http://dx.doi.org/10.1090/pspum/053/1141208

Proceedings of Symposia in Pure Mathematics
Volume 53 (1991)

# The Ubiquity of Variations of Hodge Structure

CARLOS T. SIMPSON

*Dedicated to Phillip Griffiths*

More than twenty years ago, Griffiths introduced the definition of *variation of Hodge structure* as a way of formalizing some properties of the families of Hodge structures on cohomology groups of a family of smooth projective varieties. It may have seemed, at the time, that these properties of the local systems of cohomology groups were very special to the geometric situation from which they arose. It is my goal to show that the variations of Hodge structure, in a sense slightly generalized from the original, play a central role in the theory of all local systems, or representations of the fundamental group, on a smooth projective variety.

The material in §§1, 2, and 3 is expository. In §1, which corresponds most closely to the conference talk, I will describe the basic correspondence between representations of the fundamental group and *Higgs bundles*, geometric objects on $X$. From this correspondence, it will be easy to describe which representations come from complex variations of Hodge structure—they are the fixed points of a natural $\mathbf{C}^*$ action. In §2, I will give a brief explanation of the main idea for constructing the moduli space of Higgs bundles, and from this discuss a crucial closure property which allows one to conclude that there are many fixed points of the $\mathbf{C}^*$ action. In §3, I will describe Hitchin's idea for relating the cohomology of the moduli spaces to the cohomology of the fixed point sets.

§4 contains new research, done in January 1990 (admittedly, a bit after the conference concluded!). The result is a sort of structure theorem for representations of $\pi_1(X)$ into $\mathrm{Sl}(2, \mathbf{C})$. This was prompted by discussions with D. Toledo, J. Carlson, W. Goldman, and K. Corlette. It is an analogue of topics discussed several years ago in a seminar organized by N. Boston,

---

1991 *Mathematics Subject Classification*. Primary 14D07, 32G20.

This research was partially supported by National Science Foundation grant DMS-8705757.

about work of Culler and Shalen. I would also like to thank P. Deligne for some helpful remarks.

## 1. Variations of Hodge structure and Higgs bundles

Let $X$ be a smooth projective variety. A *complex variation of Hodge structure* on $X$ is a $C^\infty$ vector bundle $V$ with decomposition $V = \bigoplus_{p+q=w} V^{p,q}$ and flat connection $D$, satisfying the axioms of Griffiths transversality

$$D = \partial + \overline{\partial} + \theta + \overline{\theta}\colon V^{p,q} \to A^{1,0}(V^{p,q}) \oplus A^{0,1}(V^{p,q}) \oplus A^{1,0}(V^{p-1,q+1}) \oplus A^{0,1}(V^{p+1,q-1}),$$

and existence of a polarization. A polarization is a hermitian form $\langle u, v\rangle$, preserved by $D$, such that $V^{p,q} \perp V^{r,s}$ for $(p, q) \neq (r, s)$, and $(-1)^p\langle u, u\rangle > 0$ for $u \in V^{p,q}$.

These objects occur naturally as descriptions of the variable Hodge structures on cohomology of varieties in algebraic families indexed by $X$ **[12]**. For example, if $f\colon Y \to X$ is a projective smooth map, define $V_x = H^i(Y_x, \mathbf{C})$, with Hodge decomposition $V_x^{p,q} = H^{p,q}(X)$. The flat connection is obtained by translating cohomology classes topologically. The notion of complex variation of Hodge structure is a slight generalization of the notion of integer variation of Hodge structure introduced by Griffiths **[13]**—the cohomology groups of a family of varieties have integer lattices preserved by the flat connection. We will ignore this aspect and concentrate on the notion of variation of Hodge structure in the abstract, as a kind of structure on a connection or representation of the fundamental group. One advantage of the notion of complex variation of Hodge structure is that such objects are direct sums of components corresponding to irreducible representations of the fundamental group [7].

The purpose of this first section is to discuss some results which indicate that the representations underlying complex variations of Hodge structure play an important role in the representation theory of $\pi_1(X)$. In discussing fundamental groups we will suppress reference to choice of base points.

From a variation of Hodge structure, one obtains a *system of Hodge bundles*, a direct sum $E = \bigoplus_{p+q=w} E^{p,q}$ of holomorphic vector bundles with endomorphism valued one-forms $\theta\colon E^{p,q} \to E^{p-1,q+1} \otimes \Omega^1_X$ such that $\theta \wedge \theta = 0$. The holomorphic bundles $E^{p,q}$ are just the bundles $V^{p,q}$ provided with holomorphic structure $\overline{\partial}$ equal to the appropriate component of the connection $D$. The maps $\theta$ are similarly components of $D$ as denoted above.

In the geometric situation of a smooth family of varieties $f\colon Y \to X$, the system of Hodge bundles associated to a variation of Hodge structure is easy to describe: $E^{p,q} = R^q f_* \Omega^p_{Y/X}$, and $\theta_x$ is obtained by cup product with the Kodaira-Spencer deformation class $\eta_x \in H^1(Y_x, T(Y_x)) \otimes (\Omega^1_X)_x$. I remember wondering, in Griffiths's class, whether one could recover the variation of Hodge structure from this holomorphic data. The answer turns out to be yes.

The systems of Hodge bundles which arise from irreducible complex variations of Hodge structure are *stable*, i.e., any subsystem of Hodge sheaves $U \subset E$ must have $\deg(U)/\operatorname{rk}(U) < \deg(E)/\operatorname{rk}(E)$ (note that this condition makes reference to a choice of hyperplane class). The condition of stability is related to Griffiths's result about degrees of the Hodge bundles **[13]**. The systems of Hodge bundles which come from variations of Hodge structures, being flat as $C^\infty$ bundles, must satisfy $c_i(E) = 0$. Using techniques of Yang-Mills theory, the construction can be inverted to give a one-to-one correspondence between irreducible complex variations of Hodge structure and stable systems of Hodge bundles with vanishing Chern classes.

This is an example of a more general statement for all representations of $\pi_1(X)$, the "nonabelian Hodge theorem." A *Higgs bundle* on $X$ is a pair $(E, \theta)$ consisting of a holomorphic bundle $E$ and an endomorphism-valued one-form $\theta: E \to E \otimes \Omega^1_X$ such that $\theta \wedge \theta = 0$. A Higgs bundle is *stable* if, for any subsheaf $U \subset E$ preserved by $\theta$, the degree divided by the rank of $U$ is strictly smaller than that of $E$ (again, this condition refers to a choice of hyperplane class).

THEOREM 1. *There is a one-to-one correspondence between the set of stable Higgs bundles with* $c_i(E) = 0$, *and the set of irreducible representations of* $\pi_1(X)$. *A representation has a structure of variation of Hodge structure if and only if the associated Higgs bundle has a structure of system of Hodge bundles, and in this case the correspondence is the one described above.*

This theorem (as well as the definitions which preceded it) is the result of work of N. Hitchin **[16]**, K. Corlette **[4]**, S. Donaldson **[8, 9, 10]**, P. Deligne (work with A. Beilinson, and various communications), K. Uhlenbeck and S. Yau **[28]**, Y. Siu [27], and myself **[24, 25]**. The simplest original version is the theorem of Narasimhan and Seshadri **[22]**.

It might be helpful to include a description, not of the proof of the theorem, but of the mechanism of the resulting correspondence. A Higgs bundle is a smooth vector bundle $E$ provided with an operator $D'' = \overline{\partial} + \theta$. If one chooses a smoothly varying hermitian metric $K$, then one obtains an operator $D'_K = \partial_K + \overline{\theta}_K$. The formulas are chosen so as to be compatible with the case of variations of Hodge structure, in which the metric $K$ is taken as the positive definite form obtained by appropriately changing the signs of the polarization on the different Hodge bundles. They are

$$(\partial_K u, v)_K + (u, \overline{\partial} v)_K = \partial(u, v)_K,$$
$$(\overline{\theta}_K u, v)_K - (u, \theta v)_K = 0.$$

Then the combination $D_K = D'_K + D''$ is a connection on $E$, with curvature $F_K = D_K^2$. One direction of the theorem consists in showing that if $E$ is stable and has vanishing Chern classes, then the equation $F_K = 0$ can be solved to find a *harmonic metric* [**22, 8, 9, 28, 16, 24**]. The resulting flat connection $D_K$ gives the representation of the fundamental group corresponding to the Higgs bundle. Note that the variations of Hodge structure are particular cases of such solutions. The other half of the theorem consists in showing that every irreducible representation occurs this way, by solving some similar equations for a metric on a flat bundle, to get a Higgs bundle [**4, 10, 27, 25**].

