Donaldson's proof of the Narasimhan-Seshadri theorem

Vamsi Pritham Pingali

July 18, 2018

1 Introduction and definitions

Suppose X is a compact Riemann surface, i.e., it is compact surface with a hyperbolic structure if $g \geq 2$. We know that $PSL(2,\mathbb{R})$ structures correspond to Riemann surfaces of genus $g \geq 2$ via $\mathbb{H}/\Gamma$. What about say $PSL(2,\mathbb{R}) \times U(n)$?

A holomorphic vector bundle E of rank r on a complex manifold X is simply a smooth complex vector bundle whose transition functions are holomorphic, i.e., $E = \frac{\cup_\alpha U_\alpha \times \mathbb{C}^r}{((x,\vec{v}_\alpha)=(x,g_{\alpha\beta}(x)\vec{v}_\beta)}$ where $g_{\alpha\beta}(z) : U_\alpha \cap U_\beta \to GL(r,\mathbb{C})$ is a holomorphic matrix-valued function of z. What are examples of such bundles ? Well, we have trivial bundles $X \times \mathbb{C}^r$ of course. The holomorphic tangent bundle (whose fibre at every point of X is generated by $\frac{\partial}{\partial z^1}, \frac{\partial}{\partial z_2}, \ldots$) is another example. In the case of Riemann surfaces, this is a line bundle.

How does one come up with more examples ? More importantly, can one classify bundles over Riemann surfaces ? Can one make the set of all bundles (the moduli space) a manifold ? Unfortunately these are too hard (and the last one is just not true). In 1963, Mumford wrote a paper where he defined some special bundles called stable bundles over Riemann surfaces, whose moduli space can be made into a smooth complex manifold. In 1965, Narasimhan and Seshadri proved a remarkable theorem that essentially described all stable bundles very nicely in terms of the fundamental group of the Riemann surface. This led to a flurry of similar results (starting with Donaldson who gave a differentio-geometric proof of the NS theorem) of the flavour of "A certain kind of a metric with good curvature properties exists if and only if and algebro-geometric obstruction is met". This is called the Kobayashi-Hitchin correspondence.

The construction employed by NS was (I believe) due to Andre-Weil. Essentially, if you take a unitary irreducible representation ρ of the fundamental group Γ, then $E = \frac{\mathbb{H}\times\mathbb{C}^r}{(z,v)=(\gamma z,\rho(\gamma)v)}$ defines a holomorphic vector bundle (exercise) on X. Is this bundle stable ? Do all stable bundles arise this way ? The answer is almost yes according to the NS theorem. Almost because this is true for the so-called degree zero bundles. We shall take Donaldson's differentio-geometric approach to this theorem.

Firstly, what is a stable bundle ? Before that, we need to know what the "degree" of a bundle means. To this end, let's define the notion of a connection ∇ on a bundle. A connection ∇ is a first order differential operator that is supposed to remind us of the directional derivative. That is, $\nabla_X s$ is the directional derivative of a section s of the bundle along the tangent vector field X. In other words, ∇ takes a section and spits out a one form tensored with sections. Naively speaking, if s is a vector of functions (in a trivilisation) $\nabla_X s = \frac{\partial}{\partial x^i} X^i s$. But this does not change correctly when a change of trivialisation changes s to $g_{\alpha\beta} s$. To correct this, we add a zeroeth order piece. Finally, a connection $\nabla_\alpha = d + A_\alpha$ where d is the exterior derivative and A is an $r \times r$ matrix of 1-forms. Under change of trivialisation, $A_\alpha = g_{\alpha\beta} A_\beta g_{\alpha\beta}^{-1} - dg_{\alpha\beta} g_{\alpha\beta}^{-1}$. Connections exist for every bundle (by a partition-of-unity) argument. The square of a connection ∇^2 is shockingly enough not a second order operator. It is a zeroeth order operator called the Curvature Θ. Indeed, $\nabla^2 s = \Theta s$ where $\Theta = dA + A \wedge A$ is the an $r \times r$ matrix of 2-forms called the curvature matrix which transforms as $\Theta_\alpha = g_{\alpha\beta} \Theta_{\alpha\beta} g_{\alpha\beta}^{-1}$. Thus $tr(\Theta_\alpha) = tr(\Theta_\beta)$ is a globally defined 2-form that is independent of the trivialisation chosen. In fact, it is a closed form whose cohomology class is independent of the connection chosen! It is called the first Chern class $c_1(E) = [tr(\frac{\sqrt{-1}}{2\pi}\Theta)]$. In particular, a trivial bundle has an obvious connection $A = 0$. Therefore $\Theta = 0$ and hence $c_1(Trivial\ bundle) = 0$. A connection whose curvature is 0 is called flat. If X is a Riemann surface, then $\int_X c_1(E)$ is an integer called the degree of E. A holomorphic vector bundle E of degree d and rank r on a Riemann surface is said to be stable if for every proper subbundle F, $\mu(F) = \frac{degree\ of\ F}{rank\ of\ F} < \mu(E) = \frac{d}{r}$. The μ is called the "slope" of the bundle. If equality is allowed to hold, then the bundle is called semi-stable. Note that if the degree and rank of E are coprime, the bundle is automatically stable.

