PUBLICATIONS MATHÉMATIQUES DE L'I.H.É.S.

CARLOS T. SIMPSON

Higgs bundles and local systems

Publications mathématiques de l'I.H.É.S., tome 75 (1992), p. 5-95.
<http://www.numdam.org/item?id=PMIHES_1992__75__5_0>

HIGGS BUNDLES AND LOCAL SYSTEMS [1]

by CARLOS T. SIMPSON

CONTENTS

[1] Prepared with the partial support of NSF grant DMS-8705757.

0. Introduction

Let X be a compact Kähler manifold. In this paper we will study a correspondence between representations of the fundamental group of X, and certain holomorphic objects on X. A *Higgs bundle* is a pair consisting of a holomorphic vector bundle E, and a holomorphic map $\theta : E \to E \otimes \Omega^1_X$ such that $\theta \wedge \theta = 0$. There is a condition of stability analogous to the condition for vector bundles, but with reference only to subsheaves preserved by the map θ. There is a one-to-one correspondence between irreducible representations of $\pi_1(X)$, and stable Higgs bundles with vanishing Chern classes. This theorem is a result of several recent extensions of the work of Narasimhan and Seshadri [39], [5], [16], [17], [18], [30], [37], [49], [52], [47]. The purpose of this paper is to discuss this correspondence in detail, to obtain some further properties, and to give some applications.

The correspondence between Higgs bundles and local systems can be viewed as a Hodge theorem for non-abelian cohomology. To understand this, let us first look at abelian cohomology: $H^1(X, \mathbf{C})$ can be thought of as the space of homomorphisms from $\pi_1(X)$ into $\mathbf{C}$, or equivalently as the space of closed one-forms modulo exact one-forms. But since X is a compact Kähler manifold, the Hodge theorem gives a decomposition

$$H^1(X, \mathbf{C}) = H^1(X, \mathcal{O}_X) \oplus H^0(X, \Omega^1_X).$$

In other words, a cohomology class can be thought of as a pair (e, ξ) with $e \in H^1(X, \mathcal{O}_X)$ and ξ a holomorphic one-form. The correspondence between Higgs bundles and local systems is analogous. If $\pi_1(X)$ acts trivially on $Gl(n, \mathbf{C})$ then the non-abelian cohomology set $H^1(\pi_1(X), Gl(n, \mathbf{C}))$ is the set of representations $\pi_1(X) \to Gl(n, \mathbf{C})$, up to conjugacy. Equivalently it is the set of isomorphism classes of C^∞ vector bundles with flat connections. The theorem stated above gives a correspondence between the set of semisimple representations and the set of pairs (E, θ) where E is a holomorphic bundle (in other words, an element of $H^1(X, Gl(n, \mathcal{O}_X))$ and θ is an endomorphism valued one-form, subject to various additional conditions.

There is a natural action of $\mathbf{C}^*$ on the set of Higgs bundles. A nonzero complex number t sends (E, θ) to $(E, t\theta)$. This preserves the conditions of stability and vanishing of Chern classes, so it gives an action on the space of semisimple representations. This $\mathbf{C}^*$ action should be thought of as the Hodge structure on the semisimplified non-abelian cohomology.

Before describing the contents of the paper, let me make a comment about the length. Several different topics are covered in different sections, and while there are some interdependencies, they do not build linearly. So the reader might well be interested in skipping to selected parts and working backwards.

The basic linear algebra of the correspondence between Higgs bundles and local systems is described in § 1. The main construction is that a metric on a Higgs bundle or flat bundle leads to an operator corresponding to a structure of the opposite kind. This operator may not satisfy the required integrability condition—the obstruction is a curvature F_K or pseudocurvature G_K. If the (pseudo-)curvature vanishes then one has a *harmonic bundle*, a bundle with both structures of Higgs bundle and flat bundle related by a metric. The main existence theorem is that if a Higgs bundle is stable, or a flat bundle is irreducible, then the equations $\Lambda F_K = \lambda I$ or $\Lambda G_K = 0$ can be solved. In the Higgs case an extra assumption, that the Chern classes vanish, is required for concluding that $F_K = 0$. In the flat case, the stronger vanishing $G_K = 0$ is automatic, which is the statement of the main Lemma 1.1. We get an equivalence between the category of direct sums of stable Higgs bundles with vanishing Chern classes, and the category of semisimple local systems, through equivalences with the intermediate category of harmonic bundles. At the end of § 1 some facts about existence of moduli spaces for Higgs bundles are stated without proofs, in the case when X is a smooth projective variety.

The classical Kähler identities for differential forms on X can be extended to the case of forms with coefficients in a harmonic bundle. The principal consequence is that the complexes of forms with coefficients in a harmonic bundle are *formal.* This provides a natural quasiisomorphism between the de Rham complex of forms with coefficients in the flat bundle, and the Dolbeault complex with coefficients in the corresponding Higgs bundle (with differential including θ). There is a natural duality statement, and also a Lefschetz decomposition for cohomology with coefficients in a semisimple local system. The next topic in § 2 is a crucial compactness property: the set of harmonic bundles with a fixed bound on the eigenvalues of θ is compact. One can conclude that the map from Higgs bundles to flat bundles is continuous. The section is closed with a brief discussion of monodromy groups and real structures. These topics are treated in greater generality in § 6, so the proofs here are redundant, but it seems worthwhile to have a straightforward introductory version.

In § 3 we will discuss a way of extending the correspondence between stable Higgs bundles and irreducible representations, to a correspondence between semistable Higgs bundles (with vanishing Chern classes) and possibly reducible representations. The reason this is possible is the formality of the complexes of forms with coefficients in harmonic bundles discussed in the previous section. We introduce some machinery, of *differential graded categories*, to carry out the argument. It is a generalization of the notion of differential graded algebra to the case when there are several different underlying objects. The theory of extensions of semisimple objects is governed by a differential graded category, and formality of the differential graded category gives a trivialization of the theory of extensions, as well as an isomorphism between the de Rham and Dolbeault theories. In order to obtain the best hypotheses, we need to make a digression and prove the theorems of Mehta and Ramanathan in the case of Higgs bundles. These say that the restriction of a semistable or stable Higgs bundle to a sufficiently general hyperplane remains semistable or stable. This allows us to prove that a semistable Higgs bundle with vanishing Chern classes is actually an extension of stable Higgs bundles (rather than stable Higgs sheaves). Finally, at the end of the section we go back to the formalism of differential graded categories, and introduce the notion of tensor product structure. The extended correspondence between semistable Higgs bundles and representations is compatible with tensor product.

There is an important class of representations of the fundamental group, the variations of Hodge structure. This class of representations was first considered by Griffiths in connection with his study of the monodromy of cohomology in smooth families of varieties [24]. Griffiths' original notion of a variation of Hodge structure with integral lattice can be weakened to the notion of complex variation of Hodge structure [8]. In § 4 we show that the representations which come from complex variations of Hodge structure can be characterized as the fixed points of the action of $\mathbf{C}^*$ on the space of semisimple representations. A consequence is that any rigid representation of the fundamental group of a compact Kähler manifold must come from a complex variation of Hodge structure. (In fact it turns out that a rigid representation must be a complex direct factor of a rational variation of Hodge structure.)

This places restrictions on which groups can occur as fundamental groups of compact Kähler manifolds. This is because there are restrictions on which groups may occur as the real Zariski closure of the monodromy group of a complex variation of Hodge structure. The groups which may occur we say are *of Hodge type*. The groups which are not of Hodge type include all complex groups, $\mathrm{Sl}(n, \mathbf{R})$, and some others listed in § 4. On the other hand, lattices in semisimple groups often provide examples of rigid representations [54], [35], [42]. A rigid lattice in a group which is not of Hodge type cannot be the fundamental group of a compact Kähler manifold. This rules out, for example, $\mathrm{Sl}(n, \mathbf{Z})$ for $n \geqslant 3$, or co-compact lattices in complex groups or other groups which are not of Hodge type. Similar topological restrictions on Kähler manifolds using harmonic maps have been obtained in many works. The first in this line was Siu [49], then Sampson [44], and others. Carlson and Toledo prove (among other things) that a discrete co-compact lattice in $\mathrm{SO}(n, 1)$ cannot be the fundamental group of a compact Kähler manifold [4]. For the groups $\mathrm{SO}(n, 1)$, this is better than our statement, which only applies when n is odd.

If X is a smooth projective variety, then one can construct an algebraic moduli space for direct sums of stable Higgs bundles with vanishing Chern classes. We will not give the construction in this paper, but only the statements. Although the moduli space is not projective, there is a proper map to a vector space. Using this proper map, the corollary about rigid representations can be extended to the statement that any representation of $\pi_1(\mathrm{X})$ can be deformed to a complex variation of Hodge structure. Using this fact, the nonexistence results for $\pi_1(\mathrm{X})$ can be extended somewhat, to groups which are semi-direct products with split quotients which are lattices ruled out as above. It seems reasonable to expect that a moduli space could be constructed in the Kähler case. These extra results would then hold in that case.

The theorem about rigid representations shows that they play a special role. It seems reasonable to make the following

Conjecture. — Rigid representations of the fundamental group of a smooth projective variety should be motivic.

To be more precise, a rigid irreducible representation should be a direct factor in the monodromy of a family of varieties. There would be several consequences of a representation's coming from a family of varieties. The first is that the representation would underly a complex variation of Hodge structure. We prove that this holds for rigid representations (Lemma 4.5). Second, a motivic representation would be a direct factor in a $\mathbf{Q}$-variation of Hodge structure, and the corresponding ℓ-adic representations would descend to a model of X over a number field. We also prove these properties for rigid representations in § 4. Another property would be integrality, because the monodromy representation of a family preserves the integral cohomology. This leads to a conjecture implied by the previous conjecture,

Any rigid irreducible representation of the fundamental group of a smooth projective variety should be defined over a ring of integers.

I do not know how to prove this (M. Larsen informs me of a simple example which demonstrates that it is not true for arbitrary discrete groups). Finally, let me remark that the properties called *absoluteness I and II* in [48] are immediate for rigid representations. So the above conjecture is actually a special case of the conjecture described in [48]. The results of § 4 may be viewed as proving the variation of Hodge structure and Galois type conjectures stated in [48], for the case of rigid representations.

The last two sections, 5 and 6, have been added after the preliminary versions of this paper were circulated. Some theorems are stated in § 5, the proofs given in § 6. The purpose is to interpret the results of the previous sections as a way of putting a Hodge structure on the fundamental group. The action of $\mathbf{C}^*$ on the space of Higgs bundles leads, via our correspondences, to an action of $\mathbf{C}^*$ on the pro-algebraic completion $\varpi_1(X, x)$ of the fundamental group. We formulate the notion of pure non-abelian Hodge structure, and show that the action of $U(1) \subset \mathbf{C}^*$ on the reductive quotient $\varpi_1^{\mathrm{red}}(X, x)$ provides an example of such a structure. We rephrase some of the results about rigid representations and variations of Hodge structure in this language. We also treat the nilpotent quotient $\varpi_1^{\mathrm{nil}}(X, x)$, which is the nilpotent completion of the fundamental group. The extended correspondence of § 3 provides an action of $\mathbf{C}^*$ here too, and we show that this action serves to define the Hodge filtration known by work of Morgan and Hain.

The proofs of the results stated in § 5 are given in § 6, using Tannakian categories. The reader should notice that the discussion is (with the exception of the part about the Hodge filtration on ϖ_1^{nil}) simply an application of the results discussed in the previous sections. We close § 6 with another application of the Tannakian formalism, to give definitions of principal objects or torsors, and to extend the correspondence to that case. The final result is a description of reductive representations with values in a real group, in terms of principal Higgs bundles with a Cartan involution. The rationale for using the somewhat complicated language of Tannakian categories here is that it demonstrates the intuitive point that one can prove existence theorems and the like in the vector ($\mathrm{Gl}(n)$) case, without having to deal with principal bundles in the beginning. Eventually, the results for principal bundles are obtained in an essentially formal way, using the information about tensor products.

The main result used in this paper, the correspondence between Higgs bundles and local systems, has the following origins. The first example of a correspondence between holomorphic objects and representations of the fundamental group was due to Narasimhan and Seshadri [39]. This was generalized to vector bundles and unitary connections in higher dimensions by Donaldson [16] [17], Mehta and Ramanathan [37], and Uhlenbeck and Yau [52]. Hitchin originated the definition of Higgs bundle in the case of objects on a curve [30]. He proved half of the correspondence in the rank two case, and the other half was provided in that case by Donaldson [18] (see also the paper by Diederich and Ohsawa [14], which treated the $\mathrm{SU}(1, 1)$ case). Hitchin considered the action of $\mathrm{U}(1) \subset \mathbf{C}^*$ on the space of Higgs bundles, and also a version of the compactness statement. Higgs bundles of the type which come from variations of Hodge structure were treated by Deligne and Beilinson (unpublished) and in [46]. For the general statement in higher ranks and higher dimensions, Corlette provides one half of the correspondence [5]. The other half is provided by [47]. A crucial lemma in the present treatment was communicated to me by Deligne. It is this lemma which allows one to apply Corlette's result and conclude that every semisimple representation comes from a Higgs bundle. A similar lemma is contained in Corlette's paper. This lemma is really a version of Siu's Bochner-type formula [49]—see also [44]. It should be noted that one of our main applications, the theorem that a rigid representation must come from a variation of Hodge structure, uses only that half of the correspondence which is provided by Corlette's paper (or Donaldson's) and this main lemma.

The Kähler identities for harmonic bundles are based on Deligne's Kähler identities for variations of Hodge structure [7]. The properties of formality and their applications discussed in § 3 are inspired by the work of Goldman and Millson [21]. The compactness result appeared as one of the main steps in Hitchin's paper [30]. The notion of group of Hodge type was essentially understood by Griffiths and Schmid and Deligne from the beginning [26] [11] [9] (and it is not clear whether the version discussed in § 4 is the optimal one). The ideas presented in § 5 were only recently fully formed. The idea of using the Tannakian formalism to put a Hodge structure on the fundamental group was partly instigated by an electronic message from K. Corlette. He contemplated using the Tannakian category of variations of mixed Hodge structure (which might lead to a theory somewhat distinct from that outlined in § 5).

Earlier versions of this paper have been circulated in preprint form. Some new material has been added since then, including everything in §§ 5 and 6. Some of the results of this paper have been announced in the paper " The ubiquity of variations of Hodge structure ".

I would like to thank the many people whose comments have led to improvements in this paper, ranging from brief comments to existence of large parts. I would like to thank K. Corlette for discussions about the overall framework into which the various existence theorems fit, which made it clear that they were two sides to the same coin. In this regard I would also like to thank W. Goldman for pointing out Hitchin's paper, and suggesting that it could probably be generalized, as well as for explaining with J. Millson their theory of deformations. I would like to thank N. Hitchin for helpful discussions about his correspondence, and about what should be done to take into account real structures. I am grateful to P. Deligne for the invaluable comments he has provided on many occasions. Principal among these was his letter stating the main lemma (1.1 below)—this vastly improved a weaker statement I had by showing that my hypothesis about vanishing of pseudo Chern classes was always true. Also important were suggestions that compatibility with tensor product, and the Lefschetz action of $\mathbf{sl}(2)$, should be considered. His communications have included helpful comments and corrections, as well as provocative ideas contributing to works beyond the present. Some years ago, W. Schmid asked me what would happen if you multiplied θ by a number t. At the time, the objects being considered were systems of Hodge bundles, so the answer was that you got an isomorphic system; in the later case of Higgs bundles, this question leads to one of the main concepts in this paper. So I would like to thank him for that comment in particular, as well as for more general encouragement. In the category of debts from many years ago, I should also like to thank the people at Harvard who fostered an environment of interest in the topics of representations of fundamental groups and harmonic maps. These include Y. Siu and his students, N. Boston and his fellow students of number theory, and many others. It is remarkable that, unbeknown to any of us, we were all working on the same things. Finally, back to more recent things, the discussion of real representations was added in response to encouragement from G. Kempf and S. Zucker. I would like to thank A. Beilinson for some useful comments about the analogy with Galois representations. I thank J. Le Potier for pointing out an error in Theorem 2 (which has been fixed in revision). And I would like to thank M. Larsen for his helpful explanations about questions of $\mathbf{Q}$-structure and other items in § 4.

1. Non-abelian Hodge theory

Let X be a compact complex manifold with a Kähler metric ω. Choose a base point x. We will describe some definitions and constructions, and then state some basic results. The discussion of the history and references for these definitions and constructions will be deferred until the statements of the main results.

We will study the representations of the fundamental group $\pi_1(X, x)$. A representation on a complex vector space V_x is the same thing as a C^∞ complex vector bundle V together with a flat connection, where V_x is equal to the fiber over x. The flat connection is a first order differential operator D which takes sections of the bundle to one-forms with coefficients in the bundle. Such an operator is a flat connection if and only if Leibniz's rule $D(av) = d(a)\,v + a\,D(v)$ holds, and it is integrable, in other words $D^2 = 0$. To understand the second condition note that D is extended to an operator on differential forms with values in the bundle, using the Leibniz formula with the usual sign which depends on the degree of the form. Another way of thinking of a flat bundle is by looking at the sheaf V^D of flat sections (those with $D(v) = 0$). This is a local system of complex vector spaces, whose monodromy representation is the one we began with. These objects depend only on the topological or smooth structure of X.

The purpose of this section is to establish a correspondence between flat bundles and another type of object which depends on the analytic structure of X. A *Higgs bundle* is a holomorphic vector bundle E together with a holomorphic map $\theta : E \to E \otimes \Omega^1_X$, such that $\theta \wedge \theta = 0$ in $\mathrm{End}(E) \otimes \Omega^2_X$. If $z_1, \ldots, z_n$ are local holomorphic coordinates, then $\theta = \Sigma \theta_i\, dz_i$, where θ_i are holomorphic endormorphisms of E. The condition that $\theta \wedge \theta = 0$ means that the matrices θ_i commute with one another. A Higgs bundle may also be thought of as a C^∞ bundle with a first order operator. The holomorphic structure of E is determined by an operator $\bar\partial$, which takes sections of E to $(0, 1)$ forms with coefficients in E, and which annihilates the holomorphic sections. The map θ is an operator of order zero taking sections to $(1, 0)$ forms with coefficients in E. Combine these to form an operator $D'' = \bar\partial + \theta$ which determines the structure of the Higgs bundle E. Conversely, such an operator defines a Higgs bundle if and only if it satisfies Leibniz's rule $D''(ae) = \bar\partial(a)\, e + a\, D''(e)$, and satisfies the integrability condition $(D'')^2 = 0$. Note that this condition contains the integrability of the holomorphic structure $\bar\partial^2 = 0$, the fact that θ is holomorphic, $\bar\partial(\theta) = 0$, and the condition $\theta \wedge \theta = 0$.

The fact that an integrable $\bar\partial$ operator is the same thing as a holomorphic structure is a consequence of the theorem of Newlander-Nirenberg. If X is a projective variety, then this may be taken one step further, by Serre's GAGA theorem. A holomorphic bundle E is in fact an algebraic vector bundle; in other words, it can be given by algebraic transition functions for a Zariski open cover. The holomorphic θ is then also algebraic. Thus if X is a projective variety, the notion of Higgs bundle is an algebraic geometric one.

CONSTRUCTIONS

In order to establish a relationship between the structures of flat bundles and Higgs bundles, we consider metrics on the underlying C^∞ vector bundles. A metric K on V or E is a positive definite hermitian inner product $(.,.)_K$ on the fibers, varying smoothly over the base. Given a frame $\{ v_i \}$ for the bundle, a metric is determined by the hermitian matrix $h_{ij} = (v_i, v_j)_K$. It is sometimes helpful to think of a metric as an isomorphism $K : V \to \overline{V}^*$ between the bundle and the dual of the complex conjugate bundle, with $K = \overline{K}^t$. The map is related to the metric by the formula $K(u)\,(\bar v) = (u, v)_K$.

If (E, D'') is a Higgs bundle with a metric K, we will define an operator D_K. Write $D'' = \bar\partial + \theta$, and first define $D'_K = \partial_K + \bar\theta_K$ as follows: ∂_K is the unique operator such that $\partial_K + \bar\partial$ preserves the metric, which means

$$(\bar\partial e, f) + (e, \partial_K f) = \bar\partial(e, f),$$

and $\bar\theta_K$ is defined by the condition that $(\theta e, f) = (e, \bar\theta_K f)$. In local coordinates, if $\theta = \Sigma \theta_i \, dz_i$ then $\bar\theta_K = \Sigma \bar\theta_i \, d\bar z_i$, where $\bar\theta_i$ is the adjoint of the matrix θ_i taken with respect to the metric. Now set $D_K = D'_K + D''$. It satisfies Leibniz's rule $D_K(ae) = a\, D_K(e) + d(a)\, e$, so it is a connection.

If (V, D) is a flat bundle with a metric K, then define an operator D''_K in the following way. Decompose $D = d' + d''$ into operators of type $(1, 0)$ and $(0, 1)$ respectively. Let δ' and δ'' be the unique operators of types $(1, 0)$ and $(0, 1)$ such that the connections $\delta' + d''$ and $d' + \delta''$ preserve the metric. Let $\partial = (d' + \delta')/2$ and $\bar\partial = (d'' + \delta'')/2$, and let $\theta = (d' - \delta')/2$ and $\bar\theta = (d'' - \delta'')/2$. In terms of a flat frame for the bundle, the operator θ may be described as follows. Suppose $\{ v_i \}$ is a frame of flat sections. Let $h_{ij} = (v_i, v_j)_K$ be the hermitian matrix for the metric. If $\theta = \Sigma \theta_k^{ij} \, v_i \otimes v_j^* \otimes dz_k$, then the coefficients θ_k^{ij} are determined by the equation

$$\sum_i \theta_k^{ij} \, h_{im} = \partial h_{jm} / \partial z_k .$$

Now set $D''_K = \bar\partial + \theta$. It satisfies the Leibniz rule $D''_K(av) = a\, D''_K(v) + \bar\partial(a)\, e$, so it is an operator of the type needed to define a Higgs bundle (but it might not satisfy the integrability condition). Note that $D'_K = \partial + \bar\theta$ is the operator in the previous paragraph associated to D''_K and the metric. It is also helpful to consider the operator

$$D^c_K = D''_K - D'_K = \delta'' - \delta'.$$

Note that $D''_K = (D + D^c_K)/2$.

FUNCTORIALITY, TENSOR PRODUCT AND DUAL

These constructions are functorial with respect to morphisms of the space X. Suppose $f: X \to Y$ is a morphism of complex manifolds. If E is a Higgs bundle on Y with metric K, then f^* E is a Higgs bundle on X, f^* K is a metric, and $f^* D_K = D_{f^*K}$. Similarly if V is a flat bundle on Y with metric K, then f^* V is a flat bundle on X with metric f^* K, and $f^* D''_K = D''_{f^*K}$.

We can define tensor products and duals of flat or Higgs bundles in an obvious way. The tensor product is just the usual tensor product of bundles on X. The structures of flat and Higgs bundles are defined as follows. If E and F are Higgs bundles, then define an operator D'' on $E \otimes F$ by

$$D''(e \otimes f) = D''(e) \otimes f + e \otimes D''(f)$$

and extend this to differential forms using Leibniz's rule. Note that as always, there will be an appropriate sign in the formulas. This operator D'' will give $E \otimes F$ a structure of Higgs bundle. In algebraic terms, $(E, \theta) \otimes (F, \varphi) = (E \otimes F, \theta \otimes 1 + 1 \otimes \varphi)$. Similarly, if V and W are flat bundles, then define an operator D by

$$D(v \otimes w) = D(v) \otimes w + v \otimes D(w);$$

it will be a flat connection.

If J, K are metrics relating Higgs operators D'' and connections D on bundles E and F, then the metric $J \otimes K$ on the tensor product $E \otimes F$, defined by

$$(e \otimes f, e' \otimes f')_{J \otimes K} = (e, e')_J \, (f, f')_K,$$

relates the corresponding Higgs operator and connection defined above. Write $D'' = \partial + \theta$, and $D''_C = \partial - \theta$. The extension of D''_C to the tensor product is given by the same formula. Then $D'_{J \otimes K}$ is characterized by the equation

$$(D'_{J \otimes K}(e \otimes f), e' \otimes f') + (e \otimes f, D''_C(e' \otimes f')) = \partial(e \otimes f, e' \otimes f').$$

However, the operator defined by $D'_{J, K}(e \otimes f) = D'_J(e) \otimes f + e \otimes D'_K(f)$ also satisfies the above equation, so these operators must be the same. On the other hand, it is clear that $D = D'' + D'_{J, K}$. Thus $D = D'' + D'_{J \otimes K}$ is the operator associated to D'' by the metric $J \otimes K$.

To define the dual, the underlying vector bundle is the dual bundle. If D'' is a Higgs operator on E, define D'' on the dual bundle E^* by the formula

$$D''(\lambda)\,(e) + \lambda(D''\, e) = \bar{\partial}(\lambda e)$$

for λ and e sections of E^* and E. In algebraic terms, the dual of (E, θ) is the Higgs bundle $(E^*, -\theta^t)$, so that the morphisms $\mathcal{O}_X \to E \otimes E^* \to \mathcal{O}_X$ are morphisms of Higgs bundles. Similarly, if (V, D) is a flat bundle, the dual connection is defined by

$$D(\lambda)\,(v) + \lambda(Dv) = d(\lambda v),$$

so the morphisms $\mathbf{C} \to V \otimes V^* \to \mathbf{C}$ are morphisms of flat bundles.

If K is a metric relating operators D'' and D on a bundle V, then as before the dual metric K^* defined by

$$(\lambda, \mu)_{K^*} = \lambda(\mu^{\dagger}_K)$$

relates the operators on the dual bundle V^* defined above. Here $\mu^{\dagger}_K$ is the element of V such that $\mu(e) = (e, \mu^{\dagger}_K)_K$.

FIRST ORDER KÄHLER IDENTITIES

Here is some motivation for the above definitions of the operators D_K and D''_K. The choice of a Kähler metric ω results in an operator Λ in the exterior algebra of forms on X. It is the adjoint of the operation of wedging with the (1, 1) form ω. See ([25] pp. 111 ff) for the description of Λ in local normal coordinates. Recall [53] that the operators ∂ and $\bar{\partial}$ on scalar forms satisfy the Kähler identities

$$(\partial)^* = \sqrt{-1}\,[\Lambda, \bar{\partial}], \qquad (\bar{\partial})^* = -\sqrt{-1}\,[\Lambda, \partial].$$

In terms of the exterior derivative and the operator $d^c = \bar{\partial} - \partial$, these become

$$(d^c)^* = -\sqrt{-1}\,[\Lambda, d], \qquad (d)^* = \sqrt{-1}\,[\Lambda, d^c].$$

If E is a Higgs bundle with metric K, the above defined operator D'_K can be characterized as the unique operator satisfying

$$(D'_K)^* = \sqrt{-1}\,[\Lambda, D''], \qquad (D'')^* = -\sqrt{-1}\,[\Lambda, D'_K].$$

The formal adjoint $(\ )^*$ uses the metric K. Similarly, if V is a flat bundle with metric, then the operator D^c_K can be characterized as the unique operator satisfying

$$(D^c_K)^* = -\sqrt{-1}\,[\Lambda, D], \qquad (D)^* = \sqrt{-1}\,[\Lambda, D^c_K].$$

These formulas can be proved by following the technique of [25].

CURVATURE

If E is a Higgs bundle with metric then the operator D_K is a connection. If V is a flat bundle with metric, then the operator D''_K is an operator of the type needed to define a Higgs bundle. But the operators do not necessarily satisfy the integrability conditions, $(D_K)^2 = 0$ or $(D''_K)^2 = 0$ respectively. The failure to satisfy the integrability is measured by a tensor.

If E is a Higgs bundle with metric K, the *curvature* is the End(E)-valued two form

$$F_K = (D_K)^2 = (D'_K\, D'' + D''\, D'_K).$$

It satisfies the Bianchi identities $D''\, F_K = D'_K\, F_K = 0$, because $(D'')^2 = 0$ and $(D'_K)^2 = 0$.

If V is a flat bundle with metric K, the *pseudocurvature* is the End(V)-valued two form

$$G_K = (D''_K)^2 = (DD^c_K + D^c_K\, D)/4.$$

It satisfies the Bianchi identities $DG_K = D^c_K\, G_K = 0$, again because $D^2 = 0$ and $(D^c_K)^2 = 0$. Note that $D^2 = 0$ implies that $(d')^2 = 0$, $(d'')^2 = 0$, and $d'\, d'' + d''\, d' = 0$. These in turn imply that $(\delta')^2 = 0$, $(\delta'')^2 = 0$, and $\delta'\, \delta'' + \delta''\, \delta' = 0$, which give $(D^c_K)^2 = 0$. Similarly we obtain the formula

$$G_K = d'\, \delta'' + \delta''\, d' - d''\, \delta' - \delta'\, d'' + d''\, \delta'' + \delta''\, d'' - d'\, \delta' - \delta'\, d'.$$

A metric K such that $F_K = 0$ on a Higgs bundle gives rise to a structure of flat bundle. A metric K such that $G_K = 0$ on a flat bundle, gives a structure of Higgs bundle. These two constructions are inverses of one another.

The equations $F_K = 0$ and $G_K = 0$ are over-determined, but there are natural intermediate equations obtained by considering only the component of the curvature or pseudocurvature pointing in the direction of the Kähler form. If E is a Higgs bundle, a metric K is called *Hermitian-Yang-Mills* if

$$\Lambda F_K = \lambda . Id$$

for a scalar constant λ which depends on the degree divided by the rank of E.

If V is a flat bundle, a metric K is called *harmonic* if

$$\Lambda G_K = 0.$$

A metric on V can be thought of as a multivalued map

$$\Phi_K : X \to Gl(n)/U(n)$$

which is equivariant with respect to the monodromy representation of the flat bundle V. The condition $\Lambda G_K = 0$ is the same as the condition obtained by Corlette [5] for the map to be harmonic. To see that this condition is equivalent to finding a critical point for the energy, proceed as follows. The derivative of the classifying map Φ_K is $\theta + \bar{\theta}$, so the energy is

$$\| \theta + \bar{\theta} \|_{L^2}^2 = \int_X (\theta, \theta) + (\bar{\theta}, \bar{\theta}).$$

Changing K infinitesimally by multiplying by $1 + h$ changes δ' by $\delta'(h)$ and δ'' by $\delta''(h)$ (to first order in h). Thus θ changes by $\delta'(h)/2$ and $\bar{\theta}$ changes by $\delta''(h)/2$. The energy changes by

$$\int_X (\theta, \delta'(h)) + (\bar{\theta}, \delta''(h)),$$

which is equal to

$$\int_X ((\delta')^* \theta + (\delta'')^* \bar{\theta}, h).$$

The Euler-Lagrange equation is $(\delta')^* \theta + (\delta'')^* \bar{\theta} = 0$. The operators δ' and δ'' are components of metric connections, so they satisfy Kähler identities. In particular, the equation becomes

$$\sqrt{-1}\Lambda(d''(\theta) - d'(\bar{\theta})) = 0.$$

But $d''(\theta) = -(d''\delta' + \delta' d'')/2$ and $d'(\bar{\theta}) = -(d'\delta'' + \delta'' d')/2$ so the equation becomes

$$\sqrt{-1}\Lambda(d'\delta'' + \delta'' d' - d''\delta' - \delta' d'') = 0.$$

This is equivalent to $\Lambda G_K = 0$ in view of the formula for G_K given above.

We are really interested in solving $F_K = 0$ or $G_K = 0$. Suppose a Higgs bundle E has a Hermitian-Yang-Mills metric K. Then it turns out that $F_K = 0$ if and only if the Chern classes of the bundle E vanish, and we will see that one only needs to consider the first and second Chern numbers. This additional vanishing of the curvature is due to the Riemann bilinear relations, which assert that

$$\int_X \mathrm{Tr}(F_K \wedge F_K) \wedge \omega^{n-2} = C_1 \| F_K \|_{L^2}^2 - C_2 \| \Lambda F_K \|_{L^2}^2$$

([47] Prop. 3.4). The quantity on the left is equal to the intersection of the second Chern character of E with a power of the Kähler class. If the degree

$$c_1(E) . [\omega]^{n-1} = \int_X \mathrm{Tr}(F_K) \wedge \omega^{n-1}$$

is equal to zero, then the Hermitian-Yang-Mills condition simply says $\Lambda F_K = 0$. Thus if $c_1(E).[\omega]^{n-1} = 0$ and $\mathrm{ch}_2(E).[\omega]^{n-2} = 0$, and if the metric is Hermitian-Yang-Mills, then $F_K = 0$. This argument originates, in the case of unitary connections on vector bundles, with Lübke [34].

One can define pseudo Chern classes for a flat bundle by integrating invariant polynomials in the pseudocurvature G_K. Similar Riemann bilinear relations hold, and one can show by an analogous argument that if $\int_X \mathrm{Tr}(G_K) \wedge \omega^{n-1} = 0$ and $\int_X \mathrm{Tr}(G_K \wedge G_K) \wedge \omega^{n-2} = 0$, and if K is a harmonic metric, then $G_K = 0$. However, there is an important difference in this case, which is that the pseudo Chern classes vanish automatically! This was pointed out to me by P. Deligne, the key being the stronger nature of the Leibniz rule for D. The history of this main lemma begins earlier, with Siu's Bochner formula for harmonic maps [49]. The conclusion $G_K = 0$ is essentially contained in Siu's argument, and appears in more explicit forms in the article of Sampson [44], and in Corlette's paper ([5] § 5) where it is used in the context of equivariant harmonic maps. The correspondence we are developing here is a further generalization of Siu's rigidity theorem, and many of our applications will be similar in spirit to those of Siu, Sampson, Corlette, and others.

Lemma **1.1** *(Siu, Sampson, Corlette, Deligne). — If* K *is a harmonic metric, then* $G_K = 0$. *Hence* V *comes from a Higgs bundle* (V, D''_K).

Proof. — Here is a direct proof not involving the pseudo Chern classes. For any metric K,

$$D^c_K = d'' - d' + 2(\theta_K - \bar{\theta}_K).$$

The conditions $(d')^2 = 0$ and $(d'')^2 = 0$ and $d'd'' + d''d' = 0$ imply that

$$D.(d'' - d') + (d'' - d').D = 0,$$

so

$$G_K = DD^c_K + D^c_K D = 2D(\theta_K - \bar{\theta}_K).$$

If K is harmonic then $D^* G_K = 0$, since $D^* = \sqrt{-1}\,[\Lambda, D^c_K]$, and $\Lambda G_K = 0$ by assumption and $D^c_K G_K = 0$ by the Bianchi identities. Now

$$\int_X |G_K|^2 = 2\int_X (D(\theta_K - \bar{\theta}_K), G_K) = 2\int_X (\theta_K - \bar{\theta}_K, D^* G_K) = 0.$$

This proves the lemma.

Here is a more conceptual argument for why the pseudo Chern classes must vanish. If we perturb the operator D'' by an operator whose symbol is ∂, then we get an operator related to a connection:

$$\nabla_\varepsilon = D''_K + \varepsilon D'_K.$$

This is not a connection, but if we resolve into parts of types $(1, 0)$ and $(0, 1)$, $\nabla_\varepsilon = \nabla'_\varepsilon + \nabla''_\varepsilon$, then we can make a connection $B_\varepsilon = \varepsilon^{-1}\nabla'_\varepsilon + \nabla''_\varepsilon$. Since B_ε

is a connection on the bundle V, its curvature B_ε^2 calculates the Chern classes of V. However, since V has a flat connection D, the Chern classes vanish. Thus for example $\int_X \mathrm{Tr}(B_\varepsilon^2 \wedge B_\varepsilon^2)\, \omega^{n-2} = 0$, and similarly for the other numbers. Decomposing according to type of forms on X, and taking into account the fact that ω has type (1, 1), we get

$$0 = \int_X \mathrm{Tr}(B_\varepsilon^2 \wedge B_\varepsilon^2)\, \omega^{n-2} = \varepsilon^{-2} \int_X \mathrm{Tr}(\nabla_\varepsilon^2 \wedge \nabla_\varepsilon^2)\, \omega^{n-2}.$$

On the other hand, ∇_ε approaches D_K'' as ε approaches zero, so ∇_ε^2 approaches G_K. This proves that

$$\int_X \mathrm{Tr}(G_K \wedge G_K)\, \omega^{n-2} = 0.$$

The same argument works for the other numbers. One could complete a second proof of the lemma from here by making explicit the argument mentioned above involving Riemann bilinear relations. $\square$

With this lemma in mind, we will use the terminology *harmonic metric* to denote a metric K on a Higgs bundle such that $F_K = 0$. Note that a Higgs bundle with harmonic metric is the same thing as a flat bundle with harmonic metric. We will use the term *harmonic bundle* to denote a C^∞ bundle provided with structures of flat bundle and Higgs bundle which are related by a harmonic metric. A choice of such metric will not be part of the data.

The notion of harmonic bundle is functorial and compatible with tensor products and duals. If E is a Higgs or flat bundle on X with metric K, and if $f: Y \to X$ is a morphism, then the curvature of the pullback metric $f^* K$ on the pullback bundle $f^* E$ is equal to the pullback of the curvature of E (this holds for either the curvature or pseudo curvature). If K is a harmonic metric for E over X, this implies that the curvature of $f^* K$ vanishes on Y, so $f^* K$ is a harmonic metric on $f^* E$ over Y. Suppose K and J are harmonic metrics on harmonic bundles E and F. Then the metric $K \otimes J$ on $E \otimes F$ relates the operators D and D'' on the tensor product. If the curvatures of E and F are zero, then the curvatures of $E \otimes F$ will be zero, because for example

$$(D \otimes 1 + 1 \otimes D)^2 = D^2 \otimes 1 + 1 \otimes D^2 + D \otimes D - D \otimes D.$$

The sign comes from passing the operators D of degree one across each other. Thus if E and F are harmonic bundles, then $E \otimes F$ is harmonic. Similarly, the curvature of the dual bundle is minus the transpose of the curvature of the bundle, so the dual of a harmonic bundle is a harmonic bundle.

NON-ABELIAN HODGE THEOREM

We say that a Higgs bundle E is *stable* if, for every subsheaf $M \subset E$ preserved by θ and with $0 < \mathrm{rk}\, M < \mathrm{rk}\, E$,

$$\frac{\deg M}{\mathrm{rk}\, M} < \frac{\deg E}{\mathrm{rk}\, E}.$$

The degree $(\mathrm{ch}_1(\mathrm{E}).[\omega]^{n-1})$ divided by the rank of E is called the *slope* of E. The corresponding notion for flat bundles is that V is irreducible. Say that a Higgs bundle E is *polystable* if it is a direct sum of stable Higgs bundles of the same slope. The corresponding notion for flat bundles is that V is semisimple, or a direct sum of irreducible local systems. Say that a Higgs bundle E is *semistable* if the inequality above holds with $\leqslant$ instead of $<$. The corresponding notion for flat bundles is vacuous (always satisfied). Now we can state the theorem of non-linear analysis which we will use.

Theorem **1.** — *(1) A flat bundle* V *has a harmonic metric if and only if it is semisimple. (2) A Higgs bundle* E *has a Hermitian-Yang-Mills metric if and only if it is polystable. Such a metric is harmonic if and only if* $\mathrm{ch}_1(\mathrm{E}).[\omega]^{\dim \mathrm{X}-1} = 0$ *and* $\mathrm{ch}_2(\mathrm{E}).[\omega]^{\dim \mathrm{X}-2} = 0$. □

The proof of this theorem is already contained in the literature. We can use this occasion to review the history of the theorem and the preceding definitions and constructions. The notion of harmonic map which enters into statement (1) originated with Eells and Sampson [19]. It was generalized to the notion of harmonic metric (or equivariant harmonic map) by Corlette, to whom the proof of statement (1) is due [5]. Corlette's work was paralleled independently by Donaldson [18] in a note after Hitchin's paper for the case of rank two bundles, and somewhat earlier for the case of holomorphic disc bundles (i.e. SU(1, 1)) by Diederich and Ohsawa [14]. The significance of statement (1) becomes clearer with the additional vanishing of the full pseudocurvature provided by Lemma 1.1, which results in the construction of Higgs bundles.

Statement (2) is the result of a succession of generalizations beginning with the theorem of Narasimhan and Seshadri [39], and continuing with Donaldson [16] [17], Uhlenbeck and Yau [52], Deligne and Beilinson (unpublished), Hitchin [30], and [47]. Hitchin's paper is an important landmark. Although he restricted his attention to the case of bundles of rank two on a Riemann surface, Hitchin obtained all of the conceptual features of the correspondence we are considering, including many of the properties to be described below.

The terminology " Higgs bundle " is a slight modification of Hitchin's calling θ the " Higgs field " in reference to the physics paper of P. W. Higgs [29]. My own motivation (and that of Deligne and Beilinson) for thinking about an additional tensor such as θ came from the Kodaira-Spencer map in Griffiths' theory of variations of Hodge structure. It is very interesting that the same object independently became manifest in elementary particle physics and in the complex analytic geometry of motives—this surely deserves further thought.