THEOREM 2. *There are coarse moduli spaces $M_{\mathrm{Higgs}}$ and $M_{\mathrm{Rep}}$ for the Higgs bundles and representations which occur in the previous theorem. They are algebraic varieties whose points parametrize Jordan equivalence classes of objects, or equally well direct sums of stable or irreducible objects. The correspondence of the theorem becomes a homeomorphism $M_{\mathrm{Higgs}} \cong M_{\mathrm{Rep}}$ of the underlying topological spaces.*

The homeomorphism is not a complex analytic map, but it is expected to be real analytic at the smooth points. The construction of $M_{\mathrm{Rep}}$, an affine variety, is standard (see, for example, [**6**]). We will give a heuristic indication of the construction of $M_{\mathrm{Higgs}}$ in the next section.

The place of the complex variations of Hodge structure in the space of all local systems is clarified by this correspondence. A representation underlies a variation of Hodge structure if and only if the associated Higgs bundle can be given a structure of system of Hodge bundles. The points in $M_{\mathrm{Higgs}}$ which have structures of systems of Hodge bundles may be characterized as the fixed points of a natural $\mathbf{C}^*$ action:

$$t\colon (E, \theta) \mapsto (E, t\theta).$$

It is easy to see that if $f\colon (E, \theta) \cong (E, t\theta)$ for a number $t$ of infinite order, then $E$ has a structure of system of Hodge bundles, by decomposing $E$ into generalized eigenspaces of $f$ and arranging them in strings with eigenvalues related by powers of $t$.

The moduli space $M_{\mathrm{Higgs}}$ is not compact, but we can still take limit points of orbits of $\mathbf{C}^*$ to find many fixed points.

PROPOSITION 3. *If $E \in M_{\mathrm{Higgs}}$, then $E_0 = \lim_{t\to 0} tE$ exists in $M_{\mathrm{Higgs}}$ and is a fixed point of $\mathbf{C}^*$.*

This was first discovered by Hitchin in his context [**16**]. We will see why it is true from the discussion of the construction of $M_{\mathrm{Higgs}}$ in the next section.

COROLLARY 4. *Any representation of $\pi_1(X)$ may be deformed to a representation underlying a complex variation of Hodge structure. In particular, a rigid representation must already underly a complex variation.*

From this corollary, one may deduce that certain groups $\Gamma$ cannot occur as fundamental groups of a smooth projective variety $X$. For example, the monodromy of a complex variation of Hodge structure must fix an indefinite hermitian form, so it lies in some $\mathrm{U}(p, q) \subset \mathrm{Gl}(n)$. If $\Gamma$ is a group with a representation which cannot be deformed into any $\mathrm{U}(p, q)$, then $\Gamma$ cannot be $\pi_1(X)$. See **[25]** for a more complete discussion.

## 2. Construction of the moduli space

Let me explain briefly the construction of the moduli space $M_{\text{Higgs}}$, leaving out the technical details. These have been circulated in **[26]**, to be published later.

A Higgs bundle on $X$ can be thought of as a vector bundle $E$ together with action of the symmetric algebra of the tangent bundle $\mathrm{Sym}^{\cdot} TX$. The one-form $\theta$ gives an action of the symmetric algebra rather than of the full tensor algebra because of the condition $\theta \wedge \theta = 0$, which says that the endomorphisms corresponding to different tangent vectors commute.

The projection of the cotangent bundle $f: T^*X \to X$ is an affine map. $T^*X$ is the relative spectrum of a sheaf of algebras, in fact, of the symmetric algebra of the tangent bundle:

$$T^*X = \mathbf{Spec}(\mathrm{Sym}^{\cdot} TX), \qquad f_*\mathscr{O}_{T^*X} = \mathrm{Sym}^{\cdot} TX.$$

A quasi-coherent sheaf $\mathscr{E}$ on $T^*X$ is therefore the same thing as a quasicoherent sheaf $E = f_*\mathscr{E}$ on $X$ together with an action of $\mathrm{Sym}^{\cdot} TX$ on $E$.

This interpretation may be applied to give a picture of a Higgs bundle $(E, \theta)$ as a sheaf $\mathscr{E}$ on $T^*X$. This picture was studied by Hitchin under the name "spectral curve" **[16, 17]**, and further by Oxbury **[23]**. A point $(x, \xi)$ in $T^*X$ is in the support of $\mathscr{E}$ if and only if the linear function $v \mapsto \xi(v)$ of $v \in T_xX$ is an eigenvalue of the operation of $T_xX$ on $E_x$ given by $\theta_x$.

The sheaves which arise from Higgs bundles this way are coherent, and in fact finite and flat over $\mathscr{O}_X$. In particular, the dimension of the support of $\mathscr{E}$ is equal to the dimension of $X$ (denote it by $d$), and the map from the support of $\mathscr{E}$ to $X$ is proper. Geometrically, the last condition means that the support of $\mathscr{E}$ does not intersect the divisor at infinity.

One can give a general construction of moduli spaces for coherent sheaves on a projective variety, similar to the construction for torsion free sheaves. Let $Z$ be a projective variety, with fixed ample bundle $\mathscr{O}_Z(1)$. A coherent sheaf $\mathscr{E}$ on $Z$ has *pure dimension* $d$ if the dimension of the support of every subsheaf of $\mathscr{E}$ is $d$. Let $p(\mathscr{E}, m)$ denote the Hilbert polynomial, equal to $h^0(\mathscr{E}(m))$ for large $m$. We may use these polynomials to give a definition of stability by comparing values for $m \gg 0$. Let $r(\mathscr{E})$ denote the highest coefficient of $p(\mathscr{E}, m)$. In the case when $d = \dim(X)$, pure dimension $d$ is the same as torsion free, and $r(\mathscr{E})$ is essentially the rank of $\mathscr{E}$. In general, we say that $\mathscr{E}$ is *p-stable* (resp. *p-semistable*) if, for any subsheaf $\mathscr{F} \subset \mathscr{E}$,

$$\frac{p(\mathscr{F}, m)}{r(\mathscr{F})} < \frac{p(\mathscr{E}, m)}{r(\mathscr{F})}$$

(resp. $\leq$), for $m \gg 0$.

PROPOSITION 5. *There is a coarse moduli space $M(P)$ for $p$-semistable coherent sheaves $\mathscr{E}$ of pure dimension $d$, with Hilbert polynomial $P$, up to the relation of Jordan equivalence. $M(P)$ is a projective variety.*

PROOF. See **[26]**.

This means that the closed points of $M(P)$ are in one-to-one correspondence with the Jordan equivalence classes of semistable sheaves (two objects are *Jordan equivalent* if the irreducible factors in their Jordan-Hölder series are the same), or equivalently with the direct sums of stable sheaves. And the coarse moduli condition means that if $\{\mathscr{E}_s\}_{s\in S}$ is an algebraic family of sheaves, indexed by a scheme $S$, such that each $\mathscr{E}_s$ is $p$-semistable with Hilbert polynomial $P$, then there is a canonical map from $S$ to $M(P)$ sending $s$ to the point representing the equivalence class of $\mathscr{E}_s$.