Connections can be used to define parallel transport. If you give me a vector $\vec{v}$ from the fibre of the vector bundle at a point $p \in X$, and a loop γ based at p, then I can solve the system of differential equations $\nabla_{\gamma'} s(\gamma(t)) = 0$ with $s(t = 0) = \vec{v}$ and come up with another vector $P_\gamma(v)$ at p which differs from identity (for small loops) by an integral involving the curvature of the connection. If we take a flat connection, P_γ depends only on the homotopy class of γ. Thus parallel transport gives a representation of the fundamental group (called the monodromy representation).

On a holomorphic vector bundle, there is a nice way to come up with a connection. Firstly, suppose we are given a hermitian metric h on E, then in a holomorphic trivialisation, h_α is a hermitian positive-definite $r \times r$ matrix. The formula $A_\alpha = h_\alpha^{-1} \partial h_\alpha$ (where $\partial f = \frac{\partial f}{\partial z^i} dz^i$ is a so-called $(1, 0)$-form) defines a connection. You can check that it transforms correctly. Its curvature turns out to be $\Theta_\alpha = \bar{\partial} A_\alpha$. This connection is called the "canonical connection associated to the given metric h" or sometimes the "Chern connection associated to h". If you take the bundle $\mathbb{H} \times \mathbb{C}^r/\Gamma$, then it admits an obvious metric whose Chern connection is flat.

Here is a property of Chern connections : If S is a sub-bundle of E, then the curvature of the Chern connection associated to the restriction of the metric to S is less than that of E. (This is in sharp contrast to intuition. The sphere is more curved than three dimensional space.) Therefore, a simple calculation shows that if you have a flat Chern connection, you are at least semi-stable.

Thus it seems reasonable to expect the NS theorem to be equivalent to saying that a degree zero indecomposable (i.e. not a direct sum of smaller bundles) bundle on a compact

RS admits a flat Chern connection if and only if it is stable. More generally,

Theorem 1.1. *An indecomposable holomorphic bundle E of rank r and degree d on a compact Riemann surface X whose tangent bundle is equipped with a metric $g = e^{-\phi} dz \otimes d\bar{z}$ is stable if and only if it admits a metric h whose Chern connection has curvature Θ satisfying $\frac{\sqrt{-1}}{2\pi}\Theta = \frac{d}{rVol(g)}\omega I$ where ω is the $(1,1)$ form given by $\omega = e^{-\phi}dz \wedge d\bar{z}$ and I is the identity endomorphism of E. This metric (called a Hermite-Einstein metric) is unique up to rescaling.*

