \zeta) = (\overline{\eta\zeta}^{-3}, \overline{\zeta}^{-1}).$$

Since $p > 1$, we see from §7 that the map $\sigma\colon U \to H^1(\Sigma; \mathscr{O})$ is injective into the real points of $H^1(\Sigma; \mathscr{O})$. On the other hand $\dim U = 2$ and $\dim H^1(\Sigma; \mathscr{O}) = p = 2$, so δ is an isomorphism. Hence any real integral class is in the image of δ, and so as in §10 we can find independent differentials θ and $\tilde{\theta}$ with periods in $2\pi i\mathbb{Z}$. Consequently, from Theorem (8.1) any curve of the form (11.1) is the spectral curve for a solution of (1.7). We now look for an example which gives a harmonic map.

We consider, as in the elliptic case, a branch locus which is invariant under the involution $\zeta \to -\zeta$. Thus $(\alpha_1, \alpha_2, \alpha_3) = (0, \alpha, -\alpha)$ and the spectral curve is

$$\eta^2 = \zeta((1+\alpha\overline{\alpha})^2\zeta^2 - (\alpha + \overline{\alpha}\zeta^2)^2). \tag{11.2}$$

Since $\zeta = 0$ is a branch point, we are looking for a *conformal* harmonic map to the sphere. Now the involution $\zeta \to -\zeta$ is covered by the transformation τ of Σ:

$$\tau(\eta, \zeta) = (i\eta, -\zeta), \tag{11.3}$$

which is of order four, and commutes with the real structure ρ.

Take for $H_1(\Sigma; \mathbb{Z})$ as a canonical basis $\{A, \rho(A), B, -\rho(B)\}$ where A is represented by the closed path in Σ which covers the straight line segment $[0, \alpha]$ in the unit disc $D \subset \mathbb{C}P^1$, and B the path which covers $[0, -\alpha]$. Then

$$\tau(A) = B, \qquad \tau(B) = -A. \tag{11.4}$$

If $\phi = \eta(1)\,d\zeta/[\eta(\zeta - 1)]$ and $\phi' = \eta(1)\,d\zeta/[\eta(\zeta+1)]$ are our two basic differentials, then (11.3) gives

$$\tau^*\phi = \phi', \qquad \tau^*\phi' = -\phi. \tag{11.5}$$

Since τ maps $\pi^{-1}(D) \subset \Sigma$ to itself, the cycles A, B are contained in $\pi^{-1}(D)$, and ϕ, ϕ' have no poles in $\pi^{-1}(D)$, we have from (11.4) and (11.5)

$$\int_A \phi' = \int_A \tau^*\phi = \int_{\tau(A)} \phi = \int_B \phi,$$
$$\int_B \phi' = \int_B \tau^*\phi = \int_{\tau(B)} \phi = -\int_A \phi,$$

and hence if (N_i, N_i') are the periods of (ϕ, ϕ') over the given basis for $H_1(\Sigma; \mathbb{Z})$, then

$$N_1' = N_3, \qquad N_3' = -N_1. \tag{11.6}$$

Since τ commutes with ρ, we also have

$$N_2' = -N_4, \qquad N_4' = N_2. \tag{11.7}$$

If the periods of θ are $2\pi i m_i$ over this basis, then since the periods over the imaginary classes $A - \rho(A)$ and $B - \rho(B)$ are zero we have $m_1 = m_2 = m$ and $m_3 = -m_4 = n$, and then the constraints (8.30) and (8.31) for a conformal harmonic map become

$$\begin{aligned} mN_3 - nN_1 + mN_4 + nN_2 + \frac{2\kappa}{\alpha^2}\eta(1) &= 2\pi i k, \\ mN_3' - nN_1' + mN_4' + nN_2' + \frac{2\kappa}{\alpha^2}\eta(-1) &= 2\pi i k. \end{aligned} \tag{11.8}$$

Using $\eta(-1) = i\eta(1)$, (11.6) and (11.7), this gives

$$\begin{aligned} m(N_3 + N_4) - n(N_1 - N_2) + \frac{2\kappa\eta(1)}{\alpha^2} &= 2\pi i k, \\ -m(N_1 - N_2) - n(N_3 + N_4) + \frac{2\kappa i\eta(1)}{\alpha^2} &= 2\pi i k', \end{aligned} \tag{11.9}$$

which are consistent only if

$$(N_1 - N_2) + i(N_3 + N_4) = -2\pi\frac{(k + ik')}{(m - in)}. \tag{11.10}$$

Conversely, suppose there exist k, k', m, $n \in \mathbb{Z}$ such that (11.10) is satisfied. Then defining κ by the first equation of (11.9) we take the corresponding differential θ with principal part given by κ and having imaginary periods. By the reciprocity law its periods $(2\pi ix, 2\pi iy)$ over (A, B) satisfy (11.9) with (m, n) replaced by (x, y). But if

$$(N_3 + N_4)^2 + (N_1 - N_2)^2 \neq 0, \tag{11.11}$$

then this linear equation for (x, y) has the unique solution (m, n), so θ has periods $(2\pi im, 2\pi in)$.

If $\tilde{\theta}$ has periods $(2\pi i\tilde{m}, 2\pi i\tilde{n})$, then there exist integers $\tilde{k}$, $\tilde{k}'$ such that

$$\tilde{m}(N_3 + N_4) - \tilde{n}(N_1 - N_2) + \frac{2\tilde{\kappa}\eta(1)}{\alpha^2} = 2\pi i k,$$
$$-\tilde{m}(N_1 - N_2) - \tilde{n}(N_3 + N_4) + \frac{2\tilde{\kappa}i\eta(1)}{\alpha^2} = 2\pi i\tilde{k}',$$

and hence just as above, we obtain

$$\frac{\tilde{k} + i\tilde{k}'}{\tilde{m} - i\tilde{n}} = \frac{k + ik'}{m - in}.$$

So if $(k + ik')$, then $(m - in)$ are mutually prime in $\mathbb{Z}[i]$, $(\tilde{k} + i\tilde{k}') = (a + ib)(k + ik')$ and $(\tilde{m} - i\tilde{n}) = (a + ib)(m - in)$, in which case $\tilde{\kappa} = (a + ib)$, and since the modulus of the torus M is from (3.10) given by $\tilde{\kappa}/\kappa = (a + ib)$, we require $b \neq 0 \in \mathbb{Z}$.

Conversely, if (11.10) and (11.11) are satisfied, then given $a + ib \in \mathbb{Z}[i]$ with $b \neq 0$ we may find a differential $\tilde{\theta}$ with periods in $2\pi i\mathbb{Z}$ and independent of θ.

