PhD Seminar WS 2018/2019
Prolegomenon to Seiberg–Witten theory

Arpan Saha

October 15, 2018

1 Why study 4-manifolds?

The grand objective we'll be working towards this semester is to show that the category of compact, connected, oriented *smooth* (CCOS) manifolds is genuinely different from the category of compact, connected, oriented *topological* (CCOT) manifolds. Now, every CCOS manifold is a CCOT manifold with additional structure, so what I mean by this statement is that:

- There are CCOT manifolds which cannot be obtained by forgetting the additional structure of some CCOS manifold i.e. they don't admit any smooth structure.
- There are CCOT manifolds which can be obtained by forgetting the additional structure of more than one non-diffeomorphic CCOS manifolds i.e. they admit multiple smooth structures.

The distinction between the two categories is obscured if we restrict ourselves to only dimensions 1, 2, 3, since in these cases, every CCOT manifold does admit a unique smooth structure up to diffeomorphisms. The fun begins in dimension 4.

In dimension 4, a 2-submanifold generically intersects other 2-submanifolds in a discrete set of points that, beacause of compactness, are finite in number and may be counted. If the submanifolds are oriented, then the intersections may be assigned a sign. Counted with sign, the number of intersection points depends only on the homology classes of the submanifolds. This allows us to define for any CCOT 4-manifold X, a $\mathbb{Z}$-valued symmetric bilinear pairing Q_X on $H_2(X;\mathbb{Z})$. This pairing in $\mathbb{Z}$, called the *intersection form*, is an invariant of CCOT manifolds. A natural question to ask is:

Question 1. *When can a given symmetric bilinear form over $\mathbb{Z}$ be realised as the intersection form of some CCOT manifold?*

For a start, Poincaré duality implies that, firstly, the matrix representing the bilinear form has to be invertible, and secondly, the inverse also has to have entries in $\mathbb{Z}$. In other words, the determininant of the matrix needs to be ± 1

i.e. the pairing is *unimodular*. This turns out to be sufficient. In fact, one can say something far sronger:

Theorem 2 (Freedman). *For every unimodular symmetric bilinear form over $\mathbb{Z}$, there is a unique simply connected CCOT 4-manifold up to homeomorphisms which realises it as its intersection form.*

This is however not the case for CCOS manifolds. Indeed, one of the results that we shall hopefully get around to proving in this seminar is:

Theorem 3 (Donaldson). *If X is a CCOS 4-manifold with Q_X definite, then Q_X is diagonalisable over $\mathbb{Z}$.*

Note that being diagonalisable over $\mathbb{Z}$ is stronger than being diagonalisable over $\mathbb{R}$, which every symmetric bilinear form is. The consequence is that there exists a wealth of CCOT 4-manifolds which do not admit any smooth structure.

On the other hand, there are CCOT 4-manifolds which admit multiple smooth structures. To prove this, we could construct simply connected CCOS manifolds that share the same intersection pairing (and hence are homeomorphic, by virtue of Theorem 2) and then showing that they are not diffeomorphic. Of course, such CCOS manifolds can't be distinguished by any of the topological invariants we know since they are homeomorphic. What is required are a new kind of invariants that are sensitive to not only the topology of X but its smooth structure as well.

2 Donaldson invariants

As motivation, it's instructive to briefly recall a classical result due to De Rham. Given a CCOS manifold X, one may introduce a Riemannian metric g on it. This allows us to define a second order differential operator called the *Laplacian* $\Delta : \Omega^\bullet(X;\mathbb{R}) \to \Omega^\bullet(X;\mathbb{R})$ acting on differential forms on X. A solution α to $\Delta\alpha = 0$ is said to be *harmonic*. The vector space of such harmonic forms shall be denoted $\mathscr{H}_\bullet(X,g)$.

Theorem 4 (De Rham). *There is a natural (vector space) isomorphism between $H^\bullet(X;\mathbb{R})$ and $\mathscr{H}_\bullet(X,g)$.*

In particular, $\dim(\mathscr{H}_\bullet(X,g))$ is independent of g and even the smooth structure on X. What we have done is introduce some auxiliary structure on X, namely the Riemannian metric g, use that extra structure to set up a system of linear PDEs $\Delta\alpha = 0$, and extract topological invariants from the space of solutions to the system of PDEs. It makes sense to hope that for some more complicated nonlinear system of PDEs, the space of its solutions would divulge information about the smooth structure in addition to the topology. This turns out to be the case.

The auxiliary structure that Donaldson introduced on a CCOS 4-manifold X consists of the following data:

- an $\mathsf{SU}(2)$-principal bundle P with connection θ satisfying $\int_X c_2(P) > 0$, where $c_2(P)$ is the second Chern class of P,

- a Riemannian metric g on X.