We apply this to our situation by letting $Z$ be a projective closure of $T^*X$. We may choose an ample $\mathscr{O}_Z(1)$ which, when restricted to the open set $T^*X$, is equal to the pullback of $\mathscr{O}_X(1)$. The Higgs bundles on $X$ correspond to coherent sheaves of pure dimension $d = \dim(X)$ on $Z$ whose support is contained in $T^*X$. Then, the notion of $p$-stability for the coherent sheaf $\mathscr{E}$ is the same as the notion of $p$-stability for the corresponding Higgs bundle $(E, \theta)$ on $X$ ($p$-stability is defined in the same way using Hilbert polynomials). It is a consequence of the nonabelian Hodge theorem that, for Higgs bundles with trivial Chern classes, the notions of $p$-stability and $p$-semistability defined using Hilbert polynomials are equivalent to the notions of stability and semistability defined just using the degree **[26]**. A similar relevant fact is that any semistable torsion free Higgs sheaf with trivial Chern classes must be locally free **[25]**.

Thus the semistable or stable Higgs bundles on $X$ with vanishing Chern classes correspond exactly to the $p$-semistable or $p$-stable coherent sheaves of pure dimension $d$ on $Z$, with appropriate Hilbert polynomial, and whose support is contained in $T^*X$. The moduli space $M_{\mathrm{Higgs}}$ is the open subset of the moduli space $M(P, Z)$ of sheaves on $Z$, determined by the condition of the support being contained in $T^*X$.

This picture makes it easy to see why Proposition 3 is true. $\mathbf{C}^*$ acts by scaling on $T^*X$, and we may assume that it acts on $Z$. This induces an action on the moduli space $M(P, Z)$, restricting to the desired action on $M_{\mathrm{Higgs}}$. Since $M(P, Z)$ is projective, the limit $\mathscr{E}_0 = \lim_{t\to 0} t\mathscr{E}$ exists and is invariant, for any point $\mathscr{E}$. If $\mathscr{E}$ is a point in $M_{\mathrm{Higgs}}$, the support of $\mathscr{E}$ is contained in $T^*X$, and as the scaling goes to zero, the support remains in a compact subset of $T^*X$. Thus the support of the limit $\mathscr{E}_0$ is also contained in $T^*X$ (and in fact $\mathscr{E}_0$ is supported along the zero section). So the limit $\mathscr{E}_0$ exists in $M_{\mathrm{Higgs}}$.

## 3. Topology of the moduli space

In this section we will discuss Hitchin's method for finding the homology of the moduli space $M_{\text{Higgs}}$. He completed the calculations for bundles of rank two on a curve. Half of his argument generalizes easily, the result saying that the Betti numbers are essentially the sums of Betti numbers of the fixed point sets of the $\mathbf{C}^*$ action, the spaces of variations of Hodge structure. The other half, the calculation of the Betti numbers for the spaces of variations of Hodge structure, seems to be very difficult. Hitchin's calculation in this step depended on the fact that bundles of rank two were considered. We will not accomplish anything in regard to this second half of the problem, so nothing new will be calculated. I just want to make the point that the homology of the whole moduli space is obtained from the homology of the spaces of variations of Hodge structure.

Another problem concerns singularities of the moduli space. Again, we will avoid this by dealing only with smooth components of the moduli space. For example, if $d$ and $r$ are relatively prime, then the moduli space of Higgs bundles of degree $d$ and rank $r$, on a curve, is smooth. Altogether this section will be concerned only with restating Hitchin's method in a general context. We will use algebraic geometric language, and are really just restating known results of Kirwan [**19**].

Suppose $M$ is a projective variety with an action of $\mathbf{C}^*$. Let $V_\alpha$ denote the components of the set of fixed points of the action. Let $S_\alpha$ and $T_\alpha$ denote, respectively, the sets of "incoming" and "outgoing" points for $V_\alpha$. In other words, $S_\alpha$ is the set of points $x$ such that $\lim_{t\to 0} tx \in V_\alpha$ and $T_\alpha$ is the set of points $x$ such that $\lim_{t\to\infty} tx \in V_\alpha$. The collections $\{S_\alpha\}$ and $\{T_\alpha\}$ are each stratifications of $M$.

Fix a linearized very ample line bundle $L$. The strata may be ordered by considering the action of $\mathbf{C}^*$ on $L$. Over each $V_\alpha$, $\mathbf{C}^*$ acts on $L$ by a character $t \mapsto t^{-k}$. Denote this $k$ by $k(\alpha)$.

LEMMA 6. *Suppose $x \in S_\alpha \cap T_\beta$. Then $k(\alpha) \le k(\beta)$ and if equality holds then $\alpha = \beta$ and $x \in V_\alpha$.*

PROOF. By looking at a single orbit, it can be reduced to the case of the standard action on $\mathbf{P}^1$, where it is easy to check.

Because of this lemma, we may define a partial ordering on the $\{\alpha\}$. Say that $\alpha \le \beta$ if $S_\alpha \cap T_\beta \ne \emptyset$. Say that a subset of indices $A \subset \{\alpha\}$ is *lower-closed* if, whenever $\alpha \in A$ and $\beta \le \alpha$, then $\beta \in A$.

Fix a lower-closed subset of indices $A$. Define an open subset $U_A$ and a closed subset $W_A$ of the space $M$ by

$$U_A = \bigcup_{\alpha\in A} S_\alpha, \qquad W_A = \bigcup_{\alpha\in A} T_\alpha.$$

We have $W_A \subset U_A$, and in fact $W_A$ is a deformation retract of $U_A$. To see this, note that there is a neighborhood of $W_A$ which retracts to $W_A$ (this is always true of closed algebraic subsets of algebraic varieties). Then, using the $\mathbf{C}^*$ action, we can move points in $U_A$ into this neighborhood.

PROPOSITION 7 (Bialynicki-Birula). *Suppose $U_A$ is smooth. Then for any $\alpha \in A$, $V_\alpha$ is smooth and $S_\alpha$ and $T_\alpha$ are algebraic geometric fiber bundles over $V_\alpha$ with fibers which are affine spaces. The dimensions of these affine spaces may be found by using the action of $\mathbf{C}^*$ on the normal bundle of $V_\alpha$ as a model.*

PROOF. See **[2]**.

THEOREM 8. *Suppose $U_A$ is smooth. Then*

$$\dim H_i(W_A) = \sum_{\alpha\in A} \dim H_i(T_\alpha, T_\alpha - V_\alpha).$$

*Thus if $t(\alpha)$ denotes the complex relative dimension of $T_\alpha$ over $V_\alpha$, then*

$$\dim H_i(U_A) = \sum_{\alpha\in A} \dim H_{i-2t(\alpha)}(V_\alpha).$$

This is a restatement of the special case $G = \mathbf{C}^*$ of the theorem of Kirwan **[19]**, based on work of Atiyah and Bott **[1]**. Their proofs were phrased in the language of Morse theory and moment maps, the main point being that the square norm of the moment map for the action of the maximal compact subgroup (in this case, $\mathrm{U}(1)$) is a perfect Morse function. We will sketch a proof in algebraic geometric language. The fact that the stratification is perfect is simply due (using resolution of singularities) to the fact that the strata are varieties fibered with fibers which are affine spaces.