Thus to find the harmonic maps of this form, we simply need to find α such that the complete elliptic integral of the third kind

$$N(\alpha) = (N_1 - N_2) + i(N_3 + N_4) = \int_{A+\rho(A)+iB+i\rho(B)} \frac{\eta(1)\, d\zeta}{\eta(\zeta - 1)},$$

where $\eta^2 = \zeta((1 + \alpha\overline{\alpha})^2\zeta^2 - (\alpha + \overline{\alpha}\zeta^2)^2)$, takes a value of the form $-2\pi(k + ik')/(m - in)$, with k, k', m, n integers. Since such complex numbers are dense, we only need to show that the image of the complex-valued function $N(\alpha)$ contains an open set. By the inverse function theorem this will follow if we can find a value α such that the Jacobian determinant

$$\Delta = \begin{vmatrix} \operatorname{Re}\frac{\partial N}{\partial r} & \operatorname{Im}\frac{\partial N}{\partial r} \\ \operatorname{Re}\frac{\partial N}{\partial \theta} & \operatorname{Im}\frac{\partial N}{\partial \theta} \end{vmatrix} \tag{11.12}$$

is nonzero, where $\alpha = re^{i\theta}$.

Now for the curve (11.2) the differential ϕ is given by

$$\phi = \frac{\sqrt{(1+r^2)^2 - 4r^2\cos^2\theta}}{\sqrt{(1+r^2)^2\zeta^2 - r^2(e^{i\theta} + e^{-i\theta}\zeta^2)^2}} \cdot \frac{d\zeta}{\zeta^{1/2}(\zeta - 1)}.$$

Thus

$$\begin{aligned}\left.\frac{\partial\phi}{\partial\theta}\right|_{\theta=0} &= \frac{-ir^2(1-r^2)(1+\zeta)(1+\zeta^2)}{((1+r^2)^2\zeta^2 - r^2(1+\zeta^2)^2)^{3/2}} \frac{d\zeta}{\zeta^{1/2}} \\ &= -ir^2(1-r^2)\frac{(1+\zeta)(1+\zeta^2)}{q(\zeta)}\frac{d\zeta}{\eta},\end{aligned} \tag{11.13}$$

where $q(\zeta) = (1+r^2)^2\zeta^2 - r^2(1+\zeta^2)^2$, and

$$\begin{aligned}\left.\frac{\partial\phi}{\partial r}\right|_{\theta=0} &= \frac{\partial}{\partial r}\frac{(1-r^2)}{\sqrt{(1+r^2)^2\zeta^2 - r^2(1+\zeta^2)^2}}\frac{d\zeta}{\zeta^{1/2}(\zeta-1)} \\ &= -r(1+r^2)\frac{(1+\zeta)(1-\zeta^2)}{q(\zeta)}\frac{d\zeta}{\eta}.\end{aligned} \tag{11.14}$$

Now

$$\begin{aligned}\rho^*\frac{(1+\zeta)(1+\zeta^2)}{q(\zeta)}\frac{d\zeta}{\eta} &= \frac{(1+\bar\zeta^{-1})(1+\bar\zeta^{-2})}{q(\bar\zeta^{-1})}\frac{d\bar\zeta}{\bar\zeta^2\bar\eta\bar\zeta^{-3}} \\ &= -\frac{\bar\zeta^2(1+\bar\zeta)(1+\bar\zeta^2)}{q(\bar\zeta)}\frac{d\bar\zeta}{\bar\eta}.\end{aligned}$$

So if $\zeta = x$ is real, then

$$\frac{\partial\phi}{\partial\theta} - \rho^*\frac{\partial\phi}{\partial\theta} = -ir^2(1-r^2)\frac{(1+x^2)^2(1+x)}{q(x)}\frac{dx}{\eta(x)}, \tag{11.15}$$

and similarly,

$$\frac{\partial\phi}{\partial r} - \rho^*\frac{\partial\phi}{\partial r} = -r(1+r^2)\frac{(1+x^2)^2(1+x)}{q(x)}\frac{dx}{\eta(x)}. \tag{11.16}$$

Now $\partial\phi/\partial\theta$ and $\partial\phi/\partial r$ are differentials of the second kind of Σ with poles at the branch points r, $-r$, r^{-1}, $-r^{-1}$ which are the zeros of $q(x)$. Representing the cycle A by the covering of the real segment $[0, r]$, indented at r, we see from (11.15) that

$$\begin{aligned}\frac{\partial}{\partial\theta}(N_1 - N_2) &= \int_{A-\rho^*A}\frac{\partial\phi}{\partial\theta} = \int_A \frac{\partial\phi}{\partial\theta} - \rho^*\frac{\partial\phi}{\partial\theta} \\ &= -ir^2(1-r^2)\int_A \frac{(1+x^2)^2(1+x)}{q(x)}\frac{dx}{\eta(x)},\end{aligned}$$

and similarly from (11.16)

$$\frac{\partial}{\partial r}(N_1 - N_2) = -r(1+r^2)\int_A \frac{(1+x^2)(1+x)(1+x^2)}{q(x)}\frac{dx}{\eta(x)}.$$

Arguing analogously for N_3 and N_4 and the cycle B, we have

$$\frac{\partial N}{\partial r} = \frac{\partial}{\partial r}((N_1 + N_2) + i(N_3 + N_4)) = -r(1+r^2)(ia - b),$$
$$\frac{\partial N}{\partial \theta} = -ir^2(1-r^2)(ia-b),$$

where

$$(ia, ib) = \int_{(A,B)} \frac{(1+x^2)^2(1+x)}{q(x)}\frac{dx}{\eta(x)}.$$

Now $\eta^2(x) = x((1+r^2)^2x^2 - r^2(1+x^2)^2) < 0$ if $0 < x < r$, so a and b are real and so the determinant (11.12) is $-(a^2+b^2)r^3(1-r^4)$ which is only zero if a and b vanish, since $0 < r < 1$. If a vanishes for *all* r, then $\partial(N_1 - N_2)/\partial r$ vanishes, so $(N_1 - N_2)$ is constant. But $N_1 - N_2$ may be written as the integral

$$\frac{1}{2}\int_0^r - \int_\infty^{1/r} \frac{(1-r^2)}{\sqrt{(1+r^2)^2x^2 - r^2(1+x^2)^2}}\frac{dx}{x^{1/2}(x-1)}$$
$$= \frac{1}{2}\int_0^r \frac{(1+x^2)(1-r^2)}{\sqrt{(1+r^2)^2x^2 - r^2(1+x^2)^2}}\frac{dx}{x^{1/2}(x-1)},$$

which is nonconstant, in fact as $r \to 0$ it is asymptotically

$$-\frac{1}{r^{1/2}}\int_0^1 \frac{dt}{\sqrt{t^2-1}t^{1/2}}.$$

Thus there exists a real point α at which the derivative of N has maximal rank, and the inverse function theorem argument may be used.

