Universiteit Utrecht

Faculteit Bètawetenschappen

Hilbert scheme of points and moduli space of instantons

Bachelor Thesis

Ludo Dekker

Mathematics and Physics

Supervisors:

Dr. M. Kool
Utrecht University

Dr . T.W. Grimm
Utrecht University

June 16, 2021

Abstract

The two main topics of this thesis are Hilbert schemes of points and moduli spaces. We will see how a set of ideals can be given the structure of a projective variety. With this structure we will use group actions to determine the Euler characteristic of the Hilbert schemes of points on a surface. Subsequently, we will explore connections on $\mathbb{R}^4$ and see how the Yang-Mills equations arise from minimizing a quantity known as action. We will then construct all solutions to these equations and see how the geometrization of these solutions gives rise to the moduli space of framed instantons. We will see how this space is some sense equal to the Hilbert scheme of points.

Acknowledgements

First of all I would like to thank both my supervisors dr. Martijn Kool and dr. Thomas Grimm for introducing me to the interesting subjects which I have studied and written about in this thesis. Although not everything went according to the original schedule I am glad we have continued this project. Secondly, I would like to thank both Chongchuo Li and Dirk van Bree for finding time to answer all of my questions and for their patience in explaining various topics to me. At last I would like to thank all of my friends and family who have helped me with finding motivation and kept life enjoyable during this pandemic.

Introduction

In mathematics, one often receives some set of objects from a problem and tries to study these objects. Most often it is necessary to equip such a set with more structure to see how our objects behave. In algebra one might see how the set of permutations can be given the structure of a group. In geometry you want to have some concept of distance between your points. In number theory you want study how addition and multiplication behave.
In this thesis, we will start off by considering the Hilbert scheme of points. Our definition will start as some set of ideals, which we will equip with a topology and we will see how we can take limits and make paths in our space. We will use these concepts to calculate the Euler characteristic of our space, which we will define in the first chapter.
The second main topic will be the Yang-Mills equations. Here one describes a physical system through the use of fields, and with the principle of least action some differential equations for connections will be derived. When you consider these equations over $\mathbb{R}^4$, one can use the ADHM construction to find all solutions to the more specific anti-self-dual equations. These solutions are called instantons. With this construction we get a way of characterizing all instantons by some linear data, which naturally comes with more structure and therefore we can use this to see our set of instantons as a space with more structure. This is called the moduli space of instantons and we will see how that space is isomorphic to the Hilbert scheme of points in some special case.

Outline of this thesis

In chapter 1 we will define the Euler characteristic of a topological space. This is a preparation for chapter 3. Then in chapter 2 we will define the Hilbert scheme of points over $\mathbb{C}^2$, where we will see the importance of monomial ideals and give a topology to this Hilbert scheme. In chapter 3 we will continue studying this Hilbert scheme and show it is path-connected and moreover calculate the Euler-characteristic. Chapter 4 will start with some mathematical preparations where we will define differential forms, and then we will define connections and derive the Yang-Mills equations. In chapter 5 we will focus on the ADHM construction, where we construct all solutions to the Yang-Mills equations. At last, in chapter 6 we will look at the set of all solutions, and study this as a geometric space, called the moduli space of instantons. Here we will see how the space of solutions to the Yang-Mills equations has the same structure as the Hilbert scheme of points for some special cases.

Contents

In this chapter we will introduce homology groups and the Euler characteristic. The Euler characteristic χ is a topological invariant that describes the shape of a topological space. We will see how we can calculate it for some examples.

1 Euler characteristic

Originally the Euler-characteristic was defined for polyhedra. One of the first applications of the Euler-characteristic was the classification of the Platonic solids [1].

Definition 1.0.1. For a polyhedron with V vertices, E edges and F faces we define the Euler characteristic χ as:

$$\chi = V - E + F \tag{1.1}$$

Leonhard Euler stated in 1758 that for all convex polyhedra we have $\chi = 2$. The Euler characteristic of a topological space can be defined in terms of homology groups. In some way we want to associate a sequence of abelian groups with our space. This sequence of groups together with a sequence of homomorphisms between them, is called a chain complex. We will then construct such a chain complex for a special set of topological spaces. Moreover we will see how a complex can be created for any topological space and we will even calculate some Euler characteristics.

1.1 Chain complex

Definition 1.1.1. A chain complex $(C_\bullet, \partial_\bullet)$ is a sequence of abelian groups C_i together with a sequence of homomorphisms $\partial_i : C_i \to C_{i-1}$ such that $\partial_n \circ \partial_{n+1} = 0$ for all $n \in \mathbb{Z}$. From $\partial_n \circ \partial_{n+1} = 0$ it follows that $\text{Im } \partial_{n+1} \subset \ker \partial_n$, and thus we can define the n^{th} homology group H_n to be $\ker \partial_n / \text{Im } \partial_{n+1}$.

We can write out a complex as follows.

$$\dots \xrightarrow{\partial_4} C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} \dots$$

Chain complexes play a big role in the definition of homology, but also appear in other situations. In the fourth chapter we will encounter differential forms and the exterior derivative d, that takes k-forms to $k+1$-forms and also satisfies $\mathrm{d}^2 = 0$. That complex is called the de Rham cohomology. Cohomology groups appear as a dual notion to homology groups.

Definition 1.1.2. A cochain complex $(C^\bullet, \partial^\bullet)$ consists of a sequence of abelian groups C^i together with homomorphisms $\partial^n : C^n \to C^{n+1}$ that satisfy $\partial^{n+1} \circ \partial^n = 0$. We define the n^{th} cohomology group H^n to be $\ker \partial^n / \text{Im } \partial^{n-1}$.

A natural way to transform a chain complex $(C_\bullet, \partial_\bullet)$ to a cochain complex is by considering the groups $\text{Hom}(C_i, G)$ for some fixed group G. We can then define a homomorphism from $\text{Hom}(C_i, G)$ to $\text{Hom}(C_{i+1}, G)$. For an element in $f \in \text{Hom}(C_i, G)$ we can compose f with ∂_{i+1} and create a map $f \circ \partial_{i+1} : C_{i+1} \to G$. In this way we have used ∂_{i+1}, that maps from C_{i+1} to C_i, to create a mapping from $\text{Hom}(C_i, G)$ to $\text{Hom}(C_{i+1}, G)$. This reverses the arrows between our groups.
We can also consider finite sequences of groups. These can be fitted in the definition of a chain complex by taking trivial groups for all other parts of the sequence. A sequence where almost all groups are trivial, is called bounded. We will now look at an example of a short sequence, which is actually some specific example of the de Rham cohomology.

Example 1.1.3. In first-year multivariable calculus class the grad, curl and div operators are often taught and it is well know that $\text{curl}(\text{grad}(\text{f})) = 0$ and $\text{div}(\text{curl}(v)) = 0$ for all $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ and $v \in C^\infty(\mathbb{R}^3, \mathbb{R}^3)$. With these two relations we can use these operators as our homomorphisms to create a chain complex:

$$0 \longrightarrow C^\infty(\mathbb{R}^3, \mathbb{R}) \xrightarrow{\text{grad}} C^\infty(\mathbb{R}^3, \mathbb{R}^3) \xrightarrow{\text{curl}} C^\infty(\mathbb{R}^3, \mathbb{R}^3) \xrightarrow{\text{div}} C^\infty(\mathbb{R}^3, \mathbb{R}) \longrightarrow 0$$

△

1.2 Simplicial homology

The original formula for the Euler characteristic was stated in terms of vertices, edges and faces. We now want to expand this idea and generalize vertices, edges and faces to bigger cells of higher dimensions. For these definitions we follow the book from Munkres [2].