PROOF OF THEOREM 8. Recall that $W_A$ is a deformation retract of $U_A$, so we may calculate the homology of $W_A$. Argue by induction on $A$. We may concentrate on the process of adding one stratum $T = T_\alpha$ to a closed space $W = W_A$, with $\alpha$ a minimal element which is not in $A$, so that $A \cup \alpha$ will still be lower-closed. Let $\overline{T}$ denote the closure of $T$, and $R$ the complement $\overline{T} - T$. The condition that $\alpha$ is minimal among those not contained in $A$ implies that $R = \overline{T} \cap W$. Let $\widetilde{T}$ denote a suitable birational transform of $\overline{T}$, and $\widetilde{R}$ the inverse image of $R$. Applying excision, the pair $(W \cup T, W)$ is the same as $(\overline{T}, R)$ or $(\widetilde{T}, \widetilde{R})$. We have a relative homology sequence

$$\cdots \to H_i(W) \to H_i(W \cup T) \to H_i(\widetilde{T}, \widetilde{R}) \to \cdots.$$

We claim that the map $H_i(W \cup T) \to H_i(\widetilde{T}, \widetilde{R})$ is surjective. This will prove the theorem. Since $\widetilde{T}$ maps to $W \cup T$, it suffices to show that

$$H_i(\widetilde{T}) \to H_i(\widetilde{T}, \widetilde{R})$$

is surjective. Geometrically this says that the relative homology classes can be closed up in $\widetilde{T}$ to give homology classes. By resolution of singularities **[15]**, we may choose the birational transform $\widetilde{T}$ in such a way that $\widetilde{T}$ is smooth and there is a map $p\colon \widetilde{T} \to V$ restricting to the given projection on $T$. Let $n$

denote the real dimension of $T$. Applying the duality between homology and cohomology (note that the relative homology is dual to compactly supported cohomology on $T$), and Poincaré duality, the map which is claimed to be surjective is the same as the restriction map

$$H^{n-i}(\tilde{T}) \to H^{n-i}(T).$$

The pullback from the cohomology of $V$ to the cohomology of $T$ is an isomorphism, since the fiber of $T$ over $V$ is topologically trivial. This isomorphism factors as

$$H^{n-i}(V) \to H^{n-i}(\tilde{T}) \to H^{n-i}(T),$$

which proves the claimed surjectivity. This proves the theorem.

We will apply the theorem in the following situation. Let $M$ denote a component of the moduli space $M_{\mathrm{Higgs}}$ which is smooth. In order to obtain such a component, one may have to consider the moduli space for Higgs bundles corresponding to groups other than $\mathrm{Gl}(n)$. For example, the moduli space of stable Higgs bundles of degree $d$ and rank $r$ on a curve corresponds to the moduli of representations into $\mathrm{P\,Sl}(r, \mathbf{C})$, with a certain characteristic class equal to $d$. If $d$ and $r$ are relatively prime, then all points are stable and this component is smooth.

Now $M$ is not projective. We may embed it in a projective closure $\overline{M}$, which might not be smooth, but where the action of $\mathbf{C}^*$ extends. Let $\{\alpha\}$ be the set of indices indicating fixed point sets $V_\alpha$. Proposition 3 implies that there is a lower-closed subset $A \subset \{\alpha\}$ such that $M = U_A \subset \overline{M}$. The fixed point sets $V_\alpha$, $\alpha \in A$, are exactly the subsets parametrizing Higgs bundles which have structure of systems of Hodge bundles. By Theorem 1, these are the moduli spaces of representations which underlie complex variations of Hodge structure. We may apply Theorem 8 to conclude

COROLLARY 9 (Hitchin). *If $M$ is a component of $M_{\mathrm{Higgs}}$ which is smooth, and if $\{V_\alpha\}_{\alpha\in A}$ is the set of connected components of the space of variations of Hodge structure in $M$, then*

$$\dim H_i(M) = \sum_{\alpha\in A} \dim H_{i-2t(\alpha)}(V_\alpha).$$

*Here $t(\alpha)$ is the complex relative dimension of the space of "outgoing" directions normal to $V_\alpha$.*

This occurred in Hitchin's paper as the basis of his calculation of the Betti numbers of the component $M(2, 1)$ of projective representations of rank two and odd degree on a Riemann surface **[16]**. The difficult part about generalizing Hitchin's calculation would be to calculate the Betti numbers of the spaces of variations of Hodge structure $V_\alpha$.

Let me indicate how to calculate the $t(\alpha)$, particularly when $X$ is a curve. This calculation will use the notion of *Dolbeault cohomology*. The Dolbeault cohomology of a Higgs bundle $(E, \theta)$ (with trivial Chern classes) is defined to be the hypercohomology

$$H^i_{\mathrm{Dol}}(X, E) = \mathbf{H}^i(E \xrightarrow{\theta} E \otimes \Omega^1_X \xrightarrow{\theta} E \otimes \Omega^2_X \to \cdots).$$

This is isomorphic to the de Rham cohomology of the associated local system [25]. If the local system has a structure of variation of Hodge structure of weight $w$, then the cohomology group $H^i_{\mathrm{Dol}}(X, E)$ will have a Hodge decomposition of weight $w + i$. The associated Higgs bundle has a structure of system of Hodge bundles $E = \bigoplus_{p+q=w} E^{p,q}$ ,and for $p + q = w + i$,

$$H^{p,q}(X, E) = \mathbf{H}^i(E^{p,q-i} \to E^{p-1,q+1-i} \otimes \Omega^1_X \to \cdots).$$

We can apply this to the tangent space of the moduli space at a smooth point corresponding to a stable Higgs bundle. Let $M$ denote a smooth component of the moduli space of semistable Higgs bundles of flat type. This could be, for example, the moduli space of Higgs bundles of relatively prime degree $d$ and rank $r$ on a Riemann surface, corresponding to structure group $\mathrm{PGl}(r)$. Or, it could be a moduli space for principal Higgs bundles for another structure group [25, 26]. In that case, $\mathrm{End}(E)$ below would be replaced by the adjoint Higgs bundle—but for simplicity we will stick to the vector case.

The Zariski tangent space to $M$ at $E$ is given by the first cohomology group $H^1_{\mathrm{Dol}}(X, \mathrm{End}(E))$. Note that although $E$ may have nonzero degree, $\mathrm{End}(E)$ has degree zero, and corresponds to a flat vector bundle. If $E$ is a fixed point of $\mathbf{C}^*$, then $E$ has a structure of Hodge bundles $E = \bigoplus E^{p,q}$. This structure corresponds to an action of $\mathbf{C}^*$ on $E$, with $t$ acting by $t^p$ on $E^{p,q}$. If $E$ is stable, this action is unique up to tensoring with a scalar (resulting in a translation of the indices $(p, q)$). The associated action on $\mathrm{End}(E)$ is uniquely determined, corresponding to a structure of system of Hodge bundles $\mathrm{End}(E) = \bigoplus_p \mathrm{End}(E)^{p,-p}$. Then $\mathbf{C}^*$ acts on the Dolbeault cohomology $H^i(X, \mathrm{End}(E))$, combining the action on $\mathrm{End}(E)$ with the action on the complex of differentials, by a character $t^m$ on $\Omega^m_X$. This action corresponds to the Hodge decomposition of $H^i(X, \mathrm{End}(E))$: $t$ acts on $H^{p,q}(X, \mathrm{End}(E))$ by $t^p$. The resulting action on the Zariski tangent space $H^1(X, \mathrm{End}(E))$ is the same as the action induced by the action of $\mathbf{C}^*$ on $M_{\mathrm{Higgs}}$ (it takes a bit of thought to check this). At a point $E$ in a fixed point set $V_\alpha$ we have the decomposition

$$T_E M = T(S_{\alpha,E}) \oplus T(V_\alpha) \oplus T(T_{\alpha,E})$$

into the tangent space of the fiber of $S_\alpha$ over the point $E$, the tangent space to the fixed point set, and the tangent space to the fiber of $T_\alpha$ over $E$.

Applying our discussion of the action on the Zariski tangent space, we have

$$T(S_{\alpha,E}) = \bigoplus_{p>0} H^{p,1-p}_{\mathrm{Dol}}(X, \mathrm{End}(E)),$$

$$T(V_\alpha) = H^{0,1}_{\mathrm{Dol}}(X, \mathrm{End}(E)),$$

and

$$T(T_{\alpha,E}) = \bigoplus_{p<0} H^{p,1-p}_{\mathrm{Dol}}(X, \mathrm{End}(E)).$$

When $X$ is a curve, we have the additional information that, if $E$ is stable, $H^0_{\mathrm{Dol}}(\mathrm{End}(E))$ and $H^2_{\mathrm{Dol}}(\mathrm{End}(E))$ are one dimensional, with Hodge types $(0, 0)$ and $(1, 1)$ respectively. The first is the statement that the only endomorphisms of $E$ are the constants, and the second follows by duality. Thus we may calculate the dimension of $H^{p,1-p}_{\mathrm{Dol}}(X, \mathrm{End}(E))$ by Riemann-Roch. That space is the first hypercohomology

$$\mathbf{H}^1(\mathrm{End}(E)^{p,-p} \to \mathrm{End}(E)^{p-1,1-p} \otimes \Omega^1_X).$$

If $p \neq 0, 1$, then the $\mathbf{H}^0$ and $\mathbf{H}^2$ are zero, so the dimension is equal to the holomorphic Euler characteristic of $\mathrm{End}(E)^{p,-p}$ minus that of $\mathrm{End}(E)^{p-1,1-p} \otimes \Omega^1_X$. If $p = 0$ or $p = 1$ then the $\mathbf{H}^0$ or $\mathbf{H}^2$ is one-dimensional, so the dimension is equal to the difference in Euler characteristics, plus one.