Note that the sign ambiguity in the choice of ϕ' (or equivalently the transformation τ) means that the argument can be applied equally to the function $M = (N_1 - N_2) - i(N_3 + N_4)$. Since we have shown that neither M nor N is identically zero, neither is (by analyticity in α) $MN = (N_1 - N_2)^2 + (N_3 + N_4)^2$. Thus condition (11.11) is also satisfied.

We may deduce therefore the following.

Proposition (11.17). *There exist conformal minimal immersions of a square torus into the* 3*-sphere whose spectral curve is of genus* 2.

Proof. We merely note that a solution to (11.10) whose existence we have shown, together with the choice $\tilde{k} + i\tilde{k}' = i(k + ik')$, $(\tilde{m} - i\tilde{n}) = i(m - in)$, gives the modulus $\tau = i$, in other words a square torus. (The only other conformal structures obtainable this way have modulus equivalent under $\mathrm{SL}(2, \mathbb{Z})$ to li for some $l \in \mathbb{Z}$, and a covering of the square torus will give such a map in any case.)

12. Genus three solutions

A spectral curve of genus three is of the form

$$\eta^2 = \prod_{i=1}^{4}(\overline{\alpha}_i\zeta^2 - (\alpha_i\overline{\alpha}_i + 1)\zeta + \alpha_i)$$

for $|\alpha_i| < 1$, $1 \le i \le 4$.

In general we now have to impose constraints on the coefficients α_i in order to obtain differentials θ and $\tilde{\theta}$ with periods in $2\pi i\mathbb{Z}$ and so solutions to (1.7), and then even more constraints to obtain a harmonic map. We shall consider here the special class of examples obtained by setting

$$\alpha_1 = \alpha, \quad \alpha_2 = \overline{\alpha}, \quad \alpha_3 = -\overline{\alpha}, \quad \alpha_4 = -\alpha$$

and hence giving the spectral curve

$$\eta^2 = ((\alpha\overline{\alpha} + 1)^2\zeta^2 - (\alpha + \overline{\alpha}\zeta^2)^2)((\alpha\overline{\alpha} + 1)^2\zeta^2 - (\overline{\alpha} + \alpha\zeta^2)^2). \tag{12.1}$$

This curve Σ is invariant by an abelian group of transformations of order 16, generated by the involutions:

$$\begin{aligned} &\tau(\eta, \zeta) = (\eta, -\zeta), \quad \nu(\eta, \zeta) = (\overline{\eta}, \overline{\zeta}), \\ &\sigma(\eta, \zeta) = (-\eta, \zeta), \quad \rho(\eta, \zeta) = (\overline{\eta\zeta}^{-4}, \overline{\zeta}^{-1}). \end{aligned} \tag{12.2}$$

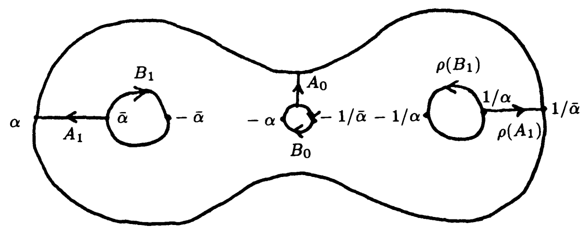
The last two are of course defined for any spectral curve, and the existence of the first one is one condition for a harmonic map to the 2-sphere.

If the four complex numbers α, $\overline{\alpha}$, $-\overline{\alpha}$, $-\alpha$ are distinct, then Σ is nonsingular and we begin by considering the action of τ and ν on the canonical basis for $H_1(\Sigma; \mathbb{Z})$:

$$\{A_0, A_1, \rho(A_1), B_0, B_1, -\rho(B_1)\}. \tag{12.3}$$

Using the fact that if C is represented by the cycle which covers the segment $[-\overline{\alpha}, -\alpha]$, then $C + A_0 = A_1$, we find

$$\begin{aligned} &\tau(A_0) = A_0, \qquad \tau(A_1) = -C = -A_1 + A_0, \\ &\tau(\rho(A_1)) = -\rho(A_1) + \rho(A_0) = -\rho(A_1) + A_0, \\ &\tau(B_0) = B_1 + B_0 - \rho(B_1), \qquad \tau(B_1) = -B_1, \\ &\tau(\rho(B_1)) = -\rho(B_1). \end{aligned} \tag{12.4}$$



Now take $x \in U \subset H^0(D; \mathscr{O}(2))$ and find, as in §10, the corresponding differential θ with principal part x and with imaginary periods. Since τ acts as -1 on U, then $\theta + \tau^*\theta$ is a *holomorphic* differential with imaginary periods and hence vanishes, so $\tau^*\theta = -\theta$.

Let the periods of θ with respect to the basis (12.3) be $2\pi i m_j$, $1 \leq j \leq 6$. Then firstly since $\rho^*\theta = -\overline{\theta}$, the periods over imaginary cycles vanish, so from the reality of the basis (12.3),

(12.5) $$m_2 = m_3 = m \quad \text{and} \quad m_5 = -m_6 = n.$$

Secondly, since $\tau^*\theta = -\theta$, (12.4) gives

(12.6) $$m_1 = -m_1 = 0 \quad \text{and} \quad m_4 = -m_4 = 0.$$

Consider next the action of the antiholomorphic involution ν on the canonical basis (12.3). We obtain immediately

$$\nu(A_0) = -A_0, \quad \nu(A_1) = -A_1, \quad \nu(\rho(A_1)) = -\rho(A_1).$$

Since ν is orientation reversing, it reverses the intersection form, so

(12.8) $$\nu(B_0) = B_0, \quad \nu(B_1) = B_1, \quad \nu(\rho(B_1)) = \rho(B_1).$$

Now the action of ν on the complex vector space U defines a real structure. If θ is a differential with ν-invariant principal part in U, then $\overline{\nu^*\theta} - \theta$ is a holomorphic differential with imaginary periods and so vanishes. Thus the periods of θ over ν-invariant classes vanish and so from (12.8)

(12.9) $$m_4 = m_5 = m_6 = 0.$$

Considering the basic differential

$$\phi = \frac{\eta(1)}{\eta} \frac{d\zeta}{(\zeta - 1)},$$

we have $\nu^*\phi = \overline{\phi}$ and so

$$\overline{N}_5 = \int_{B_1} \overline{\phi} = \int_{B_1} \nu^*\phi = \int_{\nu(B_1)} \phi = N_5,$$

and similarly for N_6. From (12.9), the constraint (8.3) for a harmonic map reads:

$$m(N_5 + N_6) + 2\lambda\eta(1)/\eta(0) = 2\pi ik. \tag{12.10}$$

But the left-hand side is real, so choosing an integer m and taking $k = 0$, we set $\lambda = -\frac{1}{2}m\eta(0)(N_5 + N_6)/\eta(1)$ and solve the constraint equation. Moreover, since τ interchanges -1 and $+1$ and $\tau^*\theta = -\theta$, the corresponding constraint (8.31) at -1 is automatically satisfied.