This is what Donaldson proved and we shall follow his proof. Here is a little more theory. Suppose E is a smooth complex vector bundle with a given metric h. A unitary gauge transformation u is simply a smooth endomorphism of E that is a unitary matrix in a trivialisation formed by orthonormal vectors. u acts on a given metric-compatible connection A (meaning that if you choose an orthonormal trivialisation, then A is a skew-hermitian matrix of one forms) via $u(A) = uAu^{-1} - duu^{-1}$. By the way, given a connection A on a bundle E, one can produce a connection on the bundle of endomorphisms $End(E)$. Indeed, if u is an endomorphism, then $\nabla_A u = du + [A, u]$ where $[,]$ is the commutator of matrices. Thus $u(A) = A - \nabla_A uu^{-1}$. The group of unitary gauge transformations is denoted by $\mathcal{G}$ and its complexification, the group of all complex gauge transformations by $\mathcal{G}_{\mathbb{C}}$. A complex gauge transformation g acts on connections compatible with a given metric h via $g(A) = A - \bar{\partial}_A gg^{-1} + (\bar{\partial}_A gg^{-1})^{\dagger}$. An interesting observation is the following. For a Chern connection on a holomorphic bundle $A = h^{-1}\partial h$, note that it has no $(0,1)$-part. Conversely, on Riemann surfaces, given any connection that is compatible with the given metric h its $(0,1)$-part $\bar{\partial}_A$ defines a holomorphic structure ! Indeed, those sections s are holomorphic that satisy $\bar{\partial}_A s = 0$. Why is this equivalent to the usual definition ? That is a theorem. So the bottom line is "Either give me a holomorphic structure and look for an appropriate HE metric h or give me a metric h_0 and look for an appropriate connection A inducing an isomorphic holomorphic structure". Given a hermitian holomorphic bundle (E, h_0), we denote by $\mathcal{O}(E)$ the complex gauge orbit of the Chern connection of h_0, i.e., these are all connections whose induced holomorphic structures are isomorphic.

2 Basic strategy

Firstly, let's look at line bundles. Line bundles have no non-trivial subbundles. So they are automatically stable. Given any metric h_0, any other metric $h = e^{-u}h_0$. Therefore, $\Theta_h = \bar{\partial}(h^{-1}\partial h) = \partial\bar{\partial}u + \Theta_0$. So we want to solve $\frac{\sqrt{-1}}{2\pi}(\Theta_0 + \partial\bar{\partial}u) = \lambda\omega$ where $\omega = \frac{\sqrt{-1}}{2\pi}\Theta_0 + d\eta$. It turns out (through a piece of analysis called Hodge theory) that every d-exact $(1,1)$-form is $\partial\bar{\partial}u$ for some function u. If the form is real, the function is $\sqrt{-1}\times$ a real function. Thus the NS theorem is proven for line bundles. Donaldson uses induction on the rank of the bundle to prove it for higher rank bundles.

The basic idea is to use calculus of variations. He came up with an "energy" functional J whose critical points are HE connections. So one needs to

1. Define over what space of connections one wants to minimise J. It is the space of rough (Sobolev) connections that are complex-gauge equivalent to a fixed one obtained by a fixed metric on the given holomorphic bundle. We call such connections $\mathcal{O}(E)$.

2. Prove that J is indeed bounded below on that space. (This will be true by construction.)

3. Since J is bounded below, there exist a sequence of connections A_i s.t. $J(A_i) \to \inf J$.

4. We need to extract a convergent subsequence somehow. A technical tool called Uhlenbeck's compactness theorem shows that there exist unitary gauge transformations g_{n_i} such that g_{n_i} acting on A_{n_i} give us a (new) subsequence of connections whose J's converge "weakly" to $\inf J$ and which converge themselves to a connection A.

5. Now a rough connection induces a rough $\bar{\partial}$ operator but this is not good enough to talk about a holomorphic structure on the bundle. However, Atiyah-Bott proved that given a rough connection, there is a smooth connection arbitarily close to it in the complex gauge orbit of the rough connection. So choose a smooth connection $\tilde{A}$ in the complex gauge orbit of A whose energy is not very different from that of A.

6. $\tilde{A}$ either induces the same holomorphic structure on E or a different one. Using stability, we rule out introducing a different holomorphic structure. (Here we use the induction hypothesis.)

7. The technique of Atiyah-Bott can be used to prove that, given a rough connection A such that its complex gauge orbit contains smooth connections inducing an indecomposable holomorphic structure, we can find a gauge unitary transformation g such that $g.A$ is smooth. Then we show that indeed the fact that it attains the minimum of J implies that it is a HE connection.