## 4. An analogue of the theorem of Culler and Shalen

The work described in this section was done in January 1990. It arose from discussions with D. Toledo, J. Carlson, W. Goldman, and K. Corlette. The theorem is motivated by work of M. Gromov [**14**], and M. Green and R. Lazarsfeld [**11**] J. Carlson and D. Toledo have already proved a local version of the theorem [**3**].

Recently it was pointed out to me that the idea that a harmonic map might factor through a Riemann surface appeared some time ago in the work of Jost and Yau [**18**]. Also I have received a preprint by Zuo which treats some similar questions in the classification of two-dimensional representations [**29**].

Culler and Shalen proved that if a group admits a nontrivial family of irreducible representations into $\mathrm{Sl}(2)$, then it has an amalgamation decomposition described by a "graph of groups" [**6**]. We will prove a geometric analogue of this theorem for fundamental groups of smooth projective varieties. The idea is that the geometric version of a graph is a Riemann surface possibly with a simple orbifold structure (an "orbicurve"—see below). We will show that if the fundamental group of a smooth projective variety $X$ admits a nonrigid representation into $\mathrm{Sl}(2)$, then there is an algebraic map from $X$ to an orbicurve and a representation of the fundamental group of the orbicurve into $\mathrm{PSl}(2)$, which when pulled back to $X$ agrees with the original representation. We avoid complications having to do with the center $\{\pm 1\}$

of $\mathrm{Sl}(2)$ by assuming that the original representation is lifted into $\mathrm{Sl}(2)$, and by constructing the representation for the orbicurve, and comparing the two, in $\mathrm{PSl}(2)$.

We will adopt a simple definition of orbifold structure for a Riemann surface, in lieu of a more abstract formulation in terms of algebraic stacks. Define an *orbicurve* $(Y, \mathbf{n})$ to be a smooth projective algebraic curve $Y$ with some marked points $s_i \in Y$ and positive integers $n_i \geq 2$ attached to the marked points.

Let $\gamma_i$ denote elements in $\pi_1(Y - \{s_i\})$ which represent loops around the points $s_i$ respectively. The fundamental group $\pi_1(Y, \mathbf{n})$ is defined to be the quotient of $\pi_1(Y - \{s_i\})$ obtained by imposing the relations $\gamma_i^{n_i} = 1$.

If $X$ is an algebraic variety, we can define the notion of an algebraic map from $X$ to an orbicurve $(Y, \mathbf{n})$. This consists of an algebraic map $\phi: X \to Y$, with the following lifting property. If $t$ is a coordinate function on a neighborhood of $s_i$ with a simple zero at $s_i$, then the pullback function $\phi^*(t)$ should have an $n_i$th root, locally on $X$ in the analytic topology. In other words, for any $x \in X$ such that $\phi(x) = S_i$, there exists a function $u$ defined on a neighborhood of $x$ such that $u^{n_i} = \phi^*(t)$. Geometrically this means that the fiber $\phi^{-1}(s_i)$ is a union of divisors each with multiplicity divisible by $n_i$. Given an algebraic map $\phi: X \to (Y, \mathbf{n})$, we obtain a map of fundamental groups $\phi_*: \pi_1(X) \to \pi_1(Y, \mathbf{n})$.

THEOREM 10. *Suppose $X$ is a smooth projective variety, and $\rho: \pi_1(X) \to \mathrm{Sl}(2, \mathbf{C})$ is a representation. Suppose that the image of $\rho$ is Zariski-dense. Then there are two possibilities*:

(a) *The representation $\rho$ is rigid, in other words any nearby representation is conjugate to it*; *or*

(b) *There exists an orbicurve $(Y, \mathbf{n})$, an algebraic map $\phi: X \to (Y, \mathbf{n})$, and a representation $\tau: \pi_1(Y, \mathbf{n}) \to \mathrm{PSl}(2, \mathbf{C})$, such that $\rho$ is the pullback of $\tau$ via $\phi$, i.e., $\rho(\gamma) = \tau\phi_*(\gamma)$ in $\mathrm{PSl}(2, \mathbf{C})$.*

REMARK. In case (a), $\rho$ comes from a variation of Hodge structure, by Corollary 4. In fact, $\rho$ is a subquotient of a $\mathbf{Q}$-variation of Hodge structure **[25]**.

REMARK. If the coefficients of $\rho$ are algebraic integers, then in case (a) $\rho$ will come from a family of abelian varieties. In this case one can interpret the conclusion as saying there is either a map to an orbicurve, or to the moduli space of abelian varieties (rather to a Shimura subvariety). This situation is closely related to that treated by Mok **[21]**.

In the case when the image of $\rho$ is discrete and cocompact, the theorem follows from results of Siu **[27]** and Carlson and Toledo **[3]**. For nondiscrete representations, their arguments give a local statement like (b). The techniques we will use to obtain a global map to a curve are more akin to the arguments of Green and Lazarsfeld. The rigidity result (a) follows from a result of Corlette **[5]**.

The following corollary may alternatively be deduced from the arguments of Siu and Carlson and Toledo.

COROLLARY 11. *Suppose $X$ is a smooth projective variety, and $F$ is a free group on two or more generators. If $\pi_1(X) \to F \to 1$ is a surjection, then there exists a map $\phi: X \to Y$ to a Riemann surface and a factorization*

$$\pi_1(X) \to \pi_1(Y) \to F.$$

PROOF. There is a nontrivial faithful family of Zariski-dense representations from $F$ into $\mathrm{PSl}(2, \mathbf{C})$. Such may be obtained, for example, by uniformizing punctured Riemann surfaces. Also, since $F$ is torsion free, any representation of the fundamental group of an orbicurve comes from a representation of the fundamental group of the underlying curve; the equation $\gamma^n = 1$ implies that $\rho(\gamma) = 1$ in $F$.

REMARK. If $X_s$ is a degenerating family of varieties, with the central fiber $X_0$ given by a union of smooth divisors with normal crossings and multiplicities one, then we can form a simplicial complex representing the combinatorics of the intersections of components of the central fiber. Vertices correspond to components, edges to intersections, and faces to multiple intersections. There is a map from $X_s$ to $X_0$, which further projects in homotopy to a map to the simplicial complex. This map will have a lifting property for paths, so we obtain a surjection from $\pi_1(X_s)$ to the fundamental group of the simplicial complex. If the fundamental group of the simplicial complex has any nonabelian free quotients (for example if the simplicial complex is a graph with two or more cycles), then we can apply the corollary to find a map from $X_s$ to a family of Riemann surfaces $Y_s$ (possibly after modifying the central fiber). So, in a sense, one does not have complete freedom to make fundamental groups by considering smooth varieties degenerating to some arrangement.