Similarly, taking $\tilde{\theta}$ as a differential with principal part anti-invariant under ν, we have $m_1 = m_2 = m_3 = 0$ and the condition:

$$n(N_3 - N_2) + 2\tilde{\lambda}\eta(1)/\eta(0) = 2\pi i\tilde{k}. \tag{12.11}$$

Here $N_3 - N_2$ is imaginary and the constraint can be satisfied by taking $\tilde{k} = 0$ and $\tilde{\lambda} = -\frac{1}{2}n\eta(0)(N_3 - N_2)/\eta(1)$. Conversely, (12.10) and (12.11) interpreted as statements of the reciprocity law show that the differentials θ and $\tilde{\theta}$ have periods in $2\pi i\mathbb{Z}$, as long as $N_3 - N_2$ and $N_5 + N_6$ are nonzero.

Now,

$$\begin{aligned} N_3 - N_2 &= \int_{\rho(A_1)} \phi - \int_{A_1} \phi = \int_{\nu(A_1)} (\nu\rho)^*\phi - \int_{A_1} \phi \\ &= \int_{A_1} (\phi + (\nu\rho)^*\phi) \quad (\text{from } (12.7)), \end{aligned}$$

$$\begin{aligned} \phi + (\nu\rho)^*\phi &= \eta(1)\left(\frac{d\zeta}{\eta(\zeta - 1)} + \frac{\zeta^3 d\zeta}{\eta(\zeta - 1)}\right) \\ &= \frac{\eta(1)}{\eta}(1 + \zeta + \zeta^2)\, d\zeta, \end{aligned}$$

so

$$N_3 - N_2 = 2\int_{\overline{\alpha}}^{\alpha} \frac{((1 + \alpha\overline{\alpha})^2 - (\alpha + \overline{\alpha})^2)(1 + \zeta + \zeta^2)\, d\zeta}{\sqrt{((1 + \alpha\overline{\alpha})^2\zeta^2 - (\alpha + \overline{\alpha}\zeta^2)^2)((1 + \alpha\overline{\alpha})^2\zeta^2 - (\overline{\alpha} + \alpha\zeta^2)^2}},$$

which, putting $\alpha = \rho e^{i\phi}$ and $\zeta = \rho e^{i\theta}$, is asymptotic as $\rho \to 0$ to the elliptic integral

$$\frac{i\sqrt{2}}{\rho} \int_{-\phi}^{\phi} \frac{d\theta}{\sqrt{\cos 2\theta - \cos 2\phi}},$$

which is nonzero.

Similarly, N_5+N_6 is asymptotic to the integral of the same holomorphic differential on an elliptic curve around an independent cycle. Since its periods cannot both be imaginary, then N_5+N_6 must be nonzero.

We have therefore proved:

Proposition (12.12). *There exist harmonic maps of a rectangular torus to the 3-sphere with spectral curve of genus 3.*

Note that the modulus of the torus is given by

$$\tau = \frac{\tilde{\lambda}}{\lambda} = \frac{n}{m}\frac{(N_3-N_2)}{(N_5+N_6)}.$$

As in Theorem (8.20), when the initial point E in the Picard variety lies in the *Prym variety* of the involution τ, we obtain a harmonic map to the 2-sphere. To understand this map more geometrically note that the vector space $U \cong \mathbb{C}$ admits the real structure ν and $\nu^*\theta = \overline{\theta}$, $\nu^*\tilde{\theta} = -\overline{\tilde{\theta}}$, so the lattice $\Gamma \subset U$ generated by the principal parts of θ and $\tilde{\theta}$ is invariant by the two reflections $\pm\nu$. If the initial point E is ν-invariant, then the harmonic map to the 2-sphere is equivariant with respect to the group of reflections generated by $\pm\nu$. In analytical terms Remark (1.12) implies that such a map corresponds to a solution of the sinh-Gordon equation on a rectangle which is invariant under reflections about its sides. An existence theorem for such solutions was given by Wente [28] who went much further and showed that the map to the 2-sphere was in fact the Gauss map of an immersed torus of constant mean curvature in $\mathbb{R}^3$. Thus our examples may be considered to be the Gauss maps of Wente's surfaces.

Actually, by a remarkable process of mathematical evolution, beginning with computer graphics of a numerical solution to Wente's problem, Abresch [1] found an analytical solution in terms of elliptic functions corresponding to a pair of distinct elliptic curves. To see the link with our point of view we note that the linearization of the equations as in §8 takes place on the abelian variety of line bundles such that $(\sigma\tau)^*L = L$. These are line bundles on the quotient curve $\Sigma/\sigma\tau = \overline{\Sigma}$ which has Euler characteristic $\frac{1}{2}(2-2g) = -2$ and hence has genus 2. Acting on $\overline{\Sigma}$ is the group of holomorphic transformations of order 4 generated by $\rho\nu$ and $\tau\rho\nu$. The involution $\rho\nu$ is induced by $\zeta \to \zeta^{-1}$ and the fixed points on Σ are the four points $\pi^{-1}\{\pm 1\}$. These give two fixed points in the quotient $\overline{\Sigma}$, hence if g_1 is the genus of $\Sigma_1 = \overline{\Sigma}/\rho\nu$, then by the Riemann-Hurwitz formula

$$2 - 2\cdot 2 = 2(2-2g_1) - 2$$

and so $g_1 = 1$, i.e., Σ_1 is an elliptic curve. Similarly, $\Sigma_2 = \overline{\Sigma}/(\tau\rho\nu)$ is elliptic and the map $\overline{\Sigma} \to \Sigma_1 \times \Sigma_2$ given by the two projections exhibits the product $\mathrm{Pic}^0(\Sigma_1)\times\mathrm{Pic}^0(\Sigma_2)$ as a covering of $\mathrm{Pic}^0(\Sigma)$, and so the equations can effectively also be linearized on the product of two elliptic curves.