8. Lastly, one proves uniqueness by taking two metrics, subtracting, integrating-by-parts etc.

3 Details of existence

Firstly, define an energy functional J on the space of smooth unitary connections as

$$J(A) = \left(\int_X \mathrm{Tr} \left[-\frac{\sqrt{-1}\Theta}{2\pi\omega} + \mu I \right]^2 \right)^{1/2}$$

Therefore, J is obviously bounded below and is 0 iff there is a HE connection. Now J depends on one derivative of A. Actually, it depends on the L^2 norm of $\Theta = dA + A \wedge A$. Note that in the proof of the uniformisation theorem we defined Sobolev spaces $W^{1,2}$ as the spaces of functions/sections of a vector bundle which are completions of the spaces of smooth functions satisfying $\int |\nabla f|^2 + |f|^2 < \infty$. These spaces are Hilbert spaces. They embed continuously in L^2. Actually, it turns out that on Riemann surfaces, the embed in L^q for all $\infty > q \geq 2$. (The Sobolev embedding theorem.) In fact, they are even better than that. If you have a bounded sequence f_i in $W^{1,2}$ then

1. A subsequence f_{n_k} converges weakly to f in $W^{1,2}$, i.e., for every linear functional F, $F(f_{n_k}) \to F(f)$.

2. A subsequence f_{m_k} converges in L^q. (The Rellich compactness lemma.)

In other words, the functional J makes sense not just for smooth connections, but for $W^{1,2}$ connections.

Since J is bounded below, there exists a sequence of $W^{1,2}$ connections A_i such that $J(A_i) \to \inf J$. At this point, we need a small observation. Suppose $A_i \to A$ in $W^{1,2}$ weakly. This means that $\Theta_i \to \Theta$ in L^2 weakly. Does this mean that $J(A_i) \to J(A)$? Not quite. Instead, one can conclude that $J(A) \leq \liminf J(A_i)$. This is because the set of α such that

$$\left(\int_X \mathrm{Tr}\left[(\alpha + \mu I)^\dagger (\alpha + \mu I) \right] \right)^{1/2} \leq J(A) - \epsilon$$

is a closed convex subset of a Hilbert space and hence one can separate a point $-\frac{\sqrt{-1}\Theta}{2\pi\omega}$ not lying on it by a hyperplane for every $\epsilon > 0$. So if (after taking a subsequence) $J(A) \geq \lim J(A_i) - \epsilon_0$, then Θ_i eventually lie on the wrong side of a hyperplane whereas, they should not (because $\Theta_i \to \Theta$ weakly). Anyway, since we have a sequence A_i such that $J(A_i) \to \inf J$ where the infimum is taking over the complex gauge orbit $\mathcal{O}(\mathcal{E})$ of a fixed connection defining the holomorphic structure of E. Ideally we want to extract a convergent subsequence of A_i and hope that the limiting connection is our desired HE connection. While we may not have such a subsequence, here is a technical, useful result due to Uhlenbeck (Uhlenbeck's compactness theorem).

Theorem 3.1. *Suppose that A_i is a sequence of $W^{1,2}$ connections with Θ_i bounded in L^2. Then there is a subsequence i_j and $W^{2,2}$ gauge transformations u_j such that $u_j(A_{i_j})$ converges weakly in $W^{1,2}$ to a connection A.*

(The rough idea of the proof is to choose the "best" connection in every gauge orbit, namely, the so-called "Coulomb gauge", and prove estimates on such "best" connections.)

So after composing with gauge transformations (gauge transforms do not change J) we see that (up to a subsequence) $J(A_i) \to \inf_{\mathcal{O}(\mathcal{E})} J$ and $A_i \to A$ in $W^{1,2}$ weakly. Thus $J(A) \leq \lim J(A_i)$. Now choose a smooth connection (which we shall also call A abusing notation a bit) lying the complex gauge orbit of A (by Atiyah-Bott) that is close to A so that its energy J is not very different from that of A.