PROOF OF THEOREM 10. We will suppose that condition (a) does not hold, i.e., $\rho$ is not rigid. There is a nontrivial algebraic family of representations, parametrized by an algebraic curve $\mathrm{Spec}(R)$ (which we may assume is defined over a finitely generated field $k \subset \mathbf{C}$),

$$\tilde{\rho}: \pi_1(X) \to \mathrm{Sl}(2, R),$$

such that $\rho$ is obtained by evaluating at a closed point. Let $\widetilde{\Gamma}$ denote the image of $\tilde{\rho}$. Let $\Gamma$ denote the image of $\widetilde{\Gamma}$ in $\mathrm{PSl}(2, R)$. The original representation $\rho$, projected into $\mathrm{PSl}(2, \mathbf{C})$, factors through $\Gamma$, so it suffices to construct a factorization

$$\pi_1(X) \to \pi_1(Y, \mathbf{n}) \to \Gamma.$$

Let $\eta$ denote a generic geometric point of $\mathrm{Spec}\, R \otimes_k \mathbf{C}$. The image of

$$\tilde{\rho}_n: \pi_1(X) \to \mathrm{PSl}(2, \mathbf{C})$$

is also equal to $\Gamma$. The moduli space $M_{\mathrm{Rep}}$ of representations of $\pi_1(X)$ is an affine algebraic variety (the "character variety" **[6]**), so the image of the algebraic curve of representations must go to infinity, in other words it is not contained in a compact subset. We may choose a family of generic geometric points $\eta_t$ going to infinity, and let $\rho_t = \tilde{\rho}_{\eta_t}$ (the projection into $\mathrm{Sl}(2, \mathbf{C})$). These are representations of $\pi_1(X)$ which go to infinity in the character variety; which factor through $\Gamma$ and in fact have image equal to $\Gamma$, and which are liftable to $\mathrm{Sl}(2, \mathbf{C})$.

Consider the Higgs bundles $E_t$ associated to the representations $\rho_t$. Choosing liftings to $\mathrm{Sl}(2, \mathbf{C})$, we may assume that the $E_t$ are Higgs bundles with $\det(E_t) = \mathscr{O}_X$ and $\mathrm{Tr}(\theta) = 0$. Let $\alpha_t^2$ denote the determinant of $\theta_t$. It is a section of the second symmetric power of the cotangent bundle, and where nonzero locally looks like the square of a one-form $\pm\alpha_t$. If we think of $(E_t, \theta_t)$ as a sheaf $\mathscr{E}_t$ on $T^*X$, then the support of $\mathscr{E}_t$ is set-theoretically the graph of the multivalued section $\pm\alpha_t \colon X \to T^*X$.

LEMMA 12. *For large values of $t$, $\alpha_t$ is not identically zero.*

PROOF. If $\alpha_t = 0$ for all values of $t$, then the supports of $\mathscr{E}_t$ would remain at the zero section. In particular, they would remain in a compact subset of $T^*X$. By the discussion of §2, the points $\mathscr{E}_t$ would remain in a compact subset of the moduli space of Higgs bundles. Since the correspondence between Higgs bundles and representations gives a topological homeomorphism of moduli spaces, the representations $\rho_t$ would remain in a compact subset of the character variety, contradicting our choice.

We may find $t$ such that $\alpha_t$ does not vanish, and such that the image of the representation $\rho_t$ is equal to $\Gamma$. We may also assume that $\rho_t$ is Zariski-dense. Replacing $\rho$ by this $\rho_t$, we have reduced to the following situation: $\rho$ is a representation $\pi_1(X) \to \mathrm{PSl}(2)$ whose image is the quotient $\Gamma$ of $\pi_1(X)$. The image is Zariski-dense. The corresponding Higgs bundle $(E, \theta)$ has the property that $\det(\theta) = \alpha^2 \neq 0$, whereas $\mathrm{Tr}(\theta) = 0$. The eigenvalues of $\theta$ are locally the square roots $\pm\alpha$. We have to show that $\rho$ factors through the fundamental group of an orbicurve.

Let $U \subset X$ denote the set of points where $\alpha^2 \neq 0$. Over $U$, the multivalued one-form $\pm\alpha$ determines an unramified double covering $W \to U$, with a single valued one-form $\tilde{\alpha}$ such that $\tilde{\alpha}^2 = \alpha^2$. For brevity we will replace the notation $\tilde{\alpha}$ simply by $\alpha$. $W$ might be disconnected, indicating that $\alpha$ is defined on $U$. In this case, some parts of the following discussion are unnecessary.

By taking the normalization of the coordinate ring of $X$ in the coordinate ring of $W$ and then resolving singularities **[15]**, we may assume that there is a smooth variety $\psi \colon Z \to X$ such that $W$ is an open set in $Z$. Let $\widetilde{E} = \psi^* E$. The support of the corresponding coherent sheaf $\widetilde{\mathscr{E}}$ on $T^*Z$ is, over $W$, a union of two connected components. Therefore it is a union of two irreducible components everywhere, and we may choose one of these components as the graph of a section. This gives an extension to a one-form

$\alpha$ on $Z$, which is an eigenvalue of $\theta$. Note that $W$ is exactly the open set of $Z$ where $\alpha$ does not vanish. The eigenspaces of $\alpha$ and $-\alpha$ are subsheaves $L \subset \widetilde{E}$ and $M \subset \widetilde{E}$ respectively. Over the open set $W$, these subsheaves are line bundles preserved by $\theta$, and in fact upon restricting to $W$ there is a decomposition of Higgs bundles

$$E_W = L_W \oplus M_W.$$

Use the one-form $\alpha$ to map $Z$ into an Albanese variety. Fix a base point $P \in Z$. Recall that

$$\mathrm{Alb}(Z) = H^0(Z, \Omega^1_Z)^*/H_1(Z, \mathbf{Z})$$

is an abelian variety, the *Albanese variety* of $Z$. There is a map $\Psi: Z \to \mathrm{Alb}(Z)$ sending $P$ to the origin, defined by integration:

$$\Psi(z)(\eta) = \int_P^z \eta.$$

The one-form $\alpha$ gives a linear function on the tangent space of $\mathrm{Alb}(Z)$. Let $D \subset \mathrm{Alb}(Z)$ be the sum of all the abelian subvarieties $B' \subset \mathrm{Alb}(Z)$ such that $\alpha|_{B'} = 0$. Let $A$ be the quotient abelian variety $A = \mathrm{Alb}(Z)/B$. Projection gives a map $\Psi: Z \to A$.

LEMMA 13. *Let $D = Z - W$. Then the image of $D$ under the map $Z \to A$ is a finite set of points.*

PROOF. Look at an irreducible component of $D$, which we may think of as the image of a smooth variety $C \to Z$ **[15]**. The one-form $\alpha$ is zero on $D$, so it pulls back to zero on $C$. Therefore the image of $\mathrm{Alb}(C) \to \mathrm{Alb}(Z)$ is a subabelian variety on which $\alpha$ vanishes. It is contained in $B$. The image of $C$ in $\mathrm{Alb}(Z)$ is contained in a translate of $\mathrm{Alb}(C)$ (the translation depends on how far $C$ is from the base point $P$). Therefore the image of $C$ in $A = \mathrm{Alb}(Z)/B$ is one point. This proves the lemma.

We would like to pass down to $U \subset X$ rather than $W \subset Z$. If $W$ is a disjoint union of two copies of $U$ then this is easy and the next few paragraphs may be ignored. Otherwise, $W$ is a Galois covering of $U$ with Galois group $\{1, \sigma\}$. The involution $\sigma$ takes $\alpha$ to $\sigma^*\alpha = -\alpha$.

LEMMA 14. *If $\eta$ is a one-form on $Z$ such that $\eta|_B = 0$, then $\sigma^*\eta = -\eta$.*

PROOF. We may assume that the resolution of singularities $Z$ is compatible with the action of $\sigma$. (This may be accomplished by resolving the singularities of the ramification locus in $X$ before taking the double cover; then there are at most conical singularities, and these can be resolved with one more blow up, which is canonical.) Therefore the space of one-forms decomposes into $+1$ and $-1$ eigenvalues of $\sigma$. The rational homology decomposes similarly, so the Albanese variety is isogenous to a product,

$$\mathrm{Alb}(Z) \sim \mathrm{Alb}^+(Z) \times \mathrm{Alb}^-(X).$$

Now $\alpha$ is a $-1$ eigenvalue. In particular, the subvariety $\mathrm{Alb}^{+}(X)$ is contained in $B$. Thus if $\eta$ vanishes on $B$, it must come from a form on $\mathrm{Alb}^{-}(Z)$, so it is a $-1$ eigenvalue.