13. An energy formula

The most fundamental invariant of a harmonic map is its *energy*—the value of the functional which defines the variational problem. This is given by

$$E = \frac{1}{2}\int_M \|df\|^2 \omega_M, \tag{13.1}$$

where the norm of $df \in \Omega^1(M; f^*(TG))$ is determined by the metric on M and TG.

In our situation, where $G = \mathrm{SU}(2)$, we used as in (1.8) the bi-invariant metric $-\mathrm{tr}(A^2)$ on G which is twice the usual metric on the 3-sphere of radius 1. If we use now instead the standard metric on S^3, then the energy may be written as

$$\begin{aligned} E &= -\frac{1}{4}\int_M \mathrm{tr}(f^{-1}df)^2\omega_M = -\frac{1}{4}\int_M \mathrm{tr}(f^{-1}df\wedge *f^{-1}df) \\ &= -\int_M \mathrm{tr}(\phi\wedge *\phi) \quad \text{from §1} \\ &= -2i\int_M \mathrm{tr}\,\Phi\wedge\Phi^*. \end{aligned} \tag{13.2}$$

This functional of course makes sense for the full gauge-theoretic equations (1.7), being essentially the L^2 length of the Higgs field.

Remark. In the geometrically significant special cases arising from a minimal surface in S^3 or a surface of constant mean curvature in $\mathbb{R}^3$, the energy can be expressed in terms of the *area*. In the first case the map is conformal and, as in (1.8), the pulled-back metric is $4\,\mathrm{tr}\,\Phi\Phi^*$ so (13.2) gives the area directly as E. In the second case the harmonic map in question is the Gauss map, whose derivative is the second fundamental form B of the surface in $\mathbb{R}^3$. Thus

$$E = \tfrac{1}{2}\int_M \|B\|^2\omega_M = \tfrac{1}{2}\int_M (H^2-2K)\omega_M,$$

where H is the mean curvature and K the Gaussian curvature. Using the Gauss-Bonnet theorem for the torus and the constancy of H, we have $E = \frac{1}{2}H^2A$, where A is the area, so taking the mean curvature to be 1, the energy is half the area.

In order to compute the energy, we shall make use of the theory of determinant line bundles due to Quillen [23], Knudsen and Mumford [17], and developed by Bismut and Freed [6].

We consider the complex manifold $M \times \mathbb{C}$ and the holomorphic rank-2 vector bundle $\tilde{V}$ of §3 obtained from the Cauchy-Riemann operator

$$d''_\zeta = d''_A - \zeta\Phi^* \qquad (\zeta \in \mathbb{C}) \tag{13.3}$$

of a solution to (1.7).

By virtue of the proper map given by projecting to $\mathbb{C}$, there is a natural holomorphic line bundle $\mathscr{L}$ over $\mathbb{C}$—the *determinant bundle* of this family. Furthermore, since the elliptic operator

$$d''_\zeta : \Omega^0(M; V_\zeta) \to \Omega^{0,1}(M; V_\zeta)$$

has index 0 for the torus M, there is also a canonical holomorphic section $\det(d''_\zeta)$ of $\mathscr{L}$ over $\mathbb{C}$ which vanishes whenever $d''_\zeta u = 0$ has a nontrivial solution u. In [23] it is shown that the determinant bundle $\mathscr{L}$ has a natural Hermitian connection, given a Hermitian form on the vector bundle $\tilde{V}$. In fact, given the universal picture of all Cauchy-Riemann operators $d''_A + B$ on a fixed C^∞ Hermitian vector-bundle V, with $B \in \Omega^{0,1}(M; \operatorname{End} V)$, Quillen showed that the *curvature* F of this connection is given by

$$F = d'd''\left(\frac{i}{2\pi}\int_M \operatorname{tr} B^* \wedge B\right), \tag{13.4}$$

where $d'd''$ is defined on the infinite-dimensional complex vector space $\Omega^{0,1}(M; \operatorname{End} V)$.

Specializing to the 1-parameter family given by (13.3), from (13.2) we see that the curvature of this connection is

$$F = d'd''\left(\frac{i\zeta\bar{\zeta}}{2\pi}\int_M \operatorname{tr}\Phi \wedge \Phi^*\right) = -\frac{E}{4\pi}\, d\zeta \wedge d\bar{\zeta}. \tag{13.5}$$

The energy E thus has an interpretation as the (constant) curvature of the determinant bundle.

If s is a nonvanishing holomorphic section of a line bundle L then, as usual, if ∇ is the unique connection compatible with the metric and holomorphic structure, we have $\nabla s = \omega s$ $(\omega \in \Omega^{1,0})$ and $d'\|s\|^2 = \langle \nabla s, s\rangle = \omega\|s\|^2$ so that the connection form is ω and hence the curvature is

$$F = d\omega = d''d' \log\|s\|^2.$$

If L is defined on a Riemann surface N with boundary ∂N and if s has n zeros (counted with multiplicity) in N but is nonvanishing on ∂N, then a straightforward and common integration gives

$$\int_N F = 2\pi i n + \int_{\partial N} d' \log \|s\|^2 = 2\pi i n + \int_{\partial N} \omega. \tag{13.6}$$

If s vanishes on ∂N with total multiplicity m, then (13.6) is modified to

$$\int_N F = 2\pi i n + \pi i m + PV \int_{\partial N} \omega \tag{13.7}$$

for an appropriate principal value integral of the singular 1-form ω.

We apply this standard procedure now to the determinant bundle $\mathscr{L}$ over the unit disc in $\mathbb{C}$ and take s to be the canonical holomorphic section $\det(d''_\zeta)$. From (13.5) we obtain

$$E/2 = 2\pi n + \pi m - iPV \int_{S^1} \omega. \tag{13.8}$$

Note that, from the work of Bismut and Freed [6], the connection and canonical section of the determinant bundle on the circle S^1 are determined by the family of operators $d''_A - e^{i\theta}\Phi^*$, and not by the embedding of the circle in a larger family. Thus the boundary term in (13.8) is intrinsically determined by $d''_A - e^{i\theta}\Phi^*$ which, from (1.7), is the $(0, 1)$ part of a flat unitary $\mathrm{SU}(2)$ connection.

Take the ramified double covering $\hat{D}$ of the unit disc D given by $\hat{D} = \pi^{-1}(D) \subset \hat{\Sigma}$. Thus $\hat{D}$ is half of the hyperelliptic curve of §4.