So either the approximate limiting connection A defines the same holomorphic structure as E or something bad has to happen which has to contradict stability somehow. Stability requires testing against a subbundle. How does one produce subbundles of vector bundles ? One way is to take an eigenspace of an endomorphism. Indeed,

Lemma 3.2. *Let E be a holomorphic bundle over X. Then either* $\inf J$ *over* $\mathcal{O}(\mathcal{E})$ *is attained in* $\mathcal{O}(\mathcal{E})$ *or there is a holomorphic bundle $F \neq E$ of the same degree and rank as E with* $\inf J|_{\mathcal{O}(\mathcal{F})} \leq \inf J|_{\mathcal{O}(\mathcal{E})}$ *and with* $Hom(E, F) \neq 0$.

Proof. Choose A_i such that $J(A_i) \to \inf_E J$. By Uhlenbeck compactness, after a subbundle one may assume that $A_i \to A$ weakly in $W^{1,2}$ and that $J(A) \leq \inf_E J$. Since a smooth approximation of A lying in the same complex gauge orbit defines a holomorphic structure E_A, the lemma will follow if we show that $Hom(E, E_A) \neq 0$. Indeed, suppose it is 0.

Note that $Hom(E, F)$ can be interpreted as the kernel of a certain elliptic $\bar{\partial}$ operator. If $Hom(E, F) = 0$ this would mean that the said operator would have trivial kernel. But A_i are all in $\mathcal{O}(E)$ and therefore the identity morphism makes sense between (E, A_i) and (E, A_j). But the limit of A_i apparently has no morphism with E which sounds strange. This is the intuition behind the following proof.

Note that given A_i, A, there is an obvious connection on $Hom(E_i, E_A) = E_i^* \otimes E_A$ with the corresponding $\bar{\partial}_{A_i,A}$ operator such that $\bar{\partial}s = 0$ correspond to $Hom(E_i, E_A)$.

By standard estimates, if $\bar{\partial}$ does not have a kernel, then $\|\bar{\partial}s\|_{L^2} \geq C\|s\|_{W^{1,2}}$ for all connections s. But by Sobolev embedding, $\|s\|_{W^{1,2}} \geq C\|s\|_L^4$. Now $\|(\bar{\partial}_{A_0,A} - \bar{\partial}_{A_0,A_i})s\|_{L^2} \leq C\|A_i - A\|_{L^4}\|s\|_{L^4}$. This implies that

$$\|\bar{\partial}_{A_0,A_i}s\|_{L^2} \geq (C_1 - C_2\|A_i - A\|_{L^4}\|\|s\|_{L^4}$$

Since $A_i \to A$ in L^4, this means that $Hom(E, E_i) = 0$ for large i contradicting the fact that E_i introduces an isomorphic holomorphic structure. $\square$

We want to rule out the possibility of $E \neq F$ by using stability. To this end, suppose we are given a matrix-valued holomorphic function whose determinant is 0 throughout in a ball. Then one can write a Jordan canonical form centred at the origin. So clearly one can locally break the matrix up as a map between two vector spaces (the non-zero eigenspaces at the origin and parts of the kernel where vanishing is at an order lower than the highest order) having a determinant that is not identically zero. This when done globally yields the following exact sequence : $0 \to \mathcal{P} \to E \to \mathcal{Q} \to 0$ and $0 \leftarrow \mathcal{N} \leftarrow F \leftarrow \mathcal{M} \to 0$ with the given morphism α between E and F and β between $\mathcal{Q}$ and $\mathcal{M}$ such that $rk(\mathcal{Q}) = rk(\mathcal{M})$ and $deg(\mathcal{Q}) \leq deg(\mathcal{M})$.

So decompose E and F as above thus producing $0 \to \mathcal{M} \to F \to \mathcal{N} \to 0$ with $\mu(\mathcal{N}) \geq \mu(\mathcal{Q}) \geq \mu(E) = \mu(F) > \mu(\mathcal{P})$. So the bundle F is unstable. The idea is to prove that this implies that its energy is pretty large (whereas it is supposed to be smaller than that of E).

Before we proceed to utilize this observation, here are some generalities about the behaviour of Chern connections on exact sequences. If

$$0 \to S \to E \to Q \to 0$$

is an exact sequence, then any Chern connection on E has the form

$$\begin{bmatrix} A_S & \beta \\ -\beta^{\dagger} & A_Q \end{bmatrix}$$

where β is End(Q,S)-valued $(0,1)$-form that represents how far E is from $S \oplus Q$. The corresponding curvature is

$$\begin{bmatrix} \Theta_S - \beta \wedge \beta^\dagger & d_A\beta \\ -d_A\beta^\dagger & \Theta_Q - \beta^\dagger \wedge \beta \end{bmatrix}$$

which means that the subbundle has smaller curvature and the quotient larger (which is why HE bundles are stable in the first place). An easy calculation shows the following lemma (which roughly states that instability increases energy).