The map $\Psi\colon W \to A$ takes $\sigma$ to the involution $-1$, up to a translation. Let $\xi = \Psi(\sigma(P)) - \Psi(P)$.

$$\Psi(\sigma(z))(\eta) = \int_P^{\sigma(P)} \eta + \int_{\sigma(P)}^{\sigma(z)} \eta = \int_P^{\sigma(P)} \eta + \int_P^{z} \sigma^* \eta$$

and $\sigma^*(\eta) = -\eta$ for one-forms $\eta$ which are pulled back from $A$, so

$$\Psi(\sigma(z)) = \xi - \Psi(z).$$

We can define an involution $\tau : A \to A$ by $\tau(a) = \xi - a$. Then $\Psi(\sigma(z)) = \tau\Psi(z)$, so in particular the map $\Psi\colon W \to A$ descends to a map

$$F\colon U \to A/\{1, \tau\}.$$

This gives a rational map from $X$ to $A/\{1, \tau\}$, so there is a birational transform $\widetilde{X} \to X$, isomorphic over $U$, and a map $F\colon \widetilde{X} \to A/\{1, \tau\}$. We may assume that $Z$ maps to $\widetilde{X}$. The inverse image in $Z$ of $\widetilde{X} - U$ is equal to the complement $D = Z - W$. Lemma 13 implies that the image $F(\widetilde{X} - U) \subset A/\{1, \tau\}$ is contained in a finite set.

By Lemma 12, the dimension of $F(\widetilde{X})$ is greater than or equal to one. There are two cases.

*Case* 1: $\dim F(\widetilde{X}) \geq 2$. We will obtain a contradiction. The basic idea is that from the point of view of the Lefschetz theorems, the complement of $U$ has effectively codimension 2 (because of Lemma 13). There is a projective curve in $U$ whose fundamental group surjects onto $\pi_1(X)$. Restricting to this curve or a double cover, the Higgs bundle decomposes, contradicting the assumed Zariski density.

We make this argument more precise. Let $n = \dim(X)$ and $d = \dim(F(\widetilde{X})) \geq 2$. By making a further birational modification we may assume that there is a projection $\widetilde{X} \to \mathbf{P}^{n-d}$, such that the map

$$\widetilde{X} \to \mathbf{P}^{n-d} \times F(\widetilde{X})$$

is generically finite. This further modification might modify $U$, but this is not a problem since we are obtaining a contradiction. Let $Y$ be the normalization of $\mathbf{P}^{n-d} \times F(\widetilde{X})$ in the function field of $\widetilde{X}$. Then we have a birational map $\widetilde{X} \to Y$. The image of $\widetilde{X} - U$ maps to a finite set in $F(\widetilde{X})$, so it has codimension $\geq 2$ in $Y$. The image of the exceptional locus also has codimension $\geq 2$. Therefore we may choose an open set $V \subset Y$, the complement of a set of codimension $\geq 2$, which is identified with its inverse image $V$ in $\widetilde{X}$, and such that $V \subset U$. By the Lefschetz theorems **[20]**, we may choose a smooth closed curve $H$ in $Y$ such that $H \subset V$ and $\pi_1(H) \to \pi_1(V) \to 1$ is a surjection. In particular, $H$ may be identified with a smooth projective curve contained in $U$. The fundamental group of $U$ surjects onto the fundamental group of $X$, so we have

$$\pi_1(H) \to \pi_1(X) \to 1.$$

On the other hand, there is a double or single cover $K \to H$, induced by the cover $W \to U$, such that $E|_K \cong L_K \oplus M_K$ is a direct sum of Higgs bundles. Since $K$ is a projective curve, this implies that $\rho|_{\pi_1(K)}$ has image contained in a torus in PGl(2). Let $\Gamma' \subset \Gamma$ be the image of $\pi_1(K)$. It is either equal to $\Gamma$, or has index two. Let $T$ be the Zariski closure of $\Gamma'$ in PSl(2). Double covers are always Galois, so $\Gamma$ normalizes $\Gamma'$, and hence $\Gamma$ is contained in the normalizer of $T$. If $T$ is nontrivial, this normalizer is a proper subgroup, contradicting Zariski density of $\Gamma$. On the other hand, if $T$ is trivial, then $\Gamma'$ is trivial and $\Gamma$ has order two, again contradicting Zariski density. So Case 1, that the dimension of the image of $F$ is greater than or equal to two, cannot happen.

*Case* 2: $\dim(F(\widetilde{X})) = 1$. We will obtain a map to an orbicurve, and factorization of the representation. The image of $F$ is a curve $Y' \subset A/\{1, \tau\}$. The image of $\widetilde{X} - U$ is contained in a finite set in $Y'$. Take the Stein factorization, and note that he image of the first map in the factorization is normal. We get a smooth projective curve $Y$ and a map $f: \widetilde{X} \to Y$.

LEMMA 15. *This factors through the original variety* $f: X \to Y$.

PROOF. Let $E$ denote the exceptional locus of the map $g: \widetilde{X} \to X$. The exceptional locus is contained in the complement of $U$, so it is contained in finitely many fibers of $f$. Choose a point $y \in Y$, and a local coordinate $t$ at $y$. This pulls back to a function $\tilde{t}$ on $\widetilde{X}$, defined on a neighborhood of the component $E'$ of the exceptional set which maps to $y$. This neighborhood is the inverse image of an open set in $X$, so $\tilde{t}$ gives a section of $g_*\mathscr{O}_{\widetilde{X}}$. However, since $X$ is normal, $g_*\mathscr{O}_{\widetilde{X}} = \mathscr{O}_X$, in other words $\tilde{t}$ descends to a function on $X$. These descended functions provide coordinates for a map from $X$ to $Y$.

So we finally have a map $f: X \to Y$ to a Riemann surface. There is a finite set $S \subset Y$ which contains the set of points where the map is not smooth, and the image of $X - U$. Let $Y^0 = Y - S$ and $V = f^{-1}(Y^0)$. $V$ is an open subset of $U$. For $y \in Y^0$, let $C_y = f^{-1}(y)$. It is a connected smooth projective variety contained in $U$. We have the fibration sequence in homotopy:

$$1 \to \pi_1(C_y) \to \pi_1(V) \to \pi_1(Y') \to 1.$$

Let $A_y$ be the single or double cover of $C_y$ induced by $W \to U$. Then the Higgs bundle $E$, restricted to $A_y$, decomposes as a direct sum of rank one Higgs bundles,

$$E|_{A_y} = L|_{A_y} \oplus M|_{A_y}.$$

Therefore the representation $\rho$ maps $\pi_1(A_y)$ into a torus in $\mathrm{Psl}(2)$. Let $G$ denote the Zariski closure of the image of $\pi_1(C_y)$ in $\mathrm{Psl}(2)$. Since $\pi_1(A_y)$ has index at most two in $\pi_1(C_y)$, $G$ is a proper subgroup of $\mathrm{Psl}(2)$ (see the argument at the end of Case 1).

If $G$ is not trivial, then its normalizer $N(G)$ is a proper subgroup of $\mathrm{Psl}(2)$ ($\mathrm{Psl}(2)$ contains no normal subgroups). But since $\pi_1(C_y)$ is normal in $\pi_1(V)$, the image of $\pi_1(V)$ is contained in the normalizer $N(G)$ of the Zariski closure of $\pi_1(C_y)$. Since the fundamental group of the open set $V$ surjects onto the fundamental group of $X$, this contradicts the assumed Zariski density of the image of $\pi_1(X)$.