By the definition of $\hat{\Sigma}$, the vector bundle $\tilde{V}$ pulled back to $M \times \hat{D}$ is an extension of the eigenspace bundle by its dual:

$$0 \to E \to \pi^* \tilde{V} \to E^* \to 0.$$

Now E has its own determinant line bundle $\mathscr{L}'$ on $\hat{D}$ and similarly E^* has a line bundle $\mathscr{L}''$. Moreover, by Serre duality $\mathscr{L}'' \cong \mathscr{L}'$. There is also, since the index of the d''-operator on E is zero, a canonical section s' of $\mathscr{L}'$. From the functorial properties of the determinant we have [17]

$$\pi^* \mathscr{L} \cong \mathscr{L}' \otimes \mathscr{L}'' \cong (\mathscr{L}')^2,$$

and under this isomorphism

$$\pi^* s = (s')^2. \tag{13.9}$$

Now consider the line bundle $E(\xi)$ on M as ξ varies in $\hat{D}$. For each ξ this is a holomorphic line bundle of degree 0 on the torus M and thus admits a *unique flat hermitian connection* compatible with the holomorphic structure. This family of hermitian structures on E gives rise to a hermitian connection on the determinant bundle $\mathscr{L}'$. Applying (13.7) to this connection over $\hat{D}$ we obtain

$$\int_{\hat{D}} F' = 2\pi i n' + \pi i m' + PV \int_{\partial \hat{D}} \omega',$$

where n' is the number of zeros of s' in $\hat{D}$, and m' is the number on the boundary.

Now from (13.9), $2n = 2n'$ and $2m = 2m'$, hence from (13.8),

$$E/2 + i \int_{\hat{D}} F' = iPV \int_{\partial \hat{D}} \omega' - iPV \int_{S^1} \omega. \tag{13.10}$$

However, as noted above, when $|\zeta| = 1$, $d''_A - \zeta\Phi^*$ is the $(0, 1)$ part of a flat unitary connection which reduces to a $\mathrm{U}(1)$ connection on $E \oplus E^*$, E being the eigenspace of the holonomy. On the boundary therefore the two families coincide. It follows then from (13.9) and the intrinsic property of the connection on the determinant bundles that $\pi^* \omega = 2\omega'$ and hence

$$PV \int_{\partial \hat{D}} \omega' = \tfrac{1}{2} PV \int_{\partial \hat{D}} \pi^* \omega = PV \int_{S^1} \omega,$$

and so from (13.10)

$$E/2 = -i \int_{\hat{D}} F'. \tag{13.11}$$

To compute the curvature F' of the determinant bundle of flat $\mathrm{U}(1)$ bundles on a torus M is straightforward. Regarding M as $\mathbb{C}$ modulo $\{1, \tau\}$ with parameter z, every flat $\mathrm{U}(1)$ connection may be written as

$$d + w d\bar{z} - \bar{w}\, dz \quad \text{for } w \in \mathbb{C},$$

where the metric is the constant metric on the trivial line bundle. From (13.4)

$$F = d'd'' \left| \frac{i w \bar{w}}{2\pi} \int_M dz \wedge d\bar{z} \right| = \frac{\operatorname{Im} \tau}{\pi}\, dw \wedge d\bar{w}. \tag{13.12}$$

Now over $\zeta \in D$, $\zeta \neq 0$ the eigenspace bundle $E(\zeta)$ over M has a flat $\mathbb{C}^*$ connection with holonomy given by $(\mu, \tilde{\mu})$ over the two generators of $\pi_1(M)$. This is the exponential of $(\log \mu, \log \tilde{\mu}) \in H^1(M; \mathbb{C})$ and the *holomorphic* structure is given by $d'' + w d\bar{z}$ with $w d\bar{z}$ representing the cohomology class which is the $(0, 1)$ part of $(\log \mu, \log \tilde{\mu})$. This means that

$$w = \frac{\tau \log \mu - \log \tilde{\mu}}{\tau - \bar{\tau}}.$$

Since $d(\log \mu) = \theta$ and $d(\log \tilde{\mu}) = \tilde{\theta}$, this gives from (13.11) and (13.12) a formula for the energy in terms of the spectral curve:

$$E/2 = -i \int_{\hat{D}} F' = -i \frac{\operatorname{Im} \tau}{\pi} \int_{\hat{D}} \frac{(\tau\theta - \tilde{\theta}) \wedge (\overline{\tau\theta} - \overline{\tilde{\theta}})}{4(\operatorname{Im}\tau)^2},$$

$$E = -\frac{i}{2\pi \operatorname{Im}\tau} \int_{\hat{D}} (\tau\theta - \tilde{\theta}) \wedge (\overline{\tau\theta} - \overline{\tilde{\theta}}). \tag{13.13}$$

Remark. Note from (3.5) and (3.10) that $\tau\theta - \tilde{\theta}$ is *holomorphic* in $\hat{D}$—the poles of $\tau\theta$ and $\tilde{\theta}$ over $\zeta = 0$ cancel.

To obtain a formula more explicit than (13.13), we may consider $\hat{D}$, half the hyperelliptic curve $\hat{\Sigma}$, as the complement of a disc in a Riemann surface of genus $\frac{1}{2}g$ if g is even and the complement of two disjoint discs in a surface of genus $\frac{1}{2}(g-1)$ if g is odd, as in §7. We now represent (as in the proof of the reciprocity law [13]) this surface as a polygon with $4k = 2g$ sides (or $4k = 2g-2$ if g is odd), which are identified in pairs. We therefore represent $\hat{D}$ as a polygon Δ with an interior disc (or two if g is odd) removed, and identifications on the outer boundary.

Now the cycle $\partial\hat{D} = \pi^{-1}(S^1)$ is invariant under the hyperelliptic involution σ, but the differentials θ and $\tilde{\theta}$ transform to their negatives, thus the periods of θ and $\tilde{\theta}$ around $\partial\hat{D}$ are zero. We may therefore write (for even g) $\theta = df$ and $\tilde{\theta} = d\tilde{f}$ for single valued meromorphic functions f and $\tilde{f}$ on Δ. In particular, we have $\tau\theta - \tilde{\theta} = d(\tau f - \tilde{f})$ and $\tau f - \tilde{f}$ is holomorphic.