Lemma 3.3. *If F is a holomorphic bundle over X such that*

$$0 \to \mathcal{M} \to F \to \mathcal{N} \to 0$$

with $\mu(\mathcal{M}) \geq \mu(F)$ then for any Chern connection A on F we have

$$J(A) \geq rk(\mathcal{M})(\mu(\mathcal{M}) - \mu(F)) + rk(\mathcal{N})(\mu(F) - \mu(\mathcal{M})) = J_0$$

with equality only if the extension splits.

The next lemma shows that assuming the induction hypothesis (on rank) stable bundles have small energy.

Lemma 3.4. *If E is a stable holomorphic bundle over X such that*

$$0 \to \mathcal{P} \to E \to \mathcal{Q} \to 0$$

, then assuming the induction hypothesis that lower rank stable bundles admit HE connections, there is a connection A on E with

$$J(A) < rk(\mathcal{P})(\mu(E) - \mu(\mathcal{P})) + rk(\mathcal{Q})(-\mu(E) + \mu(\mathcal{Q})) = J_1$$

The proof of this lemma relies on a couple of ideas (we will only sketch the proof).

1. Every holomorphic bundle on a Riemann surface admits a Harder-Narasimhan filtration, i.e., a filtration by subbundles whose succesive quotients C_i are semistable. Moreover, semistable bundles C_i admit a filtration whose successive quotients C_{ij} are stable bundles.

2. Now a connection on E can be given by the information $A_{\mathcal{P}}, A_{\mathcal{Q}}, t\beta$ where the A are connections on the sub and quotient bundles, β is the extension form, and t is a real parameter. If $t \neq 0$, all such connections define isomorphic holomorphic structures. However, when $t = 0$, the holomorphic structure on E jumps to $\mathcal{P} \oplus \mathcal{Q}$. Likewise, $\mathcal{P}$ and $\mathcal{Q}$ are themselves "perturbations" of direct sums $\oplus_{ij}C_{ij}$ of stable bundles. Thus by the induction hypothesis, they admit connections that are "perturbations" of a direct sum of HE equations. A small calculation shows that if the perturbation parameter is small, then the energy $J \leq J_1$.

These two lemmata applied to our situation of E and F as above, lead to a contradiction.

Thus the infimum of J over $\mathcal{O}(\mathcal{E})$ is attained at a $W^{1,2}$ connection A defining an isomorphic holomorphic structure. If this connection were smooth, then a simple calculus of variations argument (where A is varied along a path of complex-gauge equivalent connections) shows that indeed A is HE.

So the difficulty seems to be able to prove that there is a smooth connection unitarily gauge equivalent to A. This is accomplished by a preposterously clever argument due to Atiyah and Bott : Firstly, an open neighbourhood of the identity in the unitary gauge group acts on a connection to give a closed Hilbert submanifold of finite codimension in the space of $W^{1,2}$ connections (The idea is to compute the derivative of this map, use elliptic regularity to prove that the range of the derivative is closed, and then prove that the kernel is trivial using indecomposability. Thus the inverse function theorem on Hilbert manifolds implies that the image is an immersed submanifold). Secondly, suppose N is a finite dimensional subspace (of dimension k) transverse to the unitary gauge orbit. So given A, a small neighbourhood V of A admits a continuous map $\pi : V \to N$ near A fibring at A with fibre the Hilbert manifold. Thirdly, choose $k+1$ points in N having A has their barycentre. Note that the corresponding simplex generates the homology $H_{k-1}(N-A)$. By continuity, this will be true if we change the simplex slightly. Since smooth connections are dense, choose a simplex of smooth connections nearby generating the homology. The intersection number with the Banach submanifold of gauge equivalent connections does not change and hence there is a smooth connection in the unitary gauge orbit.