If V is a flat bundle, define $\mathrm{H}^0_{\mathrm{DR}}(\mathrm{X}; \mathrm{V})$ to be the space of sections v such that $\mathrm{D}v = 0$. If E is a Higgs bundle, define $\mathrm{H}^0_{\mathrm{Dol}}(\mathrm{X}; \mathrm{E})$ to be the space of holomorphic sections e such that $\theta e = 0$, or equivalently the space of C^∞ sections such that $\mathrm{D}'' e = 0$.

Lemma **1.2.** — *Suppose* E *is a harmonic bundle with a harmonic metric. If* e *is a section, then* $\mathrm{D}'' e = 0$ *if and only if* $\mathrm{D}e = 0$. *Thus we get a natural isomorphism* $\mathrm{H}^0_{\mathrm{DR}}(\mathrm{E}) \cong \mathrm{H}^0_{\mathrm{Dol}}(\mathrm{E})$.

Proof. — Suppose $D'' e = 0$. Then

$$(D')^* D' e = \sqrt{-1}\Lambda D'' D' e = -\sqrt{-1}\Lambda D' D'' e = 0$$

since $D' D'' + D'' D' = 0$. Therefore

$$\int_X (D' e, D' e) = \int_X ((D')^* D' e, e) = 0$$

so $De = D'' e + D' e = 0$. Similarly if $De = 0$ then $(D^c)^* D^c e = 0$ so $D^c e = 0$ so $D'' e = 0$. □

We will generalize this lemma to forms in the next section.

Corollary **1.3.** — *There is an equivalence of categories between the category of semisimple flat bundles on* X *and the category of polystable Higgs bundles with* $\mathrm{ch}_1(E).[\omega]^{\dim X - 1} = 0$ *and* $\mathrm{ch}_2(E).[\omega]^{\dim X - 2} = 0$, *both being equivalent to the category of harmonic bundles.*

Proof. — Use Theorem 1 and the above isomorphism between $H^0_{\mathrm{Dol}}(E^* \otimes F)$ and $H^0_{\mathrm{DR}}(E^* \otimes F)$ to obtain an equivalence between the morphisms of flat bundles and the morphisms of Higgs bundles. □

Theorem 1 can be interpreted as a nonabelian Hodge theorem. The first nonabelian cohomology set of $\pi_1(X)$ with coefficients in the constant group $\mathrm{Gl}(n, \mathbf{C})$ is just the set of representations $\pi_1(X) \to \mathrm{Gl}(n, \mathbf{C})$, up to conjugation. We can modify this set a little bit: define a *semisimplified* cohomology class to be a representation up to equivalence of its irreducible subquotients. The set of semisimplified cohomology classes is the same as the set of semisimple representations. The first part of the theorem says that any semisimple nonabelian cohomology class has a harmonic representative. Recall that for abelian cohomology of X, there is a parallel notion of *Dolbeault cohomology*, which is equal to cohomology of coherent sheaves [15]. The first abelian Dolbeault cohomology of X is $H^1(X, \mathcal{O}_X) \oplus H^0(X, \Omega^1_X)$. In our interpretation, the *nonabelian* Dolbeault cohomology classes are the Higgs bundles (E, θ). Note that to specify E one must give a cocycle in $H^1(X, \mathrm{Gl}(n, \mathcal{O}_X))$, while θ is an element of $H^0(X, \Omega^1_X \otimes \mathrm{End}(E))$, suggesting an analogy with the abelian Dolbeault cohomology. The second part of the main theorem constructs harmonic representatives for Dolbeault classes.

A problem suggested by this interpretation is to describe the Dolbeault cohomology in the case when $\pi_1(X)$ acts nontrivially on $\mathrm{Gl}(n)$.

MODULI SPACES

In this subsection, we will state without proof some theorems about moduli of Higgs bundles. The proofs will appear elsewhere, and will in part depend on results from this paper. Furthermore, these theorems have been proved only for the case when X is a smooth projective variety (although one expects that similar statements hold for compact Kähler manifolds). For these reasons, if a result in the following sections depends on these theorems, that will be explicitly mentioned.

Proposition **1.4.** — *Suppose* X *is a smooth projective variety. There is a quasiprojective variety* $\mathbf{M}_{\mathrm{Dol}}$ *whose points parametrize direct sums of stable Higgs bundles on* X *with vanishing Chern classes. There is a map from* $\mathbf{M}_{\mathrm{Dol}}$ *to the space of polynomials with coefficients in symmetric powers of the cotangent bundle, whose effect on points is to take a Higgs bundle* (E, θ) *to the characteristic polynomial of* θ. *This map is proper.* □

This has been proved independently in the case when X is an algebraic curve, by N. Nitsure [40].

Recall that there is a similar moduli space $\mathbf{M}_{\mathrm{B}}$ for representations of the fundamental group. Let $\mathbf{R}_{\mathrm{B}}$ denote the affine variety of homomorphisms from $\pi_1(\mathrm{X})$ into $\mathrm{Gl}(n, \mathbf{C})$ obtained by looking at generators and relations. Then $\mathbf{M}_{\mathrm{B}}$ is the affine categorical quotient of $\mathbf{R}_{\mathrm{B}}$ by the action of $\mathrm{Gl}(n, \mathbf{C})$. Again, the points of $\mathbf{M}_{\mathrm{B}}$ parametrize semisimple representations. The correspondence of Theorem 1 yields an isomorphism of sets between $\mathbf{M}_{\mathrm{B}}$ and $\mathbf{M}_{\mathrm{Dol}}$.

Proposition **1.5.** — *This map of sets is a homeomorphism of topological spaces* $\mathbf{M}_{\mathrm{B}} \cong \mathbf{M}_{\mathrm{Dol}}$. □

Example. — Suppose X is a curve of genus g, and suppose we wish to parametrize representations of rank one. The abelianization of the fundamental group has $2g$ generators, so to give a character $\pi_1(\mathrm{X}) \to \mathbf{C}^*$ is the same as to give $2g$ nonzero complex numbers. Thus

$$\mathbf{M}_{\mathrm{B}} = (\mathbf{C}^*)^{2g}.$$

On the other hand, a Higgs bundle of degree zero and rank one is just a pair consisting of a line bundle L and a scalar valued one-form θ. Thus

$$\mathbf{M}_{\mathrm{Dol}} = \mathrm{Jac}(\mathrm{X}) \times \mathrm{H}^0(\Omega^1_{\mathrm{X}}).$$

There is a decomposition $\mathbf{C}^* = \mathrm{S}^1 \times \mathbf{R}^*_+$. The homeomorphism between $\mathbf{M}_{\mathrm{B}}$ and $\mathbf{M}_{\mathrm{Dol}}$ sends $(\mathrm{S}^1)^{2g}$ isomorphically to $\mathrm{Jac}(\mathrm{X})$, because the Higgs bundles with $\theta = 0$ are exactly the unitary representations. The map is just the map sending a unitary local system to the line bundle with the same transition functions. The homeomorphism sends $\mathrm{H}^0(\Omega^1_{\mathrm{X}})$ isomorphically to $(\mathbf{R}^*_+)^{2g}$. The map is

$$\theta \mapsto \left(\ldots, \exp\left(-\int_{\gamma_i} \theta + \bar{\theta}\right), \ldots\right)$$

where γ_i form the basis for the first homology. The homeomorphism between $\mathbf{M}_{\mathrm{B}}$ and $\mathbf{M}_{\mathrm{Dol}}$ preserves tensor product, so it preserves the natural group structures of these varieties. This determines the homeomorphism in the rank one case.

2. Further properties

In this section we will discuss further properties of the correspondence between Higgs bundles and local systems.

HIGHER ORDER KÄHLER IDENTITIES

In [7] Deligne extended the Kähler identities to coefficients in variations of Hogde structure, using the operators D' and D''. This is part of the motivation for introducing these operators above. In this section we will extend the Kähler identities to harmonic bundles.

Suppose E is a harmonic bundle with a harmonic metric K relating the flat and Higgs structures. The operators D, D'', etc., can be extended to operators in the exterior algebra of C^∞ differential forms with coefficients in E. As noted above, the first order Kähler identities hold for the operators D', D'', D and D^c. Furthermore the squares of these operators are zero in the exterior algebra, so for example

$$D' D'' + D'' D' = 0.$$

Define the laplacians

$$\Delta \;= DD^* + D^* D$$
$$\Delta'' = D''(D'')^* + (D'')^* D''$$

and similarly Δ' and Δ^c. By the same arguments as in [25] [53], the previous identities imply

$$\Delta = \Delta^c = 2\Delta'' = 2\Delta'.$$

Therefore the spaces of harmonic forms with coefficients in E are all the same. Denote this space by $\mathscr{H}(E)$. A form α is harmonic if and only if $D\alpha = 0$ and $D^* \alpha = 0$, or equivalently, if and only if $D'' \alpha = 0$ and $(D'')^* \alpha = 0$. Hodge theory gives the following orthogonal decompositions of the space of E-valued forms with respect to the L^2 inner product:

$$A^\bullet(E) \;= \mathscr{H}(E) \oplus \mathrm{im}(D) \oplus \mathrm{im}(D^*)$$
$$A^\bullet(E) \;= \mathscr{H}(E) \oplus \mathrm{im}(D'') \oplus \mathrm{im}((D'')^*)$$

and similarly for D' and D^c.

Lemma **2.1** *(Principle of two types).* — *If* E *is a harmonic bundle, then*

$$\ker(D') \cap \ker(D'') \cap (\mathrm{im}(D'') + \mathrm{im}(D')) = \mathrm{im}(D' D'').$$

Proof [12] [53]. — The name notwithstanding, this does not require a decomposition into Hodge type. Suppose $\eta = D' \alpha + D'' \beta$ and $D' \eta = D'' \eta = 0$. Then $D' D'' \beta = 0$. Now apply the above decomposition with respect to D':

$$\beta = h + D' \gamma_0 + (D')^* \gamma$$

with h harmonic, and further we may assume $D'\gamma = 0$. Note that $D''h = 0$ and $D''D'\gamma_0 \in \mathrm{im}(D'D'')$ as desired. The condition $D'D''\beta = 0$ gives

$$D'D''(D')^*\gamma = 0$$

and $D'\gamma = 0$ so

$$D''\Delta'\gamma = D''D'(D')^*\gamma = 0.$$

The identity $\Delta' = \Delta''$ implies that $\Delta''D''\gamma = 0$. But the above orthogonal decomposition for D'' now implies that $D''\gamma = 0$. By the Kähler identity,

$$D''(D')^*\gamma = \sqrt{-1}\,(D''\Lambda D''\gamma + D''D''\Lambda\gamma) = 0.$$

Therefore $D''\beta \in \mathrm{im}(D'D'')$. Similarly $D'\alpha \in \mathrm{im}(D'D'')$. $\square$

Remark. — This lemma implies the seemingly stronger statement that the natural map

$$\frac{\ker(D') \cap \ker(D'')}{\mathrm{im}(D'D'')} \to \frac{\ker(D'D'')}{\mathrm{im}(D') + \mathrm{im}(D'')}$$

is bijective (pointed out to me by P. Deligne). The injectivity is simply the statement of the lemma. For the surjectivity, suppose $D'D''u = 0$. Then

$$D'u \in \ker(D') \cap \ker(D'') \cap (\mathrm{im}(D') + \mathrm{im}(D'')),$$

so by the lemma we can write $D'u = D'D''w$. Similarly, we can write $D''u = D''D'v$. Set $\beta = u - D''w - D'v$. Then $D'\beta = D''\beta = 0$, and β is equal to u modulo $\mathrm{im}(D') + \mathrm{im}(D'')$.

COHOMOLOGY

As an illustration, first consider the case of H^0. Suppose E is a harmonic bundle with a harmonic metric. Then a section e is harmonic if and only if $D''e = 0$, or equivalently if and only if $De = 0$. This again gives Lemma 2.2, the isomorphism $H^0_{DR}(E) \cong H^0_{Dol}(E)$.

There are natural cohomology functors for flat bundles and Higgs bundles, which extend the functors H^0_{DR} and H^0_{Dol}. Suppose V is a flat bundle. Let V^D be the locally constant sheaf of flat sections of V. This sheaf is resolved by the de Rham complex of sheaves of C^∞ differential forms with coefficients in V:

$$V^D \to (\mathscr{A}^0(V) \xrightarrow{D} \mathscr{A}^1(V) \xrightarrow{D} \mathscr{A}^2(V) \xrightarrow{D} \dots)$$

is a quasi isomorphism of complexes of sheaves. The sheaves of C^∞ forms are fine, so the cohomology $H^i(X; V^D)$ is naturally isomorphic to the cohomology of the complex of global sections

$$(A^\bullet(V), D) = A^0(V) \xrightarrow{D} A^1(V) \xrightarrow{D} A^2(V) \xrightarrow{D} \dots$$

Call this the de Rham cohomology $H^i_{DR}(X; V)$. Note that if V is the trivial flat bundle, then the de Rham cohomology of V is just the usual de Rham cohomology of the differential manifold X.

Suppose E is a Higgs bundle. Then we get a complex of locally free sheaves, the *holomorphic Dolbeault complex*:

$$E \xrightarrow{\theta\wedge} E \otimes \Omega^1_X \xrightarrow{\theta\wedge} E \otimes \Omega^2_X \xrightarrow{\theta\wedge} \dots$$

The condition that $\theta \wedge \theta = 0$ insures that this is a complex. Define the *Dolbeault cohomology* with coefficients in E to be the hypercohomology

$$H^i_{Dol} = \mathbf{H}^i(E \xrightarrow{\theta\wedge} E \otimes \Omega^1_X \xrightarrow{\theta\wedge} \dots).$$

The complex of sheaves of C^∞ sections of E

$$\mathscr{A}^0(E) \xrightarrow{D''} \mathscr{A}^1(E) \xrightarrow{D''} \mathscr{A}^2(E) \xrightarrow{D''} \dots$$

gives a fine resolution of the holomorphic Dolbeault complex, so $H^i_{Dol}(E)$ is naturally isomorphic to the cohomology of the complex of global sections $(A^\bullet(E, D''))$.

Remark. — Versions of the holomorphic Dolbeault complex have appeared in the work of Green and Lazarsfeld [22] [23].

Lemma **2.2** *(Formality). — Suppose* E *is a flat and Higgs bundle with harmonic metric. There are natural quasi-isomorphisms*

$$(\ker(D'), D'') \to (A^\bullet(E), D)$$
$$(\ker(D'), D'') \to (A^\bullet(E), D'')$$
$$(\ker(D'), D'') \to (H^\bullet_{DR}(E), 0)$$
$$(\ker(D'), D'') \to (H^\bullet_{Dol}(E), 0).$$

In particular, there are natural isomorphisms $H^i_{DR}(X; E) \cong H^i_{Dol}(X; E)$.

Proof. — This is due to [12]. See also ([21] section 7). In all cases on the right the cohomology is represented by the space of harmonic forms. We have to show that the same is true for $(\ker(D'), D'')$. Suppose $D'\alpha = D''\alpha = 0$. If $\alpha = D''\beta$ then the harmonic projection of α is zero. Conversely if the harmonic projection of α is zero then $\alpha = D''\beta$. By the principle of two types we may assume that $\beta = D'\gamma$ and, in particular, $D'\beta = 0$. Thus α is a coboundary. □

Corollary **2.3.** — *Let* E *be a harmonic bundle with a harmonic metric and let* T *denote the tensor algebra of* E. *The differential graded algebras* $(A^\bullet(T), D)$ *and* $(A^\bullet(T), D'')$ *are formal, in other words quasi-isomorphic to the differential graded algebra* $H^\bullet_{DR}(T) = H^\bullet_{Dol}(T)$ *with zero differential.*

Proof. — T is a direct sum of harmonic bundles with harmonic metrics so, by the above, the maps $(\ker(D'), D'') \to (A^\bullet(T), D)$, etc. are quasi-isomorphisms. These are morphisms of differential graded algebras. □

Corollary **2.4** *(Goldman and Millson). — If* V *is a semisimple flat bundle then the deformation space of representations of* $\pi_1(X)$ *is quadratic at* V.

Proof. — Goldman and Millson [21] prove that the deformation space is quadratic if the differential graded Lie algebra of forms with coefficients in End(V) is formal. Theorem 1 and the above corollary imply that this is the case if V is semisimple. □

The Serre duality theorem and the Riemann-Roch theorem apply also to the Dolbeault cohomology of Higgs bundles. The dual of a Higgs bundle $E = (E, \theta)$ is $E^* = (E^*, -\theta^*)$ where θ^* is the image of θ under the natural isomorphism $\mathrm{End}(E) \cong \mathrm{End}(E^*)$. The natural pairing

$$E \otimes E^* \to \mathcal{O}_X$$

is a morphism of Higgs bundles. Let $\chi_{\mathrm{Dol}}(X, E) = \Sigma_{i=0}^{2d}(-1)^i \dim H^i_{\mathrm{Dol}}(X, E)$.

Lemma **2.5** *(Duality and Riemann-Roch). — If* X *has dimension d then* $H^{2d}_{\mathrm{Dol}}(X, \mathcal{O}_X) = \mathbf{C}$ *and for any Higgs bundle* E, *the induced map*

$$H^i_{\mathrm{Dol}}(X, E) \otimes H^{2d-i}(X, E^*) \to \mathbf{C}$$

is a perfect pairing. If E *is a Higgs bundle on* X *of rank n, then* $\chi_{\mathrm{Dol}}(X, E) = n\chi_{\mathrm{Dol}}(X, \mathcal{O}_X)$.

Proof. — The Serre duality theorem works for complexes of sheaves. The tensor product of the dualizing sheaf Ω^d_x with the dual of the Dolbeault complex for E is the Dolbeault complex for E^*. (Note that if one changes the signs of the differentials, the complexes are still isomorphic, by operating on the ith piece by $(-1)^i$.) This proves the duality statement.

The spectral sequence for hypercohomology converges to the Dolbeault cohomology of E, but the Euler characteristic of the E_2 term is the same as that of the limit, so

$$\chi_{\mathrm{Dol}}(X, E) = \Sigma(-1)^{i+j} H^{i,j}_{\mathrm{coh}}(E \otimes \Omega^j_X) = \Sigma(-1)^j \chi_{\mathrm{coh}}(E \otimes \Omega^j_X).$$

Here the subscript " coh " denotes reference to the usual cohomology of coherent sheaves. Let ch(F) denote the exponential Chern character. Note that $\mathrm{ch}(E \otimes \Omega^j_X) = \mathrm{ch}(E)\,\mathrm{ch}(\Omega^j_X)$. Let $\mathrm{td}(T_X)$ denote the Todd class of the tangent bundle. The Grothendieck-Hirzebruch-Riemann-Roch formula says

$$\chi_{\mathrm{coh}}(E \otimes \Omega^j_X) = \deg(\mathrm{ch}(E)\,\mathrm{ch}(\Omega^j_X)\,\mathrm{td}(T_X)).$$

Thus $\chi_{\mathrm{Dol}}(X, E) = \deg(\mathrm{ch}(E)\,(\Sigma(-1)^i\,\mathrm{ch}(\Omega^i_X)\,\mathrm{td}(T_X)))$. On the other hand, one checks that

$$\Sigma(-1)^i\,\mathrm{ch}(\Omega^i_X)\,\mathrm{td}(T_X) = c_d(T_X) = e(X)$$

is the Euler class of X; it occurs only in degree $2d$. Thus

$$\chi_{\mathrm{Dol}}(X, E) = \deg(\mathrm{ch}_0(E)\,e(X)) = \mathrm{rk}(E)\deg(e(X)) = n\chi_{\mathrm{Dol}}(X, \mathcal{O}_X).$$

Surprisingly, the Riemann-Roch formula holds without any reference to (or restriction on) the Chern classes of E. □

We next show that the Lefschetz decomposition of cohomology into primitive pieces works for de Rham cohomology with coefficients in any semisimple local system, or for Dolbeault cohomology with coefficients in the corresponding Higgs bundle. Let L denote the operation of wedging with the Kähler form, $L(\alpha) = \alpha \wedge \omega$. Note that ω represents a de Rham cohomology class in $H^2_{DR}(X, C)$ as well as a Dolbeault cohomology class in $H^1(X, \Omega^1_X)$. Thus the operation L is a cup product.

Lemma **2.6** *(Lefschetz decomposition). — Suppose* E *is a harmonic bundle, on a manifold* X *of dimension d. Let* $H^i(E)$ *denote either de Rham cohomology of the flat bundle, or Dolbeault cohomology of the Higgs bundle. Then for any* $0 \leqslant i \leqslant d$,

$$L^{d-i} : H^i(E) \to H^{2d-i}(E)$$

is an isomorphism. Let $P^i(E)$ *denote the kernel of* L^{d-i+1} *acting on* $H^i(E)$, *the space of* primitive *cohomology classes. Then*

$$H^k(E) = \bigoplus L^i\, P^{k-2i}(E).$$

Proof. — The Kähler identities imply that if α is a harmonic k-form with coefficients in E, then $L(\alpha)$ and $\Lambda(\alpha)$ are also harmonic forms [25]. Consequently there is an action of $\mathbf{sl}(2)$ on the space of harmonic forms. The Lefschetz decomposition follows as in the scalar case.

From combining this lemma with the previous duality statement, the pairing

$$P^i(E) \otimes P^i(E^*) \to \mathbf{C}$$

given by

$$\alpha \otimes \beta \mapsto \int_X \alpha \wedge \beta \wedge \omega^{d-i}$$

is perfect. If ω represents the cohomology class of a hyperplane section, we obtain a Lefschetz-type theorem for restriction of cohomology classes. Namely, if Y is a codimension k complete intersection of hyperplane sections, then the restriction map

$$H^i(X, E) \to H^i(Y, E|_Y)$$

is an isomorphism for $i < d - k$ and injective for $i = d - k$. □

A COMPACTNESS PROPERTY

We will describe a compactness property for harmonic bundles. One of the main steps from Hitchin's paper [30], it is a generalization of Uhlenbeck's weak compactness theorem. In order to do this, we generalize the distance dercdasing property of variations of Hodge structure to the case of harmonic bundles where the eigenvalues of θ are bounded.

Lemma **2.7**. — *Given* C_1, *there is a constant* C_2 *such that if* V *is a harmonic bundle with harmonic metric such that all of the eigenvalues of* θ *have norm less than* C_1, *then* $|\theta| \leqslant C_2$.

Proof. — We proceed along the same lines as the proof of the distance decreasing property for variations of Hodge structure [26]. For now, assume that X is the unit disk, and try to find the estimate on a smaller disk. The curvature of the unitary bundle $(\mathrm{End}\, V) \otimes \Omega^1_X$ is

$$\nabla^2 = \mathrm{ad}(F_{\partial+\bar{\partial}}) \otimes 1 + 1 \otimes R$$

where $F_{\partial+\bar{\partial}} = (\partial + \bar{\partial})^2$ is the curvature of the metric connection $\partial + \bar{\partial}$, and R is the curvature of Ω^1_X. The statement that curvature decreases in subbundles, applied to the sub-line bundle of $(\mathrm{End}\, V) \otimes \Omega^1_X$ given by θ, can be written

$$\Delta \log |\theta|^2 \leqslant \frac{\left(2\sqrt{-1}\Lambda\nabla^2(\theta), \theta\right)}{|\theta|^2}.$$

This holds in a distributional sense everywhere.

Now note that $F_{\partial+\bar{\partial}} + \theta\bar{\theta} + \bar{\theta}\theta = 0$; this is the statement that the connection D is flat. Therefore

$$\Delta \log |\theta|^2 \leqslant -\frac{\left(2\sqrt{-1}\,\mathrm{ad}\Lambda(\theta\bar{\theta} + \bar{\theta}\theta)\,(\theta), \theta\right)}{|\theta|^2} + C_3$$

where C_3 depends on the curvature R. Now at any point, if we write $\theta = \theta_0\, dz$ and $\bar{\theta} = \bar{\theta}_0\, d\bar{z}$ in a normal coordinate z, then $2\sqrt{-1}\Lambda(\theta\bar{\theta} + \bar{\theta}\theta) = [\theta_0, \bar{\theta}_0]$ and

$$\frac{\left(2\sqrt{-1}\,\mathrm{ad}\Lambda(\theta\bar{\theta} + \bar{\theta}\theta)\,(\theta), \theta\right)}{|\theta|^2} = \frac{([[\theta_0, \bar{\theta}_0], \theta_0], \theta_0)}{|\theta_0|^2} = \frac{|[\theta_0, \bar{\theta}_0]|^2}{|\theta_0|^2}.$$

We claim that there are c_4, C_5 such that $|[\theta_0, \bar{\theta}_0]| \geqslant c_4 |\theta_0|^2 - C_5(1 + |\theta_0|)$. This claim is the generalization to Higgs bundles of the calculation of the negative curvature of the classifying space for Hodge structures [26]. We may choose an orthonormal basis for the fiber of V and write $\theta_0 = \sigma + \tau$ where σ is a diagonal matrix and τ is strictly upper triangular. Then $|\theta_0|^2 = |\sigma^2| + |\tau|^2$. By the bound for the eigenvalues of θ_0, we have a bound $|\sigma| \leqslant C$. It therefore suffices to prove that $|[\tau, \bar{\tau}]| \geqslant C|\tau|^2$. Now the argument is the same as in the case of Hodge structures; we will sketch it. Note that the matrix for $\bar{\tau}$ is the transpose complex conjugate of the matrix for τ; in particular it is lower triangular. The upper left entry of $[\tau, \bar{\tau}]$ is the square norm of the first row of τ. If this is as big as the square norm of τ, we are done. If it is much smaller than the square norm of τ, then the contribution from the first row to the next entry on the diagonal of $[\tau, \bar{\tau}]$ is small, so this second diagonal entry is approximately the square norm of the second row of τ. Continue in this fashion until reaching a row whose norm is comparable to the norm of τ. This proves the claim.

From this claim we get the estimate

$$\Delta \log |\theta|^2 \leqslant -c_4^2 |\theta|^2 + C_6,$$

where C_6 depends on the estimate for the eigenvalues of θ as well as the curvature R. Now we may apply Ahlfors' lemma [1] (the extra constant C_6 is easy to deal with, say by multiplying $|\theta|^2$ by an appropriate factor e^s) to conclude that $\sup|\theta| \leqslant C_2$. This completes the proof when X is a disk. When X is a polydisk, apply the previously obtained estimate to various disks in X. This gives the required estimate on a smaller polydisk. Now cover a compact space to obtain the proof of the lemma. □

Now, on to the compactness property. Say that a sequence $\{V_i\}$ of harmonic bundles with metrics *converges* to V with respect to a norm $\|\ \|$ if there are C^∞ isomorphisms $\psi_i : V_i \cong V$ which preserve the metric, and such that $\psi_*(D_i) - D \to 0$ and $\psi_*(D_i'') - D'' \to 0$, with respect to the norm $\|\ \|$ on sections of $A^1(\mathrm{End}(V))$.

Lemma **2.8** *(Hitchin* [30]*).* — *Suppose that* $\{V_i\}$ *is a sequence of harmonic bundles with harmonic metrics such that the coefficients of the characteristic polynomials of* θ_i *are bounded in* C^0*-norm. Then there is a harmonic bundle* V *and a subsequence* $V_{i'} \to V$ *convergent with respect to any* L^p *norm.*

Proof. — Fix p large. If the coefficients of the characteristic polynomials are bounded, than the eigenvalues of θ_i are bounded, so, by the above lemma, the $|\theta_i|$ are bounded. Now $F_{\partial_i+\bar\partial_i} + \theta_i\bar\theta_i + \bar\theta_i\theta_i = 0$, hence the curvatures $F_{\partial_i+\bar\partial_i}$ are bounded. By Uhlenbeck's weak compactness theorem [51], there is a subsequence of i and a sequence of isomorphisms $V_i \cong V$ preserving the metric, and there is a unitary connection $\partial + \bar\partial$ on V, such that $a_i \overset{\text{def}}{=} \bar\partial - \bar\partial_i \rightharpoonup 0$ weakly in L^p_1. We have $\bar\partial_i(\theta_i) = 0$, so $\bar\partial(\theta_i) = a_i(\theta_i)$, and these converge to zero strongly in L^p. Hence there is a subsequence of θ_i which converge to some θ weakly in L^p_1. The limit satisfies $\bar\partial(\theta) = 0$. If p is big enough, then the weak convergence implies that $\theta_i \to \theta$ strongly in any L^q. In particular, the equations $\theta_i \wedge \theta_i = 0$ imply that $\theta \wedge \theta = 0$. Finally, $F_{\partial_i+\bar\partial_i} \to F_{\partial+\bar\partial}$ weakly in L^p. In particular, $\bar\partial^2 = 0$. Hence $(V, \bar\partial, \theta)$ is a Higgs bundle. The limits of the operators on V_i are operators appropriately related by the metric on V. The equation $D^2 = 0$ also holds in the limit, which shows that the metric is harmonic for V. The operators on V_i converge weakly in L^p_1 to the operators on V, hence the convergence is strong in L^p. □

Remark. — This lemma is Hitchin's analytic version of the properness of the map taking a Higgs bundle to the characteristic polynomial of θ, which in his case was just the determinant of θ.

Corollary **2.9.** — *If a sequence of stable Higgs bundles* E_i *approaches a limiting* E, *and if* E *is also stable, then the monodromy representations corresponding to* E_i *approach the representation corresponding to* E.

Proof. — Choose harmonic metrics for the E_i. Then these harmonic bundles approach a limit E′. The monodromy representations corresponding to E_i approach the representation corresponding to E′, and similarly the Higgs bundles E_i approach the Higgs bundle E'. But the principle of semicontinuity implies that there must be a nonzero map between E and E'. Since E is stable and E' is semistable of the same rank, this map is an isomorphism. □

MONODROMY GROUPS

In the next two sections, we will give a preliminary discussion of various items having to do with the monodromy representation of a flat connection. These topics will be treated again in a more abstract way in the last two sections of the paper.

Suppose V is a local system on X. Let ρ denote the corresponding representation $\pi_1(X, x) \to Gl(V_x)$. The Zariski closure of the image of ρ is a group $M(V, x)$ called the *monodromy group of* V. (One might also use the terminologies *Mumford-Tate group* or *Galois group* [9], § 7.)

There is an alternate characterization based on the following principle [13]. Suppose G is a complex algebraic group, and $G \subset Gl(n, \mathbf{C})$ is a faithful representation $(n \geqslant 2)$. Denote by $T^{a,b}\,\mathbf{C}^n$ the tensor product $(\mathbf{C}^n)^{\otimes a} \otimes (\mathbf{C}^{n*})^{\otimes b}$ of powers of the standard representation and its dual. For each such tensor product, identify the set of subspaces $W \subset T^{a,b}\,\mathbf{C}^n$ which are preserved by G. Then G may be characterized as the subgroup of elements $g \in Gl(n, \mathbf{C})$ such that $g(W) \subset W$ for every preserved subspace in every tensor product. If G is reductive, then this can be simplified—one need only consider trivial subrepresentations W. In other words, if we let $(T^{a,b}\,\mathbf{C}^n)^G$ denote the subspace of elements fixed by G, then G is characterized as the subgroup of elements $g \in Gl(n, \mathbf{C})$ such that $g(v) = v$ for all $v \in (T^{a,b}\,\mathbf{C}^n)^G$, for all a, b.

These lead to characterizations of the monodromy group $M(V, x)$. For each $T^{a,b}\,V = V^{\otimes a} \otimes (V^*)^{\otimes b}$, identify the set of sub-local systems $W \subset T^{a,b}\,V$. Then $M(V, x)$ is the subgroup of all $g \in Gl(V_x)$ such that $g(W_x) \subset W_x$ for every such sub-local system W. If V is a semisimple local system, them $M(V, x)$ is reductive, and can be characterized as the set of $g \in Gl(V_x)$ such that $g(v_x) = v_x$ for every $v \in H^0_{DR}(X, T^{a,b}\,V)$.

One may use these ideas to define the *monodromy group* $M(E, x)$ for a semistable Higgs bundle E with vanishing Chern classes. It is the subgroup of $g \in Gl(E_x)$ such that, for any Higgs subbundle of degree zero $W \subset T^{a,b}\,E$, g preserves $W_x \subset T^{a,b}\,E_x$.

Lemma **2.10.** — *Suppose* $V = E$ *is a harmonic bundle, with associated flat bundle* V *and associated Higgs bundle* E. *Then* $M(E, x)$ *is reductive, and may be characterized as the group of* $g \in Gl(E_x)$ *such that* $g(v_x) = v_x$ *for all* $v \in H^0_{Dol}(X, T^{a,b}\,E)$. *Finally* $M(E, x) = M(V, x)$ *via the identification* $V_x = E_x$.

Proof. — The $T^{a,b}\,E$ are also harmonic bundles, and in particular they are direct sums of stable Higgs bundles with vanishing Chern classes. Any Higgs subbundle $W \subset T^{a,b}\,E$ of degree zero is a direct summand, coming from a harmonic subbundle [47]. In particular, these correspond exactly with the sub-local systems $W \subset T^{a,b}\,V$ (in a way which commutes with the identification $E_x \cong V_x$). Hence $M(E, x) = M(V, x)$.

The monodromy representation associated to a harmonic bundle is semisimple, so the monodromy group is reductive. We obtain the same alternate characterization for $M(E, x)$ as before, because $H^0_{DR}(T^{a,b} V)$ is equal to $H^0_{Dol}(T^{a,b} E)$ (compatibly with $E_x \cong V_x$). □

Suppose $G \subset Gl(n, \mathbf{C})$ is fixed. Define a *principal harmonic bundle* for G to be a Higgs bundle E with harmonic metric, and with identification $\beta : E_x \cong \mathbf{C}^n$, such that $\beta_* M(E, x) \subset G$, up to equivalence obtained by composing β with elements of G.

The results about moduli spaces stated at the end of § 1 work also for principal harmonic bundles—we state these again without proof. Fix a reductive algebraic subgroup $G \subset Gl(n, \mathbf{C})$. There is a moduli space $\mathbf{M}_{Dol}(G)$ whose points parametrize principal harmonic bundles for the group G. The map to the moduli space $\mathbf{M}_B(G)$ o reductive representations $\pi_1(X) \to G$ is a homeomorphism. Furthermore, the natural map $\mathbf{M}_{Dol}(G) \to \mathbf{M}_{Dol}$ is proper.

REAL REPRESENTATIONS

The existence of canonical metrics and the correspondence between Higgs bundles and local system allow us to state some analogues of the theorem of Frobenius about real representations of finite groups. Lemma 2.12 will be generalized to arbitrary reductive real algebraic groups in § 6, but it may be instructive to have a simple proof in this case.

If $E = (E, \theta)$ is a Higgs bundle, let $E_C = (E, -\theta)$. (It is the result of the action of the element $C = -1 \in \mathbf{C}^*$, cf. § 4.) Hitchin pointed out to me the importance of the action of this element in specifying the Higgs bundles corresponding to real representations.

Lemma **2.11.** — *If* E *is the Higgs bundle corresponding to a semisimple representation of the fundamental group, then the Higgs bundle* $(E_C)^*$ *corresponds to the complex conjugate representation. In particular, the character of the representation takes values in* $\mathbf{R}$ *if and only if there exists a morphism of Higgs bundles*

$$E \otimes E_C \to \mathcal{O}_X.$$

Proof. — Choose a harmonic metric K for E. Let $\overline{E}$ be the complex conjugate bundle with the operators D, D″, D′ deduced from the operators D, D′, and D″ on E respectively. Note that the sections of $\overline{E}$ are expressions of the form $\bar{e}$ where e is a section of E, and for example $D''\bar{e}$ is defined to be equal to $\overline{D'}\,\bar{e}$. Let K also denote the induced metric on $\overline{E}$, so $(\bar{v}, \bar{w}) = (v, w)$. Then the representation corresponding to $(\overline{E}, D)$ is the complex conjugate of the representation corresponding to (E, D). On the other hand, K is a harmonic metric for $\overline{E}$ and the operators D″ and D′ are those induced by this harmonic metric. The metric K gives a bilinear form $M : E \times \overline{E} \to \mathcal{O}_X$, defined by $M(u, \bar{v}) = (u, v)_K$. Let $D''_C = \bar{\partial} - \theta$ be the structural operator for the Higgs bundle E_C. The change in sign between the definitions of ∂_K and $\bar{\theta}_K$ gives the formula

$$(D''_C(u), v)_K + (u, D'(v))_K = \bar{\partial}(u, v)_K.$$

Therefore

$$M(D''_C(u), \bar{v}) + M(u, D''(\bar{v})) = \bar{\partial} M(u, \bar{v});$$

in other words M is a morphism of Higgs bundles

$$M : E_C \otimes \bar{E} \to \mathcal{O}_X.$$

The pairing is perfect, as it came from a harmonic metric, so $\bar{E} \cong (E_C)^*$. □

Remark. — We can use the pairing M_K to give Riemann bilinear relations for cohomology with coefficients in a harmonic bundle. If β is a harmonic form with coefficients in E, then $\bar{\beta}$ is a harmonic form with coefficients in $\bar{E}$ representing the complex conjugate of the de Rham cohomology class. Similarly, if α is a harmonic form decomposed according to the Hodge type of X as $\alpha = \Sigma \alpha^{p,q}$, then set $C\alpha = \Sigma(-1)^p \alpha^{p,q}$. The components $\alpha^{p,q}$ are not harmonic, but $C\alpha$ is a harmonic form for E_C. This operation C should be thought of as acting on the Dolbeault cohomology. The Riemann bilinear relations now state that if α and β are primitive harmonic i-forms, then

$$\int_X (\alpha, \beta)_K \, d\,\mathrm{vol} = (\sqrt{-1})^i (-1)^{i(i+1)/2} c_{d,i} \int_X M_K(C\alpha \wedge \bar{\beta}) \wedge \omega^{d-i}.$$

This follows from the scalar case. The $c_{d,i}$ are universal positive constants depending only on the dimension d of X and the degree i of the forms.

We can strengthen Lemma 2.11 to say when a local system can be defined over **R**.

Lemma **2.12.** — *If* E *is the Higgs bundle corresponding to a semisimple representation of the fundamental group, then the representation can be defined over* **R** *if and only if there exists a non-degenerate symmetric bilinear form*

$$S : E \otimes E_C \to \mathcal{O}_X,$$

in other words a morphism of Higgs bundles such that $S(u, v) = S(v, u)$.

Proof. — The local system corresponding to E is defined over **R** if and only if there is a semilinear involution $\sigma : E \to E$ such that $\sigma D = D\sigma$. Suppose such an involution exists. Choose a harmonic metric K_0, and define a new metric K by the formula

$$(u, v)_K = (u, v)_{K_0} + (\sigma v, \sigma u)_{K_0}.$$

The condition that σ commutes with D insures that this is still a harmonic metric. Furthermore, we have $\sigma D'' = D' \sigma$. Now define the form S by

$$S(u, v) = (u, \sigma v)_K = (v, \sigma u)_K = S(v, u).$$

It is a morphism of Higgs bundles $S : E \times E_C \to \mathcal{O}_X$ because, as above, we have

$$\begin{aligned} S(D'' u, v) + S(u, D''_C v) &= (D'' u, v) + (u, \sigma D''_C v)_K \\ &= (D'' u, v) + (u, D'_C \sigma v) = \bar{\partial} S(u, v). \end{aligned}$$

Conversely suppose S is a nondegenerate symmetric bilinear form as in the lemma. Choose a harmonic metric K for E. Define a map $\rho : E \to E$ by $S(u, v) = (u, \rho v)_K$. It is a semilinear automorphism, and $\rho D'' = D' \rho$ by the same calculation as above. However, it might not be an involution. To fix this, proceed as in [45]. We have

$$(\rho u, \rho v)_{\mathbf{K}} = (v, \rho^2 u)_{\mathbf{K}} = (\rho^2 v, u)_{\mathbf{K}}$$

so ρ^2 is self adjoint. Setting $u = v$ shows that ρ^2 is positive definite. It also commutes with D'', D' and hence D. Therefore ρ^2 is an automorphism of the harmonic bundle. We may choose a positive definite automorphism τ such that $\tau^{-2} = \rho^2$. Then set $\sigma = \rho\tau$. Note that τ is a polynomial in ρ^2, hence a polynomial in ρ, so ρ and τ commute. Therefore $\sigma^2 = 1$. Thus the local system can be defined over $\mathbf{R}$. $\square$

3. Extensions

As an illustration of the formality result (Lemma 2.2), we will extend the equivalence of categories constructed in Theorem 1 and Corollary 1.3 to extensions of irreducible objects. This is accomplished by introducing the notion of differential graded category, which generalizes the notion of differential graded algebra, and plays the role for extensions that differential graded Lie algebras play in the theory of deformations of Schlessinger-Stasheff-Deligne-Goldman-Millson [21]. In fact the following theory of extensions is just a different incarnation of their theory.