Definition 1.2.1. In $\mathbb{R}^N$ we define an n-simplex σ spanned by $a_0, \dots, a_n$ to be the set of all points x such that $x = \sum_{i=0}^n t_i a_i$ where we have $t_i \geq 0$ for all i and $\sum_{i=0}^n t_i = 1$.

For $n = 0$ we just get the point a_0 and for $n = 1$ we have the set $ta_0 + (1-t)a_1$ with $0 \leq t \leq 1$, thus a 1-simplex is an line segment.

Definition 1.2.2. For a simplex σ spanned by $a_0, \dots, a_n$ we call a simplex spanned by a subset of $\{a_0, \dots a_n\}$ a face of σ. Any face that is different from σ we call a proper face. The union of all proper faces is called the boundary of σ. We define two orderings of the vertex set of a simplex to be equal if they differ from each other by an even permutation. This creates two equivalence classes, and for a simplex σ we will call its equivalence class its orientation. An oriented simplex is a simplex σ together with an orientation of σ.

Now, with this definition that generalizes vertices, edges and points we can define a more general idea of polyhedra.

Definition 1.2.3. A simplicial complex K in $\mathbb{R}^N$ is a collection of simplexes in $\mathbb{R}^N$ such that:

1. Every face of a simplex of K is in K.
2. The intersection of any two simplexes of K is a face of each of them.

Definition 1.2.4. For a simplicial complex K we define $|K|$ to be the subset of $\mathbb{R}^N$ that is the union of the simplexes of K. We give this space a topology by giving each simplex its natural topology as a subspace of $\mathbb{R}^N$ and defining a subset A of $|K|$ to be closed if and only if $A \cap \sigma$ is closed for all $\sigma \in K$. The space $|K|$ is called the polytope of K.

We will now use this data to associate abelian groups with a simplicial complex.

Definition 1.2.5. Let K be a simplicial complex. A p-chain is function c from the set of oriented p-simplexes of K to the integers, such that:

1. $c(\sigma) = -c(\sigma')$ if σ and σ' are opposite orientations of the same simplex.
2. $c(\sigma) = 0$ for all but finitely many simplexes.

We denote the group of oriented p-chains with $C_p(K)$. If there are no p-simplexes in K we define this group to be the trivial group. For an oriented p-simplex σ spanned by $a_0, \dots, a_p$, we define $[a_0, \dots, a_p]$ to be the function c such that $c(\sigma) = 1$ and c is zero for all other simplexes.

Definition 1.2.6. We define the boundary operator $\partial_p : C_p(K) \to C_{p-1}(K)$ to be a homomorphism that takes a simplex $\sigma = [v_0, \dots, v_p]$ and maps it as follows:

$$\partial_p \sigma = \partial_p [v_0, \dots, v_p] = \sum_{i=0}^{p} (-1)^i [v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_p] \tag{1.2}$$

So on the right-hand side we have deleted v_i from the simplex.

We can now explicitly compute that $\partial_{p-1} \circ \partial_p = 0$ and then everything is prepared to define the simplicial homology groups of K. For a simplex $\sigma = [v_0, \dots, v_p]$ we see:

$$\begin{aligned} \partial_{p-1} \circ \partial_p([v_0, \dots, v_p]) &= \sum_{i=1}^{p} (-1)^i \partial_{p-1} [v_0, \dots, \hat{v}_i, \dots, v_p] \\ &= \sum_{j<i} (-1)^i (-1)^j [\dots, \hat{v}_i, \dots, \hat{v}_j, \dots] + \sum_{i<j} (-1)^i (-1)^{j-1} [\dots, \hat{v}_i, \dots, \hat{v}_j, \dots] = 0, \end{aligned} \tag{1.3}$$

where we write $\hat{v}$ for a vertex that is removed from the simplex.

Definition 1.2.7. The kernel of ∂_p is denoted by $Z_p(K)$, the image of ∂_{p+1} is denoted by $B_p(K)$. The elements of $Z_p(K)$ we call cycles and the elements of $B_p(K)$ we call boundaries. Because $\partial_p \circ \partial_{p+1} = 0$ we see that $B_p(K) \subset Z_p(K)$ and thus we can define $H_p(K) = Z_p(K)/B_p(K)$ to be the p^{th} homology group of K.

In some way these groups formalize the idea of the number of holes in $|K|$ of dimension n.

Definition 1.2.8. The k^{th} Betti number $b_k(X)$ of a space X is defined as the rank of $H_k(X)$, which is the number of linearly independent generators of $H_k(X)$.

Intuitively, we see that when a homology group is non trivial, we have an element that is a cycle but is not a boundary of some simplex. Thus it corresponds to a missing "hole" in our space. The k^{th} Betti number counts how many holes with a k-dimensional boundary we have. For instance the hole in a circle has some 1-dimensional boundary, while the hole inside of a sphere has a 2-dimensional boundary.

Example 1.2.9. We will now calculate the homology groups of a triangle $\triangle ABC$ in $\mathbb{R}^2$. A triangle has 3 vertices, 3 edges and 1 face. When we view our triangle as a simplicial complex, we see that it is equal to $\{ABC, AB, BC, CA, A, B, C\}$. Thus we see that our non-trivial groups of p-chains are $C_2 = \text{span}(ABC)$, $C_1 = \text{span}(AB, BC, CA)$ and $C_0 = \text{span}(A, B, C)$. Our boundary operators work as follows: $ABC \mapsto BC - AC + AB$, $AB \mapsto A - B$, $BC \mapsto B - C$, $CA \mapsto C - A$ and $A, B, C \mapsto 0$. So our homology groups are:

$$
\begin{aligned}
H_2 &= \ker \partial_2 / \text{Im } \partial_3 = \langle 0 \rangle / \langle 0 \rangle = 0 \\
H_1 &= \ker \partial_1 / \text{Im } \partial_2 = \langle bc - ac + ab \rangle / \langle bc - ac + ab \rangle = 0 \\
H_0 &= \ker \partial_0 / \text{Im } \partial_1 = \langle a, b, c \rangle / \langle a - b, b - c \rangle \cong \langle a \rangle \cong \mathbb{Z}
\end{aligned}
$$

And therefore we see that the zeroth Betti number is equal to 1, and all other Betti numbers are zero. $\triangle$

1.3 Other homology theories

With simplicial homology we have a way of associating a chain complex with a simplicial complex K. Moreover, it can also be proven that the homology groups only depend on the underlying topological structure of the space $|K|$ [2]. This means that these groups are topological invariants that we can assign to topological spaces which happen to be polytopes of some simplicial complex. A deficiency of these simplicial homology groups is that we cannot assign them to any topological space. There are however other homology theories that construct a chain complex on any topological space. An import theory is the theory of singular homology.

Definition 1.3.1. A singular n-simplex in a topological space X is a continuous function $\sigma : \Delta^n \to X$, where Δ^n is the polytope of a single n-simplex. We define the singular chain group $S_n(X)$ of X in dimension n to be the free abelian group generated by all of the singular n-simplexes of X. We define the boundary homomorphism $\partial_n : S_n(X) \to S_{n-1}(X)$ to be the homomorphism that sends a singular n-simplex to the alternating sum of its faces.