Therefore $G$ is trivial, so $\rho$ restricts to a trivial representation on $\pi_1(C_y)$. In other words, $\rho$ factors through a representation $\tau\colon \pi_1(Y^0) \to \mathrm{Psl}(2)$. Finally, we have to show that $Y$ can be given an orbifold structure such that $f\colon X \to (Y, \mathbf{n})$, and such that $\tau$ factors through $\pi_1(Y, \mathbf{n})$. We will use the fact that the representation $\rho = f^*\tau$ of $\pi_1(V)$ factors through $\pi_1(X)$.

Choose $s \in S$ and let $\gamma$ denote a path which goes around $s$. By intersecting generic hyperplane sections of $X$, we may find a curve $H \subset X$ such that $H \to Y$ is a ramified cover. Let $n_0$ denote the ramification index over the point $s$. There is a loop $\mu$ in $\pi_1(H \cap V)$ which projects to $\gamma^{n_0}$. On the other hand, the loop $\mu$ becomes trivial in $\pi_1(X)$. Since $f^*\tau$ factors through a representation of $\pi_1(X)$, it follows that $f^*\tau(\mu) = \tau(\gamma^{n_0}) = 1$. Thus $\tau(\gamma)$ has finite order.

Let $n = n(s)$ be the order of $\tau(\gamma)$. Let $\Delta^* \subset \Delta$ denote punctured and unpunctured neighborhoods of $y$, and let $Q^* \subset Q$ denote the inverse images in $X$. The representation $\tau$ is a finite representation of the cyclic fundamental group of $\Delta^*$. Let $\widetilde{\Delta}$ denote the ramified covering of degree $n$ corresponding to this representation. Let $\widetilde{Q}$ denote the pull-back to $X$. The fact that $f^*\tau$ factors through a representation of $\pi_1(X)$ means that the covering $\widetilde{Q}^*$ has an extension $\widetilde{Q}_1$ to an unramified covering of $Q$. If $t$ is a local coordinate on $\Delta$, then we can extract an $n$th root on $\widetilde{\Delta}$: there is a function $u$ (local coordinate also) on $\widetilde{\Delta}$ such that $u^n = t$. Pulling back, we obtain a function $f^{-1}(u)$ on $\widetilde{Q}$. In particular, the function $f^{-1}(u)$ is bounded on $\widetilde{Q}^*$. Since $\widetilde{Q}_1$ is obtained from $\widetilde{Q}^*$ by adding a compact divisor, we can extend $f^{-1}(u)$ to a holomorphic function $u_1$ on $\widetilde{Q}_1$, such that $u_1^n = f^{-1}(t)$. Since $\widetilde{Q}_1$ is unramified over the neighborhood $Q$ of the fiber $f^{-1}(s)$, the function $u_1$ provides the local analytic root extraction sought in the definition of a map to an orbicurve.

Give $Y$ the structure of orbicurve by assigning the numbers $n(s)$ to the points $s \in S$. The map $f\colon X \to Y$ becomes a map from $X$ to the orbicurve $(Y, \mathbf{n})$. Furthermore, since $n(s)$ is the order of $\tau(\gamma_s)$, the representation

$$\tau\colon \pi_1(Y^0) \to \mathrm{Psl}(2, \mathbf{C})$$

factors through

$$\tau\colon \pi_1(Y, \mathbf{n}) \to \mathrm{PSl}(2, \mathbf{C}).$$

The pullback $f^*\tau$ is representation of $\pi_1(X)$ in $\mathrm{PSl}(2, \mathbf{C})$. On the other hand, there is the open set $V \subset X$ such that $f\colon V \to Y^0$, and such that $f^*\tau = \rho$ as representations of $\pi_1(V)$. Since the fundamental group of $V$ surjects to the fundamental group of $X$, $\rho = f^*\tau$ as representations of $\pi_1(X)$. This proves Theorem 10.

## References


1. M. F. Atiyah and R. Bott, *The Yang-Mills equations over Riemann surfaces*, Philos. Trans. Roy. Soc. London Ser. A **308** (1982), 523–615.
2. A. Bialynicki-Birula, *Some theorems on actions of algebraic groups*, Ann. of Math. (2) **98** (1973), 480–497.
3. J. A. Carlson and D. Toledo, *Harmonic mappings of Kähler manifolds to locally symmetric spaces*, Inst. Hautes Ètudes Sci. Publ. Math. **69** (1989), 173–201.
4. K. Corlette, *Flat G-bundles with canonical metrics*, J. Differential Geom. **28** (1988), 361–382.
5. ——, *Rigid representations of Kählerian fundamental groups*, preprint, University of Chicago.
6. M. Culler and P. B. Shalen, *Varieties of group representations and spittings of 3-manifolds*, Ann. of Math. (2) **117** (1983), 109–146.
7. P. Deligne, *Un théorème de finitude pour la monodromie*, Discrete Groups in Geometry and Analysis, Birkhäuser, 1987, pp. 1–19.
8. S. K. Donaldson, *Anti self dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles*, Proc. London Math. Soc.(3) **50** (1985), 1–26.
9. ——, *Infinite determinants, stable bundles, and curvature*, Duke Math. J. **54** (1987), 231–247.
10. ——, *Twisted harmonic maps and self-duality equations*, Proc. London Math. Soc. (3) **55** (1987), 127–131.
11. M. Green and R. Lazarsfeld, *Higher obstructions to deforming cohomology groups of line bundles*, J. Amer. Math. Soc. **4** (1991), 87–104.
12. P. Griffiths, *Periods of integrals on algebraic manifolds.* I, Amer. J. Math. **90** (1968), 568–626.
13. ——, *Periods of integrals on algebraic manifolds.* III, Inst. Hautes Ètudes Sci. Publ. Math. **38** (1970) 125–180.
14. M. Gromov, *Sur le groupe fondamental d'une variété kählerienne*, C. R. Acad. Sci. Paris Ser. I Math. **308** (1989), 67–70.
15. H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero.* I, II, Ann. of Math. (2) **79** (1964), 109–326.
16. N. J. Hitchin, *The self-duality equations on a Riemann surface*, Proc. London Math. Soc. (3) **55** (1987), 59–126.
17. ——, *Stable bundles and integrable systems*, Duke Math. J. **54** (1987), 91–114.
18. J. Jost and S. T. Yau, *Harmonic mappings and Kähler manifolds*, Math. Ann. **262** (1983), 145–166.
19. F. Kirwan, *Cohomology of quotients in symplectic and algebraic geometry*, Math. Notes, vol. **31**, Princeton Univ. Press, Princeton, NJ, 1984.
20. S. Lefschetz, *L'analysis situs et la géométrie algébrique*, Gauthier-Villars, Paris, 1924.
21. N. Mok, *The holomorphic or anti-holomorphic character of harmonic maps into irreducible compact quotients of polydiscs*, Math. Ann. **272** (1985), 197–216.
22. M. S. Narasimhan and C. S. Seshadri, *Stable and unitary bundles on a compact Riemann surface*, Ann. of Math. (2) **82** (1965), 540–564.

23. W. M. Oxbury, *Spectral curves of vector bundle endomorphisms*, preprint, Kyoto University.
24. C. T. Simpson, *Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization*, J. Amer. Math. Soc. **1** (1988), 867–918.
25. ——, *Higgs bundles and local systems*, preprint, Princeton University.
26. ——, *Moduli of representations of the fundamental group of a smooth projective variety*, preprint, Princeton University.
27. Y. T. Siu, *Complex analyticity of harmonic maps and strong rigidity of complex Kähler manifolds*, Ann. of Math. (2) **112** (1980), 73–110.
28. K. K. Uhlenbeck and S. T. Yau, *On the existence of Hermitian-Yang-Mills connections in stable vector bundles*, Comm. Pure Appl. Math. **39-S** (1986) 257–293.
29. K. Zuo, *Two dimensional semisimple representation spaces of the fundamental groups of algebraic manifolds*. I, preprint, Max-Planck-Institut für Mathematik, 1990.

Laboratoire de Topologie et Géométrie, Université Paul Sabatier, 31062 Toulouse CEDEX, France

*E-mail address*: mestrano@frcict81.bitnet