DIFFERENTIAL GRADED CATEGORIES

A *differential graded category* (d.g.c.) is an additive $\mathbf{C}$-linear category such that for any objects U and V, Hom(U, V) is endowed with a grading $\bigoplus_{i\geq 0} \mathrm{Hom}^i(\mathrm{U}, \mathrm{V})$ and a differential d of degree one, such that $d^2 = 0$. The axioms are that the identity is an element $1 \in \mathrm{Hom}^0(\mathrm{U}, \mathrm{U})$ with $d(1) = 0$, and that composition of maps behaves like multiplication under the differential:

$$d(f.g) = d(f).g + (-1)^{|f|} f.d(g).$$

A functor between d.g.c.'s is a functor between categories which preserves the grading and differentials of the Hom complexes.

An *isomorphism* between objects in a d.g.c. will always mean a map f of degree zero, such that $d(f) = 0$, and such that there exists an inverse with the same properties. Two objects are *isomorphic* if there exists an isomorphism between them. Note that all of the structures are preserved by isomorphisms.

A natural isomorphism between two functors F and G is a collection of isomorphisms $f_{\mathrm{U}} : \mathrm{F}(\mathrm{U}) \cong \mathrm{G}(\mathrm{U})$, satisfying the usual naturality condition with respect to all elements of the Hom complexes.

We use the notation $\mathrm{Ext}^i(\mathrm{U}, \mathrm{V})$ for the i-th cohomology of the complex Hom(U, V). If $\mathscr{C}$ is a d.g.c., define an additive category $\mathrm{E}^0 \mathscr{C}$ as follows. The objects are the same as the objects of $\mathscr{C}$, but the set of morphisms from U to V is the set $\mathrm{Ext}^0(\mathrm{U}, \mathrm{V})$. Similarly, if $\mathrm{F} : \mathscr{C} \to \mathscr{D}$ is a functor between d.g.c.'s, then $\mathrm{E}^0 \mathrm{F}$ is a functor from $\mathrm{E}^0 \mathscr{C}$

to $E^0 \mathscr{D}$. The notion of natural map between functors $E^0 F$ and $E^0 G$ involves naturality conditions only for morphisms in Ext^0, so it is weaker than the notion of natural map between F and G.

Suppose $\mathscr{C}$ is a differential graded category. An *extension* in $\mathscr{C}$ is a pair of morphisms

$$M \xrightarrow{a} U \xrightarrow{b} N$$

with $a, b \in \mathrm{Hom}^0$, $ba = 0$, and $d(a) = 0$, $d(b) = 0$, such that a splitting exists: a splitting is a pair of morphisms (of degree 0)

$$M \xleftarrow{g} U \xleftarrow{h} N$$

such that $ga = 1$, $bh = 1$, $gh = 0$, and $ag + hb = 1$. Given an extension, define $\delta \in \mathrm{Hom}^1(N, M)$ by $\delta = gd(h)$. One can check that $d(\delta) = 0$ and $a\delta = d(h)$ and $\delta b = d(g)$. For any object Y one gets a long exact sequence:

$$\ldots \mathrm{Ext}^i(V, M) \to \mathrm{Ext}^i(V, U) \to \mathrm{Ext}^i(V, N) \to \mathrm{Ext}^{i+1}(V, M) \to \ldots$$

and similarly for $\mathrm{Ext}(\cdot, V)$. Also note that $[\delta]$ is a class in $\mathrm{Ext}^1(N, M)$. Two extensions are isomorphic if and only if their extension classes coincide, but the isomorphism is not unique.

We say that a d.g.c. $\mathscr{C}$ is *complete* if $\mathrm{Ext}^1(N, M)$ classifies extensions, in other words if every element of $\mathrm{Ext}^1(N, M)$ comes from an extension. We will construct the completion of a d.g.c. by the method used in [21]. Suppose $\mathscr{C}$ is a d.g.c. Define a new d.g.c. $\overline{\mathscr{C}}$ as follows. The objects of $\overline{\mathscr{C}}$ are pairs (U, η) with $U \in \mathscr{C}$ and $\eta \in \mathrm{Hom}^1(U, U)$ such that

$$d(\eta) + \eta^2 = 0.$$

The morphisms from (U, η) to (V, ξ) are the morphisms from U to V with the same grading, but a different differential. The new differential is

$$\hat{d}(f) = d(f) + \xi f - (-1)^{|f|} f\eta.$$

The condition on η and ξ insures that $\hat{d}^2 = 0$, so $\overline{\mathscr{C}}$ is a d.g.c. Define the completion $\hat{\mathscr{C}}$ to be the full subcategory of $\overline{\mathscr{C}}$ consisting of objects which are successive extensions of objects of $\mathscr{C}$.

Lemma **3.1.** — *The category $\mathscr{C}$ is a full subcategory of $\hat{\mathscr{C}}$, $\hat{\mathscr{C}}$ is complete, and every object is an extension of objects of $\mathscr{C}$. These properties uniquely characterize $\hat{\mathscr{C}}$ (up to an equivalence of d.g.c.'s which is unique up to unique natural isomorphism).*

Proof. — To see that $\hat{\mathscr{C}}$ is complete, suppose (M, η) and (N, η) are objects, and $\varphi \in \mathrm{Hom}^1(N, M)$ with $\hat{d}\varphi = 0$ represents an extension class. To construct the extension corresponding to φ take $(M \oplus N, \eta \oplus \eta + \varphi)$. Tautologically, every object of $\hat{\mathscr{C}}$ is an extension of objects of $\mathscr{C}$.

Suppose $\mathscr{D}$ is any d.g.c., and suppose F and G are two functors from $\widehat{\mathscr{C}}$ to $\mathscr{D}$. If $u : \mathrm{F}|_{\mathscr{C}} \to \mathrm{G}|_{\mathscr{C}}$ is a natural morphism, then there exists a unique extension to a natural morphism $\hat{u} : \mathrm{F} \to \mathrm{G}$. To see this, suppose for example that U is an extension of N by M (and keep the same notations as above). Then $u_{\mathrm{U}} : \mathrm{F}(\mathrm{U}) \to \mathrm{G}(\mathrm{U})$ is determined from u_{M} and u_{N} by the formula

$$u_{\mathrm{U}} = \mathrm{G}(ag + hb)\, u_{\mathrm{U}}\, \mathrm{F}(ag + hb) = \mathrm{G}(a)\, u_{\mathrm{M}}\, \mathrm{F}(g) + \mathrm{G}(h)\, u_{\mathrm{N}}\, \mathrm{F}(b).$$

The case of successive extensions can be treated by induction.

Now we show the uniqueness of $\widehat{\mathscr{C}}$. Suppose $\mathscr{D}$ is a d.g.c. with a fully faithful functor $\mathscr{C} \to \mathscr{D}$, such that $\mathscr{D}$ is complete and every object is a successive extension of objects of $\mathscr{C}$. Since $\mathscr{D}$ is complete, the inclusion $\mathscr{D} \to \widehat{\mathscr{D}}$ is surjective on isomorphism classes. One checks that a functor of d.g.c.'s which is fully faithful and surjective on isomorphism classes has an essential inverse which is a morphism of d.g.c.'s, and which is unique up to unique natural isomorphism. We may choose the essentiel inverse $\widehat{\mathscr{D}} \to \mathscr{D}$ to be the identity on $\mathscr{D}$. Composition with the canonical map $\widehat{\mathscr{C}} \to \widehat{\mathscr{D}}$ gives a functor $\widehat{\mathscr{C}} \to \mathscr{D}$ which extends the originally given one. By the previous paragraph, this extended functor is unique up to unique natural isomorphism. Again, $\widehat{\mathscr{C}} \to \mathscr{D}$ is fully faithful, and the hypothesis implies that it is surjective on isomorphism classes. Thus it has an essential inverse, unique up to unique natural isomorphism. □

A functor between d.g.c.'s is *quasi-fully-faithful* if it induces an isomorphism on Ext groups, and is a *quasi-equivalence* if it is quasi-fully-faithful and is surjective on isomorphism classes.

Lemma **3.2.** — *If* $\mathrm{F} : \mathscr{C}_1 \to \mathscr{C}_2$ *is a quasi-fully-faithful functor, then* $\mathrm{E}^0\,\mathrm{F} : \mathrm{E}^0\,\mathscr{C}_1 \to \mathrm{E}^0\,\mathscr{C}_2$ *is fully faithful. If* F *is a quasi-equivalence, then* $\mathrm{E}^0\,\mathrm{F}$ *is an equivalence of categories.*

Proof. — The first statement follows because the morphisms in $\mathrm{E}^0\,\mathscr{C}$ are the Ext^0 groups of $\mathscr{C}$. For the second, note that the isomorphism classes of $\mathrm{E}^0\,\mathscr{C}$ are the same as those of $\mathscr{C}$. □

Lemma **3.3.** — *If* $\mathrm{F} : \mathscr{C}_1 \to \mathscr{C}_2$ *is a quasi-equivalence then it induces a quasi-equivalence of completions* $\hat{\mathrm{F}} : \widehat{\mathscr{C}}_1 \to \widehat{\mathscr{C}}_2$.

Proof. — First we prove that $\hat{\mathrm{F}}$ is quasi-fully-faithful. Suppose that U and V are objects of $\widehat{\mathscr{C}}$ and U is an extension of N by M. If $\mathrm{F} : \mathrm{Ext}^i(\mathrm{V}, \mathrm{M}) \cong \mathrm{Ext}^i(\mathrm{FV}, \mathrm{FM})$ and $\mathrm{F} : \mathrm{Ext}^i(\mathrm{V}, \mathrm{N}) \cong \mathrm{Ext}^i(\mathrm{FV}, \mathrm{FN})$, then by the long exact sequence of Ext groups and the five-lemma, $\mathrm{F} : \mathrm{Ext}^i(\mathrm{V}, \mathrm{U}) \cong \mathrm{Ext}^i(\mathrm{FV}, \mathrm{FU})$. A similar statement holds if V is expressed as an extension. Then since every object of $\widehat{\mathscr{C}}_1$ is an extension of objects in $\mathscr{C}_1$, and by assumption the maps F are isomorphisms on Ext groups of objects of $\mathscr{C}_1$, one shows inductively that the maps F are isomorphisms on Ext groups of all objects of $\widehat{\mathscr{C}}_1$.

To show that $\hat{F}$ is surjective on isomorphism classes, use induction again. If U is an object of $\hat{\mathscr{C}}_2$ expressed as an extension of FN by FM, then it corresponds to an extension class $\varphi \in \mathrm{Ext}^1(\mathrm{FN}, \mathrm{FM})$. By the previous paragraph, F is an isomorphism on Ext^1, so $\varphi = \mathrm{F}\psi$ for an extension class $\psi \in \mathrm{Ext}^1(\mathrm{N}, \mathrm{M})$. Then if V is the extension given by ψ, $\mathrm{U} = \mathrm{FV}$. □

Remark. — These two lemmas imply that if F is a quasi-equivalence, then $\mathrm{E}^0\,\hat{\mathrm{F}}$ is an equivalence between $\mathrm{E}^0\,\hat{\mathscr{C}}_1$ and $\mathrm{E}^0\,\hat{\mathscr{C}}_2$.

EXAMPLES

We will give several examples of differential graded categories related to the results of § 1. The idea of completion of a differential graded category will allow us to extend the equivalence of categories given in Theorem 1, to objects which are extensions of stable or irreductible objects. Define the following d.g.c.'s.

3.4.1. $\mathscr{C}_{\mathrm{DR}}$ is the category of all flat bundles, with

$$\mathrm{Hom}^{\bullet}(\mathrm{U}, \mathrm{V}) = (\mathrm{A}^{\bullet}(\mathrm{Hom}(\mathrm{U}, \mathrm{V})), \mathrm{D}).$$

$\mathscr{C}^{s}_{\mathrm{DR}}$ is the full subcategory consisting of semisimple objects, $\mathscr{C}^{\mathrm{triv}}_{\mathrm{DR}}$ is the full subcategory of trivial objects (those isomorphic to $\mathbf{C}^n$), and $\mathscr{C}^{\mathrm{nil}}_{\mathrm{DR}}$ is the full subcategory consisting of nilpotent objects (extensions of trivial objects).

3.4.2. $\mathscr{C}_{\mathrm{Dol}}$ is the category of Higgs bundles which are extensions of stable Higgs bundles of degree zero, with $\mathrm{ch}_2 . [\omega]^{\dim \mathrm{X} - 2} = 0$; and

$$\mathrm{Hom}^{\bullet}(\mathrm{U}, \mathrm{V}) = (\mathrm{A}^{\bullet}(\mathrm{Hom}(\mathrm{U}, \mathrm{V})), \mathrm{D}'').$$

$\mathscr{C}^{s}_{\mathrm{Dol}}$ is the full subcategory consisting of semisimple objects (polystable Higgs bundles), $\mathscr{C}^{\mathrm{triv}}_{\mathrm{Dol}}$ is the full subcategory of trivial objects (those isomorphic to $\mathscr{O}^n_{\mathrm{X}}$), and $\mathscr{C}^{\mathrm{nil}}_{\mathrm{Dol}}$ is the full subcategory consisting of nilpotent objects (extensions of trivial objects).

3.4.3. $\mathscr{C}^{s}_{\mathrm{H}}$ is the category of all harmonic bundles, with

$$\mathrm{Hom}^{\bullet}(\mathrm{U}, \mathrm{V}) = (\mathscr{H}^{\bullet}(\mathrm{Hom}(\mathrm{U}, \mathrm{V})), 0).$$

Here $\mathscr{H}$ denotes the space of harmonic forms. This category is the same as the category of semisimple flat bundles with morphisms the de Rham cohomology classes and zero differential, and is the same as the category of objects of $\mathscr{C}_{\mathrm{Dol}}$ with morphisms the Dolbeault cohomology classes and zero differential.

3.4.4. $\mathscr{C}^{s}_{\mathrm{D}'}$ is the category of all harmonic bundles, with

$$\mathrm{Hom}^{\bullet}(\mathrm{U}, \mathrm{V}) = (\ker(\mathrm{D}') \subset \mathrm{A}^{\bullet}(\mathrm{Hom}(\mathrm{U}, \mathrm{V})), \mathrm{D}'').$$

3.4.5. Suppose $A^\bullet$ is any differential graded algebra. Then we can form a d.g.c. $\mathscr{C}^{\mathrm{triv}}(A^\bullet)$. The objects are of the form $\mathbf{1}^n$, and the morphisms are appropriate direct sums of $A^\bullet$. Note that $\mathscr{C}^{\mathrm{triv}}_{\mathrm{DR}} = \mathscr{C}^{\mathrm{triv}}(A^\bullet_{\mathrm{DR}})$ where $A^\bullet_{\mathrm{DR}} = (A^\bullet(X), d)$. Similarly, $\mathscr{C}^{\mathrm{triv}}_{\mathrm{Dol}} = \mathscr{C}^{\mathrm{triv}}(A^\bullet_{\mathrm{Dol}})$ where $A^\bullet_{\mathrm{Dol}} = (A^\bullet(X), \bar\partial)$.

Lemma **3.4.** — *There are quasi-equivalences of d.g.c.'s from $\mathscr{C}^s_{\mathrm{D}'}$ to $\mathscr{C}^s_{\mathrm{DR}}$, $\mathscr{C}^s_{\mathrm{Dol}}$ and $\mathscr{C}^s_{\mathrm{H}}$. There are equivalences of d.g.c.'s $\mathscr{C}_{\mathrm{DR}} = \widehat{\mathscr{C}}^s_{\mathrm{DR}}$ and $\mathscr{C}_{\mathrm{Dol}} = \widehat{\mathscr{C}}^s_{\mathrm{Dol}}$, which restrict to $\mathscr{C}^{\mathrm{nil}}_{\mathrm{DR}} = \widehat{\mathscr{C}}^{\mathrm{triv}}_{\mathrm{DR}}$ and $\mathscr{C}^{\mathrm{nil}}_{\mathrm{Dol}} = \widehat{\mathscr{C}}^{\mathrm{triv}}_{\mathrm{Dol}}$ on the subcategories of nilpotent objects.*

Proof. — The first three are due to Lemma 2.2. The last two are due to the universality of the completion $\widehat{\mathscr{C}}$, and the fact that $\mathscr{C}_{\mathrm{DR}}$ and $\mathscr{C}_{\mathrm{Dol}}$ are complete and all objects are extensions of objects of $\mathscr{C}^s_{\mathrm{DR}}$ and $\mathscr{C}^s_{\mathrm{Dol}}$. □

Remark. — We can describe the inverse functors $\widehat{\mathscr{C}}^s_{\mathrm{DR}} \to \mathscr{C}_{\mathrm{DR}}$ and $\widehat{\mathscr{C}}^s_{\mathrm{Dol}} \to \mathscr{C}_{\mathrm{Dol}}$ concretely. An object of $\widehat{\mathscr{C}}^s_{\mathrm{DR}}$ is a pair (U, η) where $U = (U, D)$ is a semisimple flat bundle, and $\eta \in A^1(X, \mathrm{End}(U))$ is a one-form with $D(\eta) + \eta \wedge \eta = 0$. This pair maps to the flat bundle $(U, D + \eta)$. The same works for the Higgs bundles.

Lemma **3.5.** — *There is an equivalence of categories between the category of flat bundles and the category of Higgs bundles which are extensions of stable bundles of degree zero, with* $\mathrm{ch}_2(E).[\omega]^{\dim X - 2} = 0$. *Furthermore these are equivalent to the category* $E^0(\widehat{\mathscr{C}}^s_{\mathrm{H}})$ *formed from the data of the simple objects and their* Ext *groups and cup products.*

Proof. — To see this, note that the categories of flat and Higgs bundles referred to are $E^0(\mathscr{C}_{\mathrm{DR}})$ and $E^0(\mathscr{C}_{\mathrm{Dol}})$ respectively. By Lemmas 3.2, 3.3 and 3.4, these are equivalent and equivalent to $E^0(\widehat{\mathscr{C}}^s_{\mathrm{D}'})$ and $E^0(\widehat{\mathscr{C}}^s_{\mathrm{H}})$. □

Remarks. — 1. The equivalence of categories constructed in this lemma is functorial with respect to morphisms of X.

2. We can interpret the last statement of the lemma as a formality statement, the generalization to nonabelian cohomology of the formality statement of [12].

In a slight abuse of notation, we can in example (3.4.5) let $\mathscr{C}^{\mathrm{nil}}(A^\bullet)$ denote the completion of $\mathscr{C}^{\mathrm{triv}}(A^\bullet)$. Then according to the lemma, $\mathscr{C}^{\mathrm{nil}}_{\mathrm{DR}} = \mathscr{C}^{\mathrm{nil}}(A^\bullet_{\mathrm{DR}})$ and $\mathscr{C}^{\mathrm{nil}}_{\mathrm{Dol}} = \mathscr{C}^{\mathrm{nil}}(A^\bullet_{\mathrm{Dol}})$, and both are quasi-equivalent to $\mathscr{C}^{\mathrm{nil}}(H^\bullet)$. In the context of example (3.4.5), everything we have done so far is exactly taken from the paper of Goldman and Millson [21].

The equivalences referred to in the lemma are obtained by equivalences with the category $E^0\,\widehat{\mathscr{C}}_{\mathrm{D}'}$, which makes them very concrete, in a manner suggested by K. Corlette. By definition, the objects of this category are pairs (U, η) where $U = (U, D, D', D'')$ is a harmonic bundle and $\eta \in A^1(X, \mathrm{End}(U))$ is a one-form such that $D'(\eta) = 0$, and $D''(\eta) + \eta \wedge \eta = 0$. The morphisms between (U, η) and (V, ξ) are maps $f\colon U \to V$ such that, first of all, $D'(f) = 0$. This implies that f is a morphism of harmonic bundles, so $D''(f) = 0$. Second, the condition that the differential of f is zero becomes $\xi f - f\eta = 0$. Thus a morphism is simply a morphism of harmonic bundles which intertwines η and ξ. A pair (U, η) goes to the Higgs bundle $(U, D'' + \eta)$, and to the flat bundle $(U, D + \eta)$. These functors are equivalences, so for any Higgs bundle (satisfying the conditions of the lemma) or flat bundle, there is a unique pair (U, η) which maps to it.

The C^∞ bundles underlying the corresponding Higgs bundles and flat bundles are naturally isomorphic (in a way which is determined by functoriality for pullback to a point)—see Lemma 6.13 below.

It is easy to see from the above interpretation that if a representation is an *extension of unitary representations* or equivalently if the corresponding Higgs bundle is an extension of stable vector bundles, then the identification between the underlying bundles preserves the holomorphic structure. This is true for unitary harmonic bundles, and adding a nilpotent η as described above changes both holomorphic structures by $\eta^{0,1}$.

THE THEOREM OF MEHTA AND RAMANATHAN

We show that the theorem of Mehta and Ramanathan [36] [37] about restriction of semistable and stable sheaves to hyperplane sections works for Higgs sheaves too. The results of this part are valid only when X is a smooth projective variety. The aim is to show that the category $\mathscr{C}_{\mathrm{Dol}}$ consists of all semistable Higgs bundles E with $\mathrm{ch}_1(E).[\omega]^{\dim X - 1} = 0$ and $\mathrm{ch}_2(E).[\omega]^{\dim X - 2} = 0$.

A Higgs sheaf is a coherent sheaf E with a map $\theta : E \to E \otimes \Omega^1_X$ such that $\theta \wedge \theta = 0$. Say that a Higgs sheaf is torsion-free if the coherent sheaf is torsion-free, and E is reflexive if it is equal to its double dual E^{**}. A Higgs sheaf is semistable or stable if it is torsion free, and if the usual condition is satisfied using the degree of a coherent sheaf.

Suppose $\mathscr{W}$ is a vector bundle on X. Then we consider the category of sheaves on X with $\mathscr{W}$-valued operators: an object is a sheaf $\mathscr{E}$ together with a map $\eta : \mathscr{E} \to \mathscr{E} \otimes \mathscr{W}$. The category of Higgs bundles is contained in the category of sheaves with Ω^1_X-valued operators. In the category of torsion free sheaves with $\mathscr{W}$-valued operators, one can make the usual definitions of semistability, stability, and so forth, by considering subsheaves $\mathscr{F} \subset \mathscr{E}$ such that $\eta\mathscr{F} \subset \mathscr{F} \otimes \mathscr{W}$. If $Y \subset X$ is a subvariety, then we may restrict $\mathscr{W}$ to Y, and if $\mathscr{E}$ is a sheaf on X with $\mathscr{W}$-valued operator, then $\mathscr{E}|_Y$ is a sheaf on Y with $\mathscr{W}|_Y$-valued operator. The arguments of Mehta and Ramanathan [36] [37] show:

Proposition **3.6.** — *If $\mathscr{E}$ is a torsion free semistable (resp. stable) sheaf on* X *with $\mathscr{W}$-valued operator, then for sufficiently general hyperplane sections* Y *of certain arbitrarily high degrees, $\mathscr{E}|_Y$ is a semistable (resp. stable) sheaf with $\mathscr{W}|_Y$-valued operator.* □

Lemma **3.7.** — *If* X *is projective and* E *is a semistable (resp. stable) torsion free Higgs sheaf then for general hyperplane sections* Y *of certain arbitrarily high degrees,* $E|_Y$ *is semistable (resp. stable).*

Proof. — For generic Y of certain arbitrarily high degrees, $E|_Y$ is semistable (resp. stable) as a sheaf with $\Omega^1_X|_Y$-valued operator. Furthermore, by applying the theorem of Mehta and Ramanathan to the canonical filtration of the sheaf E with semistable quotients, we may assume that for all generic Y of the various arbitrarily high degrees, there is an upper bound B for the degree of any subsheaf $\mathcal{F} \subset E|_Y$. Here the notion of degree should be normalized so that the degree of $\mathcal{O}_Y(1)$ is always one. Now suppose $F \subset E|_Y$ is a saturated sub-Higgs sheaf, with $\deg(F) \geqslant A = (r(F)/r(E)) \deg(E)$. Let $G = (E|_Y)/F$. Consider the exact sequence

$$0 \to E(-Y)|_Y \to E \otimes \Omega^1_X|_Y \to E \otimes \Omega^1_Y \to 0.$$

Composing the map $\theta : F \to E \otimes \Omega^1_X|_Y$ with the projection to G, we get a map $F \to G(-Y)$. Let $H \subset F$ be the kernel of this map. Then $\deg(H) \leqslant B$. Thus the image F/H has degree at least $A - B$. Let $J \subset E(-Y)|_Y$ be the inverse image of F/H. Then we have an exact sequence

$$0 \to F \to J(Y) \to F/H(Y) \to 0$$

so $\deg(J(Y)) \geqslant 2A - B + r(F/H).\deg(\mathcal{O}_Y(Y))$. But J(Y) is a subsheaf of $E|_Y$, so this number is bounded by B. If Y is taken to have high degree, this implies that $F/H = 0$, and hence $\theta : F \to F \otimes \Omega^1_X|_Y$. Thus F is a subsheaf of E with $\Omega^1_X|_Y$-valued operator, so $\deg(F)/r(F) \leqslant \deg(E)/r(E)$ (resp. $<$). This proves that $E|_Y$ is semistable (resp. stable) as a Higgs sheaf on Y. □

Corollary **3.8.** — *The tensor product of two semistable Higgs bundles is again semistable.*

Proof. — The converse of the above lemma is also true: if the restriction to a generic hyperplane section is semistable, then the Higgs bundle is semistable. This reduces the statement to the case of curves. On a curve, a semistable Higgs bundle is an extension of stable ones, and we may choose Hermitian-Yang-Mills metrics for the stable subquotients. The tensor product of these metrics is again Hermitian-Yang-Mills, so the tensor product of two stable Higgs bundles on a curve is polystable. Applied to the subquotients, this shows that the tensor product of the semistable bundles is semistable. □

In a similar vein, we can generalize the Enriques-Severi lemma of [36] and [37] to Higgs sheaves.

Lemma **3.9.** — *Suppose* M *and* N *are reflexive Higgs sheaves. If* Y *is a hyperplane section of high degree, then* $\mathrm{Hom}(M, N) \cong \mathrm{Hom}(M|_Y, N|_Y)$. *If* M *and* N *vary in bounded families then a single* Y *can be chosen.*

Proof. — This is proved for reflexive sheaves in [36]. We have to prove that if (E, θ) is a reflexive Higgs sheaf then $H^0_{\mathrm{Dol}}(E) \xrightarrow{\cong} H^0_{\mathrm{Dol}}(E|_Y)$. If Y has high degree, the map is injective. We may also assume that $H^0_{Zar}(E(-Y)|_Y) = 0$. If $e \in H^0(E|_Y)$ we may assume that e extends to a section over X. If $\theta_Y(e) = 0$ in $E|_Y \otimes \Omega^1_Y$ then by the exact sequence

$$0 \to E(-Y)|_Y \to E \otimes \Omega^1_X|_Y \to E \otimes \Omega^1_Y \to 0$$

it follows that $\theta(e) = 0$ in $E \otimes \Omega^1_X|_Y$. If Y has high degree, this implies that $\theta(e) = 0$ on X. □

Theorem **2.** — *Suppose* X *is a smooth projective variety of dimension n, with* $[\omega]$ *equal to the hyperplane class. Let* P_0 *be the Hilbert polynomial of* $\mathcal{O}_X$. *Suppose* E *is a semistable torsion free Higgs sheaf with* $ch_1(E).[\omega]^{n-1} = 0$ *and* $ch_2(E).[\omega]^{n-2} = 0$. *Make an additional assumption, either that* E *is reflexive, or that the Hilbert polynomial of* E *is* $rk(E)\, P_0$. *Then* E *is an extension of stable Higgs bundles with vanishing Chern classes.*

Proof. — If X is a curve this is clear. Suppose X is a surface. We claim more generally that if E is a torsion-free semistable Higgs sheaf of degree 0, then $ch_2(E) \geq 0$ and if equality holds, E is an extension of stable Higgs bundles. (In particular, the conclusion of the theorem will hold without the additional assumptions.) To prove the claim, suppose first that E is stable. The double dual E^{**} is a stable Higgs bundle and there is an exact sequence

$$0 \to E \to E^{**} \to S \to 0$$

where S is concentrated on a finite set of points. Thus $ch_2(S) \leq 0$ with equality only if $S = 0$. By Theorem 1, E^{**} has a Hermitian-Yang-Mills metric. This implies that $ch_2(E^{**}) \geq 0$ [16] [52] [47]. Thus $ch_2(E) \geq 0$ and if equality holds, $S = 0$ so E is a bundle. Now if E is semistable, write it as an extension of stable Higgs sheaves and apply the claimed statement to them.

To prove the theorem in higher dimensions we will use the argument of Mehta and Ramanathan [37]. Suppose $\dim(X) \geq 3$ and we have proved the theorem for smaller dimensions. Suppose E is a semistable torsion free Higgs sheaf of degree 0 with $ch_2(E).[\omega]^{n-2} = 0$. Let Y be a general hyperplane section of some high degree such that $E|_Y$ is semistable (and—in case we are using the first additional assumption—such that $E|_Y$ is reflexive). Note that $E|_Y$ satisfies the conditions of the theorem on Y. By the inductive hypothesis $E|_Y$ is an extension of stable Higgs bundles, so we may apply Lemma 3.5 to conclude that $E|_Y$ comes from a local system V_Y on Y. By the Lefschetz theorem $\pi_1(Y) \cong \pi_1(X)$, V_Y is a local system on X. By Lemma 3.5, this local system comes from a Higgs bundle E_Y on X. E_Y is an extension of stable Higgs bundles of degree zero. The correspondence of Lemma 3.5 commutes with restriction to smooth subvarieties, so $E_Y|_Y \cong E|_Y$. The Higgs bundles E_Y which come from local systems on X via Lemma 3.5 vary in a bounded family. Therefore we may choose Y so that the Enriques-Severi lemma applies to E^{**} and E_Y. Thus $E^{**} \cong E_Y$ on X, so E^{**} is an extension of stable bundles. This completes the proof in case we assume that E is reflexive.

Suppose we assume instead that the Hilbert polynomial of E is $\mathrm{rk}(E)\, P_0$. By the conclusion of the previous paragraph, the Hilbert polynomial of E^{**} is that of a flat bundle, also $\mathrm{rk}(E)\, P_0$. Now we have

$$0 \to E \to E^{**} \to T \to 0$$

and both E and E^{**} have the same Hilbert polynomial. Thus the Hilbert polynomial of T is zero, so $T = 0$, and $E = E^{**}$ is an extension of stable bundles. □

Remark. — As a consequence of this theorem, we may reinterpret $\mathscr{C}_{\mathrm{Dol}}$, as the differential graded category of all semistable Higgs bundles with $\mathrm{ch}_1(E) . [\omega]^{\dim X - 1} = 0$ and $\mathrm{ch}_2(E) . [\omega]^{\dim X - 2} = 0$.

Corollary **3.10.** — *Suppose* X *is projective and* $[\omega]$ *is the hyperplane class. There is an equivalence of categories between the category of flat bundles and the category of semistable Higgs bundles with* $\mathrm{ch}_1 . [\omega]^{\dim X - 1} = 0$ *and* $\mathrm{ch}_2 . [\omega]^{\dim X - 2} = 0$. *Furthermore these are equivalent to the category* $E^0(\widehat{\mathscr{C}}{}^s_X)$ *formed from the data of the simple objects and their* Ext *groups and cup products.* □

Continuing with the discussion at the end of the previous subsection about extensions of unitary representations, we may conclude that any semistable vector bundle with vanishing Chern classes has a holomorphic flat structure which is an extension of unitary flat bundles.

TENSOR PRODUCTS

Recall that a *tensor category* is an additive **C**-linear category with an operation $\otimes$, satisfying associativity and commutativity constraints (which entail the existence of various natural isomorphisms), and with a unit **1** (cf. [13] [43]). We would like to extend this notion to differential graded categories, and prove that it is compatible with the operation of completion. Then our equivalences constructed above will be compatible with tensor products.

If $\mathscr{C}$ is a d.g.c. then define a new d.g.c. $\mathscr{C} \boxtimes \mathscr{C}$ as follows. The objects are pairs denoted $U \boxtimes V$. The Hom complexes are

$$\mathrm{Hom}^n(U \boxtimes U', V \boxtimes V') = \bigoplus_{i+j=n} \mathrm{Hom}^i(U, V) \otimes_{\mathbf{C}} \mathrm{Hom}^j(U', V')$$

(use the notation $\alpha \boxtimes \alpha'$ for the element of $\mathrm{Hom}(U \boxtimes U', V \boxtimes V')$ corresponding to $\alpha \otimes \alpha'$ on the right hand side). The composition of maps is defined by

$$(\alpha \boxtimes \alpha')\, (\beta \boxtimes \beta') = (-1)^{|\alpha'|\,|\beta|}\, \alpha\beta \boxtimes \alpha'\, \beta'.$$

The differential is defined by

$$d(\alpha \boxtimes \alpha') = d(\alpha) \boxtimes \alpha' + (-1)^{|\alpha|}\, \alpha \boxtimes d(\alpha').$$

A *differential graded tensor category (d.g.t.c.)* is a differential graded category $\mathscr{C}$ together with a functor of differential graded categories

$$\otimes : \mathscr{C} \boxtimes \mathscr{C} \to \mathscr{C},$$

with associativity and commutativity constraints, which are natural isomorphisms denoted ψ and φ respectively, and a unit object. The definitions of the associativity and commutativity constraints involve associativity and commutativity constraints for the operation $\boxtimes$. For example there is a natural equivalence of categories

$$(\mathscr{C} \boxtimes \mathscr{C}) \boxtimes \mathscr{C} \cong \mathscr{C} \boxtimes (\mathscr{C} \boxtimes \mathscr{C})$$

both being equivalent to the category $\mathscr{C} \boxtimes \mathscr{C} \boxtimes \mathscr{C}$ of triples $U \boxtimes V \boxtimes W$, where composition of morphisms is defined using the usual sign. The associativity constraint is a natural isomorphism $\psi : (U \otimes V) \otimes W \cong U \otimes (V \otimes W)$ of functors from $\mathscr{C} \boxtimes \mathscr{C} \boxtimes \mathscr{C}$ to $\mathscr{C}$, satisfying a compatibility axiom. The commutativity constraint for the operation $\boxtimes$ must also take into account the usual sign change. This functor $\mathscr{C} \boxtimes \mathscr{C} \cong \mathscr{C} \boxtimes \mathscr{C}$ takes $U \boxtimes V$ to $V \boxtimes U$, and takes $\alpha \boxtimes \beta$ to $(-1)^{|\alpha||\beta|} \beta \boxtimes \alpha$ (it would not be a functor without the sign change). The commutativity constraint for $\otimes$ is a natural isomorphism $\varphi : U \otimes V \cong V \otimes U$ of functors from $\mathscr{C} \boxtimes \mathscr{C}$ to $\mathscr{C}$, which satisfies $\varphi_{U,V} \varphi_{V,U} = 1$, and also a compatibility with the associativity constraint. The unit object $\mathbf{1}$ for the operation $\otimes$ is an object together with natural isomorphisms $\mathbf{1} \otimes U \cong U$, satisfying the appropriate compatibility conditions.

A *tensor functor* from one d.g.t.c. $\mathscr{C}$ to another $\mathscr{D}$ is a pair consisting of a functor $F : \mathscr{C} \to \mathscr{D}$ of d.g.c.'s, and a natural isomorphism $\tau : F(U \otimes V) \cong F(U) \otimes F(V)$ compatible with the associativity and commutativity constraints. Also there should be provided a natural isomorphism $F(\mathbf{1}) \cong \mathbf{1}$.

The category $\mathscr{V}ect$ of $\mathbf{C}$-vector spaces is a d.g.t.c. in a trivial way. A *fiber functor* for a d.g.t.c. $\mathscr{C}$ is a tensor functor ω from $\mathscr{C}$ to $\mathscr{V}ect$.

The compatibility conditions mentioned above are exactly the same as the conditions for a tensor category. They are described in detail in [43] or [13], although perhaps originally due to MacLane. The category $E^0 \mathscr{C}$ is a subcategory of $\mathscr{C}$, and the constraints are morphisms in $E^0 \mathscr{C}$. The compatibility conditions are equations involving these constraints, so they are equations of morphisms in $E^0 \mathscr{C}$. The compatibility conditions may all be described by saying that the category $E^0 \mathscr{C}$, with its tensor product and constraints, should be a tensor category (and similarly for tensor functors)—making Lemma 3.11 below tautologically true. One should be careful to note, however, that the naturality conditions on the constraints are stronger than simply saying that they should hold for $E^0 \mathscr{C}$. The constraints are required to be natural with respect to all elements of the Hom complexes in $\mathscr{C}$.

Lemma **3.11.** — *If $\mathscr{C}$ is a d.g.t.c. then $E^0 \mathscr{C}$ is an additive $\mathbf{C}$-linear tensor category. If F is a tensor functor of d.g.t.c.'s then $E^0 F$ is a tensor functor. In particular, if ω is a fiber functor for $\mathscr{C}$, then $E^0 \omega$ is a fiber functor for $E^0 \mathscr{C}$.*

Proof. — See the previous paragraph. □

Lemma **3.12.** — *If $\mathscr{C}$ is a d.g.t.c., then the completion $\widehat{\mathscr{C}}$ has a canonical structure of d.g.t.c.; and if $\widehat{F}$ is a tensor functor of d.g.t.c.'s then $\widehat{F}$ has a canonical structure of tensor functor.*

Proof. — First we will show that $\overline{\mathscr{C}}$ has a structure of d.g.t.c.:

$$(\mathrm{U}, \eta) \otimes (\mathrm{V}, \mu) = (\mathrm{U} \otimes \mathrm{V}, \eta \otimes 1 + 1 \otimes \mu).$$

The object on the right is an element of $\overline{\mathscr{C}}$, as is seen by verifying that

$$d(\eta \otimes 1 + 1 \otimes \mu) + (\eta \otimes 1 + 1 \otimes \mu)^2 = (d\eta) + \eta^2) \otimes 1 \\ + 1 \otimes (d\mu + \mu^2) + (\eta \otimes 1)(1 \otimes \mu) + (1 \otimes \mu)(\eta \otimes 1) = 0.$$

The last two terms cancel because of the sign change in the definition of $\boxtimes$. To define the action of tensor product on morphisms, recall that

$$\mathrm{Hom}^{\bullet}((\mathrm{U}, \eta), (\mathrm{U}', \eta')) = \mathrm{Hom}^{\bullet}(\mathrm{U}, \mathrm{U}')$$

but with different differential. One checks that the resulting map given by functoriality of tensor product in $\mathscr{C}$

$$\mathrm{Hom}^{\bullet}((\mathrm{U}, \eta_1), (\mathrm{U}', \eta_2)) \otimes \mathrm{Hom}^{\bullet}((\mathrm{V}, \eta_3), (\mathrm{V}', \eta_4)) \\ \rightarrow \mathrm{Hom}^{\bullet}((\mathrm{U} \otimes \mathrm{V}, \eta_{13}), (\mathrm{U}' \otimes \mathrm{V}', \eta_{24}))$$

is also compatible with the new differentials. The associativity and commutativity constraints are defined using the ones from $\mathscr{C}$. Again one checks that they are morphisms killed by the new differentials as well. This is due to the fact that they are natural; for example ψ intertwines $(\eta \otimes 1) \otimes 1$ with $\eta \otimes (1 \otimes 1)$ and so forth.

Thus $\overline{\mathscr{C}}$ has a canonical structure of d.g.t.c. The compatibility conditions for the constraints remain satisfied, because they are simply equations of elements of the Hom^0 sets, which are the same as the Hom^0 sets in $\mathscr{C}$. The structure of d.g.t.c. for $\widehat{\mathscr{C}}$ is obtained by noting that if $\mathrm{M} \rightarrow \mathrm{U} \rightarrow \mathrm{N}$ is an extension, then $\mathrm{M} \otimes \mathrm{V} \rightarrow \mathrm{U} \otimes \mathrm{V} \rightarrow \mathrm{N} \otimes \mathrm{V}$ is an extension. Thus $\widehat{\mathscr{C}}$, being the full subcategory of objects which are successive extensions of objects of $\mathscr{C}$, is preserved by tensor product.

Suppose $\mathrm{F} : \mathscr{C} \rightarrow \mathscr{D}$ is a tensor functor, with associated natural isomorphism τ. Recall that $\widehat{\mathrm{F}} : \widehat{\mathscr{C}} \rightarrow \widehat{\mathscr{D}}$ is defined by

$$\widehat{\mathrm{F}}(\mathrm{U}, \eta) = (\mathrm{F}(\mathrm{U}), \mathrm{F}(\eta)).$$

Extend τ to $\widehat{\mathscr{C}}$ by setting $\widehat{\tau}_{(\mathrm{U}, \eta), (\mathrm{V}, \mu)} = \tau_{\mathrm{U}, \mathrm{V}}$ in

$$\mathrm{Hom}^0((\mathrm{F}(\mathrm{U} \otimes \mathrm{V}), \mathrm{F}(\eta \otimes 1 + 1 \otimes \mu)), (\mathrm{F}(\mathrm{U}), \mathrm{F}(\eta)) \otimes (\mathrm{F}(\mathrm{V}), \mathrm{F}(\mu))) \\ = \mathrm{Hom}^0(\mathrm{F}(\mathrm{U} \otimes \mathrm{V}), \mathrm{F}(\mathrm{U}) \otimes \mathrm{F}(\mathrm{V})).$$

Again, one must show that $\widehat{d}(\tau) = 0$. We know that $d(\tau) = 0$ so what remains is to show that τ intertwines $\mathrm{F}(\eta \otimes 1 + 1 \otimes \mu)$ with $\mathrm{F}(\eta) \otimes 1 + 1 \otimes \mathrm{F}(\mu)$. The reason for this is that τ is a natural morphism of functors.