The connection θ pulls back via local sections to locally define $\mathfrak{su}(2)$-valued 1-forms A. It is a basic result in differential geometry that the *curvature* F_A locally defined as $F_A := \mathrm{d}A + A \wedge A$ is independent of the choice of local sections and is a globally defined $\mathfrak{su}(2)$-valued 2-form on X.

Meanwhile, a metric g on any CCOS manifold of dimension n induces a signed involution $* : \Omega^\bullet(X;\mathbb{R}) \to \Omega^{n-\bullet}(X;\mathbb{R})$ called the *Hodge star*. To define it, we may choose a local oriented orthonomal basis of 1-forms α_i. Then $\Omega^k(X;\mathbb{R})$ is spanned by elements of the form $\alpha_{\pi(1)} \wedge \cdots \wedge \alpha_{\pi(k)}$ for some permutation π of $\{1, 2, \ldots, n\}$. The action of the Hodge star is given by:

$$\alpha_{\pi(1)} \wedge \cdots \wedge \alpha_{\pi(k)} \mapsto \operatorname{sgn}(\pi)\alpha_{\pi(k+1)} \wedge \cdots \wedge \alpha_{\pi(n)},$$

where $\operatorname{sgn}(\pi)$ is ± 1 depending on whether the permutation π is even or odd. The map $*$ may not seem well-defined since we can't unambiguously know π from just $\pi(1), \ldots, \pi(k)$, but it actually is, thanks to the alternating property of the wedge product.

In the case of 4-manifolds, $*$ induces an involution on $\mathfrak{su}(2)$-valued 2-forms, giving us an eigendecomposition of $\Omega^2(X;\mathfrak{su}(2))$ to a self-dual part $\Omega^2_+(X;\mathfrak{su}(2)$ belonging to eigenvalue $+1$ and an anti-self-dual part $\Omega^2_-(X;\mathfrak{su}(2)$ belonging to eigenvalue -1. In particular, the curvature F_A admits a decomposition $F_A = F_A^+ + F_A^-$ into self-dual and anti-self-dual parts.

The nonlinear analogue of the Laplace equation that Donaldson considered is the anti-self-dual Yang–Mills equation:

$$F_A^+ = 0. \tag{1}$$

This should be viewed as a system of coupled nonlinear PDEs to be solved for A.

Donaldson showed that for generic choices of g, the space M of isomorphism classes of solutions A to (1), known as the *instanton moduli space*, is an oriented smooth manifold of finite dimension equipped with certain canonical differential forms τ_i. These may be integrated over M—

$$\int_M \tau_1 \wedge \cdots \wedge \tau_k \tag{2}$$

—to yield rational invariants for smooth structures, the so-called *Donaldson invariants*.

Unfortunately, the expression in (2) is not well-defined, since the instanton moduli space M is not compact and the integrand might blow up near the "ends" of M. In fact, these "ends" arise due to the curvature F_A becoming arbitrarily concentrated at a finite set of points in X and so may be naturally identified with copies of X itself, one for each point of curvature concentration.

The issue with noncompactness can be fixed by means of a delicate regularisation procedure whereby M is first truncated to M' by introducing a cutoff boundary $\partial M'$ and then appropriate counterterms coming from integrals on the boundary $\partial M'$ are incorprated into (2) so that the result doesn't depend on how we truncate. This is enough to prove statements like Theorem 3 but quite painful to keep track of in practice. There is however a simpler way.

3 What physics has to offer

Donaldson invariants carry a natural interpretation in terms of physics. As Witten discovered, Donaldson theory is secretly a twisted version of $\mathscr{N} = 2$ supersymmetric $\mathsf{SU}(2)$ Yang–Mills theory—the correspondence between the two formulations is summarised in the table below:

Donaldson theory	**Witten's TQFT**
X	spacetime
A	$\mathsf{SU}(2)$ gauge field
M	space of susy field configurations
τ_i	physical observables $\mathscr{O}_i$
$\int_M$	localised path integral
$\int_M \tau_1 \wedge \cdots \wedge \tau_k$	correlators $\langle \mathscr{O}_1 \cdots \mathscr{O}_k \rangle$

This is all pretty neat but as things stand so far, it doesn't really help. A rose by any other name smells just as sweet and an integral over some noncompact space by any other name is just as tricky to define and compute. However, recasting Donaldson theory as a quantum field theory opens up the possibility of using certain powerful tools from physics, such as renormalisation group flow and S-duality.