Using Stokes' theorem, we may rewrite (13.13) as

$$\operatorname{Im}\tau 2\pi i E = \sum_i (\pi^i \overline{\pi}^{k+i} - \pi^{k+i} \overline{\pi}^i) - \int_{\partial\hat{D}} (\tau f - \tilde{f})(\overline{\tau\theta} - \overline{\tilde{\theta}}), \tag{13.14}$$

where π^i $(1 \le i \le 2k)$ are periods of the differential $\tau\theta - \tilde{\theta}$.

Now recall that θ satisfies the reality condition $\rho^*\theta = -\overline{\theta}$ on $\hat{\Sigma}$ and hence on $\partial\hat{D}$, which is fixed by ρ, $\theta = -\overline{\theta}$, and so

$$\int_{\partial\hat{D}} (\tau f - \tilde{f})(\overline{\tau\theta} - \overline{\tilde{\theta}}) = -\int_{\partial\hat{D}} (\tau f - \tilde{f})(\overline{\tau}\theta - \tilde{\theta}). \tag{13.15}$$

Also, the periods of θ and $\tilde{\theta}$ lie in $2\pi i\mathbb{Z}$, and so are in particular *imaginary*. Thus the periods of $\overline{\tau}\theta - \tilde{\theta}$ are the same as those of $-(\overline{\tau\theta} - \overline{\tilde{\theta}})$. Applying Stokes' theorem to the meromorphic form $(\tau f - \tilde{f})(\overline{\tau}\theta - \tilde{\theta})$ from

(13.14) and (13.15) we obtain

$$
\begin{aligned}
&2\pi i \sum \operatorname{Res}(\tau f - \tilde{f})(\bar{\tau}\theta - \tilde{\theta}) \\
&= -\sum_i (\pi^i \bar{\pi}^{k+i} - \pi^{k+i}\bar{\pi}^i) - \int_{\partial \hat{D}} (\tau f - \tilde{f})(\bar{\tau}\theta - \tilde{\theta}) \\
&= -2\pi i \operatorname{Im} \tau E.
\end{aligned} \tag{13.16}
$$

In the case of odd genus, where Δ is a polygon with two discs removed, a cut between the discs must be inserted to make f and $\tilde{f}$ single-valued and this introduces an extra term in (13.14). However, it reappears in (13.16) and cancels to give the same formula in (13.17).

Since θ and $\tilde{\theta}$ only have poles over $\zeta = 0$ in $\hat{D}$, we compute the residues directly to obtain the following:

Theorem (13.17). (i) *Let f be a nonconformal harmonic map of a torus with modulus τ to S^3, and suppose that around one point of $\pi^{-1}(0)$, the differentials θ and $\tilde{\theta}$ have expansions*

$$\theta = \left(\frac{a_{-2}}{\zeta^2} + a_0 + \cdots\right) d\zeta, \qquad \tilde{\theta} = \left(\frac{\tilde{a}_{-2}}{\zeta^2} + \tilde{a}_0 + \cdots\right) d\zeta.$$

Then the energy of the map is given by

$$E = 4i(a_0 \tilde{a}_{-2} - \tilde{a}_0 a_{-2}).$$

(ii) *Let f be a conformal harmonic map, and suppose that, if $\eta^2 = \zeta$, then θ and $\tilde{\theta}$ have expansions*

$$\theta = \left(\frac{a_{-2}}{\eta^2} + a_0 + \cdots\right) d\eta, \qquad \tilde{\theta} = \left(\frac{\tilde{a}_{-2}}{\eta^2} + \tilde{a}_0 + \cdots\right) d\eta.$$

Then the energy of the map is given by

$$E = 2i(a_0 \tilde{a}_{-2} - \tilde{a}_0 a_{-2}).$$

Example. As an example of the formula, we consider the Clifford torus. From (6.13) and (6.15) we have

$$
\begin{aligned}
a_{-2} &= 2\pi i\sqrt{-i/8}, \qquad & a_0 &= 2\pi\sqrt{-i/8}, \\
\tilde{a}_{-2} &= -2\pi i\sqrt{i/8}, \qquad & \tilde{a}_0 &= 2\pi\sqrt{i/8},
\end{aligned}
$$

which give from the second formula of (13.17),

$$E = 2\pi^2 = \text{area}.$$

Since the Clifford torus is actually the isometric product of two circles of radius $1/\sqrt{2}$, this checks with the direct calculation of the area.

14. Postscript

The motivation for generalizing the harmonic map equations to the gauge-theoretic equations (1.7) was the possibility of finding a general method of solution for which a special subclass characterized by more complicated constraints would give the harmonic maps. This aim we have achieved in Theorems (8.1) and (8.20). We gave the generalized equations themselves a geometrical interpretation in §1, namely that of describing the harmonic sections of flat S^3-bundles. Such an interpretation does not however demonstrate very well the special role of the subclass of harmonic maps. We encountered a more useful point of view in considering the elliptic solutions of §10, where the work of Hsiang and Lawson [16] showed that the gauge-theoretic equations (1.7) corresponded to equations for the *geodesics* on a surface of revolution, and the special subclass of harmonic map solutions corresponded to the *closed geodesics.*

Strangely enough, there is a general context in which equations (1.7), for any Riemann surface M, correspond to geodesic equations and the solutions giving harmonic maps become closed geodesics. The setting for this is the moduli space of gauge-equivalence classes of connections, as studied by Mitter and Viallet in [21].

Let $\mathscr{A}$ denote the space of C^∞ connections on a principal G-bundle P over a compact Riemann surface M, as in §1. The space $\mathscr{A}$ is an infinite-dimensional affine space with group of translations $\Omega^1(M; \operatorname{ad} P)$ and has a Riemannian metric defined by

$$g(\phi, \phi) = \int_M B(\phi \wedge *\phi), \tag{14.1}$$

where B is a bi-invariant metric on G.

The group $\mathscr{G}$ of gauge transformations of P acts on $\mathscr{A}$. This action is free on the irreducible connections and the tangent space to the orbit of $\mathscr{G}$ at the connection A consists of all $\phi \in \Omega^1(M; \operatorname{ad} P)$ of the form $\phi = d_A\psi$, $\psi \in \Omega^0(M; \operatorname{ad} P)$. The orthogonal complement to this tangent space is, from (13.1) the set of ϕ such that

$$\int_M B(d_A\psi \wedge *\phi) = 0 \quad \forall \psi \in \Omega^0(M; \operatorname{ad} P),$$

or equivalently all ϕ such that

$$d_A * \phi = 0. \tag{14.2}$$

The inner product (14.1) restricted to this horizontal space defines a metric on the quotient space $\mathscr{A}/\mathscr{G} = \mathscr{M}$, the moduli space of connections, at least on the dense open subspace of irreducible classes. This is a standard situation in differential geometry, and in particular geodesics on $\mathscr{M}$ lift to *horizontal* geodesics in $\mathscr{A}$, i.e., geodesics orthogonal to the $\mathscr{G}$-orbits.