Note that $\widehat{\tau}$ is compatible with the associativity and commutativity constraints, because τ is. Therefore $(\widehat{\mathrm{F}}, \widehat{\tau})$ is a tensor functor from $\widehat{\mathscr{C}}$ to $\widehat{\mathscr{D}}$. $\square$

Given a fiber functor $\omega : \mathscr{C} \rightarrow \mathscr{V}ect$, it has a natural extension to a fiber functor $\widehat{\omega} : \widehat{\mathscr{C}} \rightarrow \mathscr{V}ect$.

Finally we note that all of the categories and functors considered in examples (3.4.1)-(3.4.5) above are d.g.t.c.'s and tensor functors in obvious ways. The notions of tensor product of Higgs bundles and flat bundles provide $\mathscr{C}_{\mathrm{Dol}}$ and $\mathscr{C}_{\mathrm{DR}}$ with structures of d.g.t.c., bearing in mind the fact that the tensor product of two semistable Higgs bundles is again semistable (Corollary 3.8). Note that the unit objects are the obvious ones, either the trivial Higgs bundle $\mathcal{O}_{\mathrm{X}}$ or the trivial flat bundle $\mathbf{C}$.

Since the tensor product of two harmonic bundles is again a harmonic bundle, the category $\mathscr{C}^s_{\mathrm{D}'}$ has a structure of d.g.t.c. (the tensor product of two morphisms killed by D′ is again killed by D′). Furthermore, the subcategories $\mathscr{C}^s_{\mathrm{Dol}}$ and $\mathscr{C}^s_{\mathrm{DR}}$ of harmonic objects are preserved by tensor product. Also the category $\mathscr{C}^s_{\mathrm{H}}$ where the morphisms are the cohomology classes, is a tensor category. Even in example (3.4.5), for any d.g.a. $\mathrm{A}^\bullet$, $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}^\bullet)$ is a d.g.t.c.

The functors from $\mathscr{C}^s_{\mathrm{D}'}$ to $\mathscr{C}^s_{\mathrm{DR}}$, $\mathscr{C}^s_{\mathrm{Dol}}$, and $\mathscr{C}^s_{\mathrm{H}}$ are all tensor functors. Therefore the functors from $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{D}'}$ to $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{DR}}$, $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{Dol}}$, and $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{H}}$ are tensor functors. Recall that they were equivalences of categories. Therefore they have essential inverses which are tensor functors [43]. On the other hand, the inclusion $\mathscr{C}^s_{\mathrm{DR}} \to \mathscr{C}_{\mathrm{DR}}$ is a tensor functor. Thus we get tensor functors from $\widehat{\mathscr{C}}^s_{\mathrm{DR}}$ to $\widehat{\mathscr{C}}_{\mathrm{DR}}$ and from $\mathscr{C}_{\mathrm{DR}}$ to $\widehat{\mathscr{C}}_{\mathrm{DR}}$. These are both equivalences. Thus there is a tensor equivalence between $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{DR}}$ and $\mathrm{E}^0\,\mathscr{C}_{\mathrm{DR}}$. Similarly there is a tensor equivalence between $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{Dol}}$ and $\mathrm{E}^0\,\mathscr{C}_{\mathrm{Dol}}$. The tensor categories $\mathrm{E}^0\,\mathscr{C}_{\mathrm{DR}}$ and $\mathrm{E}^0\,\mathscr{C}_{\mathrm{Dol}}$ are just the categories of flat bundles, and of semistable Higgs bundles with vanishing Chern classes, with their respective tensor products. Thus the equivalence of categories constructed in Corollary 3.10 is an equivalence of tensor categories, as is the equivalence with the category $\mathrm{E}^0\,\widehat{\mathscr{C}}^s_{\mathrm{H}}$. This tensor category is formed from the data of the simple objects, their tensor products, the Ext groups, and the cup products.

The equivalences are compatible with pullbacks, and this extends to the tensor structure also. Suppose $f \colon \mathrm{Y} \to \mathrm{X}$ is a morphism of smooth projective varieties. If $\mathscr{C}(\mathrm{X})$ denotes any of the categories of objects on X and $\mathscr{C}(\mathrm{Y})$ denotes the corresponding category of objects on Y, then the pullback functors $f^* : \mathscr{C}(\mathrm{X}) \to \mathscr{C}(\mathrm{Y})$ are tensor functors. These pullback functors commute with all of the other functors we are considering, up to natural isomorphism compatible with the tensor structure. Thus the equivalences of tensor categories are natural with respect to pullbacks.

This allows us to define fiber functors. If $x \in \mathrm{X}$, then any of the above constructions (3.4.1)-(3.4.4), applied to the space $\{x\}$, yield canonically the d.g.t.c. $\mathscr{V}ect$. Therefore, the pullback morphism associated to the inclusion $\{x\} \hookrightarrow \mathrm{X}$ gives a functor $\omega : \mathscr{C}(\mathrm{X}) \to \mathscr{C}(x) = \mathscr{V}ect$. This will be a fiber functor, which we will refer to as the fiber functor of evaluation at x.

In example (3.4.5), if $\mathrm{A}^\bullet$ is a d.g.a. with augmentation $\omega : \mathrm{A}^0 \to \mathbf{C}$, then ω extends to a fiber functor on the d.g.t.c. $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}^\bullet)$, and hence to a fiber functor on the completion $\mathscr{C}^{\mathrm{nil}}(\mathrm{A}^\bullet)$.

4. Variations of Hodge structure

A *complex variation of Hodge structure* [24] [8] is a $\mathscr{C}^\infty$ vector bundle V with a decomposition $V = \bigoplus_{r+s=w} V^{r,s}$, a flat connection D satisfying Griffiths' transversality condition

$$D : V^{r,s} \to A^{0,1}(V^{r+1,s+1}) \oplus A^{1,0}(V^{r,s}) \oplus A^{0,1}(V^{r,s}) \oplus A^{1,0}(V^{r-1,s+1}),$$

and a parallel Hermitian form which makes the Hodge decomposition orthogonal and which on $V^{r,s}$ is positive definite if r is even and negative definite if r is odd. The monodromy representation of the flat connection D has image contained in a subgroup $U(p, q) \subset Gl(n, \mathbf{C})$. The representation is known to be semisimple [24]. By changing the sign of the polarization on alternate $V^{p,q}$, one obtains a harmonic metric K for the flat connection. Decompose $D = \overline{\theta} + \partial + \overline{\partial} + \theta$ according to the above transversality condition. Then set $D'' = \overline{\partial} + \theta$ and $D' = \partial + \overline{\theta}$. The operator D'' is the one in § 1 associated to the flat connection D using the metric K. The condition $(D'')^2 = 0$ may be obtained from $D^2 = 0$ by decomposing according to type of form in the base and type of coefficients. These operators D'' and D' were introduced by Deligne when he extended the Kähler identities to variations of Hodge structure. This is the motivation for the definitions made in § 1.

Corresponding to variations of Hodge structure, there is a type of algebraic structure for Higgs bundles. A *system of Hodge bundles* is a Higgs bundle (E, θ) with a decomposition of locally free sheaves $E = \bigoplus E^{r,s}$, such that $\theta : E^{r,s} \to E^{r-1,s+1} \otimes \Omega^1_X$. In the theorem about partial differential equations used to construct representations from Higgs bundles, the flat connections which come from variations of Hodge structure are those obtained from Higgs bundles which are systems of Hodge bundles [47]. Moreover, there is a one-to-one correspondence between the possible variations of Hodge structure on a given local system, and the structures of system of Hodge bundles on the corresponding Higgs bundle. We remark that the condition of stability for a system of Hodge bundles and the condition of stability for the resulting Higgs bundle are *a posteriori* the same. This can be seen from [47].

There is a natural action of $\mathbf{C}^*$ on the space of Higgs bundles: $t \in \mathbf{C}^*$ maps (E, θ) to $(E, t\theta)$. The action of $\mathbf{C}^*$ clearly preserves the conditions of stability and vanishing Chern classes, so by Theorem 1 it gives an action on the set of semisimple representations of the fundamental group. This action also preserves the monodromy group $M(E, x)$, and hence the Zariski closure of the image of the monodromy representation. Furthermore, Corollary 2.9 implies that on the space of semisimple representations the action of $t \in \mathbf{C}^*$ is continuous in t.

If X is a smooth projective variety, then we get an algebraic action of $\mathbf{C}^*$ on the moduli space $\mathbf{M}_{\mathrm{Dol}}$ (or $\mathbf{M}_{\mathrm{Dol}}(G)$), following from the construction of the moduli space. In that case, the homeomorphism between $\mathbf{M}_{\mathrm{Dol}}$ and $\mathbf{M}_{\mathrm{B}}$ shows that the action is continuous in both variables, t and the representation.

Hitchin originally considered this action in the form of an action of the circle $U(1) \subset \mathbf{C}^*$ [30]. The analytic picture is simplified in case of the action of $t \in U(1)$, for the harmonic metric is then left unchanged: if K is a harmonic metric for (E, θ) then K will still be a harmonic metric for $(E, t\theta)$. The operator D' changes from $\partial + \bar{\theta}$ to $\partial + \bar{t}\bar{\theta}$, but the curvature condition $F = 0$ still holds because $t\bar{t} = 1$.

If a Higgs bundle E has a structure of system of Hodge bundles, then it is a fixed point of $\mathbf{C}^*$. The automorphism of E obtained by multiplication by t^r on $E^{r,s}$ gives an isomorphism between (E, θ) and $(E, t\theta)$. The converse also holds:

Lemma **4.1.** — *If* $(E, \theta) \cong (E, t\theta)$ *for some* $t \in \mathbf{C}^*$ *which is not a root of unity, then* E *has a structure of system of Hodge bundles. If* E *is stable then this structure is unique up to translation of indices.*

Proof. — Let $f: E \to E$ be the holomorphic automorphism such that $f\theta = t\theta f$. The coefficients of the characteristic polynomial of f are holomorphic functions on X, hence constant, so the eigenvalues are constant. This gives a decomposition $E = \bigoplus_\lambda E_\lambda$ where $E_\lambda = \ker(f - \lambda)^n$. Since f is an isomorphism, $\lambda \neq 0$. Now $(f - t\lambda)^n \theta = t^n \theta(f - \lambda)^n$ so θ maps the λ generalized eigenspace E_λ to the $t\lambda$ generalized eigenspace $E_{t\lambda}$. The fact that t is not a root of unity insures that the eigenvalues break up into strings of the form $\lambda, t\lambda, \ldots, t^r\lambda$ where $t^{-1}\lambda$ and $t^{r+1}\lambda$ are not eigenvalues. Combining the eigenspaces for these eigenvalues in the appropriate fashion yields the decomposition of E. □

Corollary **4.2.** — *The representations of* $\pi_1(X)$ *which come from complex variations of Hodge structure are exactly the semisimple ones which are fixed by the action of* $\mathbf{C}^*$. □

Corollary **4.3.** — *If* X *and* Y *are compact Kähler manifolds and* $f: Y \to X$ *is a map such that the induced map* $f_*: \pi_1(Y) \to \pi_1(X)$ *is surjective, and if* V *is a flat bundle on* X *such that* $f^* V$ *comes from a variation of Hodge structure on* Y, *then* V *comes from a variation of Hodge structure on* X.

Proof. — The correspondence between flat and Higgs bundles commutes with f^*, so the action of $\mathbf{C}^*$ commutes with f^*. If the map on fundamental groups in surjective and $f^* V \cong f^* V_t$ then $V \cong V_t$. □

This corollary is particularly applicable if Y is a complete intersection of hyperplane sections. Thus if Y is such a curve on X, then any local system which restricts to a variation of Hodge structure on Y comes from a variation of Hodge structure on X. And if Y is a hyperplane surface on X, then the variations of Hodge structure on X and on Y are the same. These are Lefshetz-type theorems for variations.

Remark. — In § 3 we extended our correspondence, and hence the $\mathbf{C}^*$ action, from semisimple local systems to all local systems. There is an extension of the notion of variation of Hodge structure to the nonsemisimple case, the notion of *complex variation of mixed Hodge structure.* I do not think that the obvious generalization of Lemma 4.1 will remain true; in other words there are probably local systems fixed by the action of $\mathbf{C}^*$ but which do not underly complex variations of mixed Hodge structure. The correct generalization probably should refer to a local system and its complex conjugate. It seems to be reasonable to conjecture that if a local system and its complex conjugate are both fixed by $\mathbf{C}^*$, then they should underly complex variations of mixed Hodge structure (very nonuniquely).

Real structures. — A real variation of Hodge structure is a complex variation V with flat real structure on the local system, $V_{\mathbf{R}} \subset V$, such that $V^{p,q} = \overline{V^{q,p}}$. A real polarization means a form $S : V_{\mathbf{R}} \otimes V_{\mathbf{R}} \to \mathbf{R}$, symmetric or antisymmetric depending on the weight $w = p + q$, such that $\langle u, v \rangle = S(u, \bar{v})$ (or $\sqrt{-1}$ times it) is a complex polarization. Such may be chosen by choosing any complex polarization $\langle u, v \rangle_0$, and setting $S(u, v) = \langle u, \bar{v} \rangle_0 \pm \langle v, \bar{u} \rangle_0$.

If the complexification of a real local system is the complex local system underlying a complex variation of Hodge structure, then the real local system underlies a real variation of Hodge structure (possibly after changing the Hodge decomposition—and if the representation is not irreducible it may be a sum of real variations of different weights). The proof will be given in a slightly more general context in the next section, Lemmas 5.5 and 5.6, or you can obtain a proof by following the discussion in the back of [10]. In particular, Corollary 4.3 holds for real variations.

GROUPS OF HODGE TYPE

We will now make some remarks on the monodromy groups of variations of Hodge structure. Let W be a real algebraic group and G the associated complex group. Let σ be complex conjugation in G relative to the real form W. Recall that a *Cartan involution* is an automorphism C of W such that $C^2 = 1$, and such that $\tau = \sigma C = C\sigma$ is the complex conjugation with respect to a compact real form U of G (i.e. a compact group which has a real point in every connected component of G).

Say that W is *of Hodge type* if there is an action of $\mathbf{C}^*$ on G such that $U(1)$ preserves W and such that the element $C = -1$ is a Cartan involution. (Compare with [26], [11] § 1, and [9]).

Say that a representation $\rho : \pi_1(X) \to H$ *comes from a variation of Hodge structure* if there is a faithful representation $H \hookrightarrow Gl(n, \mathbf{C})$ such that the composite representation of $\pi_1(X)$ in $Gl(n, \mathbf{C})$ comes from a complex variation of Hodge structure.

Suppose ρ is a representation with values in a complex algebraic group H. The complex Zariski closure G is the smallest complex algebraic subgroup containing the image. We may define the *real Zariski closure* to be the smallest real algebraic subgroup W containing the image of ρ, in the real algebraic group $\mathrm{res}_{\mathbf{C}/\mathbf{R}}\, H$ obtained by restricting scalars.

Lemma **4.4.** — *Suppose that ρ is a representation which comes from a variation of Hodge structure. Let* G *be the complex Zariski closure of the monodromy group, and let* W *be the real Zariski closure. Then* W *is a real form of* G *and* W *is a group of Hodge type.*

Proof. — We will give a concrete proof here; a more abstract version of the same proof will be given in §§ 5 and 6. We may assume that ρ is a representation into $\mathrm{Gl}(n, \mathbf{C})$, and the resulting local system E is a complex variation of Hodge structure. The Zariski closure G is equal to the monodromy group $\mathrm{M}(\mathrm{E}, x)$. Choose a polarization of the variation E, and let K be the associated harmonic metric for the Higgs bundle E. Let $\mathrm{U} = \mathrm{G} \cap \mathrm{U}(\mathrm{E}_x, \mathrm{K}_x)$. Let τ denote complex conjugation in $\mathrm{Gl}(\mathrm{E}_x)$ with respect to the metric K_x. We will prove that τ preserves G and that the group of fixed points U is a compact real form. Since G is reductive, it is equal to the group of elements fixing a subspace of tensors $\mathrm{S} \subset \mathrm{T}^{a,b}(\mathrm{E}_x) = \mathrm{E}_x^{\otimes a} \otimes (\mathrm{E}_x^*)^{\otimes b}$. Furthermore we may assume that S is the space of all tensors so fixed, hence S has a Hodge structure and there is a decomposition of systems of Hodge bundles

$$\mathrm{T}^{a,b}(\mathrm{E}) = \mathrm{S} \otimes \mathcal{O}_{\mathrm{X}} \oplus \mathrm{F}$$

with F not containing any trivial subobject. In particular, $\mathrm{G} = \mathrm{M}(\mathrm{E}, x)$ preserves the subspace F_x. Now the harmonic metric K on E induces a harmonic metric on the tensor product, and it follows that the direct sum $\mathrm{S} \otimes \mathcal{O}_{\mathrm{X}} \oplus \mathrm{F}$ is an orthogonal direct sum of bundles with harmonic metrics. Any element g of $\mathrm{Gl}(\mathrm{E}_x)$ may be written uniquely as a product $g = u.\exp(y)$ where $u \in \mathrm{U}(\mathrm{E}_x, \mathrm{K}_x)$, and y is in the Lie algebra $\mathrm{End}(\mathrm{E}_x)$ with $\tau(y) = -y$. This decomposition is compatible with tensor products and orthogonal direct sums. Thus if $g \in \mathrm{G}$, then the elements u and y preserve the decomposition $\mathrm{T}^{a,b}(\mathrm{E}_x) = \mathrm{S} \oplus \mathrm{F}_x$ (and act trivially on S), so $u \in \mathrm{G}$ and y is in the Lie algebra of G. Since $\tau(g) = u.\exp(-y)$, the conjugation τ preserves G. The set of fixed points U is compact and a real form of G. It meets every connected component because any element $g = u.\exp(y)$ may be joined to an element $u \in \mathrm{U}$ by a path of elements $u.\exp(ty)$ which remain in G. This proves that U is a compact real form of G.

Next notice that the subspace of tensors S is compatible with the decomposition into Hodge type. Therefore the action of $\mathbf{C}^*$ obtained by multiplication by t^r on $\mathrm{E}^{r,s}$ preserves S, so it normalizes G. Thus we get an action of $\mathbf{C}^*$ on G. The elements of $\mathrm{U}(1)$ preserve the polarization and the metric K, so they preserve U. The element C, which is the image of -1, yields a new complex conjugation $\sigma = \mathrm{C}\tau$ and hence a new real form W. By the fact that the polarization and the metric are related by a change of sign on alternate $\mathrm{E}^{r,s}$, the group W is the intersection of G with the group of elements preserving the polarization form. But the monodromy representation preserves the polarization, so the image of $\pi_1(\mathrm{X})$ is contained in W. As G is the complex Zariski closure, W must be the real Zariski closure. Finally, W is by construction a group of Hodge type. □

We will list several facts of a standard nature related to the question of whether a real algebraic group is of Hodge type. Suppose W is a real algebraic group, and G is its complexification.

4.4.1. *If* W *is of Hodge type, then it is reductive.*

This is because the complexification G has a compact real form. □

4.4.2. W *is of Hodge type if and only if it has a Cartan involution which is an inner automorphism, given by an element of the connected component of* G.

Since G is reductive, the map of $G^0 \to \mathrm{Aut}(G)^0$ is surjective. If W is of Hodge type, then the Cartan involution is in the connected component of Aut(G), so it comes from an element of G^0.

Suppose C in the connected component $\mathrm{Aut}(G)^0$ gives the Cartan involution for W. Let U be the compact real form defined by C. Then C is contained in the compact group $\mathrm{Aut}(U)^0$, and we can let T be a maximal torus in $\mathrm{Aut}(U)^0$ containing C. The equation $C^2 = 1$ implies that there is an algebraic one parameter subgroup $U(1) \to \mathrm{Aut}(U)$ such that -1 maps to C. This provides the structure of group of Hodge type. □

Remark. — If W is an algebraically connected group, then it has a Cartan involution, and any two Cartan involutions are conjugate by an inner automorphism [28].

4.4.3. *If* W *is of Hodge type, the center of* W *is compact.*

The condition that the Cartan involution is inner means in particular that C fixes the center. Therefore the center of W is the same as the center of the compact real form U. □

4.4.4. *If* W *is a group of Hodge type, then it has a compact maximal abelian subgroup.*

Lift the map $U(1) \to G^{\mathrm{ad}}$ to a map $U(1) \to G$ (possibly taking a cover of $U(1)$). It maps into the maximal compact subgroup U and the real subgroup W. Let T be a maximal torus in U containing the image of $U(1)$. Then T contains the lift of the Cartan involution, so T is fixed by the Cartan involution. Therefore T is contained in W. It is a maximal abelian subgroup because if $w \in W$ commutes with T, it is fixed by C so it is also in U, hence it is in T because T is maximal abelian in U. □

4.4.5. *Suppose* W *is an algebraically connected reductive real group, which contains a compact maximal abelian subgroup* T. *Then* W *is of Hodge type.*

Choose a Cartan involution for W. The fixed points form a maximal compact subgroup. Since all maximal compact subgroups are conjugate, we may assume that T is perserved by C. As T is a Cartan subgroup, we may look at the root system of W with respect to T to write

$$\mathbf{g} = \mathbf{t} \oplus \bigoplus_{\alpha} \mathbf{g}_\alpha$$

where the α are characters of T, and the spaces $\mathbf{g}_\alpha$ are one dimensional. The Cartan involution C preserves this root space decomposition, and $C^2 = 1$, so C acts by ± 1 on $\mathbf{g}_\alpha$. We claim that there is a linear function d from the root lattice to $\mathbf{Z}$, such that C acts by $(-1)^{d(\alpha)}$ on $\mathbf{g}_\alpha$. This is because a set of simple roots γ forms a basis for the root lattice. Set $d(\gamma)$ equal to 0 or 1 depending on whether C acts by 1 or -1 on $\mathbf{g}_\gamma$, and extend this to a linear function on the root lattice. Then since C acts trivially on $\mathbf{t}$, and $[\mathbf{g}_\gamma, \mathbf{g}_{-\gamma}] \neq 0$ in $\mathbf{t}$, C must act by $(-1)^{d(\gamma)} = (-1)^{d(-\gamma)}$ on $\mathbf{g}_{-\gamma}$. Now the vectors in $\mathbf{g}_\gamma$ and $\mathbf{g}_{-\gamma}$ generate $\mathbf{g}$, and in general if $[\mathbf{g}_\alpha, \mathbf{g}_\beta] = \mathbf{g}_{\alpha+\beta}$ then the action of C on $\mathbf{g}_{\alpha+\beta}$ is determined by the action on $\mathbf{g}_\alpha$ and $\mathbf{g}_\beta$. Therefore C acts by $(-1)^{d(\alpha)}$ on $\mathbf{g}_\alpha$. This linear function d is given by an imaginary element $\widetilde{c}$ of the Lie algebra $\mathbf{t}_\mathbf{C}$, and we can set $\widetilde{C} = \exp(\widetilde{c}\pi i)$. Conjugation by this element is the Cartan involution of the Lie algebra, and hence of the connected group W. □

4.4.6. *If* W *is of Hodge type, then the algebraic connected component is also of Hodge type. However, the converse may not be true.*

If the Cartan involution of the algebraic connected component is an inner automorphism, this need not remain true for the nonconnected group W. Consider the following example. Let W^0 be the group SU(3, 3), the group of matrices A with determinant one such that

$$A^t Q \bar{A} = Q,$$

where Q is the diagonal matrix diag$(1, 1, 1, -1, -1, -1)$. There is an automorphism m of W^0 given by $m(A) = \bar{A}$. Let W be the semidirect product $W = W^0 \times \{1, m\}$ with two components. The Cartan involution of W^0 is given by the element $C = Q$ in PSU(3, 3). However Q has determinant -1, so in order to lift C to an element $\widetilde{C}$ in SU(3, 3) we must choose a number a with $a^6 = -1$ and set

$$\widetilde{C} = \mathrm{diag}(a, a, a, -a, -a, -a).$$

Now $\widetilde{C}^2 = a^2 I$ is in the center of W^0. However, it is not in the center of W, as it does not commute with m. There is no way to find C in the connected component of Aut(W) such that $C^2 = 1$ in Aut(W). This example shows that a group W might not be of Hodge type, even though its connected component W^0 is of Hodge type. □

4.4.7. *A compact group is automatically of Hodge type.* □

4.4.8. *Any group which is a product of groups, is of Hodge type if and only if all of its components are.*

The Cartan involution of the product of groups is equal to the product of the Cartan involutions. It is inner if and only if all of the components are inner. □

4.4.9. *If two connected groups are isogenous, then one is of Hodge type if and only if the other is.*

The conditions that the Cartan involutions are inner are the same. □

7

4.4.10. *A complex group considered as a real group is never of Hodge type.*

If a complex group is considered as a real group, the Cartan involution interchanges the two factors of the complexification, so it cannot be an inner automorphism. □

S. Zucker points out that the condition of Hodge type for a connected real group is just the condition that the rank of the group is equal to the rank of the maximal compact subgroup (4.4.4 and 4.4.5 above). He directs me, for example, to Corollary 1.6 of [3]; an earlier discussion is in § 1 of [26].

One can interpret the condition that the Cartan involution is an inner automorphism in terms of the Dynkin diagram. The group of outer automorphisms modulo inner automorphisms is equal to the group of automorphisms of the Dynkin diagram, so a connected group is of Hodge type if and only if its Cartan involution fixes the Dynkin diagram. In particular, if a connected absolutely simple real group is not of Hodge type, it must be of type $\mathbf{a}_n$, $\mathbf{c}_n$, or $\mathbf{e}_6$ (as may also be verified by the list below). This comment was pointed out to me by Deligne.

We will divide the groups of Hodge type into two categories, those of Hermitian type, and those not of Hermitian type. A group is of *Hermitian type* if W/K is a Hermitian symmetric space, where K is the maximal compact subgroup. A group is of Hermitian type if and only if it has a structure of group of Hodge type such that the weights of $\mathbf{C}^*$ acting on the Lie algebra are $-1, 0, 1$. See for example ([11] § 1.5).

Now we may give a list of which connected simple groups are of Hodge type, by consulting Helgason's book [28]. The classification is due to E. Cartan, and Helgason's treatment is based on results of Kac. The noncompact real forms of simple complex Lie algebras are listed in Tables II and III on pages 514-515. As a consequence of Theorem 5.16 in that section (one checks that if the Cartan involution is inner in the complex group, then it is inner in the corresponding real form), the Cartan involution is inner for the entries in the first column of Table II (labeled $k = 1$) and for the entries of Table III. The groups which are not of Hodge type are those in the second column of Table II (labeled $k = 2$). The groups in Table III are those of Hermitian type. Thus we compile the following lists:

Simple groups which are not of Hodge type:

$\mathrm{Sl}(n, \mathbf{R})$ $n \geqslant 3$	Any complex group
$\mathrm{SU}^*(2n)$ $n \geqslant 3$	$\mathbf{e}_{6(6)}$
$\mathrm{SO}(p, q)$ p, q odd	$\mathbf{e}_{6(-26)}$

Groups which are of Hermitian type:

$\mathrm{SU}(p, q)$	$\mathrm{S}p(n, \mathbf{R})$
$\mathrm{SO}^*(2n)$ $n \geqslant 3$	$\mathbf{e}_{6(-14)}$
$\mathrm{SO}(p, 2)$	$\mathbf{e}_{7(-25)}$

Groups which are of Hodge type but not Hermitian type:

$\mathrm{SO}(p, 2q)$ $q \geqslant 2$	$\mathbf{e}_{8(8)}$
$\mathrm{Sp}(p, q)$	$\mathbf{e}_{8(-24)}$
$\mathbf{e}_{6(2)}$	$\mathbf{f}_{4(4)}$
$\mathbf{e}_{7(7)}$	$\mathbf{f}_{4(-20)}$
$\mathbf{e}_{7(-5)}$	$\mathbf{g}_{2(2)}$

RIGID REPRESENTATIONS

A representation of the fundamental group into an algebraic group G is called *rigid* if every nearby representation is conjugate to it. In other words, the set theoretic orbit of the representation in the representation space $\mathrm{Hom}(\pi_1(X), G)$ under the action of G is an open subset. If the Zariski closure of the monodromy group is not equal to G, then there is a slightly weaker notion: a reductive representation is *properly rigid* if it is rigid as a representation into the Zariski closure of the image.

Lemma **4.5.** — *Any properly rigid reductive representation comes from a complex variation of Hodge structure. Consequently, the real Zariski closure* W *of the image of the monodromy representation is a real form of the complex Zariski closure, and* W *is a group of Hodge type.*

Proof. — Replace G by the Zariski closure of the image, and fix a faithful linear representation of G. Let E be the corresponding flat bundle, which has a harmonic metric by Theorem 1. Choose a sequence $t_i \to 1$ in $\mathbf{C}^*$, and assume that the t_i are not roots of unity. Then the representations corresponding to $(E, t_i\,\theta)$ converge to the representation corresponding to (E, θ) (by Corollary 2.9 or the remark below). Furthermore, the monodromy groups of these representations are all the same as the monodromy group of E, so these representations are contained in G. The representation (E, θ) is rigid as a representation in G, so for some i, the representation $(E, t_i\,\theta)$ is conjugate to the representation (E, θ). Hence the Higgs bundles are isomorphic. Now Lemma 4.1 implies that (E, θ) comes from a variation of Hodge structure. □

Remark. — We may choose the numbers t_i in the unit circle. In this case, a harmonic metric K for (E, θ) is also a harmonic metric for $(E, t_i\,\theta)$. It is therefore easy to see that the representations $(E, t_i\,\theta)$ approach the representation (E, θ). Furthermore, this obviates the need to refer to the second part of Theorem 1.

Corlette [5] has given many other results showing that under certain conditions a flat bundle must be of Hodge type. The conditions are on the structure group G and the rank of the harmonic map into G/K. The latter he verifies by the nonvanishing of certain characteristic classes. Corlette has also proved a result which is partially converse to the result about rigid representations. He proves that certain variations of Hodge structure which have nonvanishing volume invariant, are rigid [6].

DEFORMATION TO A VARIATION OF HODGE STRUCTURE

In the algebraic case, there is a moduli space which, although not projective, has a proper map to a vector space. Using this we may strengthen the previous results about rigid representations as follows.

Theorem **3**. — *Suppose* X *is a smooth projective variety, and* G *is a reductive complex algebraic group. Any representation* $\rho : \pi_1(X) \to G$ *can be deformed to a representation which comes from a variation of Hodge structure.*

Proof. — First of all, any representation can be deformed to a reductive representation. This is easy to see in case $G = Gl(n, C)$: one puts the representation in block upper triangular form, and conjugates by a diagonal matrix so as to make the strictly upper triangular parts go to zero in the limit. The same type of argument works for any reductive G. Let $M \subset G$ denote the monodromy group, and assume it is not reductive. A theorem of Morozov provides a one-parameter subgroup $\mathbf{C}^* \to G$ such that $\lim_{t \to 0} \mathrm{Ad}(t)\,(m)$ exists for $m \in M$, and such that this limit is the identity matrix for some $m \neq 1$. In particular, the dimension of the monodromy group is reduced in the limit. Continue until the monodromy group is reductive.

Now suppose we are given a reductive representation. Treat first the case $G = Gl(n, \mathbf{C})$. Using the previous paragraph and Corollary 4.2, we just have to prove that in any component of the moduli space $\mathbf{M}_{\mathrm{Dol}}$, there is a fixed point of the action of $\mathbf{C}^*$. Refer to the proper map $f : \mathbf{M}_{\mathrm{Dol}} \to C^k$ given by Proposition 1.4. If we try to take the limit of tE as $t \to 0$, then the corresponding characteristic polynomials $f(t\mathrm{E})$ approach the trivial one. In particular by the properness of f, there is a limiting E_0. The limit is unique, and hence is preserved by $\mathbf{C}^*$. If G is another group, choose a faithful representation $G \subset Gl(n, \mathbf{C})$, and refer to the discussion following Lemma 2.10. Note that the action of $\mathbf{C}^*$ preserves the monodromy group $M(E, x)$, and hence there is an action on the moduli space $\mathbf{M}_{\mathrm{Dol}}(G)$. The rest of the argument works equally well because the inclusion $G \subset Gl(n, \mathbf{C})$ induces a proper map $\mathbf{M}_{\mathrm{Dol}}(G) \to \mathbf{M}_{\mathrm{Dol}}(Gl(n, \mathbf{C}))$. This proves the theorem. □

Remark. — This theorem depends on the construction of the moduli space, and hence is currently valid only for smooth projective varieties. This strengthening of the earlier result about rigid representations is the motivation for the construction of the moduli spaces.

Corollary **4.6**. — *Let* W *be a real reductive algebraic group, with complex form* G. *Suppose* X *is a smooth projective variety and* $\rho : \pi_1(X) \to W$ *is a representation. Assume that there exists a subgroup* $\Gamma \subset \pi_1(X)$ *with* $\rho|_\Gamma$ *rigid and Zariski dense as a representation in* G. *Then* W *is a group of Hodge type.*

Proof. — Deform ρ in G to a variation of Hodge structure ρ'. Since Γ is reductive, the orbit of $\rho|_\Gamma$ is closed in the representation space $\mathrm{Hom}(\Gamma, G)$. Since Γ is rigid, the orbit is open. Thus the orbit is a connected component of the representation space of Γ. In particular, $\rho'|_\Gamma$ is isomorphic to $\rho|_\Gamma$. Thus W is contained in the Zariski closure of ρ'; but this Zariski closure is a real form of G, so it is equal to W and W is a group of Hodge type. □

NONEXISTENCE STATEMENTS FOR FUNDAMENTAL GROUPS

We will now review some corollaries of these facts, based on known examples of rigid representations.

Suppose W is a reductive real algebraic group with no compact factor. A *lattice* is a discrete subgroup $\Gamma \subset W$ such that W/Γ has finite volume. A lattice Γ is called *uniform* if W/Γ is compact, and *non-uniform* otherwise; and Γ is *irreducible* if for any projection onto a factor $f: W \to H$, $f(\Gamma)$ is not discrete.

Let G be the complex form of W. Any lattice $\Gamma \subset W$ is Zariski dense in G. Say that Γ is *rigid* if the inclusion $\Gamma \to G$ is a rigid representation.

The following facts are known.

4.7.1. If Γ is uniform and irreducible, and if W is not isomorphic to $\mathrm{Sl}(2, \mathbf{R})$, then Γ is rigid [54].

4.7.2. If W has real rank $\geqslant 2$ and Γ is an irreducible lattice, then Γ is arithmetic and rigid [35] [42]. For example, $\mathrm{Sl}(n, \mathbf{Z})$ is rigid for $n \geqslant 3$.

4.7.3. If Γ is an arithmetic lattice in a group of $\mathbf{Q}$-rank one, if W has no component isogenous to $\mathrm{Sl}(2, \mathbf{R})$, and if W is not isogenous to $\mathrm{Sl}(2, \mathbf{C})$, then Γ is rigid [42].

4.7.4. If Γ is a nonarithmetic lattice in a simple group of $\mathbf{R}$-rank one, and if W is not isogenous to $\mathrm{Sl}(2, \mathbf{R})$ or $\mathrm{Sl}(2, \mathbf{C})$, then Γ is rigid [20].

Lemma **4.7.** — *If* W *is not of Hodge type, and if* $\Gamma \subset W$ *is a lattice covered by one of the above instances, then* Γ *is rigid. Hence* Γ *cannot be the fundamental group of a compact Kähler manifold.*

Proof. — Otherwise Lemma 4.5 would provide a contradiction. □

If W is a group of Hermitian type, then $K\backslash W$ is a hermitian symmetric space. If $\Gamma \subset W$ is a uniform lattice, then $K\backslash W/\Gamma$ is a smooth projective variety with fundamental group Γ. D. Toledo has pointed out that in many cases when Γ is a non-uniform lattice, there is still a compact two-dimensional hyperplane section of $K\backslash W/\Gamma$ which has fundamental group Γ.

This leaves open the question of whether the other cases not ruled out above can occur. For example, if W is a Hodge group but not of hermitian type, like $SO(p, 2q)$, can a lattice Γ be the fundamental group of a compact Kähler manifold? There are no obvious ways of producing compact varieties which have these fundamental groups, so it is an interesting question whether they exist. J. Carlson and D. Toledo considered this question and have some partial results [4]. They have proved that compact quotients $K\backslash SO(p, q)/\Gamma$ are not homotopic to compact Kähler manifolds, and also that if $\Gamma \subset SO(n, 1)$ is discrete and co-compact, then Γ is not the fundamental group of a compact Kähler manifold. This latter result is stronger than ours when n is even, because in that case $SO(n, 1)$ is a Hodge group.

These corollaries can provide examples of compact complex manifolds which are not homotopic the Kähler manifolds. If G is a complex group and $\Gamma \subset G$ is a uniform lattice, then G/Γ is a compact complex manifold; Γ is rigid and G is not of Hodge type, so Γ cannot be the fundamental group of a compact Kähler manifold. In particular, G/Γ is not homotopic to a Kähler manifold.

For smooth projective varieties, we can extend the non-existence results to some groups which contain lattices. This paragraph uses the result about deformation to a variation of Hodge structure, and hence depends on the construction of the moduli space. Let W be a reductive real algebraic group which is not of Hodge type. Let Υ be a group with a representation in W. Suppose Υ contains a subgroup Γ which maps isomorphically to a lattice covered by statements 4.7.1-4.7.4. Then Υ cannot be the fundamental group of a smooth projective variety. This is an application of Corollary 4.6. A consequence is that if Γ is a lattice covered by statements 4.7.1-4.7.4, in a group which is not of Hodge type, then the fundamental group of a smooth projective variety cannot contain Γ as a split quotient. For example it cannot be a free or direct product of Γ with any other group.

RIGID ℓ-ADIC REPRESENTATIONS

Suppose X is a variety over a field K, with a K-valued point. Then we have an exact sequence of fundamental groups

$$1 \to \pi_1^{\mathrm{alg}}(X \otimes_K \overline{K}) \to \pi_1^{\mathrm{alg}}(X) \to \mathrm{Gal}(\overline{K}/K) \to 1.$$

Here π_1^{alg} is the profinite algebraic fundamental group of a scheme. The K-valued point gives a splitting, so $\mathrm{Gal}(\overline{K}/K)$ acts on $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$. In particular $\mathrm{Gal}(\overline{K}/K)$ acts on the space of continuous ℓ-adic representations of $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$. The fixed points of this action are the smooth ℓ-adic sheaves which are defined over K. These are the ones which have some relationship with arithmetic. Any representation coming from geometry (i.e. from a family of varieties over X) would be defined over a number field in this way. Beilinson pointed out to me that the action of $\mathbf{C}^*$ on the moduli space of complex representations is analogous to this action of the Galois group on the set of ℓ-adic representations in the arithmetic case.

As a demonstration of this philosophy, we obtain a result about rigid ℓ-adic representations which is analogous to the corollaries about rigid complex representations. Make the assumption that there is a subgroup $\Gamma \subset \pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$ such that the fundamental group is equal to the profinite completion of Γ. This is true, for example, if X is of finite type over a field K of characteristic zero: choose $K \subset \mathbf{C}$, and take $\mathbf{C}$ to be the image of the usual fundamental group of the complex variety $X \otimes_K \mathbf{C}$.

Let $\mathcal{O}_\lambda$ be the ring of integers in a finite extension E_λ of $\mathbf{Q}_\ell$. The continuous representations of $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$ in $\mathrm{Gl}(n, \mathcal{O}_\lambda)$ are the same as the $\mathcal{O}_\lambda$ points of the representation space of Γ. Say that a continuous representation ρ of $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$ is *rigid* if the representation restricted to Γ is rigid in the usual sense, as a representation with coefficients in $\overline{\mathbf{Q}}_\ell$.

Theorem **4.** — *Suppose* ρ *is an absolutely irreducible rigid representation of* $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$ *with coefficients in* $\mathcal{O}_\lambda$. *Then there exists a finite extension* L *of* K *and a representation* $\rho' : \pi_1^{\mathrm{alg}}(X \otimes_K L) \to \mathrm{PGl}(n, \mathcal{O}_\lambda)$ *which restricts to* ρ *on* $\pi_1^{\mathrm{alg}}(X \otimes_K \overline{K})$.