Very similar to the proof of the simplicial homology we can show that $\partial_{n+1} \circ \partial_n = 0$. An import result in homology theory is that the simplicial and singular homology groups of a polytope of a simplicial complex are isomorphic [3]. It should be noted that quite a lot of topological spaces are the polytope of some simplicial complex. In particular it is proven in [4] that every smooth manifold is a polytope of some simplicial complex. Using the Betti numbers of a space we can now define the Euler characteristic of a space X.

Definition 1.3.2. For a topological space X we define its Euler characteristic χ as follows:

$$
\chi(X) = \sum_{i=0}^{\infty} (-1)^i b_k(X) \tag{1.4}
$$

Where $b_k(X)$ denotes the k^{th} Betti number, which was defined as the rank of $H_k(X)$.

For make this sum well-defined, we need two things. First of all, we want the sum to converge. Since all Betti numbers are integers, we want this to be a finite sum. For this thesis we will only be considering the Euler characteristic of manifolds. For a manifold M of dimension n it is a well known result that $b_k(M) = 0$ for $k > n$. Thus our sum will always converge. Second of all, we have the requirement that all Betti numbers should actually be finite. For compact manifolds this is always the case, but we will see that this can become a problem for other manifolds.
We can compare this definition 1.3.2 to the first definition 1.0.1 of the Euler characteristic for polyhedra. At 1.0.1 we also had an alternating sum, however the rank of a homology group is not always equal to the number of n-simplexes in a complex. For example we have seen a triangle $\triangle ABC$ where only the zeroth Betti number was nonzero and equal to 1. If we fill in our original equation for the Euler characteristic we get $\chi = 3 - 3 + 1 = 1$, which fortunately agrees with the value obtained from the new definition. We will now prove that these definitions correspond for polytopes of simplicial complexes.

Theorem 1.3.3. *For a polytope of a finite simplicial complex we have $\chi = \sum_i (-1)^i k_i$, where k_i is the number of i-simplexes.*

Proof. From 1.2.5 we see that k_i is equal to the dimension of C_i. Let n be the highest number such that C_n is nontrivial. We calculate:

$$\begin{aligned}\chi &= \sum_{i=0}^{n}(-1)^i \mathrm{rank}(\ker\partial_i/\mathrm{Im}\ \partial_{i+1}) = \sum_{i=0}^{n}(-1)^i(\mathrm{rank}\ker\partial_i - \mathrm{rank\ Im}\ \partial_{i+1}) \\ &= \sum_{i=0}^{n}(-1)^i \mathrm{rank}\ker\partial_i + \sum_{j=1}^{n+1}(-1)^j \mathrm{rank\ Im}\ \partial_j \\ &= \mathrm{rank}\ker\partial_0 + (-1)^{n+1}\mathrm{rank\ Im}\ \partial_{n+1} + \sum_{i=1}^{n}(-1)^i(\mathrm{rank}\ker\partial_i + \mathrm{rank\ Im}\ \partial_i)\end{aligned} \tag{1.5}$$

We know that ∂_0 maps to the trivial group, so $\ker\partial_0 = C_0$. We also notice that the domain of ∂_{n+1} is the trivial group, implying that its image is also trivial and thus that rank Im $\partial_{n+1} = 0$. Since all these maps are $\mathbb{Z}$-linear we also know that for a map with domain C_i that $\mathrm{rank}\ker\partial_i + \mathrm{rank\ Im}\ \partial_i = \mathrm{rank}\ C_i$. So our last expression simplifies to $\sum_{i=0}^{n} \mathrm{rank}\ C_i$, which is exactly what we wanted to prove. □

In this thesis we want to talk about the Euler characteristic of a non-compact space. Definition 1.3.2 for the Euler characteristic will not work for non-compact spaces, since the Betti numbers then do not have to exist. However there exists a cohomology theory named cohomology with compact support that will also make sense for some non-compact spaces. For locally compact spaces we can use the cohomology with compact support to define the Euler characteristic with compact support χ_c. It is proven in [5] and [6] that χ and χ_c correspond for compact spaces. In that way χ_c extends the definition of the Euler characteristic. We do however have to be careful, there are some properties of the Euler characteristic for compact spaces that are not true any more for the Euler characteristic with compact support. One important example is that for compact space the Euler characteristic is homotopy-invariant, for the Euler characteristic with compact support this is not the case. We will now state some important properties for the Euler characteristic with compact support from [7] and [8].

Theorem 1.3.4. *For a proper smooth submersion $f : X \to Y$ with X, Y smooth manifolds, Y connected and generic fiber F we have*

$$\chi(X) = \chi(Y) \cdot \chi(F). \tag{1.6}$$

If we have a locally compact space X with M, N open or closed in X we have

$$\chi(M \cup N) = \chi(M) + \chi(N) - \chi(M \cap N). \tag{1.7}$$

If we have a fibration $p : E \to B$ with fiber F where B is path-connected and the fibration is orientable over a field K, then the Euler characteristic with coefficients in K satisfies

$$\chi(E) = \chi(F) \cdot \chi(B). \tag{1.8}$$

These properties can be used to calculate the Euler characteristic of many spaces. The last property is a more general form of the first one, and will become useful later in chapter 3 3.

1.4 Calculating Euler characteristics

We now have found ways to calculate the Euler characteristics of compact spaces and moreover have some tools to relate the Euler characteristic of non-compact spaces to smaller spaces. We will now calculate the Euler characteristics of some well-known spaces. We will start off by proving a lemma for the Cartesian product of spaces.

Lemma 1.4.1. *For two spaces M, N which have some Euler characteristic we have $\chi(M \times N) = \chi(M) \cdot \chi(N)$.*

Proof. We will use the first property from 1.3.4. Take $X = M \times N$ and take the function $f : X \to M$ to be the projection to the first coordinate. Projections are examples of smooth submersions. For each point m in M we have the fiber $F = f^{-1}(\{m\}) = \{m\} \times N \cong N$. Therefore we see that $\chi(M \times N) = \chi(M) \cdot \chi(N)$. □

Example 1.4.2. First of all, we will calculate $\chi(\mathbb{R})$. We can use the second property in 1.3.4, by observing that $\mathbb{R} = (-\infty, 0) \cup \{0\} \cup (0, \infty)$. Thus we can use the union property twice to see that $\chi(\mathbb{R}) = \chi((-\infty, 0)) + \chi(\{0\}) + \chi((0, \infty))$. Now we can use that the Euler characteristic of a single point is 1 and since $\mathbb{R}$ is homeomorphic to $(-\infty, 0)$ and $(0, \infty)$ the equation reduces to $\chi(\mathbb{R}) = 2\chi(\mathbb{R}) + 1$, therefore we have $\chi(\mathbb{R}) = -1$. With the product property we see $\chi(\mathbb{R}^n) = (-1)^n$ and because $\mathbb{C}$ is homeomorphic to $\mathbb{R}^2$ we have $\chi(\mathbb{C}^n) = 1$. △

Example 1.4.3. We can see the circle S^1 as a simplicial complex with two edges that meet at both sides. Then we have two edges and two vertices, so by 1.3.3 we see that $\chi = 2 - 2 = 0$. △

Example 1.4.4. For the torus $T = \mathbb{C}^*$ we can calculate its Euler characteristic in two ways. First of all we have the homeomorphism $\mathbb{C}^* \to S^1 \times \mathbb{R}$ the map $z \mapsto (z/|z|, \log(|z|))$. Thus we see that $\chi(\mathbb{C}^*) = \chi(S^1)\chi(\mathbb{R}) = 0 \times -1 = 0$. We can also use the second property in 1.3.4 and see that $\chi(\mathbb{C}) = \chi(\mathbb{C}^*) + 1$. △

We also have all the tools to prove Euler's statement that all convex polyhedra have $\chi = 2$, by recognizing that all convex polyhedra are homeomorphic to S^2. In particular we see that all convex polyhedra should have the same Euler characteristic as the tetrahedron, where we can count 4 vertices, 6 edges and 4 faces to conclude that $\chi = 4 - 6 + 4 = 2$.