Now $\mathscr{M}$ is just an affine space with a flat metric, so geodesics are straight lines. The geodesic joining connections A_1 and A_2 is

$$\nabla_{A_1} + 2t\phi, \qquad 0 \leq t \leq 1,$$

where $\nabla_{A_2} = \nabla_{A_1} + 2\phi$.

If $d_{A_1}\phi = 0$ then from (14.2) the geodesic is horizontal at $t = 0$, and since

$$(d_{A_1} + 2t\phi) * \phi = d_{A_1} * \phi + 2t[\phi \wedge *\phi] = d_{A_1} * \phi = 0,$$

the geodesic is horizontal for all t.

If A_1 and A_2 are *flat*, then setting $A = \frac{1}{2}(A_1 + A_2)$ we have the equations

$$d_A * \phi = (d_{A_1} + \phi) * \phi = 0 \quad \text{and} \quad (d_A + \phi)^2 = (d_A - \phi)^2 = 0,$$

which as in §1 become equations (1.7), setting $\phi = \Phi - \Phi^*$. Thus, the gauge-theoretic equations (1.7) may be interpreted as the equations for a geodesic whose endpoints lie on the moduli space of *flat* G-connections. This is actually a finite-dimensional subspace of the infinite-dimensional space $\mathscr{M}$.

The solutions which give a harmonic map are those for which the endpoints are not just flat but *trivial* and thus from this point of view correspond to closed geodesics emanating from the equivalence class of the trivial connection. This is of course a reducible connection and hence actually a singular point of $\mathscr{M}$.

This observation does not provide much help in solving the equations but does show the natural context of equations (1.7) from both a gauge-theoretic and differential-geometric point of view.

Acknowledgements. Parts of this work were carried out while the author was a Visiting Professor in the Department of Mathematics and the Institute of Theoretical Physics at Stony Brook in 1983/84, and a visitor at the Institut des Hautes Études Scientifiques in the spring of 1987. Thanks are owed to both institutions for their support and hospitality.

References

[1] U. Abresch, *Constant mean curvature tori in terms of elliptic functions*, J. Reine Angew. **374** (1987) 169–192.

[2] M. Adler & P. Van Moerbeke, *Linearization of Hamiltonian systems, Jacobi varieties and representation theory*, Advances in Math. **38** (1980) 318–379.

[3] M. F. Atiyah, *Riemann surfaces and spin structures*, Ann. Sci. École Norm. Sup. (4) **4** (1971) 47–62.
[4] M. F. Atiyah & R. Bott, *The Yang-Mills equations over Riemann surfaces*, Philos. Trans. Roy. Soc. London Ser. A **308** (1982) 523–615.
[5] M. F. Atiyah & R. S. Ward, *Instantons and algebraic geometry*, Comm. Math. Phys. **55** (1977) 117–124.
[6] J. M. Bismut & D. S. Freed, *The analysis of elliptic families.* I, *Metrics and connections on determinant bundles*, Comm. Math. Phys. **106** (1986) 159–176.
[7] R. L. Bryant, *Conformal and minimal immersions of compact surfaces into the* 4-*sphere*, J. Differential Geometry **17** (1982) 455–473.
[8] H. I. Choi & R. Schoen, *The space of minimal embeddings of a surface into a three-dimensional manifold of positive Ricci curvature*, Invent. Math. **81** (1985) 387–394.
[9] C. Delaunay, *Sur la surface de révolution dont la courbure moyenne est constante*, J. Math. Pures Appl. (1) **6** (1841) 309–320.
[10] J. Eells & J. C. Wood, *Harmonic maps from surfaces to complex projective spaces*, Advances in Math. **49** (1983) 217–263.
[11] L. P. Eisenhart, *Differential geometry*, Ginn & Co., Boston, 1909.
[12] J. D. Fay, *Theta functions on Riemann surfaces*, Lecture Notes in Math., Vol. 352, Springer, Berlin, 1973.
[13] P. Griffiths & J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978.
[14] N. J. Hitchin, *On the construction of monopoles*, Comm. Math. Phys. **89** (1983) 145–190.
[15] ——, *The self-duality equations on a Riemann surface*, Proc. London Math. Soc. **55** (1987) 59–126.
[16] W. Y. Hsiang & H. B. Lawson, *Minimal submanifolds of low cohomogeneity*, J. Differential Geometry **5** (1971) 1–38.
[17] F. F. Knudsen & D. Mumford, *The projectivity of the moduli space of stable curves.* I: *Preliminaries on* "det" *and* "div", Math. Scand. **39** (1976) 19–55.
[18] H. B. Lawson, *Complete minimal surfaces in* S^3, Ann. of Math. (2) **92** (1970) 335–374.
[19] H. P. McKean, *The sine-Gordon and sinh-Gordon equations on the circle*, Comm. Pure Appl. Math. **34** (1981) 197–257.
[20] H. P. Mckean & P. van Moerbeke, *The spectrum of Hill's equation*, Invent. Math. **30** (1975) 217–274.
[21] P. K. Mitter & C. M. Viallet, *On the bundle of connections and the gauge orbit manifold in Yang-Mills theory*, Comm. Math. Phys. **79** (1981) 457–472.
[22] W. Nahm, *The construction of all self-dual monopoles by the ADHM method*, in Monopoles in Quantum Field Theory (N. S. Craigie, P. Goddard & W. Nahm, eds.), World Scientific, Singapore, 1982.
[23] D. Quillen, *Determinants of Cauchy-Riemann operators over a Riemann surface*, Functional Anal. Appl. **19** (1985) 31–34.
[24] E. A. Ruh & J. Vilms, *The tension field of the Gauss map*, Trans. Amer. Math. Soc. **149** (1970) 569–573.
[25] W. Schmid, *Variation of Hodge structure: the singularities of the period mapping*, Invent. Math. **22** (1973) 211–319.
[26] G. Segal & G. Wilson, *Loop groups and equations of KdV type*, Inst. Hautes Études Sci. Publ. Math. **61** (1985) 5–65.
[27] K. K. Uhlenbeck, *Harmonic maps into Lie groups* (*Classical solutions of the chiral model*), J. Differential Geometry **30** (1989) 1–50.
[28] H. C. Wente, *Counterexample to a conjecture of H. Hopf*, Pacific J. Math. **121** (1986) 193–243.

St. Catherine's College, Oxford