Proof. — We work in the profinite topology of the Galois group, and the λ-adic topologies of varieties over $\mathcal{O}_\lambda$. If $\gamma \in \Gamma$, and g is an element near the identity in $\mathrm{Gal}(\overline{K}/K)$, then $g\gamma g^{-1}$ is near γ. If ρ is a continuous representation, this means that $\rho(g\gamma g^{-1})$ is a matrix near $\rho(\gamma)$. The space of representations of Γ in $\mathrm{Gl}(n, \mathcal{O}_\lambda)$ is the space of $\mathcal{O}_\lambda$-valued points of the scheme $\mathbf{Rep}(\Gamma)$ of representations of Γ over $\mathbf{Z}$. Choosing a finite set of generators for Γ, we get an embedding of $\mathbf{Rep}(\Gamma)$ into a product of copies of the group scheme $\mathrm{Gl}(n)$. We have seen that if ρ is a point of $\mathbf{Rep}(\Gamma)$ $(\mathcal{O}_\lambda)$ and g is near the identity in $\mathrm{Gal}(\overline{K}/K)$, then $\rho \circ \mathrm{Ad}(g)$ is near ρ in the topology of the product of copies of $\mathrm{Gl}(n, \mathcal{O}_\lambda)$. On the other hand, the condition that the representation ρ is rigid and absolutely irreducible means that the orbit of ρ under conjugation by $\mathrm{Gl}(n, \overline{\mathbf{Q}})$ is a connected component of $\mathbf{Rep}(\Gamma)$ $(\overline{\mathbf{Q}})$. But two points which are sufficiently close in the λ-adic topology must lie in the same connected component. Therefore $\rho \circ \mathrm{Ad}(g)$ is conjugate to ρ for g sufficiently near the identity. By going to the finite extension L, we may assume that $\rho \circ \mathrm{Ad}(g)$ is conjugate to ρ for all g. Since ρ is irreducible, the element $P(g)$ such that $\rho \circ \mathrm{Ad}(g) = \mathrm{Ad}(P(g)) \circ \rho$ is uniquely determined up to scalars, in other words it is an element of $\mathrm{PGl}(n, E_\lambda)$ (the rationality over E_λ is due to the uniqueness). We need to prove that P is continuous.

Let Z be the reduced irreducible scheme over $\mathcal{O}_\lambda$ which is the closure of the orbit of ρ. We have a map $f: \mathrm{PGl}(n) \otimes E_\lambda \to Z \otimes E_\lambda$ which is an isomorphism. Let $\varphi : Z \otimes E_\lambda \to \mathrm{PGl}(n) \otimes E_\lambda$ be the inverse. Embed $\mathrm{PGl}(n)$ in an affine space $\mathbf{A}^N_{\mathcal{O}_\lambda}$. Then the map φ is given by a vector of regular functions $(a_1, \ldots, a_N)$ on $Z \otimes E_\lambda$. Write $a_i = \lambda^{-\alpha} b_i$ with b_i regular functions on Z. In particular, $b_i(\rho) = \lambda^\alpha u_i$ for $u_i \in \mathcal{O}_\lambda$. If $\rho' \sim \rho$ modulo (λ^β) then $b_i(\rho') \sim b_i(\rho)$ modulo (λ^β), so $a_i(\rho') \sim a_i(\rho)$ modulo $(\lambda^{\beta-\alpha})$. In other words, $\varphi(\rho') \sim \varphi(\rho)$ modulo $(\lambda^{\beta-\alpha})$. This proves that the inverse φ is continuous in the λ-adic topology. Therefore $P(g) = \varphi(\rho \circ \mathrm{Ad}(g))$ is continuous in g. If g is sufficiently near the identity, then $P(g)$ is in $\mathrm{PGl}(n, \mathcal{O}_\lambda)$. Combining this representation P of $\mathrm{Gal}(\overline{K}/K)$ with the representation ρ, we get a representation of all of $\pi_1^{\mathrm{alg}}(X \otimes_K L)$ in $\mathrm{PGl}(n, \mathcal{O}_\lambda)$, to complete the proof. $\square$

One way of obtaining a rigid representation is to take a rigid representation of Γ with coefficients in some finite extension E of $\mathbf{Q}$. Then for all primes ℓ except finitely many, this representation will have coefficients integral at ℓ. Take such an ℓ, and then by completion we get a rigid representation of Γ in $\mathrm{Gl}(n, \mathcal{O}_\lambda)$. This gives a continuous representation of $\pi_1^{\mathrm{alg}}(\mathrm{X} \otimes_{\mathrm{K}} \overline{\mathrm{K}})$. By the theorem, this representation descends to the arithmetic fundamental group over a finite extension L of K. This is the arithmetic analogue of the statement that the representation comes from a complex variation of Hodge structure. Theorem 4 may be viewed as proving a version of the " Galois-type " conjecture in [48] for rigid representations, in the same way that Lemma 4.5 gives the " variation of Hodge structure " conjecture of [48] for rigid representations.

$\mathbf{Q}$-STRUCTURE

Recall that a $\mathbf{Q}$-variation of Hodge structure is a complex variation of Hodge structure with a local system $\mathrm{V}_{\mathbf{Q}}$ contained in the local system $\mathrm{V}_{\mathbf{C}}$ of flat sections, such that $\mathrm{V}_{\mathbf{C}} = \mathrm{V}_{\mathbf{Q}} \otimes_{\mathbf{Q}} \mathbf{C}$, and subject to the following additional conditions. First, $\mathrm{V}^{p,q} = \overline{\mathrm{V}^{q,p}}$ where the complex conjugation is taken with respect to the real structure $\mathrm{V}_{\mathbf{Q}} \otimes_{\mathbf{Q}} \mathbf{R}$. Second, the polarization should be defined over $\mathbf{Q}$, in the sense that there is a symmetric or antisymmetric bilinear form $\mathrm{S} : \mathrm{V}_{\mathbf{Q}} \times \mathrm{V}_{\mathbf{Q}} \to \mathbf{Q}$, such that the hermitian form $\langle u, v \rangle = \mathrm{S}(u, \bar{v})$ polarizes the complex variation of Hodge structure.

A $\mathbf{Q}$-variation of Hodge structure may be tensored with $\mathbf{C}$ to obtain a complex variation of Hodge structure, and this may then be broken down into irreducible components. We call these components the complex direct factors of the $\mathbf{Q}$-variation.

Theorem **5.** — *Suppose* $\rho : \pi_1(\mathrm{X}) \to \mathrm{Gl}(n, \mathbf{C})$ *is a properly rigid irreducible representation. Then* ρ *is a complex direct factor of a* $\mathbf{Q}$*-variation of Hodge structure.*

Proof. — The idea is that if ρ is rigid, then it can be defined over a number field and all of its Galois conjugates are rigid. Adding these together gives a representation with coefficients in $\mathbf{Q}$. By Lemma 4.5, all of the summands come from complex variations of Hodge structure. By choosing everything carefully, we get a rational variation.

Suppose $\rho : \pi_1(\mathrm{X}) \to \mathrm{Gl}(n, \mathbf{C})$ is a properly rigid irreducible representation. Then, possibly after conjugating ρ, there is a number field $\mathrm{K}/\mathbf{Q}$ such that $\rho : \pi_1(\mathrm{X}) \to \mathrm{Gl}(n, \mathrm{K})$. Let $\mathrm{L} \subset \mathrm{K}$ be the field generated over $\mathbf{Q}$ by the traces $\mathrm{Tr}(\rho(\gamma))$.

For every embedding $\sigma : \mathrm{K} \to \mathbf{C}$ the resulting representation $\sigma\rho$ is properly rigid and hence comes from a variation of Hodge structure. For each embedding σ, let $\bar{\sigma}$ denote the complex conjugate embedding. If V is a complex variation of Hodge structure, then the flat hermitian polarization form

$$\mathrm{V} \otimes \overline{\mathrm{V}} \to \mathbf{C}$$

provides an isomorphism of representations $\overline{\mathrm{V}} \cong \mathrm{V}^*$. Since $\sigma\rho$ comes from a variation of Hodge structure, we get $\bar{\sigma}\rho \cong \sigma\rho^*$. In particular, $\sigma^{-1}\bar{\sigma}\rho \cong \rho^*$. Thus

$$\sigma^{-1}\bar{\sigma}\mathrm{Tr}(\rho(\gamma)) = \mathrm{Tr}(\rho^*(\gamma))$$

is independent of σ. Note that the trace of the dual of a matrix is equal to the trace of the inverse, so the $\mathrm{Tr}(\rho^*(\gamma))$ are in the field L. Hence there is a well-defined automorphism C of L which is equal to $\sigma^{-1}\bar{\sigma}$, in other words induces complex conjugation, for every embedding $\sigma : \mathrm{L} \to \mathbf{C}$. Thus L is either totally real (in the case $\mathrm{C} = 1$) or a purely imaginary extension of a totally real field F. These cases are distinguished by whether $\rho \cong \rho^*$ or not.

M. Larsen explained the following lemma to me.

Lemma **4.8.** — *There is an extension* L′ *of* L *which is the composite of* L *(or a purely imaginary extension thereof, if* L *is totally real) and a totally real extension* F′ *of* F, *such that after conjugation we may assume* ρ *has coefficients in* L′.

Proof. — There is a central simple algebra A over L with the property that ρ can be realized in a field L′ if and only if $\mathrm{A} \otimes \mathrm{L}' = \mathrm{M}_n(\mathrm{L}')$—see [2], [33]. We analyze what happens to the element A in the Brauer group, upon making a field extension. There are finitely many primes λ of L where A does not split. Let N be a number divisible by all of these primes. Suppose L″ is an extension such that $\mathrm{A} \otimes_{\mathrm{L}} \mathrm{L}''$ splits. Let f be a monic polynomial over $\mathbf{Z}$ such that L″ is contained in the splitting field of f. Let g be a polynomial of degree one smaller than f, having all real roots. Set $h = f + \mathrm{N}^k g$ for large k. Note that $h \sim f$ modulo N^k, but h has all real roots. Let L′ be the splitting field over L of h, or if L is totally real, let L′ be a purely imaginary quadratic extension of the splitting field. It is a field of the desired type. The local extensions L'_λ contain the L''_λ for λ dividing N; thus $\mathrm{A} \otimes \mathrm{L}'$ splits at all primes of L′. Since L′ has no real place at infinity, $\mathrm{A} \otimes \mathrm{L}'$ splits at the infinite places too. Therefore $\mathrm{A} \otimes \mathrm{L}'$ splits globally ([55] XI-2). □

Now assume that we have chosen L′ totally imaginary and a realization $\rho : \pi_1(\mathrm{X}) \to \mathrm{Gl}(n, \mathrm{L}')$. Let V be the corresponding local system of L′-vector spaces on X. We may choose a Hermitian form

$$\langle \ , \ \rangle : \mathrm{V} \times \mathrm{V} \to \mathrm{L}';$$

in other words $\langle v, w \rangle = \mathrm{C}\langle w, v \rangle$, and $a\langle v, w \rangle = \langle av, w \rangle = \langle v, (\mathrm{C}a)\, w \rangle$.

Let $\mathrm{F}' \subset \mathrm{L}'$ be the totally real subfield. Set $\mathrm{W}_{\mathrm{F}'} = \mathrm{res}_{\mathrm{L}'/\mathrm{F}'}\mathrm{V}$, so $\mathrm{W}_{\mathrm{L}'} = \mathrm{V} \oplus \mathrm{CV}$ with complex conjugation $(u, \mathrm{C}v) \mapsto (v, \mathrm{C}u)$. We get a form $\mathrm{S} : \mathrm{W}_{\mathrm{F}'} \times \mathrm{W}_{\mathrm{F}'} \to \mathrm{F}'$ as follows:

$$\mathrm{S}(u + \mathrm{C}u, v + \mathrm{C}v) = \langle u, \mathrm{C}v \rangle + \langle v, \mathrm{C}u \rangle$$

(resp. $$j\mathrm{S}(u + \mathrm{C}u, v + \mathrm{C}v) = \langle u, \mathrm{C}v \rangle - \langle v, \mathrm{C}u \rangle$$

where j is a purely imaginary number). Thus S is symmetric (resp. antisymmetric).

For each embedding $\sigma : \mathrm{L}' \to \mathbf{C}$ we get a local system V_σ, and $\mathrm{V}_{\bar\sigma} = \overline{\mathrm{V}}_\sigma$. For each of half of the embeddings σ, choose a complex variation of Hodge structure $\mathrm{V}_\sigma \otimes \mathrm{C}^\infty = \bigoplus \mathrm{V}_\sigma^{p,q}$. Then for the other half of the embeddings $\bar\sigma$, set $\mathrm{V}_{\bar\sigma}^{p,q} = \overline{\mathrm{V}}_\sigma^{q,p}$. Now for any embedding $\tau : \mathrm{F} \to \mathbf{R}$ we get a real variation of Hodge structure $\mathrm{W}_\tau = \mathrm{W}_{\mathrm{F}'} \otimes \mathbf{R}$, with $\mathrm{W}_\tau^{p,q} = \mathrm{V}_\sigma^{p,q} + \mathrm{V}_{\bar\sigma}^{p,q}$ where σ and $\bar\sigma$ are the two extensions of τ to L′.

Choose the Hodge types (p, q) for the components to all have the same weight w. If w is even, choose the form S to be symmetric; if w is odd, choose S antisymmetric.

By the way, if the rank of the representation is 2, it is either unitary or else we may choose the Hodge types to be $(1, 0)$ and $(0, 1)$. To deal with the unitary factors, we can add together two copies of everything, so if the rank is two, we can obtain a Hodge decomposition with Hodge types $(1, 0)$ and $(0, 1)$.

The hermitian form $\langle\ ,\ \rangle$ gives a form on V_σ which is a real multiple of a polarization. This is because the complex local system V_σ is irreducible, so there is only one flat hermitian form up to scalars, and the scalars must be real by the hermitian condition. Thus the form S_τ induced by S is either a polarization or the negative of a polarization for the real variation of Hodge structure W_τ. Cure this problem by multiplying S by a number $f \in F'$. We have to choose f to have a specified sign in each embedding τ. This is possible because F' spans $\bigoplus_\tau \mathbf{R}$, so there is a point which is in the specified quadrant. Thus we may assume that S_τ is a polarization for every τ.

Now let $U_{\mathbf{Q}} = \mathrm{res}_{F'/\mathbf{Q}} W_{F'}$. Then

$$U_{\mathbf{Q}} \otimes \mathbf{R} = \bigoplus_{\tau : F' \to \mathbf{R}} W_\tau.$$

Thus $U_{\mathbf{Q}}$ is a real variation of Hodge structure, polarized by the form ΣS_τ. This form is equal to the form

$$\mathrm{Tr} \circ S : U_{\mathbf{Q}} \times U_{\mathbf{Q}} \to \mathbf{Q}.$$

Thus U is a polarized $\mathbf{Q}$-variation of Hodge structure and our original representation ρ was isomorphic to one of the irreducible summands of $U \otimes \mathbf{C}$. This completes the proof. $\square$

If the traces $\mathrm{Tr}(\rho(\gamma))$ are all algebraic integers, then ρ preserves a lattice $\Lambda \subset V$ [2]. In this case, $\mathrm{res}_{\mathcal{O}_L/\mathbf{Z}}(\Lambda)$ is an integer lattice $U_{\mathbf{Z}} \subset U_{\mathbf{Q}}$ preserved by the monodromy, in other words U is an integral variation of Hodge structure. Note that our previous result about rigid ℓ-adic representations applies at every prime ℓ. So in this case, the representation ρ has most of the attributes of a motive. This is in fact the case if the dimension is two:

Corollary **4.9.** — *If ρ is a rigid two-dimensional irreducible representation of $\pi_1(X)$ such that the traces of all the elements are algebraic integers, then ρ is a direct summand of the monodromy representation of a family of abelian varieties.* $\square$

Proof. — If the dimension is two, then the Hodge types can be chosen as $(1, 0)$ and $(0, 1)$. An integer variation of Hodge structure U with those Hodge types comes from a family of abelian varieties. $\square$

Corollary **4.10.** — *If $\pi_1(X)$ does not have an infinite cyclic quotient and does not decompose as a nontrivial amalgamated product, then every two-dimensional irreducible representation is a direct factor of the monodromy representation of a family of abelian varieties.*

Proof. — This is a consequence of the theory of Bass and Serre on groups acting on trees. If $\pi_1(X)$ has no infinite cyclic quotient and does not decompose as a nontrivial amalgamated product, then every two dimensional representation is rigid, and furthermore has the property that the traces are integers [2]. Now apply the previous corollary. □

5. The Hodge structure on the fundamental group

In this section we will indicate how the Hodge theory for nonabelian cohomology developed above may be interpreted as Hodge theory for the fundamental group of X. In order to sharpen the exposition, we will defer the main proofs until the next section. The reader should bear in mind that the theorems follow in an essentially formal way from the information in the previous sections.

The Hodge structure on the first homology group $H^1(X, \mathbf{Z})$ is actually a structure on the vector space $H^1(X, \mathbf{C})$, which one may think of as being the *universal complex vector space* containing the abelian group $H^1(X, \mathbf{Z})$. Similarly, in order to speak of a Hodge structure for the fundamental group, we must consider the *universal complex pro-algebraic group* containing $\pi_1(X, x)$. This will be denoted by $\varpi_1(X, x)$. It is a projective limit of complex algebraic groups. More precisely it is the *pro-algebraic completion* of $\pi_1(X, x)$, defined by

$$\varpi_1(X, x) = \varprojlim (G, \rho)$$

where the inverse limit runs over the directed system of representations $\rho : \pi_1(X, x) \to G$ for complex algebraic groups G, such that the image of ρ is assumed (for convenience) to be Zariski dense in G. An arrow $(G, \rho) \to (G', \rho')$ in this directed system consists of a homomorphism $f: G \to G'$ such that $f\rho = \rho'$.

The completion $\varpi_1(X, x)$ is characterized by the pro-universal property that for any representation $\rho : \pi_1(X, x) \to G$ into a complex algebraic group, there is a unique extension to $\rho : \varpi_1(X, x) \to G$ making the diagram

$$\begin{tikzcd} \pi_1(X, x) \arrow[r] \arrow[d] & G \\ \varpi_1(X, x) \arrow[ur] & \end{tikzcd}$$

commute. The reader might refer to several papers by Hochschild on the topic of pro-affine groups ([31] and papers in subsequent issues of the same journal; and [32]).

We will consider some quotients of $\varpi_1(X, x)$. The *reductive quotient* $\varpi_1^{\mathrm{red}}(X, x)$ is the inverse limit over Zariski dense representations $\rho : \pi_1(X, x) \to G$ where G is a reductive complex algebraic group. The *nilpotent quotient* $\varpi_1^{\mathrm{nil}}(X, x)$ is the inverse limit over representations where G is a nilpotent group.

The group $\varpi_1(X, x)$ may be given several topologies, corresponding to the various topologies on complex algebraic groups. The *projective limit topology* is the projective limit of the discrete topologies on the algebraic groups G. The neighborhoods of the identity are simply the inverse images $\rho^{-1}(1)$ for representations $\rho : \varpi_1(X, x) \to G$ into algebraic groups. The *pro-algebraic topology* is the inverse limit of the Zariski topologies on the algebraic groups G, whereas the *pro-analytic topology* is the inverse limit of the usual analytic topologies on the complex algebraic groups G. The neighborhoods of the identity are sets of the form $\rho^{-1}(W)$ for representations ρ into algebraic groups G, and neighborhoods W of the identity in G (in the Zariski or analytic topology respectively).

Our basic structure for the fundamental group is an action of $\mathbf{C}^*$ on the completion $\varpi_1(X, x)$. In order to state his properly, we must recall that there is an action of $\mathbf{C}^*$ on the set of representations $\rho : \pi_1(X, x) \to Gl(n, \mathbf{C})$. A representation is equivalent to a local system, an object V in $\mathscr{E}_{DR}$, together with an isomorphism $\beta : V_x \cong \mathbf{C}^n$. The equivalence of categories of Theorem 1, Lemma 3.5, and Corollaries 1.3 and 3.10 identifies the set of representations (V, β) with the set of (E, θ, β), where (E, θ) denotes a semistable Higgs bundle with vanishing Chern classes, and $\beta : E_x \cong \mathbf{C}^n$. The abstract discrete group $\mathbf{C}^*$ acts on this set by the prescription

$$t(E, \theta, \beta) = (E, t\theta, \beta).$$

Theorem **6.** — *There is a unique action of the discrete group* $\mathbf{C}^*$ *on* $\varpi_1(X, x)$, *each* t *acting by a homomorphism of pro-algebraic groups, such that if* $\rho : \varpi_1(X, x) \to Gl(n, \mathbf{C})$ *is the representation corresponding to* (E, θ, β), *then* $\rho \circ t$ *is the representation corresponding to* $(E, t\theta, \beta)$.

The idea to be conveyed by this chapter is that the action of $\mathbf{C}^*$ on $\varpi_1(X, x)$ constitutes the data of the *Hodge structure on the fundamental group*. We will support this contention in two ways, related to the reductive and nilpotent quotients. First we will formulate the notion of a pure Hodge structure for a pro-reductive group, modeled on the notion of group of Hodge type defined in the previous section. We will show that the action of $U(1) \subset \mathbf{C}^*$ provides such a Hodge structure for the reductive quotient $\varpi_1^{red}(X, x)$. The restrictions on the fundamental group of a compact Kähler manifold obtained in the previous section may also be derived from the fact that $\varpi_1^{red}(X, x)$ has a pure Hodge structure. Second, we will briefly recall that Morgan, Hain, and Deligne have already shown $\varpi_1^{nil}(X, x)$ to carry a natural mixed Hodge structure. Our $\mathbf{C}^*$ action splits the Hodge filtration they have defined. Thus their mixed Hodge structure may be recovered by taking the filtration associated to the grading given by the $\mathbf{C}^*$-action. Doing these two things will not finish the formulation of the notion of Hodge structure for $\varpi_1(X, x)$. The first is analogous in the abelian case to formulating the notion of a pure Hodge structure, while the second is analogous to formulating the notion of unipotent mixed Hodge structure. But the structure of $\varpi_1(X, x)$ should combine pure and mixed Hodge structures, in other words it should be analogous to a mixed Hodge structure where the quotients are nontrivial pure Hodge structures. This problem will be left for the future.

A *pure non-abelian Hodge structure* is an affine group scheme (pro-algebraic group) $\mathscr{G}$ defined over $\mathbf{R}$, a finitely generated subgroup $\Gamma \subset \mathscr{G}_{\mathbf{R}}$, and an action of the abstract discrete group $\mathrm{U}(1)$ on $\mathscr{G}_{\mathbf{R}}$ by homomorphisms of the pro-algebraic group, subject to the following axioms which will be elaborated below.

1. Γ is Zariski dense in $\mathscr{G}$;
2. The map $\mathrm{U}(1) \times \Gamma \to \mathscr{G}^{\mathrm{an}}$ is continuous;
3. The element $\mathrm{C} = -1$ in $\mathrm{U}(1)$ is a Cartan involution of $\mathscr{G}_{\mathbf{R}}$.

The first condition means that for any surjective representation $\mathscr{G} \to \mathrm{H} \to 1$, the image of Γ is Zariski dense in H. In the second condition, $\mathrm{U}(1)$ is given the usual analytic topology, Γ is given the discrete topology, and $\mathscr{G}^{\mathrm{an}}$ refers to the group $\mathscr{G}$ endowed with its pro-analytic topology. The third condition needs more explanation. Let σ denote the antilinear involution of $\mathscr{G}$ given by complex conjugation with respect to the real form $\mathscr{G}_{\mathbf{R}}$. Then $\tau = \mathrm{C}\sigma = \sigma\mathrm{C}$ is an antilinear involution whose fixed points are a real form $\mathscr{G}^{\tau}$. The Cartan condition is that $\mathscr{G}^{\tau}$ is a compact real form of $\mathscr{G}$. By this we mean that $\mathscr{G}^{\tau}$ is compact in the pro-analytic topology, and that every path-connected component of $\mathscr{G}$ (in the pro-analytic topology) contains a point of $\mathscr{G}^{\tau}$. In particular, for any algebraic quotient $\mathscr{G} \to \mathrm{H} \to 1$, the image of $\mathscr{G}^{\tau}$ is a compact real form of H, meeting every connected component of H.

As a consequence, the group $\mathscr{G}$ is pro-reductive, an inverse limit of reductive groups. Indeed, any quotient admits a compact real form. Furthermore, since Γ is Zariski dense in $\mathscr{G}$, the real structure of $\mathscr{G}$ is determined: if σ' were another real structure fixing Γ, then the homomorphism $\sigma\sigma'$ would fix Γ and so it would be the identity. Thus we need not specify $\mathscr{G}_{\mathbf{R}}$ when speaking of a nonabelian Hodge structure.

Theorem 7. — *Suppose* X *is a compact Kähler manifold. The action of* $\mathrm{U}(1) \subset \mathbf{C}^*$ on $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ *gives a pure nonabelian Hodge structure, with finitely generated subgroup being the image of* $\pi_1(\mathrm{X}, x)$.

This amounts to two statements: that the map

$$\mathrm{U}(1) \times \pi_1(\mathrm{X}, x) \to \varpi_1^{\mathrm{red}}(\mathrm{X}, x)^{\mathrm{an}}$$

is continuous; and that the element $\mathrm{C} = -1$ in $\mathrm{U}(1)$ acts as a Cartan involution for $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}$, in other words, $\mathrm{C}\sigma = \sigma\mathrm{C} = \tau$ is an involution giving a compact real form $\mathscr{U}(\mathrm{X}, x) \overset{\mathrm{def}}{=} \varpi_1^{\mathrm{red}}(\mathrm{X}, x)^{\tau}$.

One might ask whether $\pi_1(\mathrm{X}, x)$ injects into $\varpi_1(\mathrm{X}, x)$. The answer turns out to be negative: D. Toledo has recently constructed an example of a smooth projective variety whose fundamental group is not residually finite, in which case $\pi_1(\mathrm{X}, x) \to \varpi_1(\mathrm{X}, x)$ is not injective.

We can now reinterpret the results of the previous section in terms of the abstract framework of a nonabelian Hodge structure $\Gamma \subset \mathscr{G}$.

Lemma **5.1.** — *If* G *is an algebraic group with a pure Hodge structure, then the real form* $G_{\mathbf{R}}$ *is a group of Hodge type in the sense of the previous section.*

Proof. — The only point which is not obvious is that the action of U(1) comes from an algebraic action of $\mathbf{C}^*$. We have $\Gamma \subset G$, Zariski dense, such that $U(1) \times \Gamma \to G$ is continuous. The group Aut(G) embeds in the affine variety Hom(Γ, G) as a closed subvariety. Its connected components are orbits under the adjoint action of G; the fact that Γ is Zariski dense means that these orbits are stable in the sense of geometric invariant theory, hence are closed. Now the action of U(1) gives a homomorphism $U(1) \to \mathrm{Aut}(G)$, but by hypothesis the composition $U(1) \to \mathrm{Hom}(\Gamma, G)$ is continuous. Therefore $U(1) \to \mathrm{Aut}(G)$ is a continuous group homomorphism, so it extends to a one-parameter subgroup $\mathbf{C}^* \to \mathrm{Aut}(G)$.

Conversely if $G_{\mathbf{R}}$ is a group of Hodge type, then it is easy to choose a finitely generated Zariski dense subgroup $\Gamma \subset G_{\mathbf{R}}$, to obtain a pure Hodge structure. □

The restrictions obtained in the last section, on which groups can be of Hodge type, become restrictions on which groups can have pure Hodge structures.

Suppose that $\Gamma \subset \mathcal{G}$ is a pure nonabelian Hodge structure on a pro-reductive group, and that $\rho : \mathcal{G} \to H \to 1$ is a surjective representation to a complex algebraic group. If ker(ρ) is fixed by U(1), then we obtain an action of U(1) on H. Together with the finitely generated subgroup $\mathrm{im}(\Gamma) \subset H$, this gives a pure Hodge structure for H: the image of the compact real form $\mathcal{G}^\tau$ is a compact real form H^τ, and therefore by twisting, the image of the real form $\mathcal{G}_{\mathbf{R}}$ is a real form $H_{\mathbf{R}}$.

Recall that a representation $\rho : \Gamma \to H$ is said to be *rigid* if any morphism $\rho' : \Gamma \to H$, nearby in the space of representations, is conjugate by an element $h \in H$, $\rho'(\gamma) = h\rho(\gamma)h^{-1}$. We may now state the analogue of Lemma 4.5 for abstract nonabelian Hodge structures. The reader may compare with the number theoretic analogue, Theorem 4.

Lemma **5.2.** — *Suppose* $\Gamma \subset \mathcal{G}$ *is a pure nonabelian Hodge structure. If* $\rho : \mathcal{G} \to H \to 1$ *is a surjective representation onto a complex algebraic group such that* $\Gamma \to H$ *is rigid, then* ker(ρ) *is fixed by* U(1), *so* $\mathrm{im}(\Gamma) \subset H$ *is a pure Hodge structure.*

Proof. — The continuity of $U(1) \times \Gamma \to \mathcal{G}$ implies that the representation $\rho \circ t$ is near the representation ρ. Hence there is an element h_t in H such that $\rho(t\gamma) = h_t\,\rho(\gamma)\,h_t^{-1}$ for $\gamma \in \Gamma$. The fact that Γ is Zariski dense implies that $\rho(tg) = h_t\,\rho(g)\,h_t^{-1}$ for all $g \in \mathcal{G}$. Now if g is in ker(ρ) then $h_t\,\rho(g)\,h_t^{-1} = 1$ so tg is in ker(ρ). By the previous discussion, H has a pure Hodge structure. □

Corollary **5.3.** — *If* Γ *is one of the groups listed above Lemma* 4.7 *and if* $\mathcal{G}$ *is the reductive completion of* Γ, *then* $\mathcal{G}$ *does not have a pure Hodge structure.* □

Corollary **5.4.** — *If* $\pi_1(X, x) \to H$ *is a rigid Zariski dense representation into a reductive group, then* H *has a pure Hodge structure with image of* $\pi_1(X, x)$ *as the finitely generated subgroup.*

In particular, the image of $\pi_1(X, x)$ *is contained in a unique real form* $H_{\mathbf{R}}$, *and* $H_{\mathbf{R}}$ *is a group of Hodge type. If* Γ *is one of the groups listed in* 4.7, *then* Γ *cannot occur as* $\pi_1(X, x)$ *for any compact Kähler manifold* X. $\square$

A *Hodge representation of weight* w is a vector space V, a representation $\rho : \mathcal{G} \to \mathrm{Gl}(V)$, and a vector space decomposition $V = \bigoplus_{p+q=w} V^{p,q}$, such that if we let U(1) act on V by the prescription $t(v) = r^p v$ for $v \in V^{p,q}$, then this is compatible with the action on $\mathcal{G}$ in the sense that $\rho(t\gamma)\,(t(v)) = t(\rho(\gamma)\,(v))$. A *polarization* of a Hodge representation is an indefinite Hermitian form $\langle u, v \rangle$ on V, satisfying the following conditions: the image of $\mathcal{G}_{\mathbf{R}}$ should be contained in the subgroup of transformations which preserve the form $\langle u, v \rangle$, the decomposition $V = \bigoplus V^{p,q}$ should be orthogonal, and if the form is twisted in the usual way, the result should be a positive definite form $(u, u) = (-1)^p \langle u, u \rangle > 0$ for $u \in V^{p,q}$.

A *real Hodge representation* is a Hodge representation V with a real structure $V_{\mathbf{R}}$ such that $\rho : \mathcal{G}_{\mathbf{R}} \to \mathrm{Gl}(V_{\mathbf{R}})$, and such that $\overline{V}^{p,q} = V^{q,p}$. A *real polarization* is a bilinear form $S(u, v)$ defined over $\mathbf{R}$ such that $\langle u, v \rangle = (\sqrt{-1})^w S(u, \bar{v})$ is a polarization in the above sense. Here $w = p + q$ is the weight of the Hodge structure.

Lemma 5.5. — *Suppose* $\rho : \mathcal{G} \to \mathrm{Gl}(V)$ *is a representation. It has a structure of Hodge representation if and only if the kernel if fixed by* U(1). *This condition always holds if* $\mathcal{G}$ *is an algebraic group. If the representation is irreducible, then a structure of Hodge representation is unique up to shifting indices. If* $V = \bigoplus V^{p,q}$ *is a structure of Hodge representation, then a polarization always exists, and again is unique up to a scalar if* ρ *is irreducible. If* $\rho : \mathcal{G}_{\mathbf{R}} \to \mathrm{Gl}(V_{\mathbf{R}})$ *is an irreducible real representation such that* $\ker(\rho)$ *is fixed by* U(1), *then it has a structure of real Hodge representation, and a real polarization exists.*

Proof. — Factor ρ through a surjection $\mathcal{G} \to H \to 1$ where H is a reductive algebraic group. If the kernel of ρ is fixed by U(1), then U(1) acts on H, and we obtain a pure nonabelian Hodge structure on H (the image of $\mathcal{G}^\tau$ is a compact real form H^τ, and twisting by C we obtain a real form $H_{\mathbf{R}}$, the image of $\mathcal{G}_{\mathbf{R}}$; the image of Γ is the finitely generated subgroup). We will show in general that if $\rho : H \to \mathrm{Gl}(V)$ is a representation of an algebraic Hodge group, then it has a structure of Hodge representation. The connected component of Aut(H) is the adjoint group H^0/Z, where Z is the center of H. The action of U(1) lifts to a map $U(1)^\sim \to H$, where $U(1)^\sim \to U(1)$ is a covering with kernel contained in Z. Considering the characters of $U(1)^\sim$ acting on V, we obtain a decomposition $V = \bigoplus V^p$, for rational numbers p, such that if we let $t \in U(1)$ act on V^p by t^p, then the representation ρ intertwines this with the action on $\mathcal{G}$. We may assume that V is irreducible, so Z acts by a single character, and hence the indices differ by integers. We may translate indices to integers without affecting the intertwining condition, to obtain the desired Hodge decomposition of V. To construct the polarization, choose a positive definite H^τ-invariant form (u, v) on V, and set $\langle u, v \rangle = (-1)^p\,(u, v)$ for

$u, v \in V^{p,q}$. This works because the action of $U(1)^{\sim}$ factors through H^{τ}, consequently preserving the forms.

Suppose ρ is a real representation. It factors as $\mathscr{G}_{\mathbf{R}} \to H_{\mathbf{R}} \to Gl(V_{\mathbf{R}})$, where the first map is a surjection onto a real algebraic group $H_{\mathbf{R}}$. If $\ker(\rho)$ is fixed by $U(1)$, then $U(1)$ acts on $H_{\mathbf{R}}$, so again $U(1)$ maps to the connected component of the automorphism group $(H/Z)_{\mathbf{R}}$. Let $H'_{\mathbf{R}}$ be the subgroup of $H_{\mathbf{R}}$ on which all characters vanish. Then H is isogenous to $H' \times Z$, so $H'_{\mathbf{R}} \to (H/Z)_{\mathbf{R}}$ is a finite covering (at least over the identity component). Hence we may lift the action of $U(1)$ to a map $\varphi : U(1)^{\sim} \to H'_{\mathbf{R}} \subset H_{\mathbf{R}}$. This gives a decomposition $V = \bigoplus V^p$, where for $t \in U(1)$, $\rho\varphi(t)$ acts by t^p on V^p. Note that the p may be rational numbers. Now suppose $v \in V^p$. Then

$$\rho\varphi(t)\,\bar{v} = \overline{\rho\varphi(t)\,v} = \overline{(t^p\,v)} = t^{-p}\,\bar{v}.$$

Thus $\bar{v} \in V^{-p}$. If r is any rational number, then

$$V_{(r)} = \bigoplus_{p \in \mathbf{Z}} V^{p+r}$$

is a complex direct summand of the representation V, since as noted above, the p differ by integers in an irreducible complex representation. Furthermore $\overline{V}_{(r)} = V_{(-r)}$. Since V is irreducible over $\mathbf{R}$, there are three possibilities. If $V = V_{(0)}$ then we obtain a real Hodge representation of even weight. If $V = V_{(1/2)} = V_{(-1/2)}$, we obtain a real Hodge representation of odd weight. If $V = V_{(r)} \oplus V_{(-r)}$ with $2r \notin \mathbf{Z}$, then we may arrange the Hodge types of $V_{(r)}$ arbitrarily. These determine the Hodge types of the complex conjugate $V_{(-r)}$, so we may obtain a Hodge representation of any weight. To find a real polarization, use the usual trick of adding a complex polarization to its complex conjugate (as in Lemma 2.12). □

Say that a surjection $\mathscr{G} \to H \to 1$ *comes from a Hodge representation* if there exists a faithful representation V of H such that the composed representation $\mathscr{G} \to Gl(V)$ has a structure of Hodge representation. It follows from Lemma 5.5 that a surjection $\mathscr{G} \to H \to 1$ comes from a Hodge representation if and only if the kernel is fixed by $U(1)$, in which case H becomes a pure Hodge structure with subgroup the image of Γ.

Lemma **5.6.** — *In the case of the Hodge structure on $\varpi_1^{\mathrm{red}}(X, x)$, there is a natural one-to-one correspondence $V \leftrightarrow V_x$ between complex variations of Hodge structure on X and Hodge representations of $\varpi_1^{\mathrm{red}}(X, x)$. Furthermore this induces one-to-one correspondences between polarizations of the Hodge representation and polarizations of the complex variation of Hodge structure, and between real Hodge representations and real variations of Hodge structure.*

Proof. — If V is a complex variation of Hodge structure, with $\beta : V_x \cong \mathbf{C}^n$, let $\rho : \varpi_1^{\mathrm{red}}(X, x) \to Gl(n, \mathbf{C})$ be the corresponding representation. The corresponding Higgs bundle (E, θ, β) is a system of Hodge bundles, $E = \bigoplus E^{p,q}$. The resulting action of $U(1)$ on E intertwines θ with $t\theta$, so by Theorem 6, the action $\beta t_x \beta^{-1}$ on $\mathbf{C}^n$ maps ρ to $\rho.t$. Thus the decomposition $\beta(E_x^{p,q}) = \beta(V_x^{p,q})$ of $\mathbf{C}^n$ provides ρ with a structure of Hodge representation. Conversely, if ρ has a structure of Hodge representation, let (E, θ, β) be the corresponding Higgs bundle. The Hodge structure on $\mathbf{C}^n$ gives an isomorphism of representations between ρ and $\rho.t$, and hence an isomorphism of Higgs bundles $(E, \theta) \cong (E, t\theta)$, compatible with the action on $\mathbf{C}^n$ by the identifications $\beta : E_x \cong \mathbf{C}^n$. Decomposing E according to characters of U(1) (or using the argument of Lemma 4.1), E becomes a system of Hodge bundles $E = \bigoplus E^{p,q}$, and the corresponding local system V becomes a complex variation of Hodge structure.

If $\langle u, v \rangle$ is a polarization of a complex variation of Hodge structure V, then it polarizes the Hodge structure $\bigoplus V_x^{p,q}$. The form $\langle u, v \rangle$ is preserved by $\pi_1(X, x)$ and hence by its real Zariski closure $\varpi_1^{\mathrm{red}}(X, x)_{\mathbf{R}}$, so we obtain a polarization of the Hodge representation. Conversely, a polarization of the Hodge representation V_x is by definition fixed by $\varpi_1^{\mathrm{red}}(X, x)_{\mathbf{R}}$, and hence by the subgroup $\pi_1(X, x)$, so it extends to a locally constant hermitian form $\langle u, v \rangle$ on the local system V. Compare this with a known polarization $\langle u, v \rangle_0$ of the variation of Hodge structure. There is an endomorphism h, locally constant and self-adjoint with respect to the known polarization, such that $\langle u, v \rangle = \langle hu, v \rangle_0$. The condition that h is a polarization of V_x means that h_x is positive definite and of type (0, 0). Hence h is of type (0, 0) everywhere, due to the principle that a constant tensor which has pure type at one point, has the same pure type everywhere. Furthermore, h is positive definite everywhere, since its eigenvalues are constant. Thus a polarization of the Hodge representation extends to a polarization of the variation of Hodge structure.

Similarly, given a real structure $V_{\mathbf{R}}$ for the Hodge representation V_x, by definition preserved by $\varpi_1^{\mathrm{red}}(X, x)_{\mathbf{R}}$, it is preserved by the subgroup $\pi_1(X, x)$. Thus $V_{\mathbf{R}}$ extends to a local subsystem of the local system V. Choose a known complex polarization $\langle u, v \rangle$, The $V_{\mathbf{R}}$ is determined by a locally constant tensor $h \in V^* \otimes V^*$ by the rule $\langle u, \bar{v} \rangle = h(u, v)$. The condition $\overline{V^{p,q}} = V^{q,p}$ is equivalent to the condition that h have pure type $(-w, -w)$. Again, if this holds at one point, it holds everywhere. □

Let us turn now to the mixed Hodge structure on $\varpi_1^{\mathrm{nil}}(X, x)$. A nilpotent group is determined by its Lie algebra; Morgan and Hain construct a mixed Hodge structure on Lie $\varpi_1^{\mathrm{nil}}(X, x)$ [38] [27]. They construct a Hodge filtration F^p, decreasing and indexed by negative numbers p. It is compatible with the Lie bracket in the sense that $[F^p, F^q] \subset F^{p+q}$. The weight filtration is defined by the derived series : $W_0 = \mathrm{Lie}\, \varpi_1^{\mathrm{nil}}(X, x)$, and for $k < 0$, W_k is defined inductively to be $[\mathrm{Lie}\, \varpi_1^{\mathrm{nil}}(X, x), W_{k+1}]$. The Lie algebra Lie $\varpi_1^{\mathrm{nil}}(X, x)$ is complete with respect to the weight filtration. The Hodge filtration is a full filtration in the topological sense that $\bigcup F^p$ is dense. Morgan and Hain show that the Hodge filtration F, its complex conjugate $\bar{F}$ taken with respect to the real structure Lie $\varpi_1^{\mathrm{nil}}(X, x)_{\mathbf{R}}$, and the weight filtration W, define an **R**-*mixed Hodge structure* with weights concentrated in $k \leqslant 0$. Notice that the weight filtration and real structure are defined group theoretically, so the only analytic information involved is the Hodge filtration.