2 Hilbert scheme

The main topic of this chapter will be defining the Hilbert scheme of points. We will see how we can use the structure of the Grassmanian to give structure to our set. Furthermore we will also see how the Hilbert scheme is related to sets of points.

2.1 The Hilbert scheme

Although we will be mostly interested in the Hilbert scheme $\mathrm{Hilb}^n(\mathbb{C}^2)$, we will start off by studying a simpler space: $\mathrm{Hilb}^n(\mathbb{C})$.

Definition 2.1.1. The Hilbert scheme $\mathrm{Hilb}^n(\mathbb{C})$ as a set is defined as the set of ideals I in $\mathbb{C}[x]$ such that $\mathbb{C}[x]/I$ is a vector space of dimension n over $\mathbb{C}$.

Since $\mathbb{C}[x]$ is a principal ideal domain, we know that each ideal is generated by a polynomial $f(x)$. Moreover we can also see that $\dim(\mathbb{C}[x]/(f(x))) = \deg(f)$. The generator for such an ideal is unique up to multiplication with a constant. We see that every ideal is in bijective correspondence with the roots of the polynomial f, since those exactly fix a polynomial up to multiplication with a constant. We define the set $S^n\mathbb{C}$ as the set of arrangements of n points modulo the symmetric group acting on these points by interchange. We see that $\mathrm{Hilb}^n(\mathbb{C}) \cong S^n\mathbb{C}$. Although $\mathrm{Hilb}^n(\mathbb{C})$ has an abstract definition, it is actually an algebraic description of a set that feels more natural. The set $S^n\mathbb{C}$ can easily describe physical situations where you have similar objects appearing with a different parameter. However, from a mathematical point of view the Hilbert scheme is preferred.
We can make a map from $\mathrm{Hilb}^n(\mathbb{C})$ to $\mathbb{C}^n$ by taking the monic generator of an ideal, which is unique, and send it to its non-leading coefficients. This map is bijective, since every monic polynomial also generates a unique ideal and we can inverse the map. Although we have not given $\mathrm{Hilb}^n(\mathbb{C})$ the structure of a topological space yet, this can be done and then this map actually forms a homeomorphism, and we can actually calculate that $\chi(\mathrm{Hilb}^n(\mathbb{C})) = \chi(\mathbb{C}^n) = 1$.
We will now continue to the object that is more interesting to us. The Hilbert scheme of points over $\mathbb{C}^2$.

Definition 2.1.2. The Hilbert scheme $\mathrm{Hilb}^n(\mathbb{C}^2)$ of n points in the plane is defined as a set by the ideals $I \subset \mathbb{C}[x, y]$ such that $\mathbb{C}[x, y]/I$ has dimension n as a vector space over $\mathbb{C}$. For an ideal $I \subset \mathbb{C}[x, y]$ we call the dimension of $\mathbb{C}[x, y]/I$ the colength of I.

A first remark we can make is that both definitions for these Hilbert schemes of points only refer to the set of objects. The complete structure of these objects was developed by Alexander Grothendieck[9]. In general a Hilbert scheme is a scheme that is the parameter space for the closed subschemes of some projective space. But the complete definition of Hilbert schemes goes beyond this thesis. Fogarty proved that $\mathrm{Hilb}^n(X)$ is smooth and irreducible for a smooth irreducible surface X over $\mathbb{C}$ and that this variety comes equipped with a Hilbert-Chow morphism that is a resolution of singularities of $\mathrm{Sym}^n(X)$ [10]. Finally the construction of the Hilbert scheme of points of a surface we are using here is due to Haiman [11].
We have seen that for the Hilbert scheme over the complex line, we have correspondence between ideals and sets of points. We can also try to make such a correspondence for our Hilbert scheme over $\mathbb{C}^2$. But multivariable polynomials do not have some standard form whence we can read the roots. For instance we can look at the ideal $I = (x^2 - y, y^2)$, we can calculate that it is of colength 4. There is no direct way to think of 4 points that we can associate with this ideal. One might recognize that $(0, 0)$ is a root of every polynomial in I, this is however the only point that is a root of every polynomial in I. The ideal (x^2, y^2) has the same property however, thus in the sense of points there is no trivial way to uniquely assign ideals to sets of points.
We can also try work the other way around. For a set of points in $\mathbb{C}^2$, does there exist a natural ideal we can associate it with? It turns out that for a set of n different points the ideal that vanishes on all of those points is an ideal of colength n. If we have a set of n different points with different x-coordinates, we can use Lagrange interpolation to find an polynomial $y = P(x)$ of degree $n-1$ such that all points are on that curve. Furthermore we can take the polynomial $Q(x) = \prod_i (x - x_i)$, where x_i are the first coordinates of our n points. The ideal $I = (y - P(x), Q(x))$ vanishes exactly on the n different points we started with. Moreover we see that $1, x, \dots, x^{n-1}$ form a basis of $\mathbb{C}[x, y]/I$, and thus this is an ideal with colength n. If some x-coordinates

of points coincide, the Lagrange polynomial does not exists. We could then try to switch x and y around, or if that is still not possible we can rotate our coordinates such that all points have a different first coordinate. For instance we can consider the points $(0, 0)$ and $(-1, 1)$. The Lagrange polynomial $y = -x$ passes through both points, and $x(x + 1)$ gives our x-coordinates, thus $(y + x, x^2 + x)$ is the ideal that vanishes on both points. An ideal that vanishes on n different points is called a generic ideal. We will later see a way to associate a set of points with every ideal, but it will not be a nice bijective correspondence. This mapping is called the Hilbert-Chow morphism.
A special kind of ideals in $\mathrm{Hilb}^n(\mathbb{C}^2)$ are the monomial ideals.

Definition 2.1.3. A monomial is a polynomial of the form $x^a y^b$ for some non-negative integers a, b. A monomial ideal is an ideal that is generated by monomials.

For a monomial ideal we can look at the monomials that are not contained in the ideal. By definition of an ideal we know that if $x^a y^b \in I$ and we have both $a' \geq a$ and $b' \geq b$, then $x^{a'} y^{b'}$ is also an element of I. Because we want the colength n of an ideal to be finite, there are also finitely many monomials that are not contained in I, since they form a basis of $\mathbb{C}[x, y]/I$. We can visualize these monomials as a partition of n.

Definition 2.1.4. A partition of a positive integer n is a non-increasing sequence $(\lambda_i)_{i\in\mathbb{N}}$ of natural numbers where only finitely many terms are non-zero and the sum of all terms is n.

We will see that there exists a bijection between the partitions of n and the monomial ideals of colength n. We will illustrate this by example and then prove the actual statement.