The basic idea behind renormalisation group flow is that physical models are typically meaningful only above a certain cutoff length scale and if we try making predictions about what will happen when we probe using pulses of wavelengths shorter than this cutoff, we will get nonsensical answers. For instance, the field amplitude at some point as predicted by the model may become infinite. We should therefore model reality not by a single QFT but by a whole family of them parametrised by the length cutoff above which we wish to probe. The family thus interpolates between the limiting QFT we get as we zoom infinitely inwards, i.e. the *ultraviolet* (UV) theory, and the one we get as we zoom infinitely outwards, i.e. the *infrared* (IR) theory.

(Why not just work with the UV theory, you might ask, since, in principle, it contains all the information we could hope to get? Because that is like simulating the behaviour of individual molecules to solve fluid dynamics problems. Not only is it computationally intractable, but we may not have access to initial data fine-grained enough to make the model work in the first place. And even if we could do all that, at the end of the day, what each individual molecule is doing is not something we care about when the question is "Is this plane going to fly?")

In a topologically twisted QFT (TQFT), the correlators of the physical observables don't depend on the metric, so one could in principle compute them in the UV or in the IR; the answer would be the same. Donaldson–Witten theory has the additional property that it is "asymptotically free." This means that the physicists' usual strategy of perturbative analysis is reliable in the "weakly coupled" UV regime but not in the "strongly coupled" IR . Indeed, this was what made it possible to show that Donaldson invariants coincided with the correlators in the TQFT that Witten defined. On the other hand, the fact that the IR theory is strongly coupled made it hard to say anything about it at first glance.

This is where S-duality comes in. S-duality is a supersymmetric generalisation of the electric-magnetic duality of Maxwell theory, which exchanges electric charges quantised in units of the fundamental electric coupling e with magnetic monopoles quantised in units of the fundamental magnetic coupling $g \sim 1/e$. Note that when the electric coupling is strong (i.e. $e \gg 1$), the magnetic coupling is weak (i.e. $g \ll 1$), and vice versa. A similar statement holds for S-duality in general, which in fact stands for *strong-weak* duality.

Seiberg and Witten discovered that the strongly coupled IR limit of twisted $\mathscr{N} = 2$ supersymmetric $\mathsf{SU}(2)$ Yang–Mills theory is dual to a weakly coupled theory of a $\mathsf{U}(1)$ gauge field coupled to a monopole on X equipped with the following auxiliary structure:

- a Riemannian metric g,
- an imaginary-valued 2-form ω self-dual with respect to g.

The dynamics of these gauge fields is described by the Seiberg–Witten equations:

$$\begin{aligned} D^+_{\hat{A}}\Phi &= 0, \\ F^+_{\hat{A}} &= \sigma(\Phi, \Phi) + \omega. \end{aligned} \tag{3}$$

For sake of completeness, here is a list of what all the letters stand for, presented without explanation:

- $\hat{A}$ is the $\mathsf{U}(1)$ gauge field,
- Φ is the monopole, which is a left-handed Weyl spinor,
- $F^+_{\hat{A}}$ is the self-dual part of the curvature $F_{\hat{A}} := \mathrm{d}\hat{A}$,
- $D^+_{\hat{A}}$ is the Dirac operator on left-handed Weyl spinors induced by A,
- σ is a certain sesquilinear form on the left-handed Weyl spinor bundle taking values in the space of self-dual 2-forms.

The upshot of all this is that the curvature concentration problem that plagued Donaldson theory is now no longer there. So, the space of isomorphism classes of solutions $(\hat{A}, \Phi)$ of (3) is compact and in fact, for generic choices of (g, ω), a CCOS manifold itself. Which means that we can do integration on it and compute Donaldson invariants of any CCOS 4-manifold of our choice without having to resort to technical acrobatics!

4 Plan of the seminar

We shall adopt an ahistorical approach in our seminar and take the Seiberg–Witten equations as god-given (which is tehnically true). There will be *13 sessions* in all, excluding the organisational meetings today and at the end, when we'll be soliciting feedback from all the participants before voting on the topic of next semester. The rough distribution of the content shall be:

- *Session 1*: recalling background material from differential topology that we'll be needing for the rest of the seminar.
- *Sessions 2-6*: setting up the Seiberg–Witten equations and making sense of the various entities that enter the equations.
- *Sessions 6-10*: defining and proving that the moduli space of solutions of the Seiberg–Witten equations is actually a CCOS manifold of finite dimension.
- *Sessions 11-13*: using Seiberg–Witten theory to deduce interesting things about 4-manifolds in general and in specific examples.

In particular, most of the things I mentioned today do not really feature in the plan. In case you are interested in knowing more about Donaldson theory proper and the physical context behind Seiberg–Witten theory, please consult the additional references on the seminar homepage:

`https://www.math.uni-hamburg.de/home/saha/phd-ws2018.html`