From Theorem 6, $\mathbf{C}^*$ acts on $\varpi_1^{\mathrm{nil}}(X, x)$, and hence on the Lie algebra.

Theorem **8.** — *Our action of* $\mathbf{C}^*$ *on* Lie $\varpi_1^{\mathrm{nil}}(\mathrm{X}, x)$ *induces the Hodge filtration of Morgan and Hain in the following sense. Let* $\mathrm{H}^r \subset$ Lie $\varpi_1^{\mathrm{nil}}(\mathrm{X}, x)$ *denote the subspace of elements h such that* $t(h) = t^r h$ *for* $t \in \mathbf{C}^*$. *Then*

$$\mathrm{F}^p = \bigoplus_{r \geqslant p} \mathrm{H}^r.$$

Consequently, the $\mathbf{C}^*$*-action determines the mixed Hodge structure.*

One should be careful that, unlike the reductive case, the action of $\mathrm{U}(1) \subset \mathbf{C}^*$ does not preserve the real form $\varpi_1^{\mathrm{nil}}(\mathrm{X}, x)_{\mathbf{R}}$. This reflects the fact that the mixed Hodge structure is not $\mathbf{R}$-split.

An aspect of this theory as further developed by Hain and Zucker is that unipotent variations of mixed Hodge structure on X are identified with mixed Hodge theoretic representations of Lie $\varpi_1^{\mathrm{nil}}(\mathrm{X}, x)$. This was the motivation for the analogous result in the reductive case, Lemma 5.6 above. Hain and Zucker also discuss the dependence on the base point, showing that as x varies the Lie $\varpi_1^{\mathrm{nil}}(\mathrm{X}, x)$ fit together into a pro-unipotent variation of mixed Hodge structure over X, which is in some sense the universal variation of mixed Hodge structure.

The same should be true for the reductive case. At the present I will formulate this as a problem for further study: there should be an appropriate notion of *pure variation of nonabelian Hodge structure*, such that if (X_t, x_t) varies in an analytic family indexed by a space T, then the $\varpi_1^{\mathrm{red}}(\mathrm{X}_t, x_t)$ fit together into a variation of nonabelian Hodge structure over T. One might then proceed to study degenerations of nonabelian Hodge structures and monodromy of $\pi_1(\mathrm{X}, x)$, following by analogy the abelian case.

In lieu of treating these topics, we will finish with a description of the classifying map for the dependence on the base point. It will map to a universal version of Corlette's harmonic maps. Let $\mathscr{K}(\mathrm{X}, x)$ be the subgroup of elements k in $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}$ which are fixed by the action of $\mathrm{U}(1)$. It is compact, since it is contained in the compact real form $\mathscr{U}(\mathrm{X}, x)$. The adjoint action of $\mathscr{K}(\mathrm{X}, x)$ on $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}$ fixes the nonabelian Hodge structure. Let $\widetilde{\mathrm{X}}$ denote the pointed universal cover of X determined by the base point x. A point in $\widetilde{\mathrm{X}}$ may be thought of as a pair (y, P) with $y \in \mathrm{X}$ and P a homotopy class of paths from x to y.

Theorem **9.** — *There is a natural* $\pi_1(\mathrm{X}, x)$*-equivariant map*

$$\Phi : \widetilde{\mathrm{X}} \to \varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}/\mathscr{K}(\mathrm{X}, x)$$

such that if (y, P) *is a point in* $\widetilde{\mathrm{X}}$, *then the pullback of the Hodge structure on* $\varpi_1^{\mathrm{red}}(\mathrm{X}, y)$ *by parallel transport along the path* P *is conjugate to the Hodge structure on* $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ *by the element* $\Phi(y, \mathrm{P})$.

If $\rho : \varpi_1^{\mathrm{red}}(\mathrm{X}, x) \to \mathrm{Gl}(n, \mathbf{C})$ *is a representation such that* $\rho(\mathscr{U}(\mathrm{X}, x))$ *preserves the standard metric on* $\mathbf{C}^n$, *then*

$$\rho \circ \Phi : \widetilde{\mathrm{X}} \to \mathrm{Gl}(n)/\mathrm{U}(n)$$

is the harmonic map referred to in Theorem 1.

The map Φ should be interpreted as the classifying map for the variation of non-abelian Hodge structure describing how $\varpi_1^{\mathrm{red}}(X, x)$ varies with the base point x. I have not yet made precise the notion of Griffiths transversality for Φ, but surely it is straightforward.

6. Tannakian considerations

In this final section we will show how the results of the first sections imply the theorems stated in the previous section. This is mostly formal, and we will adopt the mechanism of Tannaka duality as the most natural way of using the relevant information about tensor products. At the end of this section, we will discuss the notion of principal object or *torsor*, using the ideas of Tannaka duality.

TANNAKA DUALITY

An algebraic group may be recovered from its category of representations, and this process may be used to construct a group if given a category with the right properties.

Let us briefly recall the notion of neutral Tannakian category [50] [43] [13]. A *tensor category* is a category provided with a functorial binary operation denoted as tensor product of objects. An *associative and commutative* tensor category is a tensor category provided with additional natural isomorphisms expressing associativity and commutativity of the tensor product (we call these isomorphisms *constraints*, even though they really represent additional structure—for example the commutativity constraint is a natural isomorphism $U \otimes V \cong V \otimes U$). The associativity and commutativity constraints are required to satisfy some axioms. The simplest way to describe the axioms is to say that any natural automorphism of some object, canonically produced by composition of the constraints, should be the identity. In the future we call such axioms *canonical axioms*. A *unit* is an object $\mathbf{1}$ provided with natural isomorphisms $\mathbf{1} \otimes V \cong V$ satisfying canonical axioms. A functor F between associative and commutative tensor categories with unit is a functor provided with natural isomorphisms $F(U \otimes V) \cong F(U) \otimes F(V)$, satisfying canonical axioms.

A *neutral Tannakian category* is an associative and commutative tensor category with unit, which is *abelian*, which is *rigid* (in other words, duals exist), such that $\mathrm{End}(\mathbf{1}) = \mathbf{C}$, and which is provided with an *exact faithful fiber functor* ω to the tensor category of vector spaces. If G is an affine group scheme over $\mathbf{C}$, then the category $\mathrm{Rep}(G)$ of representations of G is a neutral Tannakian category. It has a natural fiber functor ω_G which sends a representation to the underlying vector space. The group G is recovered as the group $\mathrm{Aut}^{\otimes}(\omega_G, \mathrm{Rep}(G))$ of tensor automorphisms of the fiber functor. Conversely, given a neutral Tannakian category $(\mathscr{E}, \omega)$ set $G = \mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$. Then $\mathscr{E}$ is recovered as the category $\mathrm{Rep}(G)$.

We will describe the basic facts in some detail; the proofs may be found in [50] [43] [13]. Suppose $\mathscr{E}$ is a neutral Tannakian category, with fiber functor ω. Making no reference to the tensor structure, let the algebra of endomorphisms of the fiber functor be denoted by $\mathrm{End}(\omega, \mathscr{E})$. Its elements are collections $\{f_V\}$ with $f_V \in \mathrm{End}(\omega(V))$, such that for any morphism $a: V \to W$, $\omega(a) f_V = f_W\, \omega(a)$. For example, $f_{V\oplus W} = f_V \oplus f_W$. The algebra $\mathrm{End}(\omega, \mathscr{E})$ is a projective limit of finite-dimensional algebras (see below), so it is endowed with a projective limit topology. We will consider this algebra as the primary object of study.

The tensor structure gives $\mathrm{End}(\omega, \mathscr{E})$ a structure of continuous Hopf algebra (cf. [27]). The diagonal

$$\Delta^{\otimes} : \mathrm{End}(\omega, \mathscr{E}) \to \mathrm{End}(\omega, \mathscr{E}) \mathbin{\hat{\otimes}} \mathrm{End}(\omega, \mathscr{E})$$

is defined by $\Delta^{\otimes}\{f_V\} = \{f_{V\otimes W}\}$, where the completed tensor product on the right is viewed as a subset of collections $\{f_{V,W}\}$ with $f_{V,W} \in \mathrm{End}(\omega(V)) \otimes \mathrm{End}(\omega(W))$. The augmentation $\varepsilon : \mathrm{End}(\omega, \mathscr{E}) \to \mathbf{C}$ is given by $\varepsilon\{f_V\} = f_1$.

Let $\mathbf{aut}^{\otimes}(\omega, \mathscr{E})$ and $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ be the closed pro-algebraic subsets of $\mathrm{End}(\omega, \mathscr{E})$ defined by the conditions of respecting the tensor structure, in the Lie algebra or Lie group sense respectively. Thus $\mathbf{aut}^{\otimes}(\omega, \mathscr{E})$ consists of those elements $\{f_V\}$ of $\mathrm{End}(\omega, \mathscr{E})$ such that

$$f_1 = 0 \qquad f_{V\otimes W} = f_V \otimes 1 + 1 \otimes f_W,$$

whereas $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ consists of those elements $\{f_V\}$ such that

$$f_1 = 1 \qquad f_{V\otimes W} = f_V \otimes f_W.$$

The condition that duals exist in $\mathscr{E}$ implies that any element $\{f_V\}$ satisfying the conditions for inclusion in $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ will automatically consist entirely of automorphisms—there is no need to include a condition of invertibility.

These subsets may be defined using the co-algebra structure of $\mathrm{End}(\omega, \mathscr{E})$. The Lie algebra $\mathbf{aut}^{\otimes}(\omega, \mathscr{E})$ is just the set of Lie-like elements h, those with $\varepsilon(h) = 0$ and $\Delta^{\otimes}(h) = h \otimes 1 + 1 \otimes h$. Similarly, the group $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ is the set of group-like elements g, those with $\varepsilon(g) = 1$ and $\Delta^{\otimes}(g) = g \otimes g$.

The algebra $\mathrm{End}(\omega, \mathscr{E})$ has a structure of a projective limit of finite-dimensional algebras. If $S \subset \mathscr{E}$ is a subset of objects, let $\mathscr{E}_S$ denote the category generated by objects in S and their direct sums; let $\mathscr{E}_{S,\otimes}$ denote the bigger category generated also by tensor products. Then

$$\mathrm{End}(\omega, \mathscr{E}) = \varprojlim \mathrm{End}(\omega, \mathscr{E}_S),$$

the projective limit being indexed by the directed set of subsets S. Note that $\mathrm{End}(\omega, \mathscr{E}_S) \subset \prod_{V\in S} \mathrm{End}(\omega(V))$ is a finite dimensional algebra. In particular, we get a topology on $\mathrm{End}(\omega, \mathscr{E})$, the projective limit topology.

The subsets $\mathbf{aut}^{\otimes}(\omega, \mathscr{E})$ and $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ are projective limits of closed algebraic subsets inside the projective limit $\mathrm{End}(\omega, \mathscr{E})$. For example,

$$\mathrm{Aut}^{\otimes}(\omega, \mathscr{E}) = \varprojlim \mathrm{Aut}^{\otimes}(\omega, \mathscr{E}_{S,\otimes}),$$

and $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E}_{S,\otimes})$ is a closed algebraic subset of $\mathrm{End}(\omega, \mathscr{E}_S)$ (one might want to assume that S contains the duals of all its objects). In particular, these subsets are closed in the projective limit topology. Also, $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ has a structure of projective limit of algebraic varieties, hence it has a structure of *pro-algebraic affine group scheme*. The coordinate ring of $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ is the continuous dual of $\mathrm{End}(\omega, \mathscr{E})$ (so it is a direct limit of finite dimensional spaces) [13]. A *representation* of $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ into an algebraic group G will always mean a representation which factors through an algebraic representation of one of the algebraic quotients.

The pro-algebraic group $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ is the object which is dual to the Tannakian category in the duality statement of Tannaka-Grothendieck-Saavedra [50] [43] [13]. The theorems (which we state without proof) are as follows: the category $\mathscr{E}$ is isomorphic to the category of representations of $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$; conversely, if G is a pro-algebraic group, its category of representations is a neutral Tannakian category Rep(G) with a natural fiber functor ω_G, and $G = \mathrm{Aut}^{\otimes}(\omega_G, \mathrm{Rep}(G))$; and also $\mathbf{aut}^{\otimes}(\omega, \mathscr{E})$ is the Lie algebra of $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$.

The map from G to $\mathrm{Aut}^{\otimes}(\omega_G, \mathrm{Rep}(G))$ sends a group element $g \in G$ to the natural automorphism $\{f_V\}$ of ω_G defined by setting f_V equal to the action of g on the vector space $\omega_G(V)$ underlying the representation V. One can think of $\mathrm{Aut}^{\otimes}(\omega_G, \mathrm{Rep}(G))$ as being the group of " virtual group elements ", namely systems of transformations of representation spaces which behave like group elements with respect to natural transformations and tensor products. One half of the duality theorem says that for complex affine group schemes, all virtual group elements are actual group elements. In the case of a discrete group, the group of " virtual group elements " is the pro-algebraic completion (see Lemma 6.1).

The notion of monodromy group can be defined in the context of a Tannakian category. Suppose V is an object of a neutral Tannakian category $(\mathscr{E}, \omega)$. Then the *monodromy group* $M(V, \omega)$ of V at the fiber functor ω is defined to be the image of the map $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E}) \to \mathrm{Gl}(\omega(V))$. If $\mathscr{E}_{V,\otimes}$ denotes the subcategory of all objects in $\mathscr{E}$ which are subquotients of tensor products of V and V^*, then $M(V, \omega) = \mathrm{Aut}^{\otimes}(\omega, \mathscr{E}_{V,\otimes})$. The monodromy group may also be defined in the manner of § 2: $M(V, \omega)$ is the subgroup of those elements g in $\mathrm{Gl}(\omega(V))$ which preserve subspaces $\omega(U) \subset T^{a,b}\,\omega(V)$ whenever $U \subset T^{a,b}\,V$ is a subobject in $\mathscr{E}$. The object V is called *reductive* if $M(V, \omega)$ is a reductive group. This is equivalent to the condition that V is semisimple, a direct sum of simple objects.

PROOF OF THEOREM 6

Lemma **6.1.** — *Suppose* Γ *is a finitely generated group. The tensor category* $\mathrm{Rep}(\Gamma)$ *of representations of* Γ, *with its obvious fiber functor* ω, *is a neutral Tannakian category. The group* $\mathrm{Aut}^{\otimes}(\omega, \mathrm{Rep}(\Gamma))$ *is naturally isomorphic to the pro-algebraic completion* $\mathscr{G}$ *of* Γ. *If* $\mathrm{Rep}^s(\Gamma)$ *and* $\mathrm{Rep}^{\mathrm{nil}}(\Gamma)$ *denote the subcategories of semisimple and nilpotent representations, the quotients* $\mathrm{Aut}^{\otimes}(\omega, \mathrm{Rep}^s(\Gamma))$ *and* $\mathrm{Aut}^{\otimes}(\omega, \mathrm{Rep}^{\mathrm{nil}}(\Gamma))$ *are, respectively, the pro-reductive and pro-nilpotent quotients* $\mathscr{G}^{\mathrm{red}}$ *and* $\mathscr{G}^{\mathrm{nil}}$.

Proof. — The Tannakian conditions are easy to check, and the rest follows from the Tannaka duality theorem stated above. The details are left to the reader. □

The differential graded categories $\mathscr{C}$ in examples (3.4.1)-(3.4.5) in § 3 give rise to tensor categories $\mathscr{E} = \mathrm{E}^0 \mathscr{C}$, for example $\mathscr{E}_{\mathrm{DR}} = \mathrm{E}^0 \mathscr{C}_{\mathrm{DR}}$ is the category of local systems on X, and $\mathscr{E}_{\mathrm{Dol}} = \mathrm{E}^0 \mathscr{C}_{\mathrm{Dol}}$ is the category of semistable Higgs bundles with vanishing Chern classes. These have fiber functors ω_x, evaluation at the base point x.

The operation of taking the monodromy representation provides an equivalence of tensor categories between the category of local systems $\mathscr{E}_{\mathrm{DR}}$ and the category of representations $\mathrm{Rep}(\pi_1(\mathrm{X}, x))$. This equivalence is naturally compatible with the fiber functors, ω_x and the standard ω. The above lemma then implies that $(\mathscr{E}_{\mathrm{DR}}, \omega_x)$ is a neutral Tannakian category, and that $\mathrm{Aut}^{\otimes}(\omega_x, \mathscr{E}_{\mathrm{DR}}) = \varpi_1(\mathrm{X}, x)$ is the pro-algebraic completion of the fundamental group.

The categories $\mathscr{E}_{\mathrm{DR}}$, $\mathscr{E}_{\mathrm{Dol}}$, $\mathrm{E}^0 \widehat{\mathscr{C}}_{\mathrm{D}'}$, and $\mathrm{E}^0 \widehat{\mathscr{C}}_{\mathrm{H}}$ constructed in examples (3.4.1)-(3.4.4) in § 3 are all naturally equivalent tensor categories, and this natural equivalence preserves fiber functors ω_x. Therefore all are Tannakian, and for example $\varpi_1(\mathrm{X}, x) = \mathrm{Aut}^{\otimes}(\omega_x, \mathscr{E}_{\mathrm{Dol}})$. From now on, simply let $\mathscr{E}$ denote one of these mutually equivalent categories. The category of semisimple objects $\mathscr{E}^s$ is equivalent to the category of harmonic bundles on X, and the category of nilpotent objects $\mathscr{E}^{\mathrm{nil}}$ is equivalent to the category $\mathrm{E}^0 \mathscr{C}^{\mathrm{nil}}(\mathrm{A}^{\bullet})$ constructed from the algebra of forms on X in example (3.4.5) of § 3. By Lemma 6.1, the Tannaka duals of the categories $\mathscr{E}^s$ and $\mathscr{E}^{\mathrm{nil}}$ are, respectively, the pro-reductive completion $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ and the pro-nilpotent completion $\varpi^{\mathrm{nil}}(\mathrm{X}, x)$.

We should make precise the notion of a group acting on a category, say in the specific case which occurs below. An action of the discrete group $\mathbf{C}^*$ on the category $\mathscr{E}$ consists of a collection of functors $a_t : \mathscr{E} \to \mathscr{E}$ indexed by $t \in \mathbf{C}^*$, together with natural isomorphisms of functors $a_s\, a_t \cong a_{st}$, satisfying the canonical axioms that the two isomorphisms $a_r\, a_s\, a_t \cong a_{rst}$ are the same. In the case of an action by tensor functors, there are additional canonical axioms expressing compatibility between the constraints of the group action and the constraints of the tensor functors. We make the convention that such considerations are implicit in discussing group actions.

There is a natural action of the discrete group $\mathbf{C}^*$ on the category of Higgs bundles $\mathscr{E}_{\mathrm{Dol}}$, by tensor functors which preserve the fiber functor:

$$a_t : (\mathrm{E}, \theta) \mapsto (\mathrm{E}, t\theta).$$

These are tensor functors because $\theta_{\mathrm{E}\otimes\mathrm{F}} = \theta_{\mathrm{E}} \otimes 1 + 1 \otimes \theta_{\mathrm{F}}$, and multiplying by t preserves this linear relation. The fiber functor ω_x is preserved because the underlying bundle is fixed, $\omega_x(\mathrm{E}, t\theta) = \mathrm{E}_x = \omega_x(\mathrm{E}, \theta)$. By transport of structure, $\mathbf{C}^*$ acts on the algebra $\mathrm{End}(\omega_x, \mathscr{E}_{\mathrm{Dol}})$, the Lie algebra $\mathbf{aut}^{\otimes}(\omega_x, \mathscr{E}_{\mathrm{Dol}})$, and the group $\mathrm{Aut}^{\otimes}(\omega_x, \mathscr{E}_{\mathrm{Dol}})$. This gives the requisite action on $\varpi_1(\mathrm{X}, x)$. This completes the proof of Theorem 6.

Some more insight can be gained as follows. We can define an action of $\mathbf{C}^*$ on the differential graded category $\mathscr{C}_{\mathrm{Dol}}$. Suppose U and V are objects of $\mathscr{C}_{\mathrm{Dol}}$. Let $a_t\,\mathrm{U}$ and $a_t\,\mathrm{V}$ denote the same underlying bundles with new operators $\mathrm{D}_t'' = \bar\partial + t\theta$. These are the new Higgs bundles resulting from the action of $\mathbf{C}^*$ described previously. Define the map

$$a_t : \mathrm{Hom}^\bullet(\mathrm{U}, \mathrm{V}) \to \mathrm{Hom}^\bullet(t\mathrm{U}, t\mathrm{V})$$

by

$$a_t(f) = t^p f \quad \text{for } f \in \mathrm{A}^{p,q}(\mathrm{X}, \mathrm{Hom}(\mathrm{U}, \mathrm{V})) = \mathrm{A}^{p,q}(\mathrm{X}, \mathrm{Hom}(t\mathrm{U}, t\mathrm{V})).$$

Note that $\mathrm{D}_t''(a_t(f)) = t(\mathrm{D}''(f))$, so this defines a functor $a_t : \mathscr{C}_{\mathrm{Dol}} \to \mathscr{C}_{\mathrm{Dol}}$. Furthermore, a_t has a structure of tensor functor. The natural isomorphism

$$\tau_{\mathrm{U},\mathrm{V}} : a_t\,\mathrm{U} \otimes a_t\,\mathrm{V} \to a_t(\mathrm{U} \otimes \mathrm{V})$$

is given by the identity map on the underlying bundles. Note that $\theta_{\mathrm{U}\otimes\mathrm{V}} = \theta_{\mathrm{U}} \otimes 1 + 1 \otimes \theta_{\mathrm{V}}$, and $\theta_{a_t\mathrm{U}} = t\theta_{\mathrm{U}}$ and similarly for V, so $\theta_{a_t(\mathrm{U}\otimes\mathrm{V})} = t\theta_{\mathrm{U}\otimes\mathrm{V}} = \theta_{a_t\mathrm{U}\otimes a_t\mathrm{V}}$. We have an action of $\mathbf{C}^*$ by tensor functors on $\mathscr{C}_{\mathrm{Dol}}$ (the composition of tensor functors $a_s\,a_t$ is naturally isomorphic to a_{st}, with canonical axioms satisfied). This action is compatible with pullbacks, and it is trivial over a point, so the action preserves the fiber functor ω, evaluation at a point x. In other words there are natural isomorphisms $\omega(a_t\,\mathrm{U}) \cong \omega(\mathrm{U})$. These are again just the identity maps. The composition of these natural isomorphisms for s and t is equal to the natural isomorphism for st, so we have an action of $\mathbf{C}^*$ by " tensor functors which preserve the fiber functor ". The action of $\mathbf{C}^*$ preserves the subcategory $\mathscr{C}^s_{\mathrm{Dol}}$, and the action on the full category is obtained from this restricted action by the process of completion (the second paragraph of the proof of Lemma 3.1 shows that the required uniqueness is satisfied, for extending an action by functors to the completion of a d.g.c.). Be careful to note, however, that the quasi-equivalence between $\mathscr{C}^s_{\mathrm{Dol}}$ and $\mathrm{C}^s_{\mathrm{H}}$ may not be compatible with the $\mathbf{C}^*$-action, because it goes through the category $\mathscr{C}^s_{\mathrm{D}'}$. Compatibility will only hold for subcategories of objects (such as the trivial object) where D' is compatible with the $\mathbf{C}^*$-action. □

Caution. — The action of $\mathbf{C}^*$ on $\mathscr{C}_{\mathrm{Dol}}$ permutes the objects in an essential way. Thus the action on $\mathrm{End}(\omega, \mathrm{E}^0\,\mathscr{C}_{\mathrm{Dol}})$ will not be algebraic, but rather more like the action on a space such as $\prod_{t\in\mathbf{C}^*}\mathbf{C}$. This problem is most serious on the category of semisimple objects, which we discuss in the next section. But then in the following section, we will discuss nilpotent objects, and there it will be a consequence of our argument that the action of $\mathbf{C}^*$ on $\mathrm{End}(\omega, \mathrm{E}^0\,\mathscr{C}^{\mathrm{nil}}_{\mathrm{Dol}})$ will be pro-algebraic, a projective limit of algebraic actions.

REAL AND CARTAN STRUCTURES

A *real structure* for a Tannakian category is an antilinear tensor functor

$$\sigma : \mathscr{E} \to \mathscr{E}$$

(in other words $\sigma : \mathrm{Hom}(\mathrm{U}, \mathrm{V}) \to \mathrm{Hom}(\sigma(\mathrm{U}), \sigma(\mathrm{V}))$ is antilinear) satisfying the condition that σ be an involution, meaning that there is provided a natural isomorphism

$\mu : \sigma\sigma \cong 1$ satisfying the canonical axiom that the two maps $\sigma\sigma\sigma \cong \sigma$ are equal. A *real structure* for the fiber functor ω is a natural isomorphism of tensor functors

$$\varphi_\sigma : \omega\sigma \cong \iota\omega$$

such that the two resulting maps $\iota\omega\sigma \cong \omega$ are equal. Here we are using the notation ι for the canonical involution $\iota : \mathscr{V}ect \to \mathscr{V}ect$, $\iota(\mathbf{1}^n) = \mathbf{1}^n$ and $\iota(f) = \bar{f}$ for an $m \times n$ complex matrix $f : \mathbf{1}^n \to \mathbf{1}^m$.

Given $\mathbf{R}$-structures for a Tannakian category $\mathscr{E}$ and its fiber functor ω, there is an $\mathbf{R}$-structure for the pro-algebraic group $G = \mathrm{Aut}^\otimes(\omega, \mathscr{E})$, such that if V is an object of $\mathscr{E}$ corresponding to a representation of G, then $\sigma(V)$ corresponds to the complex conjugate representation. The real structure of G is given by an involution σ, defined by

$$\sigma(g)_V = \iota(g_{\sigma(V)}),$$

or more precisely

$$\sigma(g)_V = \iota(\varphi_{\sigma, V}\, g_{\sigma(V)}\, \varphi_{\sigma, V}^{-1}),$$

for an element $g = \{ g_V \}$ of $\mathrm{Aut}^\otimes(\omega, \mathscr{E})$. The real points of the group are then elements g such that $\sigma(g) = g$, in other words such that $g_{\sigma(V)} = \iota(g_V)$. More concretely this says that g should act in the complex conjugate representation, as the complex conjugate of the way it acts in the original representation.

Caution. — While it is true that all fiber functors defined over $\mathbf{C}$ are isomorphic, the real structures for a fiber functor are not all isomorphic. The real forms of the algebraic group $\mathrm{Aut}^\otimes(\omega, \mathscr{E})$ given by two different real structures for ω might not be isomorphic.

Say that a pair (τ, φ_τ) consisting of a real structure for $\mathscr{E}$ and a real structure for a fiber functor ω, is *compact* if the associated real form is a compact real form of $\mathrm{Aut}^\otimes(\omega, \mathscr{E})$, meaning that it is compact and it meets every path-connected component (in the pro-analytic topology).

Recall that a *Cartan involution* of a real algebraic group (G, σ) is an involution $C : G \to G$, with $C^2 = 1$, such that $\tau = C\sigma = \sigma C$ is an antilinear involution defining a compact real form of G. We can make some similar definitions in the dual Tannakian context.

An *involution* of a Tannakian category $\mathscr{E}$ is an exact $\mathbf{C}$-linear tensor functor $C : \mathscr{E} \to \mathscr{E}$ together with a natural isomorphism $CC \cong 1$ satisfying the canonical axiom that the two isomorphisms $CCC \cong C$ are equal. Say that C *commutes* with an antilinear involution σ if there is given a natural isomorphism $C\sigma \cong \sigma C$, such that the three isomorphisms $C\sigma C\sigma \cong 1$ are equal. If this is the case, then $\tau = C\sigma$ has a structure of antilinear involution, in other words it gives a new real structure.

A C-*structure* for a fiber functor ω is a natural tensor isomorphism $\varphi_C : \omega C \cong \omega$ such that the two isomorphisms $\omega CC \cong \omega$ are equal. As before, this leads to an involution C, now linear, of $\mathrm{Aut}^\otimes(\omega, \mathscr{E})$. Suppose that $\mathscr{E}$ has a real structure σ and a Cartan involution C, which commute. Suppose ω is a fiber functor. Say that a C-structure φ_C and a real structure φ_σ *commute* if the two maps $\omega\sigma C \to \iota\omega$ are equal (this refers to the structure of isomorphism $C\sigma \cong \sigma C$). If φ_C and φ_σ commute, then we may compose them to get a real structure φ_τ for ω relative to the real structure $\tau = C\sigma$ on $\mathscr{E}$. Say that (C, φ_C) is a *Cartan involution* for (σ, φ_σ) if C and σ commute, and φ_C and φ_σ commute, and if the resulting real structure (τ, φ_τ) is compact.

A *Cartan triple* for $(\mathscr{E}, \omega)$ is a triple of pairs $((C, \varphi_C), (\sigma, \varphi_\sigma), (\tau, \varphi_\tau))$ where (C, σ, τ) is a triple of involutions of $\mathscr{E}$, with C linear and σ, τ antilinear, and φ_C, φ_σ, and φ_τ are commuting structures for ω, with (τ, φ_τ) equal to the composition of (C, φ_C) and (σ, φ_σ), and (τ, φ_τ) compact. A *Cartan triple* for a group G is a triple (C, σ, τ) where C is an involution commuting with real structures σ and τ, and $\tau = C\sigma$ is a compact real form of G.

Lemma **6.2.** — *Cartan triples for* $(\mathscr{E}, \omega)$ *are in natural one to one correspondence with Cartan triples for* $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$.

Proof. — One has to check that two pairs commute if and only if the corresponding involutions of $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E})$ commute. □

The following lemma is the standard fact about Cartan structures.

Lemma **6.3** *(E. Cartan). — Suppose* G *is a reductive algebraic group. Suppose either that* C *is an involution, or that* σ *is an antilinear involution. Suppose* H_0 *is a compact real subgroup which is preserved by* C *or* σ. *Then there is a compact real form* τ *of* G *which is preserved by* C *or* σ, *and such that* H_0 *is contained in the compact group* G^τ. *Thus,* C *or* σ *may be extended to a Cartan structure* (C, σ, τ) *such that* $H_0 \subset G^\tau$. *The compact real form* τ *is unique up to conjugation by an element of* G^C *or* G^σ.

Proof. — The proof is well known—see, for example, [28]. We will discuss it briefly. Let ν denote either C or σ. For existence, choose a compact real form $U' = G^{\tau'}$ containing H_0. Set $N = \nu\tau'\nu\tau'$; it is semisimple with positive real eigenvalues in its action on the Hopf algebra $\mathrm{End}(\omega_G, \mathrm{Rep}(G))$. Setting $\tau = N^{1/4}\tau' N^{-1/4} = N^{1/2}\tau'$ gives the desired compact real form. For the uniqueness statement, suppose τ and τ' are two compact real forms preserved by ν, and let U and U′ denote the compact subgroups. They are conjugate, $gUg^{-1} = U'$ with $g \in G$ unique up to multiplying on the right by an element of U. From the Cartan decomposition $G = \exp(i\mathbf{u})\,U$ we may assume that $g = \exp(y)$ with $y \in i\mathbf{u}$. In particular, $\tau(g) = g^{-1}$. Apply the involution ν to the equation $gUg^{-1} = U'$, to find $\nu(g)\,U\nu(g)^{-1} = U'$; by the uniqueness of g this implies that $\nu(g) = gu$ with $u \in U$. Apply τ to the equation $\nu(g) = gu$ to obtain $\nu(g)^{-1} = g^{-1}u$ (since τ commutes with ν and $\tau(g) = g^{-1}$). Therefore $g^{-1}\nu(g) = g\nu(g)^{-1}$, so $\nu(g^2) = g^2$. Note that ν acts on the Lie algebra $\mathbf{u}$ and hence on $i\mathbf{u}$, so the last equation translates to $\exp(2\nu(y)) = \exp(2y)$. But the map $\exp : i\mathbf{u} \to G$ is injective, so $2\nu(y) = 2y$. Therefore $\nu(y) = y$ so $\nu(g) = g$, $g \in G^\nu$ as desired. □

Question. — Is this lemma true for reductive pro-algebraic groups? The main question is whether a compact real form exists. This will be true in our example $\varpi_1^{\mathrm{red}}(X, x)$, but I don't know if it is true in general.

PROOF OF THEOREM 7

The pro-algebraic fundamental group $\varpi_1(X, x)$ has a natural real structure. The involution σ is the unique involution which fixes the original group $\pi_1(X, x)$, or to put it another way, $\varpi_1(X, x)_{\mathbf{R}}$ is the real algebraic completion of $\pi_1(X, x)$. The action of σ on the category $\mathscr{E}$ of local systems is defined by $\sigma(V, D) = (\overline{V}, \overline{D})$.

We restrict our attention now to the reductive quotient of the fundamental group. The subcategory $\mathscr{E}^s$ of semisimple objects is isomorphic to the category of harmonic bundles, and $\varpi_1^{\mathrm{red}}(X, x)$ is the group associated to the Tannakian category $\mathscr{E}^s$ with fiber functor evaluation at x.

Lemma **6.4.** — *If* K *is a harmonic metric on a harmonic bundle* (V, D, D', D''), *then* $\overline{K}$ *is a harmonic metric on the complex conjugate bundle* $(\overline{V}, \overline{D}, \overline{D}'', \overline{D}')$. *If* $t \in U(1)$ *then* K *is a harmonic metric on the bundle* (V, D_t, D'_t, D''_t) *where* $D''_t = \overline{\partial} + t\theta$, $D'_t = \partial + t^{-1}\overline{\theta}$, *and* $D_t = D'_t + D''_t$.

Proof. — The operator $\overline{D}'$ is a Higgs operator on the complex conjugate bundle $\overline{V}$, and the operator related to it by the metric $\overline{K}$ is $\overline{D}''$. Since all of these operators are integrable and they are related by the metric, $\overline{K}$ is harmonic (cf. Lemma 2.11). Note that $\overline{t\theta} = \overline{t}\overline{\theta} = t^{-1}\overline{\theta}$. Therefore D'_t is the operator associated to D''_t and the metric. The curvature is $D'_t D''_t + D''_t D'_t = \partial\overline{\partial} + \overline{\partial}\partial + \theta\overline{\theta} + \overline{\theta}\theta + t\partial(\theta) + t^{-1}\partial(\overline{\theta})$. This is the same as the original curvature $D' D'' + D'' D'$ except that the components of type (p, q) have been multiplied by t^p. Thus, if the original curvature vanishes, then so does the new curvature, so K is a harmonic metric for the operator D''_t. □

Corollary **6.5.** — *The action of* $U(1) \subset \mathbf{C}^*$ *on* $\varpi_1^{\mathrm{red}}(X, x)$ *preserves the real form* $\varpi_1^{\mathrm{red}}(X, x)_{\mathbf{R}}$.

Proof. — The σ complex conjugate of (V, D'_t, D''_t) is $(\overline{V}, \overline{D}''_t, \overline{D}'_t)$. But for example

$$\overline{D}'_t = \overline{\partial + t^{-1}\overline{\theta}} = \partial_{\overline{V}} + t\theta_{\overline{V}} = (\overline{D}')_t.$$

Therefore the action of $U(1)$ commutes with the involution σ. The natural isomorphism expressing this commutativity is just the identity $\overline{V} = \overline{V}$ on underlying bundles, so the action of $U(1)$ preserves the σ-structure of the fiber functor ω_x. The corresponding action on the group $\varpi_1^{\mathrm{red}}(X, x)$ commutes with the involution σ there. □

Lemma **6.6.** — *Let* C *denote the element* $-1 \in U(1)$. *There is an antilinear involution* $\tau = \sigma C = C\sigma$ *of the category* $\mathscr{E}^s$, *defined by* $\tau(V, D, D', D'') = (\overline{V}, \overline{D}_C, \overline{D}''_C, \overline{D}'_C)$. *The fiber functor* ω *has a natural real structure with respect to this involution also. The resulting real form* $\mathscr{U}(X, x)$ *of the group* $\varpi_1^{\mathrm{red}}(X, x)$ *is a compact real form (in particular, it meets every path-component).*

If $\varpi_1^{\mathrm{red}}(X, x) \to Gl(V_x)$ *is a representation, and if* K *is a harmonic metric on the associated harmonic bundle* V, *then* $g \in \varpi_1^{\mathrm{red}}(X, x)$ *is contained in* $\mathscr{U}(X, x)$ *if and only if* g *preserves* K_x.

Conversely, a metric K_x *on* V_x *comes from a harmonic metric* K *on* V *if and only if* K_x *is preserved by every real point* $u \in \mathcal{U}(X, x)$.

Proof. — Note that $C\sigma = \sigma C$ by the previous corollary, so τ is well defined and is an involution. Suppose (V, D, D', D'') is a harmonic bundle. Applying both statements of the previous lemma, we see that $\tau(V, D, D', D'')$ is a harmonic bundle. The involution τ commutes with the fiber functor, in the sense that canonically $\omega\tau(V) = \iota\omega(V)$, where ι denotes complex conjugation of the vector space $\omega(V)$. Also τ commutes with tensor product and dual (because σ obviously does, and C does). Therefore τ provides real structures for $\mathscr{E}^s$ and for the fiber functor, so it gives a real form $\mathcal{U}(X, x)$.

Here is an interpretation of this real form. If K is a harmonic metric, we can think of K as a C^∞ isomorphism between V and $\overline{V}^*$. In fact, the definition of the relationship between the operators D′ and D″, and the metric, says exactly that K is an isomorphism of harmonic bundles from V to $\tau(V^*)$. Suppose that $g = \{ g_V \}$ is an element of $\varpi_1^{\mathrm{red}}(X, x)$. The condition that $g \in \mathcal{U}(X, x)$ is that for all V, $g_{\tau(V)} = \iota g_V$. The fact that g is natural with respect to the morphism K means that $g_{\tau(V^*)} = K_x g_V K_x^{-1}$. Compatibility of g with the tensor structure implies that $g_{V^*} = {}^t g_V^{-1}$. Therefore the condition that g lies in $\mathcal{U}(X, x)$ may be interpreted as ${}^t\bar{g}_V^{-1} = K_x g_V K_x^{-1}$. This is exactly the condition that g preserves the metric K_x. This proves the first statement of the second paragraph, and also immediately implies that $\mathcal{U}(X, x)$ is an inverse limit of compact groups—hence, by Tychonoff's theorem, $\mathcal{U}(X, x)$ is compact.

Next we will show by a standard type of argument, that $\mathcal{U}(X, x)$ meets every path-component of $\varpi_1^{\mathrm{red}}(X, x)$. Suppose $g \in \varpi_1^{\mathrm{red}}(X, x)$ is some element. The element $h = g^{-1}\tau(g)$ is positive self adjoint in any representation, with respect to any harmonic metric. Thus for any reductive representation V and any real number t there is a unique automorphism h_V^t characterized by the conditions that the eigenspaces are the same as those of h_V and the eigenvalues are raised to the power t (remaining positive numbers). This is compatible with direct sums, tensor products, and isomorphisms, so it defines an element $h^t \in \varpi_1^{\mathrm{red}}(X, x)$. As t varies this path of elements gives a map $\mathbf{R} \to \varpi_1^{\mathrm{red}}(X, x)$ which is continuous with respect to the pro-analytic topology, because in any algebraic image it is continuous with respect to the usual topology. All of the elements h^t commute with each other. Note that $\tau(h) = h^{-1}$, so

$$\tau(gh^{1/2}) = \tau(g)\, h^{-1/2} = \tau(g)\, h^{-1}\, h^{1/2} = gh^{1/2}.$$

Therefore the element $gh^{1/2}$ is in $\mathcal{U}(X, x)$. But it is joined to the element g by the continuous path $t \mapsto gh^t$, for $t \in [0, 1/2]$. Thus we have shown that $\mathcal{U}(X, x)$ has a point in every path-component, so it is a compact real form of $\varpi_1^{\mathrm{red}}(X, x)$.

Finally, suppose K_x is a metric preserved by all $g \in \mathcal{U}(X, x)$. Choose a harmonic metric H for V. There is a unique endomorphism h_x of V_x which is self adjoint and positive definite for H_x, and such that $K_x = H_x h_x^2$. Since H_x and K_x are preserved by $\mathcal{U}(X, x)$, we have $uh_x u^{-1} = h_x$, and since $\mathcal{U}(X, x)$ is a compact real form of $\varpi_1^{\mathrm{red}}(X, x)$ (meeting every path-component), this implies that $gh_x g^{-1} = h_x$ for every $g \in \varpi_1^{\mathrm{red}}(X, x)$.