Example 2.1.5. The sequence $(4, 2, 1, 0, \dots)$ forms a partition of 7. We associate this with the monomial ideal $I_{4+2+1} = (x^4, x^2y, xy^2, y^3)$. We visualize this by drawing the Young diagram.

$$\begin{array}{|c|c|c|c|} \cline{1-1} y^2 \\ \cline{1-2} y & xy \\ \hline 1 & x & x^2 & x^3 \\ \hline \end{array} \tag{2.1}$$

We see that the width of the different layers corresponds to the partition, and the monomials in the boxes are exactly the monomials that are not contained in I. △

Theorem 2.1.6. *There is a bijection between the partitions λ of n and the monomial ideals $I \subset \mathbb{C}[x, y]$ of length n given by*

$$\lambda \mapsto I_\lambda := (\{x^r y^s | r \geq \lambda_s\}) \tag{2.2}$$

with inverse

$$I \mapsto \lambda(I) = (\lambda_i)_{i\in\mathbb{N}} := (\min\{j | x^j y^i \in I\})_{i\in\mathbb{N}}. \tag{2.3}$$

Proof. First of all we will prove these maps are well-defined. We start off by remarking that by definition I_λ is a monomial ideal. Furthermore we note that for a partition λ of n we have $\#\{(r, s) | r < \lambda_s\} = \sum \lambda_s = n$. Thus there are exactly n monomials that are not contained in I_λ, since I_λ is a monomial ideal it follows that these monomials span $\mathbb{C}[x, y]/I_\lambda$ and thus the colength of I_λ is n. A proof for the fact that the monomials outside of a monomial ideal span its quotient space will be seen later in 2.2.3.
We know for an ideal I that if $x^j y^i \in I$, we also have that $x^j y^{i+1} \in I$, therefore we see that the sequence $\lambda(I)$ is non-increasing. Additionally $\min\{j | x^j y^i \in I\}$ is equal to the number of monomials of the form $x^a y^i$ for some a that are not elements of I. Hence we see that $\sum_i \min\{j | x^j y^i \in I\}$ is equal to the number of monomials that are not elements of I, since I is a monomial ideal of colength n we see that this sum must be equal to n. Therefore our sequence also sums to n, and will be zero from some point on.
Now just have to check that these maps are inverses of each other.

$$\lambda(I_\lambda) = (\min\{j | x^j y^i \in (\{x^r y^i | r \geq \lambda_i\})\})_{i\in\mathbb{N}} = (\lambda_i)_{i\in\mathbb{N}} \tag{2.4}$$

$$I_{(\min\{j|x^j y^i \in I\})_{i\in\mathbb{N}}} = (\{x^r y^s | r \geq \min\{j | x^j y^s \in I\}\}) = (\{x^r y^s | \exists c \in \mathbb{N} : x^{r-c} y^s \in I\}) = I \tag{2.5}$$

We can conclude that these two maps are indeed inverses of each other and therefore form a bijection. □

2.2 Gröbner basis

Before we continue working with the Hilbert scheme, we will do some preparations to work with multivariable polynomials. For polynomials in one variable you can look at the leading term and in a quotient ring $\mathbb{C}[x]/I$ you can write elements with a representative of lowest possible degree. Such concept are not unambiguously defined for multivariable polynomials. For instance the leading term of $x^2 + y^2$ is not clear without defining an order first. Another problem we have for ideals in $\mathbb{C}[x, y]$ is that there is no standard way to write their generators.
First of all we have to choose some ordering of all monomials. We will focus on an ordering for two variables. Most of the things we mention here will also be applicable to situations with more variables. There is some freedom in the choice of this ordering, since there is no fundamental difference between x and y, but we have to choose one to be smaller. For the theory of Gröbner bases we do want some properties in the ordering. The first property we require is that for all monomials M, N, P we have that

$$M \leq N \iff MP \leq NP. \tag{2.6}$$

The second property we require is that $M \leq MP$ for all monomials M, P. Similarly to the order with one variable these requirements guarantee us that multiplying things will in general increase the size of a polynomial. A possible ordering for monomials would be to start off by comparing the degree in x, and if that is equal compare the degree in y. This is called the lexicographic order and it can easily be checked that satisfies our two requirements. We however will be working with the graded reverse lexicographic order. We will first compare the total degree of a monomial, and if those are the same we will compare the degree of x. We then get the following order:

$$1 < y < x < y^2 < xy < x^2 < y^3 < xy^2 < x^2y < x^3 < y^4 < \dots$$

An advantage of this ordering is that most of the time calculations will be a bit shorter. This also has the advantage that there are only finitely many monomials smaller than any specific monomial M.
With this ordering every polynomial can be rewritten in a standard form, we see that x^2+xy+y^2 is decreasing in terms, just like it is natural to write a polynomial as $2x^2 - 3x + 1$, instead of $-3x + 1 + 2x^2$.

Definition 2.2.1. Let $\mathbb{K}$ be some field and fix a term order on a polynomial ring R over $\mathbb{K}$. A Gröbner basis G of an ideal I in R is a generating set of I such that the ideal generated by the leading terms of polynomials in I equals the ideal generated by the leading terms of G. Equivalently we have the property that the leading term of any polynomial in I is divisible by the leading term of some polynomial in G. A Gröbner basis is termed reduced if the leading coefficient of each element in G is 1 and no monomial in any element of G is generated by the leading terms of the other elements of the basis.

Reduced Gröbner bases are unique for any given ideal and any monomial order [12]. We will see that the ideals generated by the leading terms of some other ideal come back later. We define the initial ideal $\text{in}(I)$ to be the ideal generated by the leading terms of the elements of I. Notice that any initial ideal is a monomial ideal by construction.
We will now prove a lemma from [12], that will become useful later.

Lemma 2.2.2. *Let V_m be the vector subspace inside of $\mathbb{C}[x, y]$ spanned by the monomials of degree at most m. Given any colength n ideal I, the image of V_m spans $\mathbb{C}[x, y]/I$ as a vector space for $m \geq n$*

Proof. We look at the set M of monomials that are not contained in $\text{in}(I)$. If those are linearly dependent in $\mathbb{C}[x, y]/I$, then we would have a non-zero sum of these monomials that is contained in I. That would imply that the leading term of that sum is also in $\text{in}(I)$, which is impossible. Therefore they form a linearly independent set in $\mathbb{C}[x, y]/I$. Moreover for each polynomial $p \in \mathbb{C}[x, y]$ that has a leading term that is not in M, we can subtract some element of I to get a smaller leading term. We can repeat this process until we get a polynomial whose leading term is in M. We can then subtract the leading term and look at the rest of the polynomial. With this algorithm we keep on reducing the leading term of our polynomial. Since there are only finitely many monomials smaller than the original leading term of p, this algorithm will reach 0 at some point. We can reverse the process and write $p = m + i$ for $i \in I$ and $m \in \langle M \rangle$, since each steps either subtracts an element of I or an element of $\langle M \rangle$. In conclusion we see that $p + I = m + I$, thus every

polynomial in $\mathbb{C}[x,y]/I$ can be written as a linear combination of the monomials in M. Therefore M forms a basis of $\mathbb{C}[x,y]/I$.
Since $\dim\mathbb{C}[x,y]/I = n$, we see $|M| = n$. We have already seen that if $x^a y^b \notin \mathrm{in}(I)$, then also $x^{a'} y^{b'} \notin \mathrm{in}(I)$ for $a' \leq a$ and $b' \leq b$, therefore we see that the highest possible degree for an element of M is $n-1$, and thus all element of M are also contained in V_m for $m \geq n$. □

Corollary 2.2.3. *The colength of I is equal to the colength of in(I).*

Proof. We have seen that the set M of monomials that are not contained in $\mathrm{in}(I)$ forms a basis in $\mathbb{C}[x,y]/I$. This set also forms a basis in $\mathbb{C}[x,y]/\mathrm{in}(I)$, since every polynomial in $\mathbb{C}[x,y]$ can naturally be written as a sum of monomials, some elements of $\mathrm{in}(I)$ and some not elements of $\mathrm{in}(I)$. Since this set forms a basis in both quotient spaces, we see that their dimensions are the same. □

2.3 Topology of the Hilbert scheme

For more structure on our set of ideals we want to view it as a subset of some variety with more structure, and try to inherit some of this structure. All elements of $\mathrm{Hilb}^n(\mathbb{C}^2)$ are ideals I with some property for the linear space $\mathbb{C}[x,y]/I$. This incentivizes us to look at $\mathrm{Hilb}^n(\mathbb{C}^2)$ like it is a collection of linear subspaces. This brings us to the Grassmannian. For some n-dimensional vector space V we write $Gr_k(V)$ as the space of k-dimensional linear subspaces of V.

Definition 2.3.1. The Grassmannian $\mathrm{Gr}_k(V)$ is the space of all k-dimensional subspaces of some vector space V. When V is a real (or complex) vector space with dimension n it can be shown that $\mathrm{Gr}_k(V)$ is a compact smooth (or complex) manifold of real (or complex) dimension $k(n-k)$. We will write $\mathrm{Gr}^k(V)$ for the set of codimension k subspaces of V [13].

We now want to find a vector space V that contain all our vector spaces from the Hilbert scheme. We can view $\mathbb{C}[x,y]/I$ as a subspace of $\mathbb{C}[x,y]$ by choosing a basis B of $\mathbb{C}[x,y]/I$, and then choosing any representative of each element in B. However, we can not choose $\mathbb{C}[x,y]$ as our V, since the Grassmannian is defined for a finite-dimensional vector space. In our choice for the representative we can still fix this problem, by choosing the representatives with a small degree, we can replace $\mathbb{C}[x,y]$ by some finite dimensional space. We will use our lemma 2.2.2. From now on we will choose $m = n$ for the rest of our proofs, however there do exist situations where a higher choice of m can be beneficial [12]. We want to use the structure of $\mathrm{Gr}^n(V_m)$ to define structure on $\mathrm{Hilb}^n(\mathbb{C}^2)$.

Definition 2.3.2. By $\mathbb{P}^n$ we denote the projective n-space over $\mathbb{C}$. A projective variety over $\mathbb{C}$ is a subset $W \subset \mathbb{P}^n$ such that W is the zero-set of some family of homogeneous polynomials in $\mathbb{C}[x_0, \ldots, x_n]$. A quasi-projective variety is a locally closed subset of a projective variety.

Theorem 2.3.3. *The Grassmanian $Gr^k(V)$ is a projective variety for every finite linear space V.*

This is proven in [12].

Theorem 2.3.4. *The Hilbert scheme $Hilb^n(\mathbb{C}^2)$ is a quasiprojective variety in $Gr^n(V_m)$.*

Proof. We will sketch the idea of this proof, but will not fill in all details. A full proof can be found in [12] We start off by making an open covering of $\mathrm{Hilb}^n(\mathbb{C}^2)$. For a partition λ of n we define U_λ to be the set of ideals I such that the monomials outside of I_λ map to a vector basis of $\mathbb{C}[x,y]/I$. We have seen before that for every ideal I we can construct similar bases for $\mathbb{C}[x,y]/I$ and $\mathbb{C}[x,y]/\mathrm{in}(I)$, and since all monomial ideals are in bijective correspondence with the partitions we conclude that each I is an element of U_λ for some partition λ.
This condition of having a certain basis corresponds to the non-vanishing of some Plücker coordinate. Combined with some other remarks it will follow that U_λ is locally closed. Now we have an open cover of $\mathrm{Hilb}^n(\mathbb{C}^2)$ in $\mathrm{Gr}^n(V_m)$ that is locally closed, thus $\mathrm{Hilb}^n(\mathbb{C}^2)$ is a quasi-projective variety. □

2.4 Hilbert-chow morphism

We have seen that for $\mathrm{Hilb}^n(\mathbb{C})$ there is a natural map to the n-fold symmetric product of $\mathbb{C}$. For $\mathbb{C}^2$ we have a somewhat similar idea, but especially the n-fold symmetric product of $\mathbb{C}^2$ is not a "nice" space any more. This subsection will not be relevant to the next chapter, but it will explain some ideas that we will see later for the moduli space of some system we will solve.

Definition 2.4.1. Let M be some set and m a positive integer. In M^m we define the relation $\approx$ such that $x \approx y$ if and only if there is a permutation $\theta : \{1, \ldots, m\} \to \{1, \ldots, m\}$ such that $y_i = x_{\theta(i)}$. We will denote the set $M^m/\approx$ as $M^{(m)}$ and this set is called the symmetric product of M.

Theorem 2.4.2 ([14])**.** *The set $(\mathbb{C}^n)^{(m)}$ is a manifold only for $n = 1$.*

This case of $n = 1$ exactly corresponds with the symmetric product of $\mathbb{C}$ which we saw to be isomorphic to $\mathrm{Hilb}^n(\mathbb{C})$. For the other cases we get singularities in our space. Now as stated before it was proven by Fogarty that $\mathrm{Hilb}^n(\mathbb{C}^2)$ is a resolution of singularities of $\mathrm{Sym}^n(\mathbb{C}^2)$. This resolution is given by the Hilbert-Chow Morphism. The complete definition of algebraic varieties and irreducibility can be found in [15].

Definition 2.4.3. For two irreducible varieties X, Y a birational map $g : X \to Y$ is a rational map such that there is a rational map $Y \to X$ inverse to f. A resolution of singularities of an algebraic variety V is a non-singular variety W with a proper birational map $W \to V$.

The exact definition of a rational map f is a bit tedious, but is a kind of partial function between algebraic varieties. It has some open $U \subset X$ as its domain and can locally be written in coordinates using rational functions. For algebraic varieties we use the Zariski topology, which has the property that every non-empty open set is dense in the total variety, provided the total variety is irreducible. If we have birational map between two varieties X, Y we call those birationally equivalent. This means that we can identify X minus some lower dimensional subset with Y minus some lower dimensional subset.

Example 2.4.4. The circle X with equation $x^2 + y^2 - 1 = 0$ in $\mathbb{C}^2$ is birationally equivalent with the line $y = 0$. We can define the maps:

$$\begin{aligned} f(t) &= \left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right) \\ g(x, y) &= \frac{1-y}{x} \end{aligned} \tag{2.7}$$

We see that these maps are indeed given by rational functions. The fact that g is not defined in $(0, -1)$ is no problem for a rational map, since it is just some zero dimensional variety that is missing from the domain. We can also see that f is not defined for $t \in \{-i, i\}$, again however this is not a problem. △

The theory of resolutions is concerned with the question whether every algebraic variety V has a resolution. A singular point of a variety is a point where its tangent space may not be regularly defined. For instance the curve defined by $y^2 - x^2(x + 1) = 0$ has a singular point at the origin, since it crosses itself. A non-singular variety is a variety without any singular points.

Theorem 2.4.5 ([16])**.** *Every variety over a field with characteristic 0 has a resolution*

Now we go back to our case. The Hilbert-Chow morphism is a resolution, thus we should get some mapping from ideals in $\mathrm{Hilb}^n(\mathbb{C}^2)$ to a set of points in $\mathrm{Sym}^n(\mathbb{C}^2)$.