Therefore h_x is the value at x of an endomorphism h of the harmonic bundle V. Now let K be the metric defined by $(u, v)_{\mathrm{K}} = (hu, hv)_{\mathrm{H}}$. It restricts to $\mathrm{K}_x = \mathrm{H}_x h_x^2$ at x. This proves the last statement of the lemma. □

Corollary **6.7.** — *The category* $\mathscr{E}^s = \mathscr{E}^s_{\mathrm{DR}} = \mathscr{E}^s_{\mathrm{Dol}}$ *has a natural Cartan triple* $(\mathrm{C}, \sigma, \tau)$, *and if* ω_x *is a fiber functor of evaluation at* x, *then* $(\mathscr{E}^s, \omega)$ *has a natural Cartan triple* $((\mathrm{C}, \varphi_{\mathrm{C},x}), (\sigma, \varphi_{\sigma,x}), (\tau, \varphi_{\tau,x}))$. □

Finally we will complete the proof of Theorem 7. The above discussion shows that the action of $\mathrm{C} = -1$ on the category $\mathscr{E}^s$ is a Cartan involution with respect to the real structure σ, and hence the action of $\mathrm{C} = -1$ on the group $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}$ is a Cartan involution. The image of $\pi_1(\mathrm{X}, x)$ in $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ is Zariski dense by definition. To finish the proof we must verify the continuity axiom. For any representation $\rho : \varpi_1^{\mathrm{red}}(\mathrm{X}, x) \to \mathrm{Gl}(n, \mathbf{C})$, corresponding to a flat bundle (V, D) with $\mathrm{V}_x \cong \mathbf{C}^n$, the composed map

$$\mathrm{U}(1) \times \pi_1(\mathrm{X}, x) \to \varpi_1^{\mathrm{red}}(\mathrm{X}, x) \to \mathrm{Gl}(n, \mathbf{C})$$

is given concretely as follows. A pair (t, γ) is sent to the monodromy around γ of the flat bundle $(\mathrm{V}, \mathrm{D}_t)$, using the notation of Lemma 6.4. The action of t preserves the harmonic metric, and the new connection is $\mathrm{D}_t = \partial + \overline{\partial} + t\theta + t^{-1}\overline{\theta}$. This connection on the fixed bundle V varies continuously with t, so the resulting monodromy around γ varies continuously (using the analytic topology on $\mathrm{Gl}(n, \mathbf{C})$). This continuity for any representation ρ implies that the map

$$\mathrm{U}(1) \times \pi_1(\mathrm{X}, x) \to \varpi_1^{\mathrm{red}}(\mathrm{X}, x)$$

is continuous in the pro-analytic topology on $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$. □

PROOF OF THEOREM 9

The metric connection $\partial + \overline{\partial}$ on a harmonic bundle is compatible with tensor products. Therefore if P is a path from x to y, then parallel transport along P using this connection defines an isomorphism $\Phi_{\mathrm{P}} : \omega_x \cong \omega_y$ of fiber functors. This isomorphism preserves the $\mathrm{U}(1)$ action, since the action of $t \in \mathrm{U}(1)$ does not change $\partial + \overline{\partial}$. It also preserves the compact real form, because $\partial + \overline{\partial}$ is a unitary connection for any harmonic metric; and therefore Φ_{P} preserves the real structure σ too. The map Φ_{P} provides an isomorphism between the Hodge structures $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ and $\varpi_1^{\mathrm{red}}(\mathrm{X}, y)$ except for the fact that Φ_{P} does not preserve the finitely generated subgroup $\pi_1(\mathrm{X}, x)$. On the other hand, let T_{P} denote parallel transport along P using the flat connection D. This again defines an isomorphism $\mathrm{T}_{\mathrm{P}} : \omega_x \cong \omega_y$, which, however, is not compatible with the action of $\mathrm{U}(1)$. Composing we obtain an automorphism $\mathrm{T}_{\mathrm{P}}^{-1}\Phi_{\mathrm{P}}$ of the fiber functor ω_x, hence an element $\Phi(y, \mathrm{P}) = \mathrm{T}_{\mathrm{P}}^{-1}\Phi_{\mathrm{P}} \in \varpi_1^{\mathrm{red}}(\mathrm{X}, x)_{\mathbf{R}}$. If P' is a different path from x to y, then the element $\Phi(y, \mathrm{P}')$ will be different:

$$\Phi(y, \mathrm{P}') = \mathrm{T}_{\mathrm{P}'}^{-1}\mathrm{T}_{\mathrm{P}}\,\Phi(y, \mathrm{P})\,\Phi(x, \mathrm{P}^{-1}\mathrm{P}').$$

The element $T_{P'}^{-1} T_P$ on the left is in $\pi_1(X, x)$. The element on the right is an element of $\mathscr{K}(X, x)$ because the connection $\partial + \bar{\partial}$ is unaffected by the action of $U(1)$, so the parallel translation $\Phi(x, P^{-1} P')$ is an automorphism of the fiber functor which is fixed by $U(1)$. So after dividing out by $\mathscr{K}(X, x)$ on the right, we obtain a $\pi_1(X, x)$-equivariant map.

$$\Phi : \widetilde{X} \to \varpi_1^{\mathrm{red}}(X, x)_{\mathbf{R}}/\mathscr{K}(X, x).$$

Suppose $\rho : \varpi_1^{\mathrm{red}}(X, x) \to \mathrm{Gl}(n, \mathbf{C})$ is a representation such that $\rho(\mathscr{U}(X, x))$ preserves the standard metric on $\mathbf{C}^n$. Then there is a harmonic metric K for the resulting local system V, which is equal to the standard metric on $V_x \cong \mathbf{C}^n$. The parallel transport Φ_P from V_x to V_y takes K_x to K_y, and therefore the element $\rho(\Phi(y, P)) = T_P^{-1} \Phi_P$ takes the standard metric on $\mathbf{C}^n$ to the pullback of the harmonic metric K_y. The element $\rho(\Phi(y, P))$ is by definition a representative for the image of y in the harmonic map. $\square$

PROOF OF THEOREM 8

We have discussed the reductive part of the fundamental group above; now we will discuss the nilpotent part $\varpi_1^{\mathrm{nil}}(X, x)$. This will not complete an exhaustive discussion, because it leaves the question of what happens to the part of the group detected by representations which are extensions of nontrivial semisimple representations. That topic must be left for another time.

Morgan and Hain have defined a mixed Hodge structure on the nilpotent completion of $\pi_1(X, x)$, which is to say, on the completed group algebra $\mathbf{C}\pi_1(X, x)^\wedge$ and on the Malcev Lie subalgebra [38] [27]. In the present case, X is compact, so the weight filtration is the same as the filtration by powers of the augmentation ideal. Thus the data of the mixed Hodge structure is essentially contained in the Hodge filtration F_p (nontrivial only for negative p) and the real structure $\mathbf{R}\pi_1(X, x)^\wedge$. We would like to show a basic compatibility, that our action of $\mathbf{C}^*$ determines the Hodge filtration. On the other hand, it will probably be impossible to detect the real structure through $\mathscr{E}_{\mathrm{Dol}}$.

The completion of the group algebra $\mathbf{C}\pi_1(X, x)^\wedge$ is a Hopf algebra, and the Malcev Lie subalgebra is defined to be the subspace of Lie-like elements using the Hopf algebra structure. The Malcev Lie group is the exponential of the Malcev Lie algebra. The following proposition and its corollary show that we obtain this Hopf algebra, and hence the Malcev nilpotent completion of the fundamental group of X, by looking at the category of nilpotent flat bundles $\mathscr{E}_{\mathrm{DR}}^{\mathrm{nil}}$.

Proposition **6.8.** — *Let* $\omega : \mathscr{E}_{\mathrm{DR}}^{\mathrm{nil}} \to \mathscr{V}ect$ *be the fiber functor of evaluation at a point* $x \in X$. *Then* $\mathrm{End}(\omega, \mathscr{E}_{\mathrm{DR}}^{\mathrm{nil}})$ *is naturally isomorphic to* $\mathbf{C}\pi_1(X, x)^\wedge$, *the completion of the group algebra of the fundamental group with respect to the augmentation ideal* J. *The projective limit topology corresponds to the adic topology of the completion. Furthermore, the structures of Hopf algebra are the same.*

Proof. — There is an obvious natural map $\pi_1(X, x) \to \mathrm{End}(\omega, \mathscr{E}_{\mathrm{DR}}^{\mathrm{nil}})$, sending $\gamma \in \pi_1(X, x)$ to $\{\gamma_V\}$, where γ_V is the action of γ on $\omega(V)$ (the fiber of V at x) by parallel translation around γ. In fact, the category $\mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}}$ is naturally isomorphic to the category of nilpotent representations of $\pi_1(X, x)$ (with ω corresponding to the obvious fiber functor).

We get a map $\mathbf{C}\pi_1(X, x)^{\wedge} \to \mathrm{End}(\omega, \mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}})$, well defined and continuous on the completion, because the action of the group algebra on any finite collection of nilpotent representations factors through a power of the augmentation ideal. The map is also injective, because any nonzero element of $\mathbf{C}\pi_1(X, x)/J^k$ acts nontrivially on some nilpotent representation, namely the representation $\mathbf{C}\pi_1(X, x)/J^k$ itself.

Suppose V is a nilpotent representation (abusing notation we denote the vector space $\omega(V)$ also by V). Let $A_V \subset \mathrm{End}(V)$ denote the subalgebra generated by the image of $\pi_1(X, x)$. We claim that if $\{f_W\}$ is an element of $\mathrm{End}(\omega, \mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}})$, then $f_V \in A_V$. To prove this, note that $A_V \subset V \otimes V_0^*$ is the subrepresentation generated by the identity matrix, under the action of $\pi_1(X, x)$ on the first factor only (V_0^* denotes the dual vector space but with trivial group action). Therefore, $f_{V\otimes V_0^*}(1) \in A_V$. Now $V \otimes V_0^*$ is really a direct sum of copies of V, so $f_{V\otimes V_0^*} = f_V \otimes 1$ is the same as left multiplication by f_V. Therefore, left multiplication by f_V maps the identity matrix into A_V, so $f_V \in A_V$.

Recall that $\mathrm{End}(\omega, \mathscr{E}) = \varprojlim \mathrm{End}(\omega, \mathscr{E}_S)$. The previous paragraph implies that the image of $\mathbf{C}\pi_1(X, x)$ in any of the quotients $\mathrm{End}(\omega, \mathscr{E}_S)$ is equal to the image of $\mathrm{End}(\omega, \mathscr{E})$. The completion with respect to the augmentation ideal may be written as an inverse limit over the same directed set: for each $S \subset \mathscr{E}$, let $k(S)$ denote the maximum length of a minimal nilpotent filtration for objects in S (and hence in $\mathscr{E}_S$). Then $\mathbf{C}\pi_1(X, x)^{\wedge} = \varprojlim \mathbf{C}\pi_1(X, x)/J^{k(S)}$. Replace the directed system $\mathrm{End}(\omega, \mathscr{E}_S)$ by the directed system of images of $\mathrm{End}(\omega, \mathscr{E})$. Then the map between the directed systems is surjective at each stage. Therefore the map between the inverse limits is surjective. We saw above that it was injective, so it is an isomorphism. To see that the topologies are the same, note that if an element is close to the origin in $\mathrm{End}(\omega, \mathscr{E})$, meaning that it acts trivially on a large collection of objects, then in particular it acts trivially on the representation $\mathbf{C}\pi_1(X, x)/J^k$ for some large k, so it is close to the origin in $\mathbf{C}\pi_1(X, x)^{\wedge}$.

Finally we have to show that the structures of continuous Hopf algebra are the same (cf. [27]). The comultiplication in $\mathbf{C}\pi_1(X, x)^{\wedge}$ is given by $\Delta(\gamma) = \gamma \otimes \gamma$ for group elements γ. Recall that the comultiplication of $\mathrm{End}(\omega, \mathscr{E})$ was defined by $\Delta^{\otimes}\{f_V\} = \{f_{V\otimes W}\}$. Thus if $\gamma \in \pi_1(X, x)$, $\Delta^{\otimes}\{\gamma_V\} = \{\gamma_{V\otimes W}\}$. But $\gamma_{V\otimes W} = \gamma_V \otimes \gamma_W$ in $\mathrm{End}(\omega(V)) \otimes \mathrm{End}(\omega(W))$, so the comultiplications are the same. The augmentations are the same since γ acts by 1 on the unit object **1**. □

Corollary **6.9.** — *Under the above identification,* $\mathbf{aut}^{\otimes}(\omega, \mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}})$ *corresponds to the Malcev Lie algebra, and* $\mathrm{Aut}^{\otimes}(\omega, \mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}})$ *to the* **C**-*Malcev completion of the fundamental group.*

Proof. — See [27]. The Malcev Lie algebra and the Malcev completion are the subsets of Lie and group-like elements of $\mathbf{C}\pi_1(X, x)^{\wedge}$, respectively. □

The action of $\mathbf{C}^*$ on $\mathscr{E}^{\mathrm{nil}}_{\mathrm{Dol}}$ gives an action on $\mathscr{E}^{\mathrm{nil}}_{\mathrm{DR}}$, and hence by Proposition 6.8, an action of $\mathbf{C}^*$ on $\mathbf{C}\pi_1(X, x)^{\wedge}$ preserving the Malcev Lie algebra and Malcev group. This action on the Malcev Lie algebra is the one referred to in Theorem 8.

Our reason for introducing differential graded categories was to deal with the case of objects which are extensions of nontrivial irreducible objects. However, for nilpotent objects (extensions of trivial objects), the essential information comes from differential graded algebras. In following the definitions to describe our action of $\mathbf{C}^*$ on the nilpotent completion of the fundamental group, it will often be easier to work at the level of d.g.a.'s.

Recall that $\mathscr{C}_{\mathrm{DR}}^{\mathrm{triv}}$ is equivalent to the d.g.t.c. $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}_{\mathrm{DR}}^{\bullet})$ derived from the d.g.a. $\mathrm{A}_{\mathrm{DR}}^{\bullet}$. This completes to give an equivalence $\mathscr{C}_{\mathrm{DR}}^{\mathrm{nil}} \cong \mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{DR}}^{\bullet})$, and hence an isomorphism

$$\mathbf{C}\pi_1(\mathrm{X}, x)^{\wedge} \to \mathrm{End}(\omega, \mathscr{E}_{\mathrm{DR}}^{\mathrm{nil}}) \cong \mathrm{End}(\omega, \mathrm{E}^0\, \mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{DR}}^{\bullet})).$$

We can describe this concretely as follows. An object in $\mathrm{E}^0\, \mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{DR}}^{\bullet})$ is a pair $(\mathbf{1}^n, \eta)$ where $\eta \in \mathrm{A}^1(\mathrm{X}) \otimes \mathrm{M}_n$ is a nilpotent $n \times n$ matrix valued one form such that $d(\eta) + \eta \wedge \eta = 0$. If γ is an element of the fundamental group $\pi_1(\mathrm{X}, x)$, then it maps to the endomorphism of $\omega(\mathbf{1}^n, \eta) = \mathbf{C}^n$ given by transport around γ using the connection η, which we can expand in a series of iterated integrals:

$$\gamma \mapsto 1 + \int_{\gamma} \eta + \int_{\gamma} \eta \mid \eta + \dots$$

(see below). The series terminates because η is nilpotent.

There is an action of $\mathbf{C}^*$ on the d.g.a. $\mathrm{A}_{\mathrm{Dol}}^{\bullet}$, obtained by letting t act on $\mathrm{A}^{p,q}$ by t^p. The $\mathbf{C}^*$ action on $\mathrm{A}_{\mathrm{Dol}}^{\bullet}$ provides an action on $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet})$, and by completion, an action on $\mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet})$ (recall that this d.g.t.c. was defined to be the completion of previous one). Recall that we defined an action of $\mathbf{C}^*$ on the d.g.t.c. $\mathscr{C}_{\mathrm{Dol}}$ of Higgs bundles. The category $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet})$ is equal to the subcategory $\mathscr{C}_{\mathrm{Dol}}^{\mathrm{triv}} \subset \mathscr{C}_{\mathrm{Dol}}$ of trivial Higgs bundles. The identification $\mathscr{C}^{\mathrm{triv}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet}) = \mathscr{C}_{\mathrm{Dol}}^{\mathrm{triv}}$ preserves the $\mathbf{C}^*$ action, and since the extension of a functor between d.g.t.c.'s to a functor between completions is unique up to unique natural isomorphism, the identification $\mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet}) \cong \mathscr{C}_{\mathrm{Dol}}^{\mathrm{nil}}$ preserves the $\mathbf{C}^*$ action.

The quasi-equivalence between $\mathscr{C}_{\mathrm{Dol}}^{\mathrm{nil}}$ and $\mathscr{C}_{\mathrm{DR}}^{\mathrm{nil}}$ was defined to be the completion of the quasi-equivalence between $\mathscr{C}_{\mathrm{Dol}}^{\mathrm{triv}}$ and $\mathscr{C}_{\mathrm{DR}}^{\mathrm{triv}}$, which in turn is equal to the quasi-equivalence provided by the quasi-isomorphism of d.g.a.'s

$$\mathrm{A}_{\mathrm{Dol}}^{\bullet} = (\mathrm{A}^{\bullet}(\mathrm{X}), \bar{\partial}) \to (\ker(\partial), \bar{\partial}) \to (\mathrm{A}^{\bullet}(\mathrm{X}), d) = \mathrm{A}_{\mathrm{DR}}^{\bullet}.$$

The $\mathbf{C}^*$ action on $\mathrm{A}_{\mathrm{Dol}}^{\bullet}$ induces an action of $\mathbf{C}^*$ on

$$\mathrm{E}^0\, \mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{Dol}}^{\bullet}) = \mathrm{E}^0\, \mathscr{C}^{\mathrm{nil}}(\ker(\partial)) = \mathrm{E}^0\, \mathscr{C}^{\mathrm{nil}}(\mathrm{A}_{\mathrm{DR}}^{\bullet}),$$

and hence on $\mathbf{C}\pi_1(\mathrm{X}, x)^{\wedge}$. From our discussion so far, this is the action of $\mathbf{C}^*$ referred to in Theorem 8.

In order to prove that this $\mathbf{C}^*$ action corresponds to the known Hodge filtration, we must choose one of the ways of defining this Hodge filtration. The most convenient seems to be the method of Hain, using the bar construction, which we now recall.

Chen and Hain associate to every d.g.a. $A^\bullet$ with augmentation denoted ω, the *bar complex*

$$B^\bullet_\omega A^\bullet = \bigoplus_k \bigotimes^k \ker(\omega).$$

The elements are sums of tensor products written with bars $\langle a_1 \mid a_2 \mid \ldots \mid a_k \rangle$, where $\omega(a_i) = 0$. The total degree of $\langle a_1 \mid \ldots \mid a_k \rangle$ is defined to be $\Sigma \deg(a_i) - k$. The differential is defined by

$$\delta \langle a_1 \mid \ldots \mid a_k \rangle = \Sigma \pm \langle a_1 \mid \ldots \mid da_i \mid \ldots \mid a_k \rangle \pm \langle a_1 \mid \ldots \mid a_i a_{i+1} \mid \ldots \mid a_k \rangle.$$

Then $H^0 B^\bullet_\omega A^\bullet$ has a structure of Hopf algebra, and its dual $(H^0 B^\bullet_\omega A^\bullet)^*$ is a continuous Hopf algebra [27]. In the case when $A^\bullet = A^\bullet_{DR} = (A^\bullet(X), d)$, and the augmentation ω is given by evaluation at $x \in X$, they define a map $\mathscr{I} : \pi_1(X, x) \to (H^0 B^\bullet_\omega A^\bullet_{DR})^*$ by iterated integrals:

$$\mathscr{I}(\gamma)\,(\langle a_1 \mid \ldots \mid a_k \rangle) = \int_\gamma a_1 \mid \ldots \mid a_k.$$

To define the iterated integrals, represent γ by a path $\gamma(t) : [0, 1] \to X$, and let $\Delta^k \subset [0, 1]^k$ be the tetrahedron defined by $0 \leqslant t_1 \leqslant \ldots \leqslant t_k \leqslant 1$. Then

$$\int_\gamma a_1 \mid \ldots \mid a_k \overset{\text{def}}{=} \int_{\Delta^k} \gamma^* a_1(t_1) \wedge \ldots \wedge \gamma^* a_k(t_k).$$

Chen and Hain show that this map gives an isomorphism

$$\mathscr{I} : \mathbf{C}\pi_1(X, x)^\wedge \overset{\cong}{\to} (H^0 B^\bullet_\omega A^\bullet_{DR})^*$$

of continuous Hopf algebras. The filtration $F^p A^\bullet_{DR} = \bigoplus_{r \geqslant p} A^{r,s}(X)$ induces a filtration on $(H^0 B^\bullet_\omega A^\bullet_{DR})^*$ and hence on $\mathbf{C}\pi_1(X, x)^\wedge$. This is the Hodge filtration on the fundamental group as defined by Hain.

For any augmented d.g.a. $A^\bullet$, we can define a map

$$T : (H^0 B^\bullet_\omega A^\bullet)^* \to \mathrm{End}(\omega, E^0 \mathscr{C}^{\mathrm{nil}}(A^\bullet))$$

as follows. For $\lambda \in (H^0 B^\bullet_\omega A^\bullet)^*$ and $(\mathbf{1}^n, \eta) \in E^0 \mathscr{C}^{\mathrm{nil}}(A^\bullet)$, set

$$T(\lambda)_{(\mathbf{1}^n, \eta)} = \lambda(1 + \langle \eta \rangle + \langle \eta \mid \eta \rangle + \ldots) \in M_n = \mathrm{End}(\omega(\mathbf{1}^n, \eta)).$$

When we bar a number of elements $\eta \in A^1(X) \otimes M_n$, it means to multiply the matrices, combining by bar the coefficient one forms. Note that $T(\lambda)_{(\mathbf{1}^n, \eta)}$ is well defined, since the equation $d\eta + \eta \wedge \eta = 0$ implies that $\delta(1 + \langle \eta \rangle + \langle \eta \mid \eta \rangle + \ldots) = 0$. Furthermore, $T(\lambda)$ is a natural endomorphism, for if

$$s : (\mathbf{1}^n, \eta) \to (\mathbf{1}^m, \xi), \qquad ds + \xi s - s\eta = 0,$$

then

$$\begin{aligned} &\pm\, \delta(\Sigma \langle \xi \mid \ldots \mid \xi \mid (s - \omega(s)) \mid \eta \mid \ldots \mid \eta \rangle) \\ &\qquad = (\Sigma \langle \xi \mid \ldots \mid \xi \rangle)\, \omega(s) - \omega(s)\, (\Sigma \langle \eta \mid \ldots \mid \eta \rangle) \end{aligned}$$

is zero in $H^0 B^\bullet_\omega A^\bullet_\omega$, so

$$\omega(s)\, T(\lambda)_{(\mathbf{1}^n, \eta)} = T(\lambda)_{(\mathbf{1}^m, \xi)}\, \omega(s).$$

One should be able to show directly that T is a morphism of continuous Hopf algebras. This will be automatically verified in our cases because of the following lemma.

Lemma **6.10.** — *The composition*

$$\mathbf{C}\pi_1(X, x)^\wedge \xrightarrow{\mathscr{I}} (H^0\, B^\bullet_\omega\, A^\bullet_{DR})^* \xrightarrow{T} \mathrm{End}(\omega, E^0\, \mathscr{C}^{\mathrm{nil}}(A^\bullet_{DR}))$$

is equal to the transport map defined in Proposition 6.8. *In particular, the above map* T *is an isomorphism of continuous Hopf algebras in the case of* $A^\bullet_{DR}$.

Proof. — This follows immediately from the Picard-Chen expansion: the monodromy of a connection $d + \eta$ on the trivial bundle $\mathbf{1}^n$ is equal to

$$1 + \int_\gamma \eta + \int_\gamma \eta \mid \eta + \dots$$

which is the same as $T \circ \mathscr{I}(\gamma)_{(\mathbf{1}^n, \eta)}$. □

Now we can complete the comparison between our $\mathbf{C}^*$ action and the Hodge filtration. The quasi-isomorphisms of augmented d.g.a.'s

$$A^\bullet_{Dol} \leftarrow (\ker(\partial), \bar\partial) \rightarrow A^\bullet_{DR}$$

lead to a diagram

$$\begin{tikzcd}
(H^0\, B^\bullet_\omega\, A^\bullet_{Dol})^* \arrow[r] \arrow[d] & (H^0\, B^\bullet_\omega \ker(\partial))^* \arrow[d] & (H^0\, B^\bullet_\omega\, A^\bullet_{DR})^* \arrow[l] \arrow[d] \\
\mathrm{End}(\omega, E^0\, \mathscr{C}^{\mathrm{nil}}(A^\bullet_{Dol})) \arrow[r] & \mathrm{End}(\omega, E^0\, \mathscr{C}^{\mathrm{nil}}(\ker(\partial))) & \mathrm{End}(\omega, E^0\, \mathscr{C}^{\mathrm{nil}}(A^\bullet_{DR})). \arrow[l]
\end{tikzcd}$$

Hain shows that the cohomology of the bar complex is invariant under quasi-isomorphisms, so the maps along the top are isomorphisms. Similarly, quasi-isomorphic d.g.a.'s give quasi-equivalent d.g.c.'s $\mathscr{C}^{\mathrm{triv}}(A^\bullet)$, and hence quasi-equivalent completions $\mathscr{C}^{\mathrm{nil}}(A^\bullet)$. The functor E^0 transforms quasi-equivalences to equivalences, so the maps along the bottom are isomorphisms. The vertical map at the right is an isomorphism by the previous lemma, so all of the maps are isomorphisms. Therefore all of the algebras including the above and $\mathbf{C}\pi_1(X, x)^\wedge$ are canonically isomorphic; let H denote this algebra. Then $\mathbf{C}^*$ acts on H via its action on $A^\bullet_{Dol}$. Note that the isomorphism T between $(H^0\, B^\bullet_\omega\, A^\bullet_{Dol})^*$ and $\mathrm{End}(\omega, E^0\, \mathscr{C}^{\mathrm{nil}}(A^\bullet_{Dol}))$ is compatible with the $\mathbf{C}^*$ action, since the construction T is natural. Let H^p denote the subspace of H on which $\mathbf{C}^*$ acts by the character t^p. By looking at $(H^0\, B^\bullet_\omega\, A^\bullet_{Dol})^*$ it is clear that $\bigoplus H^p$ is dense in H. The filtration $F^p = \bigoplus_{r \geqslant p} H^r$ is equal to the filtration induced on $(H^0\, B^\bullet_\omega\, A^\bullet_{Dol})^*$ by the Hodge filtration of $A^\bullet_{Dol}$. But this is equal to the Hodge filtration of H, in other words the filtration induced by the Hodge filtration of $A^\bullet_{DR}$. This is because the quasi-isomorphisms of d.g.a.'s considered above are filtered quasi-isomorphisms, and the differentials are strictly compatible with the Hodge filtrations. Hain proves that these imply that the filtrations on the cohomology of the bar complex are preserved [27]. Thus we have proved Theorem 8. □

Remark. — Our $\mathbf{C}^*$ action preserves the Malcev Lie algebra, and hence induces the Hodge filtration there also.

Question. — We have obtained a canonical splitting of the Hodge filtration on the fundamental group. Deligne constructed a canonical splitting of the Hodge filtration of any mixed Hodge structure [10], so it is natural to ask, are they the same?

PRINCIPAL OBJECTS

This section formalizes the intuition that one does not need to worry about principal objects—once a theory is developed for vector objects, with information about tensor products, then the notion of principal object is recovered as the notion of a functor from Rep(G) to the Tannakian category of objects in question.

Suppose G is an algebraic group. We will discuss the notion of principal object, or G-torsor, from three points of view. From the group theoretic standpoint, a G-torsor is just a representation $\rho : \varpi_1(X, x) \to G$, up to conjugacy. Such a representation can be reinterpreted in terms of Tannakian categories as a tensor functor $p : \mathrm{Rep}(G) \to \mathscr{E}$. A more intrinsic description is that a G-torsor is a principal G-bundle P on X together with some additional structure. We will draw the connections between these viewpoints, and use them to include additional information such as real or Cartan structures.

First we will define the *groupoid of representations in* G to be the groupoid whose objects are representations $\rho : \varpi_1(X, x) \to G$. The set of isomorphisms from ρ to ρ' is the set of all group elements $g \in G$ such that $g\rho = \rho' g$.

A G-*torsor* in $\mathscr{E}$ is an exact fully faithful tensor functor $p : \mathrm{Rep}(G) \to \mathscr{E}$. There is a functor from the groupoid of representations in G to the groupoid of G-torsors as defined in the previous sentence. It sends a representation $\rho : \varpi_1(X, x) \to G$ to the tensor functor

$$p = \rho^* : \mathrm{Rep}(G) \to \mathscr{E} = \mathrm{Rep}(\varpi_1(X, x)).$$

Lemma **6.11.** — *This functor* $\rho \mapsto p = \rho^*$ *is an equivalence between the groupoid of representations in* G *and the groupoid of* G-*torsors.*

Proof. — If $p : \mathrm{Rep}(G) \to \mathscr{E}$ is an exact fully faithful tensor functor, then $\omega_x p$ is a fiber functor for Rep(G). We may choose an isomorphism $\varphi : \omega_G \cong \omega_x p$. This yields a representation $\rho = (p, \varphi)^* : \varpi_1(X, x) \to G$, and an isomorphism $\rho^* \cong p$. Suppose ρ_1 and ρ_2 are two representations. If $f = \{f_V : \rho_1^*(V) \to \rho_2^*(V)\}$ is a natural map of tensor functors on Rep(G), then applying the fiber functor ω_x, and using $\omega_x \rho_i^* \cong \omega_G$, we obtain a natural automorphism $\omega_x(f)$ of the tensor functor ω_G. This corresponds to a unique group element, $\omega_x(f) = g$. The condition that f maps $\rho_1^*(V)$ to $\rho_2^*(V)$ means that the element g intertwines the representations ρ_1 and ρ_2. Conversely, any group element intertwining ρ_1 and ρ_2 gives rise to a natural map $f : \rho_1 \to \rho_2$. This gives an identification between the isomorphisms of objects ρ, and tensor isomorphisms of functors p, as desired. □

In view of this lemma, we will use interchangeably the concepts of representation in G and G-torsor. These concepts can also be expressed in terms of principal objects on X, using the notion of a family of fiber functors.

We begin with a general definition. A *continuous*, C^∞, *complex analytic*, *locally constant*, or *complex algebraic family of fiber functors* on a Tannakian category $\mathscr{E}$, indexed by a topological space, C^∞ manifold, reduced complex analytic space, topological space, or reduced complex algebraic variety Y, is an exact faithful tensor functor ω_Y from $\mathscr{E}$ to the category of (continuous, C^∞, complex analytic, locally constant, or algebraic) complex vector bundles on Y, such that for each $y \in Y$, the functor ω_y obtained by composing the functor " fiber at y " with ω_Y is a fiber functor for $\mathscr{E}$.

One way to construct a family of fiber functors is as follows. Suppose G is an algebraic group. Suppose P is a continuous, C^∞, complex analytic, locally constant, or algebraic principal G-bundle on Y. Then for any representation $V \in \mathrm{Rep}(G)$, we can form the vector bundle $P \times_G V$ on Y. If we set $\omega_Y(V) = P \times_G V$ then ω_Y becomes a continuous, C^∞, complex analytic, locally constant, or algebraic family of fiber functors on the Tannakian category Rep(G). The following converse is a generalization of a result of M. Nori, who proved it in the algebraic case [41].

Lemma **6.12.** — *If* G *is an algebraic group and* ω_Y *is a continuous*, C^∞, *complex analytic, locally constant, or algebraic family of fiber functors on* Rep(G), *then there is a continuous*, C^∞, *complex analytic, or complex algebraic principal* G-*bundle* P *on* Y *and a natural isomorphism* $\omega_Y(V) \cong P \times_G V$ *of families of fiber functors. Any natural tensor isomorphism between two families of fiber functors* $P \times_G V \cong P' \times_G V$ *comes from a unique isomorphism of principal bundles* $P \cong P'$. *And* P_y *is naturally identified with the set of tensor isomorphisms* $\omega_G \cong \omega_y$.

Proof. — The proof will be the same in all of the cases. Define the fiber of the principal bundle P at a point $y \in Y$ to be the set P_y of natural tensor isomorphisms $\omega_G(V) \cong \omega_y(V)$. We have to describe how these fit together into a bundle over Y.

Choose an object U in Rep(G) such that any representation is a subquotient of a tensor product $T^{a,b}U = U^{\otimes a} \otimes (U^*)^{\otimes b}$. Fix a frame $\omega_G(U) \cong \mathbf{C}^n$. Then P_y may be identified as the subset of the set of isomorphisms of vector spaces $\varphi : \mathbf{C}^n \cong \omega_y(U)$ defined by equations expressing the conditions of naturality and compatibility with tensor product. (These equations say, in particular, that for any subrepresentation $W \subset T^{a,b}U$, the isomorphism φ maps $\omega_G(W) \subset T^{a,b}\mathbf{C}^n$ to $\omega_y(W) \subset T^{a,b}\omega_y(U)$—there are further equations saying that the resulting transformations of subquotients $\varphi : \omega_G(W/W') \to \omega_y(W/W')$ are natural with respect to morphisms, and compatible with tensor product.) Locally, we may choose a frame $\beta : \omega_Y(U) \cong \mathbf{C}^n \times Y$ varying in the appropriate fashion. Then P is identified with a subset of the bundle $\mathrm{Gl}(n, \mathbf{C}) \times Y = \mathrm{I}so_Y(\mathbf{C}^n \times Y, \mathbf{C}^n \times Y)$. It is a right principal G-bundle, in other words G acts on $\mathrm{Gl}(n, \mathbf{C})$ on the right (by precomposition via the representation U with its chosen frame) and for each $y \in Y$, P_y is a coset. Thus we obtain a map

$Y \to \mathrm{Gl}(n, \mathbf{C})/G$. The main contention is that this map is a continuous, C^∞, complex analytic, locally constant, or complex algebraic morphism. The principal bundle P with then be the pullback of the natural bundle of cosets over $\mathrm{Gl}(n, \mathbf{C})/G$.

To prove this contention, we use " Plücker coordinates ". Recall that there is a one dimensional representation $L \subset T^{a,b} U$ such that $G \subset \mathrm{Gl}(n, \mathbf{C})$ is characterized as the subgroup of elements which preserve the line $\omega_G(L) \subset T^{a,b} \mathbf{C}^n$ [13]. Let $\mathbf{P}^N$ denote the projective space of lines in $T^{a,b} \mathbf{C}^n$, and let L_0 denote the point corresponding to $\omega_G(L)$. Then $\mathrm{Gl}(n, \mathbf{C})$ acts on $\mathbf{P}^N$ and the stabilizer of L_0 is G. The orbit $\mathrm{Gl}(n, \mathbf{C})\, L_0$ is a locally closed subvariety of $\mathbf{P}^N$, and the map $\mathrm{Gl}(n, \mathbf{C})/G \to \mathrm{Gl}(n, \mathbf{C})\, L_0$ is an isomorphism.

A map $\varphi : A \to B$ of vector bundles is *strict* if locally there are direct sum decompositions $A = A_0 \oplus A_1$ and $B = B_0 \oplus B_1$ such that φ factors through an isomorphism $\varphi : A_1 \cong B_1$. If ω_Y is a family of fiber functors, then the image $\omega_Y(f)$ of any morphism is a strict morphism, because of the conditions that ω_Y and each ω_y are exact and faithful. Apply this to the inclusion $f : L \to T^{a,b} U$. We obtain a family of lines $\beta\omega_Y(L) \subset T^{a,b} \mathbf{C}^n$ varying in the appropriate way, hence a morphism to $\mathbf{P}^N$ (continuous, C^∞, etc.). The fact that all fiber functors over $\mathbf{C}$ are isomorphic [13] means that the morphism $Y \to \mathbf{P}^N$ maps Y into the orbit $\mathrm{Gl}(n, \mathbf{C})\, L_0$. This provides our morphism from Y to $\mathrm{Gl}(n, \mathbf{C})/G$. For each $y \in Y$, P_y is contained in the set of elements of $\mathrm{Gl}(n, \mathbf{C})$ which map $\omega_G(L)$ to $\beta\omega_y(L)$. Since both are G cosets, P_y is equal to that set of elements. Thus P is equal to the pullback of the natural bundle of cosets over $\mathrm{Gl}(n, \mathbf{C})/G$, which gives P the structure of principal bundle in the appropriate sense (continuous, C^∞, etc.). Changing the trivialization β causes the map $Y \to \mathrm{Gl}(n, \mathbf{C})/G$ to change by the same matrix-valued function. The structure of principal bundle remains the same.

For each $y \in Y$ we have a natural isomorphism of fiber functors $\omega_y(V) \cong P_y \times_G V$. To see that these isomorphisms vary nicely with y, express V as a subquotient of a tensor product $T^{c,d} U$. Note that tautologically from our construction, the isomorphisms $\omega_y(U) \cong P_y \times_G U$ vary in the appropriate way.

If P is a principal bundle and we set $\omega_Y(V) = P \times_G V$, then P is equal to the principal bundle provided by the above construction. This is because P embeds naturally in the bundle of isomorphisms from $\mathbf{C}^n$ to $\omega_Y(U)$, and in each fiber is equal to the principal bundle constructed above. This completes the proof. □

Lemma **6.13**. — *Let $\mathscr{E} = \mathscr{E}_{DR} = \mathscr{E}_{Dol}$. The fiber functors ω_x, evaluation at x, form a C^∞ family ω_X of fiber functors. This family has a structure of locally constant family, coming from $\mathscr{E}_{DR}$; and two structures of algebraic family, one coming from $\mathscr{E}_{DR}$ and the other coming from $\mathscr{E}_{Dol}$.*

Proof. — There are obvious functors ω_{DR} and ω_{Dol} from $\mathscr{E}$ to the categories of flat vector bundles on X and analytic vector bundles on X, respectively. The first gives the flat family of fiber functors. Since a flat bundle may be considered as an analytic vector bundle, both functors give analytic families of fiber functors. Serre's GAGA theorem says that analytic vector bundles are the same as algebraic vector bundles, so we obtain two algebraic families of fiber functors. To complete the proof, we must show that the two corresponding C^∞ families of fiber functors are isomorphic. Recall that the isomorphism $\mathscr{E}_{\mathrm{Dol}} \cong \mathscr{E}_{\mathrm{DR}}$ is obtained from the diagram of quasi-equivalences of differential graded tensor categories

$$\mathscr{C}_{\mathrm{DR}} \cong \widehat{\mathscr{C}}^s_{\mathrm{DR}} \leftarrow \widehat{\mathscr{C}}^s_{\mathrm{D}'} \rightarrow \widehat{\mathscr{C}}^s_{\mathrm{Dol}} \cong \mathscr{C}_{\mathrm{Dol}}.$$

Each of these categories has a natural C^∞ family of fiber functors ω_{X}. For the completions, note that the completion of a C^∞ family of fiber functors ω on $\mathscr{C}$ yields a C^∞ family of fiber functors on $\widehat{\mathscr{C}}$, defined by $\omega(\mathrm{U}, \eta) = \omega(\mathrm{U})$. There are natural isomorphisms between these families of fiber functors (intertwining with the above quasi-equivalences). For example,

$$\omega_{\mathrm{X}, \mathscr{C}^s_{\mathrm{D}'}} \cong \omega_{\mathrm{X}, \mathscr{C}^s_{\mathrm{DR}}}$$
$$\omega_{\mathrm{X}, \mathscr{C}^s_{\mathrm{D}'}} \cong \omega_{\mathrm{X}, \mathscr{C}^s_{\mathrm{Dol}}}$$

by our harmonic bundle construction (the C^∞ bundles underlying the flat and Higgs bundles of a harmonic bundle are the same). These complete to give

$$\omega_{\mathrm{X}, \widehat{\mathscr{C}}^s_{\mathrm{D}'}} \cong \omega_{\mathrm{X}, \widehat{\mathscr{C}}^s_{\mathrm{DR}}}$$
$$\omega_{\mathrm{X}, \widehat{\mathscr{C}}^s_{\mathrm{D}'}} \cong \omega_{\mathrm{X}, \widehat{\mathscr{C}}^s_{\mathrm{Dol}}}.$$

This proves the lemma. □

Suppose now that $p : \mathrm{Rep}(\mathrm{G}) \rightarrow \mathscr{E}$ is a representation in G, or G-torsor. Then, from the above lemma, we obtain a family of fiber functors $\omega_{\mathrm{X}} \circ p$ on $\mathrm{Rep}(\mathrm{G})$. By the previous lemma, there is a canonical principal G-bundle P such that $\omega_{\mathrm{X}} p(\mathrm{V}) = \mathrm{P} \times_{\mathrm{G}} \mathrm{V}$. This P is a C^∞ principal bundle, with a structure of flat principal bundle, and two algebraic structures P_{DR} and $\mathrm{P}_{\mathrm{Dol}}$. A point of P_x corresponds to an isomorphism of tensor functors $\varphi : \omega_{\mathrm{G}} \cong \omega_x p$. If such a point is chosen, then we get a representation $\rho_\varphi : \varpi_1(\mathrm{X}, x) \rightarrow \mathrm{G}$, hence an element in the groupoid of G-torsors.