Definition 2.4.6. The Hilbert-Chow morphism $f : \mathrm{Hilb}^n(\mathbb{C}^2) \to \mathrm{Sym}^n(\mathbb{C}^2)$ can be calculated as follows. For a point $p \in \mathbb{C}^2$ we define $I_p := I + (x - p_x, y - p_y)^n$. We send an ideal I to the set that contains p exactly $\dim \mathbb{C}[x, y]/I_p$ times.

Example 2.4.7. Consider the ideal $I = (x^2 + x, y^2 + x - 1)$. We can check that its generators form a Gröbner basis and thus $\mathrm{in}(I) = (x^2, y^2)$. We also know that $\mathbb{C}[x, y]/I$ is spanned by monomials outside of $\mathrm{in}(I)$, and thus we see $1, x, y, xy$ form a basis for $\mathbb{C}[x, y]/I$ and $I \in \mathrm{Hilb}^4(\mathbb{C}^2)$.
If we have a point $p \in \mathbb{C}^2$ with $p_x \notin \{0, -1\}$, then we can calculate that $I + (x - p_x, y - p_y)^4 = \mathbb{C}[x, y]$, since $(x - p_x)^4$ and $x^2 + x$ share no roots, and thus we can use Euclid's algorithm to find a non-zero constant in I_p.

If $p_x = 0$, we can use $(x^2-x+1)(x^2+x)-x^4 = x$ to see that $x \in I_p$, thus y^2-1 is also in I_p. Again we see that if p_y is not a root of y^2-1, then we get a non-zero constant in I_p. We can check that $I_{(0,1)}$ and $I_{(0,-1)}$ both have colength 1. For $p_x = -1$ we can find with similar reasoning that p_y should be $\pm\sqrt{2}$. Both corresponding ideals also have colength 1 and thus we send I to the set $\{(0,1),(0,-1),(-1,\sqrt{2}),(-1,-\sqrt{2})\}$. △

Example 2.4.8. The ideals (x^3, y), (x^2, xy, y^2) and (x, y^3) all have colength 3 and are all mapped to the set with the origin three times. △

It is possible to interpret this construction geometrically. The addition of two ideals corresponds to intersecting their zero-sets. Therefore we are looking how much of a zero every points in $\mathbb{C}^2$ is of an ideal. It can be shown that $\frac{\mathbb{C}[x,y]}{I} = \bigoplus_{p\in\mathbb{C}^2} \frac{\mathbb{C}[x,y]}{I_p}$.With this equation we can compare dimensions and verify that the Hilbert-Chow morphism always gives a set of n points.

3 Paths in the Hilbert scheme

In this section we will see how we can move points in our Hilbert scheme. Most importantly we will see the flat limit of an ideal. Furthermore we will prove that the Hilbert scheme is connected and we will finally calculate $\chi(\text{Hilb}^n(\mathbb{C}^2))$.

3.1 Flat limit

In this paragraph we will define the flat limit. In this context we see it as a parametrized family of ideals in $\text{Hilb}^n(\mathbb{C}^2)$ that will contain some starting ideal I, and will also contain $\text{in}(I)$. We will try to make an ideal I "flow" to its initial ideal. For this to happen we want the non-leading terms of the reduced Gröbner basis of I to vanish. We will start with an example.

Example 3.1.1. We have seen before that the ideal $I = (x^2 + x, y^2 + x - 1)$ has colength 4 and the written generators form the reduced Gröbner basis, therefore we also have $\text{in}(I) = (x^2, y^2)$. We want to multiply x and y by some power of t such that the non-leading terms vanish when t goes to 0. If we have $x \mapsto t^a x$ and $y \mapsto t^b y$ we get the ideal

$$I_t = (t^{2a}x^2 + t^a x, t^{2b}y^2 + t^a x - 1) = (x^2 + t^{-a}x, y^2 + t^{a-2b}x - t^{-2b}). \tag{3.1}$$

So we want to choose a, b such that $-a > 0$, $a - 2b > 0$,$-2b > 0$. If we can find such a, b the non-leading terms will vanish when t goes to 0. We can choose $a = b = -1$ in this case, and we check these satisfy our conditions. Therefore the family $I_t = (x^2 + tx, y^2 + tx - t^2)$ creates a path from I to $\text{in}(I)$, since $t = 1$ gives us I and $t = 0$ gives us $\text{in}(I)$. △

Here we have to be a bit careful, since we are taking a limit of ideals. Also the map $x \mapsto t^{-1}x$ makes no sense for $t = 0$, but because ideals are closed under scaling this mapping does seem to work when we apply it to the whole ideal at the same time. If we had chosen $a = b = 1$ and take the limit from t to 0, the only surviving term is 1, but that would mean that $I_0 = \mathbb{C}[x, y]$. Then this family would not be a subset of $\text{Hilb}^4(\mathbb{C}^2)$ any more and our limit point would not exist. The correct terminology for such a family to be continuous is for such a family to be flat.

Theorem 3.1.2 ([12])**.** *We view a family I_t as an ideal in $\mathbb{C}[x, y][t]$. If $\mathbb{C}[x, y][t]/I_t$ is a finitely generated free $\mathbb{C}[t]$-module with dimension n then I_t is a flat family.*

A first remark we make here, is that we cannot multiply or divide by t anymore once we fix our family of ideals as an ideal in $\mathbb{C}[x, y][t]$. In our example we did divide to make our leading terms free of t.

Theorem 3.1.3. *For $I \in \text{Hilb}^n(\mathbb{C}^2)$ we can create a flat family of ideals I_t such that $I_0 = \text{in}(I)$ and $I_1 = I$.*

Proof. We take our original ideal I and we map $x \mapsto t^a x$ and $y \mapsto t^b y$, where $a = -2n - 1$ and $b = -2n$. We know all elements of the reduced Gröbner basis should at most have a leading term of degree n. Now we compare two monomials $p_1 = x^{m_1}y^{n_1}$ and $p_2 = x^{m_2}y^{n_2}$ that could appear in elements of the reduced Gröbner basis. We see that our mapping will send $p_i \mapsto t^{am_i + bn_i}$ for $i \in \{1, 2\}$. Now assume that $p_1 > p_2$, we will show that the exponent of t will be greater in front of the smaller monomial.
Since we use the graded reverse lexicographic order, we can separate $p_1 > p_2$ in two cases. For the first case we have that the degree of p_1 is greater than the degree of p_2, thus $m_1 + n_1 > m_2 + n_2$. Since these are all integers we see that $m_1 + n_1 \geq m_2 + n_2 + 1$. Furthermore we know that the total degree of both monomials can be at most n, thus $m_2 \leq n$ and $m_1 - m_2 > -2n$. We can use these inequalities to see that

$$am_1 + bn_1 = -2nm_1 - m_1 - 2nn_1 \leq -2nm_2 - m_1 - 2nn_2 - 2n < -2nm_2 - m_2 - 2nn_2 = am_2 + bn_2. \tag{3.2}$$

Now in the other case we have $m_1 + n_1 = m_2 + n_2$ and $m_1 > m_2$. We can then directly see that

$$am_1 + bn_1 = -2nm_1 - m_1 - 2nn_1 < -2nm_2 - m_2 - 2nn_2 = am_2 + bn_2. \tag{3.3}$$

We conclude that $p_1 > p_2$ implies $am_1 + bn_1 < am_2 + bn_2$. Thus the factor t in front of the leading term has the smallest exponent. We can thus multiply or divide by a power of t such that the leading term of the polynomial has no factor of t and all other monomials have some positive power of t in front of it. We can write all of the polynomials in the Gröbner basis this way and choose those polynomials to generate an ideal I_t in $\mathbb{C}[x,y][t]$. Since the leading terms have no factor of t, we see that $\mathbb{C}[x,y][t]/I_t$ has to have dimension n as a module over $\mathbb{C}[t]$. This is because we can still reduce all polynomials in the same way to lower degree in x and y as we could do in $\mathbb{C}[x,y]/I$, since our leading terms did not change. The only difference is that we will gather some factors of t, but those do not matter since we see our space as a module over $\mathbb{C}[t]$. We conclude that this is a flat family. By construction we can also see that for $t=1$ we get I back and for $t=0$ we get $\text{in}(I)$. □

This construction of the family of ideals can also be seen as a group action of $\mathbb{C}^*$ working on an ideal. We will later use this idea to calculate the Euler characteristic of $\text{Hilb}^n(\mathbb{C}^2)$.