We would like to obtain a converse to this construction, showing how to define a G-torsor in terms of a principal bundle. This can be done with $\mathscr{E}_{\mathrm{DR}}$ as follows. If given a G-torsor p, the principal bundle P constructed above has a flat structure P_{DR}, since $\omega_{\mathrm{X}, \mathrm{DR}} \circ p$ is a flat family of fiber functors on $\mathrm{Rep}(\mathrm{G})$. Conversely, given any flat principal G-bundle P, we obtain a G-torsor $p : \mathrm{Rep}(\mathrm{G}) \rightarrow \mathscr{E}_{\mathrm{DR}}$ defined by $p(\mathrm{V}) = \mathrm{P} \times_{\mathrm{G}} \mathrm{V}$. Furthermore if P′ is another flat principal bundle, the natural tensor isomorphisms $p \cong p'$ are in one to one correspondence with the isomorphisms $\mathrm{P} \cong \mathrm{P}'$ of flat principal G-bundles, by the last statement of Lemma 6.12 applied in the locally constant case. Thus the category of G-torsors in $\mathscr{E}_{\mathrm{DR}}$ is equivalent to the category of flat principal G-bundles P_{DR}.

We can also define a G-torsor as a principal bundle with some additional Higgs structure. If given a G-torsor p, the algebraic principal bundle $\mathrm{P}_{\mathrm{Dol}}$ has a *Higgs field*: a functorial section θ_V of $\mathrm{End}(P \otimes_G V) \otimes \Omega^1_X$, compatible with tensor product in the infinitesimal sense. For each x, the set of endomorphisms of $P_x \times_G V$, functorial in V and infinitesimally compatible with tensor product, is equal to the Lie algebra $\mathbf{aut}^{\otimes}(\omega_x \circ p, \mathrm{Rep}(G))$. As x varies, these Lie algebras fit together into the principal bundle $\mathrm{ad}(P) = P \times_G \mathbf{g}$, where $\mathbf{g} = \mathbf{aut}^{\otimes}(\omega_G, \mathrm{Rep}(G))$ is the Lie algebra of G. Thus we get a section θ of $\mathrm{ad}(P) \otimes \Omega^1_X$. If V is any representation, then θ_V is deduced from θ by the action of ad(P) on $P \times_G V$. By looking at a faithful representation V, it follows that θ is a holomorphic section of ad(P), and that $[\theta, \theta] = 0$. Conversely, given a holomorphic principal G-bundle P and a holomorphic section θ of $\mathrm{ad}(P) \otimes \Omega^1_X$ such that $[\theta, \theta] = 0$ in $\mathrm{ad}(P) \otimes \Omega^2_X$, we get a functor $p : V \mapsto P \times_G V$ from Rep(G) to the category of Higgs bundles on X.

We can define a *Higgs* G-*torsor* to be a tensor functor p from Rep(G) to the category of Higgs bundles on X, faithful and exact, taking all morphisms to strict morphisms. This is equivalent to the data of a *principal Higgs bundle* consisting of a holomorphic principal G-bundle P, and a section θ of $\mathrm{ad}(P) \otimes \Omega^1_X$ such that $[\theta, \theta] = 0$. Say that (P, θ) is *semistable* if there exists a faithful representation V of G such that $P \times_G V$ is a semistable Higgs bundle. Recall that we get Chern classes for the principal bundle P (one for each invariant polynomial on the Lie algebra of G). If P has vanishing Chern classes then for any representation W, the bundle $P \times_G W$ has vanishing Chern classes. In particular, any subquotient of a tensor power of V leads to a bundle of degree zero. Since any representation can be expressed as a subquotient of $T^{a,b}V$, and tensor powers of a semistable Higgs bundle are semistable, this means that for any representation W, the Higgs bundle $P \times_G W$ is semistable and therefore an object of $\mathscr{E}_{\mathrm{Dol}}$. This shows that a G-torsor in $\mathscr{E}_{\mathrm{Dol}}$ is the same thing as a principal Higgs bundle which is semistable and has vanishing Chern classes.

(*Caution.* — Without assuming at least the vanishing of those Chern classes of P which arise from invariant linear functionals, the semistability of $P \times_G V$ doesn't imply semistability of bundles associated to other representations.)

Our equivalence of categories $\mathscr{E}_{\mathrm{Dol}} \cong \mathscr{E}_{\mathrm{DR}}$ provides an equivalence between the notions of flat G-torsor, and of semistable Higgs G-torsor with vanishing Chern classes. It is compatible with pullbacks, so the fibers of corresponding objects at a point x are naturally identified.

If P is a G-torsor in $\mathscr{E}_{\mathrm{Dol}} \cong \mathscr{E}_{\mathrm{DR}}$, and φ is a point in P_x, then we obtain a representation $\varpi_1(X, x) \to G$, because φ amounts to an identification between the fiber functors ω_G and $\omega_x \circ p$. Choice of a different point φg results in a representation conjugate by the inner automorphism $\mathrm{Ad}(g)$.

Say that a G-torsor P as above is *reductive* if the image of the representation of $\varpi_1(X, x)$ is a reductive group. Let $\mathscr{E}^s$ denote the category of harmonic bundles (isomorphic to $\mathscr{E}^s_{\mathrm{Dol}}$ and $\mathscr{E}^s_{\mathrm{DR}}$). Then a G-torsor is reductive if and only if the image of the functor p is contained in the subcategory $\mathscr{E}^s$.

REAL AND CARTAN TORSORS

In what follows, we will consider the relation between principal objects, and the various structures C, σ, τ for $\varpi_1^{\mathrm{red}}(X, x)$. In order to simplify notation, let the letter ν stand for one of the following types of structures: an involution C, an antilinear involution σ or τ, or a Cartan triple (C, σ, τ) with $C = \sigma\tau = \tau\sigma$. If G is a group with ν structure, let G^ν denote the fixed points of ν; in the last case this means $G^C \cap C^\sigma \cap G^\tau$. If $\nu = C$ then G^C is an algebraic subgroup of G, whereas if $\nu = \sigma$ or $\nu = \tau$, G^ν is a real form. We will usually use the notation τ for a compact real form, in keeping with the notation of a Cartan triple.

The Tannakian categories we will consider have structures ν. A *ν-structure* for a tensor functor $f: \mathscr{F} \to \mathscr{F}'$ is a natural isomorphism of tensor functors $\psi_f : f\nu \cong \nu f$ such that the two isomorphisms $\nu f\nu \cong f$ are equal. If $\nu = (C, \sigma, \tau)$, then ψ consists of three isomorphisms $\psi_{f,C}$, $\psi_{f,\sigma}$, and $\psi_{f,\tau}$. In this case they should satisfy natural compatibility conditions. These may be phrased as saying that the various automorphisms of functors such as $Cf\sigma$, $Cf\tau$ or $\sigma f\tau$, constructed using ψ and the commutativity constraints $C\sigma \cong \sigma C$, etc., should be equal to the identity.

A *family of ν-structures* for a family of fiber functors ω is a natural tensor isomorphism $\psi : \omega\nu(V) \cong \nu\omega(V)$ compatible with the continuous, C^∞, analytic, locally constant, or algebraic structure of ω (and subject to the constraints described above). Note that this only makes sense in some combinations. If $\nu = C$ then, on the category Vect(X) of vector bundles over X, ν is defined to be the identity. However, if $\nu = \sigma$ or $\nu = \tau$, then ν is defined on the category Vect(X) to be the complex conjugation ι. This only makes sense for continuous, C^∞, or locally constant vector bundles. If U is such a vector bundle, then ιU is defined to be the vector bundle whose transition functions are the complex conjugates of the transition functions of U. There is a canonical antilinear identification between U and ιU. Thus if $\nu = \sigma$, $\nu = \tau$, or $\nu = (C, \sigma, \tau)$, only the notions of continuous, C^∞, or locally constant family of ν structures make sense, In practice, if $\nu = \sigma$ then the notion of locally constant family is useful, but if $\nu = \tau$ or $\nu = (C, \sigma, \tau)$, then our families will be C^∞ only.

Suppose G is a reductive group with ν-structure. Then the canonical fiber functor ω_G has a canonical ν-structure ψ_G. Suppose ω is a C^∞ family of fiber functors with ν-structure on the category Rep(G). It will not always be the case that (ω_x, ψ_x) are isomorphic to (ω_G, ψ_G) as fiber functors with ν-structure. Therefore we make the following definition: a family of ν-structures for fiber functors ω is *good* if $(\omega_x, \psi_x) \cong (\omega_G, \psi_G)$ for all $x \in X$.

Lemma **6.14**. — *Assume that* X *is connected. Suppose* ω *is a family of fiber functors. The data of a ν-structure ψ is the same as the data of an involution ν induced on the corresponding principal bundle* P. *The ν-structure is good if and only if this involution has fixed points in at least one fiber* P_x. *If this is the case, then there are fixed points in all of the fibers, and the fixed points* P^ν *form a principal* G^ν*-bundle, with* $P = P^\nu \times_{G^\nu} G$.

Proof. — The ν-structure ψ_x induces an involution of $P_x = \mathrm{Iso}^{\otimes}(\omega_G, \omega_x)$ by the following rule. Suppose $p : \omega_G \to \omega_x$, and suppose $u \in \omega_G(V)$. Then

$$\nu(p)\,(u) = \nu\psi_x p \psi_G^{-1}\,\nu(u).$$

Note that $\nu(u) \in \nu\omega_G(V)$, so $\psi_G^{-1}\,\nu(u) \in \omega_G(\nu V)$, p maps this element to $\omega_x(\nu V)$, ψ_x maps it to an element of $\nu\omega_x(V)$, and ν then takes it to $\omega_x(V)$. The involution of P_x varies smoothly with x, so it gives an involution of the principal bundle P. The fixed points of this involution are exactly the isomorphisms of fiber functors with ν-structure $(\omega_G, \psi_G) \cong (\omega_x, \psi_x)$. Conversely, given an involution ν of P, we obtain a family of ν-structures ψ_x by the same formula.

We have to show that if an involution ν has fixed points in one fiber, then it has fixed points in all fibers, and the set of fixed points P^{ν} is smooth over X. Suppose now that ν is an involution (not a Cartan triple). By covering X, it suffices to prove the lemma in the case that P has a trivialization $P \cong P_0 \times X$ (of right G spaces). Fix a base point $q \in P_0$. Let N denote the space of involutions ν of P_0 compatible with the right action of (G, ν). If $\nu \in N$, then there is a unique group element h such that $\nu(qg) = qh\nu(g)$ for all $g \in G$. The condition that ν is an involution is that $h\nu(h) = 1$, and N may be identified with the set of $h \in G$ satisfying this condition. Let $\mathrm{Aut}(P_0)$ denote the group of automorphisms of the G-space P_0. This group acts on N by transport of structure. The choice of base point gives an isomorphism $\mathrm{Aut}(P_0) \cong G$, and with this identification the action of G by transport of structure is given by the formula $g(h) = gh\nu(g)^{-1}$. The stabilizer of an element h is the set of fixed points of the twisted ν-structure for G,

$$\nu_h(g) = h\nu(g)\,h^{-1},$$

We claim that the orbits under the action of G are open. If h' is a point in N near h, we can write $h' = uh$, with $uh\nu(u)\,\nu(h) = 1$. Since $h\nu(h) = 1$, this gives

$$u\nu_h(u) = 1.$$

The element u is near the identity, so we may write $u = \exp(Z)$, $Z = \log(u)$ in the Lie algebra $\mathbf{g}$ of G. The involution ν_h acts on $\mathbf{g}$, and we have

$$\nu_h(Z) = \log(\nu_h(u)) = \log(u^{-1}) = -\log(u) = -Z.$$

Set $\omega = \exp(Z/2)$. Then still $\nu_h(w) = w^{-1}$, so

$$w\nu_h(w)^{-1} = w^2 = u.$$

The point h is translated by the action of w to

$$w(h) = wh\nu(w)^{-1} = w\nu_h(w)^{-1}\,h = uh = h'.$$

This proves that the orbit of any $h \in N$ contains all nearby points. The set N is therefore a disjoint union of the orbits, which are its connected components.

The family of involutions of the bundle $P = P_0 \times X$ corresponds to a map from X to N. Since X is connected, the image is contained in one orbit. Thus all of the involutions ν_x are conjugate by automorphisms of the G-space P_0. In particular, the condition that there exists a fixed point is independent of x. To see that the sets of fixed points P^{ν}_{x} fit together into a smooth principal G^{ν} bundle, use the same argument as in Lemma 6.12. Let N^0 denote the component of N corresponding to involutions with fixed points. Then $N^0 \cong G/G^{\nu}$, so there is a universal bundle of cosets over N^0. The bundle P^{ν} is the pull-back of this bundle of cosets by the map $X \to N^0$.

To finish the proof in the case where ν is a Cartan triple, apply the above proof to the case of the involution C, to obtain a principal bundle P^C. Then apply the same proof to the involution σ of P^C to obtain the principal bundle P^{ν}. □

Corollary **6.15.** — *If ψ is a ν-structure on a family of fiber functors ω on* Rep(G), *then the fiber functors with ν-structure (ω_x, ψ_x) and (ω_y, ψ_y) are isomorphic for any $x, y \in X$.*

Proof. — Choose an isomorphism $\omega_G \cong \omega_x$. Let ν' be the involution of G determined by the structure ψ_x, so $(\omega_G, \psi'_G) \cong (\omega_x, \psi_x)$. This involution of G yields an involution ν' of the principal bundle P, and the condition of the previous sentence implies that it has fixed points in the fiber P_x. Therefore there are fixed points in all fibers, so all (ω_y, ψ_y) are isomorphic to (ω_G, ψ'_G). □

The Tannakian category $\mathscr{C} = \mathscr{C}_{DR} = \mathscr{C}_{Dol}$ has ν-structures for $\nu = C$ or $\nu = \sigma$, and the subcategory of semisimple objects $\mathscr{C}^s$ has a Cartan structure $\nu = (C, \sigma, \tau)$. The family of fiber functors ω_X has a family of ν-structures in each of these cases. These induce ν-structures on $\varpi_1(X, x)$. If $\nu = \tau$ or $\nu = (C, \sigma, \tau)$, then the subcategory $\mathscr{C}^s$ of reductive objects has a ν-structure, and again ω_X has a family of ν-structures. The corresponding reductive quotient $\varpi_1^{red}(X, x)$ has a ν-structure.

Now we may proceed with our definition of torsors. The *groupoid of* (G, ν)*-torsors* has for objects the representations $\rho : \varpi_1(X, x) \to G$ compatible with ν. If $\nu = \tau$ or $\nu = (C, \sigma, \tau)$, then the representation ρ is assumed to factor through the reductive quotient $\varpi_1^{red}(X, x)$. The isomorphisms between ρ and ρ' are the elements g of G^{ν} such that $g\rho = \rho' g$.

The corresponding definition in terms of Tannakian categories is as follows. A (G, ν)-torsor in $\mathscr{C}$ (or $\mathscr{C}^s$) is a tensor functor $p : \mathrm{Rep}(G) \to \mathscr{C}$ together with a ν-structure $\psi_p : p\nu \cong \nu p$ such that the associated family of fiber functors with ν-structure $(\omega_x p, \psi_x \psi_p)$ on Rep(G), is good with respect to (G, ν). Recall that this means that $(\omega_x p, \psi_x \psi_p) \cong (\omega_G, \psi_G)$, and by Corollary 6.15 it need only be checked at one point x.

Given a representation ρ compatible with ν, one obtains a functor $p = \rho^*$ with ν-structure ψ_p. Conversely, the condition of goodness assures that, given a functor p, we may choose an isomorphism $(\omega_x p, \psi_x \psi_p) \cong (\omega_G, \psi_G)$, and so obtain a representation $\rho : \varpi_1(X, x) \to G$ compatible with ν-structures. The morphisms of functors p compatible with ν-structure ψ_p are the same as the morphisms of representations ρ in the previously defined groupoid, because the group of tensor automorphisms of (ω_G, ψ_G) is equal to G^{ν} (a tensor automorphism of ω_G is given by a group element g, and the condition of compatibility with ψ_G is that $g_V = \nu(\psi_{G,V}\, g_{\nu(V)}\, \psi_{G,V}^{-1})$; the latter expression is the definition of $\nu(g)_V$, so the condition of compatibility with ψ reduces to $g = \nu(g)$). This shows that our two notions of (G, ν)-torsor are equivalent, extending Lemma 6.11.

Finally, we will establish the connection with principal bundles. Suppose p is a (G, ν)-torsor. Then we obtain a good family of fiber functors $\omega_x p$ with ν-structures $\psi_x \psi_p$, on $\mathrm{Rep}(G)$. By Lemma 6.14, there is a unique principal G^ν-bundle P^ν, with the principal G-bundle corresponding to p equal to $P = P^\nu \times_{G^\nu} G$. The bundle P^ν is the set of fixed points of an involution ν of P. The relationship between P^ν and the additional structures of P depends on what ν is, so we examine each case separately.

If $\nu = \tau$, assume that G^τ is a compact real form. The representation ρ factors through $\varpi_1^{\mathrm{red}}(X, x)$. If V is a representation of G, then $P \times_G V$ is an element of $\mathscr{E}^s$, in other words a harmonic bundle. In this case, the principal bundle P^τ is a *harmonic reduction* of P, in other words for any G^τ-invariant metric K on a representation V, the resulting metric $P^\tau \times_{G^\tau} K$ is a harmonic metric for the harmonic bundle $P \times_G V$. Conversely, given any reduction of structure group P^τ for P from G to G^τ, with the above harmonic property for a faithful representation V, there will be a corresponding structure of (G, τ)-torsor on the G-torsor p. It is given by the isomorphism

$$\tau(P^\tau \times_{G^\tau} V) \cong P^\tau \times_{G^\tau} \tau(V).$$

That this is an isomorphism of harmonic bundles may be checked by interpreting a G^τ-invariant metric K as an isomorphism between V^* and $\tau(V)$. The harmonic condition says that $P^\tau \times_{G^\tau} K$ gives an isomorphism between harmonic bundles $P^\tau \times_{G^\tau} V^*$ and $\tau(P^\tau \times_{G^\tau} V)$, which implies that the above isomorphism is also an isomorphism of harmonic bundles. Note also that if this harmonic condition is satisfied for one representation V, then it is satisfied for any other embedded in a tensor power of V.

A G-torsor p has a harmonic reduction if and only if it is reductive. For, if $\rho : \varpi_1(X, x) \to G$ factors through $\varpi_1^{\mathrm{red}}(X, x)$, then the image $\rho U(X, x)$ of the compact real form of $\varpi_1^{\mathrm{red}}(X, x)$ defined by τ is a compact subgroup of G. Any compact subgroup of G is conjugate to a subgroup of the compact real form G^τ, so after conjugation we may assume that ρ is compatible with τ.

The notion of harmonic reduction can no doubt be defined in terms of the differential geometry of the principal bundle, with operators D, D″, etc., related by the reduction of structure group.

Suppose $\nu = \sigma$, and fix a real form G^σ of G. Then a (G, σ)-torsor is a representation $\rho : \varpi_1(X, x)^\sigma \to G^\sigma$. Since the real form σ of $\varpi_1(X, x)$ is just the real pro-algebraic closure of $\pi_1(X, x)$, a (G, σ)-torsor is the same thing as a representation $\pi_1(X, x) \to G^\sigma$. In terms of principal bundles, the principal bundle P associated to ρ has a locally constant structure. The family of σ-structures for the fiber functors ω_x on $\mathscr{E}_{\mathrm{DR}}$ is a locally constant family, so the principal bundle P^σ associated to the (G, σ)-torsor is a locally constant principal G^σ-bundle. Conversely, given any locally constant G^σ-bundle P^σ, the monodromy representation $\rho : \pi_1(X, x) \to G^\sigma$ (which depends on choice of a point φ in P^σ_x)

will be a (G, σ)-torsor. Given two principal bundles with points chosen, an isomorphism $f: P^\sigma \to P^{\sigma'}$ is given by an element $g \in G^\sigma$, defined by $\varphi' = f(\varphi)\, g$. This element intertwines the representations ρ and ρ'. Thus the groupoid of (G, σ)-torsors is the same as the groupoid of locally constant principal G^σ bundles P^σ.

Suppose $\nu = C$. Fix an involution C of G. A (G, C)-torsor is a representation $\rho : \varpi_1(X, x) \to G$ which is compatible with C, up to conjugacy by G^C. Associated to the underlying G-torsor is a principal bundle P, which we now consider as an algebraic principal bundle by the algebraic structure of the family of fiber functors on $\mathscr{E}_{\mathrm{Dol}}$. There is an involution C of P defined as follows. If V is a representation of G, then the representation CV is the vector space of elements denoted Cv, $v \in V$. A group element acts by $g(Cv) = C(C(g)\, v)$. The functorial isomorphism $\psi : p(CV) \cong Cp(V)$ is an isomorphism $\psi : P \times_G CV \cong C(P \times_G V)$ of Higgs bundles. Note that the vector bundle underlying $C(P \times_G V)$ is the same as that underlying $P \times_G V$. The isomorphism ψ is given by $\psi(\varphi, Cv) = (C\varphi, v)$, which defines the involution $\varphi \mapsto C\varphi$ of P. Since the C-structures of $\mathscr{E}_{\mathrm{Dol}}$ and the family of fiber functors are defined algebraically, the involution C of P is algebraic, and the set of fixed points P^C is an algebraic principal G^C-bundle.

The Higgs structure of the principal bundle P is a section θ of $P \times_G \mathbf{g} \otimes \Omega^1_X$. Write this section locally as $(\varphi, \theta_\varphi)$ with $\theta_\varphi \in \mathbf{g} \otimes \Omega^1_X$. This serves to define the Higgs structure θ_V of a bundle $P \times_G V$, by the formula $\theta_V(\varphi, v) = (\varphi, \theta_\varphi v)$. Now we can translate the condition that the isomorphism ψ is an isomorphism of Higgs bundles. Recall that the operation C applied to $P \times_G V$ changes the sign of θ. Therefore the condition that ψ preserves θ is that

$$\psi(\theta_{CV}(\varphi, Cv)) = -\theta_V\, \psi(\varphi, Cv),$$

or

$$\psi(\varphi, \theta_\varphi(Cv)) = -\theta_V(C\varphi, v) = (C\varphi, -\theta_{C\varphi}\, v).$$

The involution C of G acts on the Lie algebra $\mathbf{g}$, and $\theta_\varphi(Cv) = C(C(\theta_\varphi)\,(v))$. Therefore

$$\psi(\varphi, \theta_\varphi(Cv)) = \psi(\varphi, C(C(\theta_\varphi)\,(v)) = (C\varphi, C(\theta_\varphi)\,(v)).$$

Thus our condition becomes

$$C(\theta_\varphi) = -\theta_{C\varphi}.$$

If φ lies in the reduction of structure group P^C, then $C\varphi = \varphi$ so this condition becomes $C(\theta_\varphi) = -\theta_\varphi$. Decompose the Lie algebra $\mathbf{g}$ as $\mathbf{g} = \mathbf{k} \oplus \mathbf{p}$ according to the eigenvalue of C (this is the Cartan decomposition). The fixed points $\mathbf{k}$ form the Lie algebra of G^C, and $\mathbf{p}$ is the -1 eigenspace of C. They are representations of G^C. Our condition above says that if $\varphi \in G^C$, then $\theta_\varphi \in \mathbf{p} \otimes \Omega^1_X$. In other words, θ is a section of $P^C \times_{G^C} \mathbf{p} \otimes \Omega^1_X$. Conversely, this condition is sufficient for ψ to be a morphism of Higgs bundles.

Thus we have arrived at the principal bundle version of the notion of (G, C)-torsor. A *Higgs* (G, C)-*torsor* is an algebraic principal G^C-bundle P^C together with a section θ of $P^C \times_{G^C} \mathbf{p} \otimes \Omega^1_X$, such that $[\theta, \theta] = 0$. This results in a Higgs G-torsor $P = P^C \times_{G^C} G$, and the conditions of semistability and vanishing of Chern classes need only be stated in terms of P. A(G, C)-torsor in $\mathscr{E}$ is a Higgs (G, C)-torsor P^C such that the associated Higgs G-torsor P is semistable and has vanishing Chern classes. This means that the Chern classes of the principal bundle P vanish and, for some faithful representation V of G, $P \times_G V = P^C \times_{G^C} V$ is a semistable Higgs bundle.

Finally we must treat the case where $\nu = (C, \sigma, \tau)$. Fix a Cartan structure (C, σ, τ) for G. A (G, C, σ, τ)-torsor is a representation $\rho : \varpi_1^{\mathrm{red}}(X, x) \to G$ which is compatible with Cartan structures, up to conjugacy by $G^{C, \sigma, \tau}$. Such a torsor results in a C^∞ reduction of structure group to a $G^{C, \sigma, \tau}$-bundle $P^{C, \sigma, \tau}$, for the G-torsor P. From this we obtain reductions

$$P^\tau = P^{C, \sigma, \tau} \times_{G^{C, \sigma, \tau}} G^\tau,$$

$$P^\sigma = P^{C, \sigma, \tau} \times_{G^{C, \sigma, \tau}} G^\sigma,$$

$$P^C = P^{C, \sigma, \tau} \times_{G^{C, \sigma, \tau}} G^C,$$

which are, respectively, a harmonic reduction, a locally constant G^σ-bundle, and a Higgs (G, C)-torsor. Conversely, suppose we are given a reduction $P^{C, \sigma, \tau}$ which results in a harmonic reduction, a locally constant G^σ bundle, and a Higgs (G, C)-torsor. Then, by choosing a point φ in $P_x^{C, \sigma, \tau}$, we obtain a representation $\rho : \varpi_1^{\mathrm{red}}(X, x) \to G$. It is in turn compatible with all of the structures τ, σ, and C, so it is a (G, C, σ, τ)-torsor. As usual, the isomorphisms between P and P′ which map a reduction $P^{C, \sigma, \tau}$ to a reduction $P^{C, \sigma, \tau'}$ are seen to correspond to elements of $G^{C, \sigma, \tau}$ which intertwine the representations ρ and ρ', if one bears in mind that the representations come from choices of points φ and φ'.

Say that a (G, σ)-torsor or (G, C)-torsor is *reductive* if the representation ρ factors through the reductive quotient $\varpi_1^{\mathrm{red}}(X, x)$. This is equivalent to requiring that the functor p has image in $\mathscr{E}^s$, the category of semisimple objects. In terms of principal bundles, this means that for a representation V, the object $P \times_G V$ should be a direct sum of irreducible objects. This condition need only be checked for one faithful representation V. In terms of representations of the fundamental group, it means that the Zariski closure of the image of $\pi_1(X, x)$ should be a reductive group. In terms of Higgs bundles, it means that for a faithful representation V, the Higgs bundle $P^C \times_{G^C} V$ should be a direct sum of stable Higgs bundles of degree zero (and as usual, the Chern classes of P vanish). Note that if P is a (G, C, σ, τ)-torsor, then the associated (G, σ)-torsor and (G, C)-torsor are reductive. We can now state the main theorem about real structures for reductive principal objects. It generalizes to any reductive group Lemma 2.12, which concerned the case $G^\sigma = \mathrm{Gl}(n, \mathbf{R})$.

Theorem **10.** — *Suppose that* (C, σ, τ) *is a Cartan structure for* G. *The natural functors from the category of* (G, C, σ, τ)*-torsors, to the categories of reductive* (G, C)*-torsors or reductive* (G, σ)*-torsors, are bijective on sets of isomorphism classes.*

Proof. — Let $\nu = \mathrm{C}$ or $\nu = \sigma$; the proof is the same in either case. Suppose $\rho : \varpi_1^{\mathrm{red}}(\mathrm{X}, x) \to \mathrm{G}$ is a representation compatible with ν. The compact real form $\mathscr{U}(\mathrm{X}, x)$ of $\varpi_1^{\mathrm{red}}(\mathrm{X}, x)$ maps to a compact subgroup $\mathrm{U} \subset \mathrm{G}$. This subgroup is preserved by the involution ν. Therefore by Lemma 6.3 it can be extended to a compact real form τ' preserved by ν, with $\mathrm{U} \subset \mathrm{G}^{\tau'}$. This gives a Cartan structure (ν, τ') for G. The uniqueness statement of Lemma 6.3 shows that this Cartan structure is conjugate to the original one (ν, τ) by an element g of G^{ν}. After conjugating ρ by g, it preserves Cartan structures. Therefore, ρ is isomorphic as a (G, ν)-torsor to a $(\mathrm{G}, \mathrm{C}, \sigma, \tau)$-torsor.

To prove essential injectivity, suppose ρ and ρ' are two $(\mathrm{G}, \mathrm{C}, \sigma, \tau)$-torsors, and suppose the corresponding (G, ν)-torsors are isomorphic, so there is an element $g \in \mathrm{G}^{\nu}$ such that $g\rho = \rho' g$. We would like to replace g by an element of $\mathrm{G}^{\mathrm{C}, \sigma, \tau}$. Set $h = \tau(g)\, g^{-1}$. The Cartan decomposition for the group G considered as real group with maximal compact $\mathrm{U} = \mathrm{G}^{\tau}$, says that we may write $g = \exp(y)\, u$ with $y \in i\mathbf{u}$ and $u \in \mathrm{U}$. Then $\tau(g) = \exp(-y)\, u$ so $h = \exp(-2y)$. Since ν commutes with τ, and g is an element of G^{ν}, $\nu(h) = h$. Since $\exp : i\mathbf{u} \to \mathrm{G}$ is injective, $\nu(y) = y$. Thus $\nu(u) = u$, so $u \in \mathrm{G}^{\mathrm{C}, \sigma, \tau}$. Apply τ to the equation $g\rho g^{-1} = \rho'$ to get $\tau(g)\, \rho\tau(g)^{-1} = \rho'$, which implies $h\rho'\, h^{-1} = \rho'$. This can be rewritten as $\rho'(\gamma)\, h\rho'(\gamma)^{-1} = h$ for $\gamma \in \varpi_1^{\mathrm{red}}(\mathrm{X}, x)$. If $\gamma \in \mathscr{U}(\mathrm{X}, x)$, then $\rho'(\gamma) \in \mathrm{U}$, so $\rho'(\gamma)$ acts on $i\mathbf{u}$. Therefore we have $\exp(-2\rho'(\gamma)\, y\rho'(\gamma)^{-1}) = \exp(-2y)$. Again since $\exp : i\mathbf{u} \to \mathrm{G}$ is injective this implies that $\rho'(\gamma)\, y\rho'(\gamma)^{-1} = y$. This is true for all γ in the compact real form $\mathscr{U}(\mathrm{X}, x)$, so it is true for all γ. In particular, $\exp(y)$ commutes with ρ'. Now the equation $\exp(y)\, u\rho(\gamma)\, u^{-1} \exp(-y) = \rho'(\gamma)$ implies that $u\rho u^{-1} = \rho'$. Thus we may replace g by the element $u \in \mathrm{G}^{\mathrm{C}, \sigma, \tau}$. If two representations are isomorphic as (G, ν)-torsors, then they are also isomorphic as $(\mathrm{G}, \mathrm{C}, \sigma, \tau)$-torsors. □

Corollary **6.16.** — *There is a one to one correspondence between the isomorphism classes of reductive* (G, C)*-torsors and the isomorphism classes of reductive* (G, σ)*-torsors.* □

In interpreting this corollary, one should note that the notion of reductive (G, C)-torsor is an algebraic one: it is a principal Higgs bundle (P, θ) consisting of a G^{C} bundle P^{C} together with a section θ in $\mathrm{H}^0(\mathrm{P}^{\mathrm{C}} \times_{\mathrm{G}^{\mathrm{C}}} \mathbf{p} \otimes \Omega^1_{\mathrm{X}})$ such that $[\theta, \theta] = 0$, such that the Chern classes of $\mathrm{P} = \mathrm{P}^{\mathrm{C}} \times_{\mathrm{G}^{\mathrm{C}}} \mathrm{G}$ vanish, and for some faithful representation V of G, the resulting Higgs bundle $\mathrm{P}^{\mathrm{C}} \times_{\mathrm{G}^{\mathrm{C}}} \mathrm{V}$ is a direct sum of stable factors of degree zero. The notion of reductive (G, σ)-torsor is a topological one, simply meaning a semisimple representation of $\pi_1(\mathrm{X}, x)$ in the real group $\mathrm{G}_{\mathbf{R}} = \mathrm{G}^{\sigma}$.

REFERENCES

[1] L. V. Ahlfors, An extension of Schwarz's lemma, *Trans. Amer. Math. Soc.*, **43**, no. 3 (1938), 359-364.

[2] H. Bass, Groups of integral representation type, *Pacific J. of Math.*, **86** (1980), 15-51.

[3] A. Borel and W. Casselman, L^2-cohomology of locally symmetric manifolds of finite volume, *Duke Math. J.*, **50** (1983), 625-647.

[4] J. A. Carlson and D. Toledo, Harmonic mappings of Kähler manifolds to locally symmetric spaces, *Publ. Math. I.H.E.S.*, **69** (1989), 173-201.

[5] K. Corlette, Flat G-bundles with canonical metrics, *J. Diff. Geom.*, **28** (1988), 361-382.

[6] K. Corlette, Rigid representations of Kählerian fundamental groups, *J. Diff. Geom.*, **33** (1991), 239-252.

[7] P. Deligne, letter to J.-P. Serre (1968).

[8] P. Deligne, Un théorème de finitude pour la monodromie, *Discrete Groups in Geometry and Analysis*, Birkhauser (1987), 1-19.

[9] P. Deligne, La conjecture de Weil pour les surfaces K3, *Invent. Math.*, **15** (1972), 206-222.

[10] P. Deligne, Theorie de Hodge, II, *Publ. Math. I.H.E.S.*, **40** (1971), 5-58.

[11] P. Deligne, Travaux de Shimura, *Séminaire Bourbaki, Lect. Notes in Math.*, **244** (1971), 123-165.

[12] P. Deligne, P. Griffiths, J. Morgan and D. Sullivan, Real homotopy theory of Kähler manifolds, *Invent. Math.*, **29** (1975), 245-274.

[13] P. Deligne, J. S. Milne, A. Ogus and K. Shih, *Hodge Cycles, Motives, and Shimura Varieties, Lect. Notes in Math.*, **900** (1982), Heidelberg, Springer.

[14] K. Diederich and T. Ohsawa, Harmonic mappings and disc bundles over compact Kähler manifolds, *Publ. R.I.M.S.*, **21** (1985), 819-833.

[15] P. Dolbeault, Formes différentielles et cohomologie sur une variété analytique complexe, I, *Ann. of Math.*, **64** (1956), 83-130.

[16] 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.

[17] S. K. Donaldson, Infinite determinants, stable bundles, and curvature, *Duke Math. J.*, **54** (1987), 231-247.

[18] S. K. Donaldson, Twisted harmonic maps and self-duality equations, *Proc. London Math. Soc.*, **55** (1987), 127-131.

[19] J. Eells and J. H. Sampson, Harmonic mappings of Riemannian manifolds, *Amer. J. Math.*, **86** (1964), 109-160.

[20] H. Garland and M. S. Raghunathan, Fundamental domains for lattices in (**R**-)rank 1 semisimple Lie groups, *Ann. of Math.*, **92** (1970), 279-326.

[21] W. Goldman and J. Millson, The deformation theory of representations of fundamental groups of compact Kähler manifolds, *Publ. Math. I.H.E.S.*, **67** (1988), 43-96.

[22] M. Green and R. Lazarsfeld, Deformation theory, generic vanishing theorems, and some conjectures of Enriques, Catanese, and Beauville, *Invent. Math.*, **90** (1987), 389-407.

[23] M. Green and R. Lazarsfeld, *Higher obstructions to deforming cohomology groups of line bundles*, Preprint (1990).

[24] P. Griffiths, Periods of integrals on algebraic manifolds I, II, *Amer. J. Math.*, **90** (1968), III, *Publ. Math. I.H.E.S.*, **38** (1970).

[25] P. Griffiths and J. Harris, *Principles of Algebraic Geometric*, John Wiley & Sons (1978).

[26] P. Griffiths and W. Schmid, Locally homogeneous complex manifolds, *Acta Math.*, **123** (1969), 253-302.

[27] R. Hain, The de Rham homotopy theory of complex algebraic varieties, I, K-*Theory*, **1** (1987), 271-324.

[28] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, New York, Academic Press (1978).

[29] P. W. Higgs, Broken symmetries and the masses of gauge bosons, *Phys. Rev. Lett.*, **13**, no. 16 (1964), 508-509.

[30] N. J. Hitchin, The self-duality equations on a Riemann surface, *Proc. London Math. Soc.* (3), **55** (1987), 59-126.

[31] G. Hochschild and G. D. Mostow, On the algebra of representative functions of an analytic group, *Amer. J. Math.*, **83** (1961), 111-136.

[32] G. Hochschild, Coverings of pro-affine algebraic groups, *Pacific J. Math.*, **35** (1970), 399-415.

[33] P. B. Kronheimer, M. J. Larsen and J. Scherk, Casson's invariant and quadratic reciprocity, *Topology*, **30** (1991), 335-338.

[34] M. Lubke, Chernklassen von Hermite-Einstein-vektorbundeln, *Math. Ann.*, **260** (1982), 133-141.

[35] G. A. Margulis, Discrete groups of motions of manifolds of nonpositive curvature, *Proc. Inter. Cong. Math. Vancouver* (1974), v. 2, 21-34; transl. *Amer. Math. Soc. Transl.*, **109** (1977), 33-45.

[36] V. B. Mehta and A. Ramanathan, Semistable sheaves on projective varieties and their restriction to curves, *Math. Ann.*, **258** (1982), 213-224.

[37] V. B. Mehta and A. Ramanathan, Restriction of stable sheaves and representations of the fundamental group, *Invent. Math.*, **77** (1984), 163-172.

[38] J. Morgan, The algebraic topology of smooth algebraic varieties, *Publ. Math. I.H.E.S.*, **48** (1978), 137-204; and **64** (1985), 185.

[39] M. S. Narasimhan and C. S. Seshadri, Stable and unitary bundles on a compact Riemann surface, *Ann. of Math.*, **82** (1965), 540-564.

[40] N. Nitsure, Moduli spaces of semistable pairs on a curve, *Proc. London Math. Soc.* **62** (1991), 275-300.
[41] M. V. Nori, On the representations of the fundamental group, *Compositio Math.*, **33** (1976), 29-41.
[42] M. S. Ragunathan, Cohomology of arithmetic subgroups of algebraic groups, I, *Ann. of Math.*, **86** (1967), 409-424; II, *Ann. of Math.*, **87** (1968), 279-304.
[43] N. Saavedra Rivano, *Catégories tannakiennes*, *Lect. Notes in Math.*, **265**, Heidelberg, Springer-Verlag (1972).
[44] J. H. Sampson, Applications of harmonic maps to Kähler geometry, *Contemp. Math.*, **49** (1986), 125-133.
[45] J.-P. Serre, *Linear Representations of Finite Groups*, New York, Springer-Verlag (1977).
[46] C. T. Simpson, *Systems of Hodge bundles and uniformization*, doctoral dissertation, Harvard University (1987).
[47] C. T. Simpson, Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization, *Journal of the A.M.S.*, **1** (1988), 867-918.
[48] C. T. Simpson, Transcendental aspects of the Riemann-Hilbert correspondence, *Illinois J. of Math.*, **34** (1990), 368-391.
[49] Y. T. Siu, Complex analyticity of harmonic maps and strong rigidity of complex Kähler manifolds, *Ann. of Math.*, **112** (1980), 73-110.
[50] T. Tannaka, Über den Dualitätssatz der nichtkommutativen topologischen Gruppen, *Tohoku Math. J.*, **45** (1938), 1-12.
[51] K. K. Uhlenbeck, Connections with L^p bounds on curvature, *Comm. Math. Phys.*, **83** (1982), 31-42.
[52] K. K. Uhlenbeck and S. T. Yau, On the existence of Hermitian-Yang-Mills connections in stable vector bundles, *Comm. Pure and Appl. Math.*, **39-S** (1986), 257-293.
[53] A. Weil, *Introduction à l'étude des variétés kähleriennes*, Paris, Hermann (1952).
[54] A. Weil, Discrete subgroups of Lie group I, *Ann. of Math.*, **72** (1960), 369-384; II, *Ann. of Math.*, **75** (1962), 578-602; Remarks on the cohomology of groups, *Ann. of Math.*, **80** (1964), 149-157.
[55] A. Weil, *Basic Number Theory*, New York, Springer-Verlag (1967).

Laboratoire de Topologie et Géométrie
Université Paul-Sabatier, UFR-MIG
118, route de Narbonne
31062 Toulouse Cedex

Manuscrit reçu le 4 mai 1990.
Révisé le 9 avril 1991.