3.2 Connectedness

To show the Hilbert scheme is connected, we will construct a path between any two ideals. We will follow some ideas from [12]. We have already seen that we can use the flat limit to make a path from any ideal I to its initial ideal, and therefore to some monomial ideal. So all for us left to do is connect the monomial ideals by at path.

Definition 3.2.1. For a partition λ of n we have seen there exists a monomial ideal I_λ. For all monomials $x^h y^k$ outside of I_λ we can look at the point $(h,k) \in \mathbb{N}^2 \subset \mathbb{C}^2$. The ideal I'_λ that vanishes on all of these n points is called the distraction of I_λ.

For a monomial ideal $I_\lambda = \langle x^{a_1}y^{b_1}, \dots, x^{a_m}y^{b_m}\rangle$ we can consider the polynomials

$$f_i = x(x-1)(x-2)\dots(x-a_i+1)y(y-1)(y-2)\dots(y-b_i+1) \tag{3.4}$$

We see that $\langle f_1,\dots,f_m\rangle \subset I'_\lambda$, since every polynomial vanishes on the given points in $\mathbb{N}^2$. Therefore we see that $\langle f_1,\dots,f_m\rangle$ has a colength of at least n. But by construction we can also see that the leading terms of the f_i are exactly the $x^{a_i}y^{b_i}$, which are the generators of I_λ. Therefore we have a colength of at least n. In conclusion we can see that $\langle f_1,\dots,f_m\rangle = I'_\lambda$. But we can also see that $\text{in}(\langle f_1,\dots,f_m\rangle) = I_\lambda$, because we know the leading terms of all generators. Therefore we see that $\text{in}(I'_\lambda) = I_\lambda$, and from the flat limit we see that there must be a path connecting these ideals. By construction we also know that I'_λ is a generic ideal.

Theorem 3.2.2. *There exists a path between any two generic ideals.*

Proof. We will start off by creating a path between two generic ideals of which only one vanishing point differs. Let I be the ideal that vanishes on $S \cup \{p_I\}$ and J a generic ideal that vanishes on $S \cup \{p_J\}$. There must exist some path $\gamma : [0,1] \to \mathbb{C}^2$ such that $\gamma(0) = p_I$, $\gamma(1) = p_J$ and γ does not intersect S. Then we can create the path I_t as the ideal that vanishes on $S \cup \{\gamma(t)\}$. This is well-defined since there is a unique colength n ideal that vanishes on n points. Moreover this is continuous since γ is continuous.
Now for any two generic ideals we can repeat the construction above and move all points one by one. □

Theorem 3.2.3. *The space* $\text{Hilb}^n(\mathbb{C}^2)$ *is connected*

Proof. For any two ideals I, J we will construct a path between them. We saw before that there exists a path from I to $\text{in}(I)$, from where we can make a path to its distraction. Then we can do the same for J, and the two distractions are both generic, thus there is a path between them. Combining all of these path we get a path from I to J. Thus $\text{Hilb}^n(\mathbb{C}^2)$ is path-connected, and therefore it is also connected. □

3.3 Euler characteristic

The goal of this subsection is to calculate the Euler characteristic of $\text{Hilb}^n(\mathbb{C}^2)$. We will use the Białynicki-Birula decomposition of $\text{Hilb}^n(\mathbb{C}^2)$ to make this calculation. To talk about such a decomposition we have to have a group action of the algebraic torus $\mathbb{C}^*$ that works on our space.
We have seen that with the flat limit that we can make create a group action of $\mathbb{C}^*$ on $\text{Hilb}^n(\mathbb{C}^2)$ by defining $t \cdot I = \{f(t^a x, t^b y) | f \in I\}$.Here we choose the a and b we found for the flat limit in 3.1.3, thus $a = -2n-1$ and $b = -2n$. Since $t \in \mathbb{C}^*$ we see that the leading coefficients will not change for polynomials under this action, thus we indeed get a new element in $\text{Hilb}^n(\mathbb{C}^2)$.

Theorem 3.3.1 (Białynicki-Birula [17])**.** *Consider a smooth variety X over $\mathbb{C}$ with an action of $\mathbb{C}^*$. Let $F_1, \ldots, F_r$ be the components of $X^{\mathbb{C}^*}$. Let $X_i := \{x \in X | lim_{t\to 0} t \cdot x \in F_i\}$; Then all X_i are locally closed in X. Furthermore we have for each i that the limit map $X_i \to F_i$ is regular and it is an affine bundle.*

Thus we are interested in the fixed points of this torus action. We can take the reduced Gröbner basis of an ideal I, and notice that $t \cdot I$ is generated by polynomials that have the same leading terms as I, but only in the smaller terms change due to t. Therefore we see that these new polynomials form a reduced Gröbner basis for $t \cdot I$. But two ideals are the same if and only if they have the same reduced Gröbner basis, thus all generators cannot have any smaller terms than the leading term. Thus we see that all fixed points of $\mathbb{C}^*$ are monomial ideal. And it can also easily be verified that all monomial ideals are indeed fixed points. We have already seen that there are finitely many monomial ideals, thus they are isolated and they all forms their own component of $\mathbb{C}^*$. Therefore there are exactly $p(n)$ different X_i, where $p(n)$ is the number of partitions of n.

Theorem 3.3.2. *The Euler characteristic χ of $Hilb^n(\mathbb{C}^2)$ is equal to the number of partitions of n.*

Proof. We will use the third property of 1.3.4. We can use theorem 3.3.1 to see that the limit map is regular and an affine bundle. Because it is an affine bundle it is also a fibration. And our limit map has a codomain of one point, which is path-connected. Moreover it also orientable, once again because our codomain only has one point. Because we have an affine bundle, we see that our fiber is homeomorphic to $\mathbb{C}^k$ for some k. We can conclude that

$$\chi(X_i) = \chi(\mathbb{C}^k)\chi(\{I\}) = 1 \tag{3.5}$$

Where we have also used that the Euler characteristic of a single point is 1. Now we can use the second property of 1.3.4 to calculate:

$$\chi(\mathrm{Hilb}^n(\mathbb{C}^2)) = \sum_{i=1}^{p(n)} \chi(X_i) = \sum_{i=1}^{p(n)} 1 = p(n) \tag{3.6}$$

□

4 Yang Mills Instantons and the ADHM construction

In this section we will define connections and see how we can define the action. We will start with a mathematical preparation before we can define connections. We will then see how the Yang-Mills equations arise from minimization and reformulate the problem to the anti-self-dual equations. We mostly follow the notes from [18].

4.1 Differential forms

To formulate the Yang-Mills equations, we will have to work with differential forms. More specifically we will work with vector-valued differential forms. We will first define ordinary differential forms, and later expand that definition to vector-valued differential forms.

Definition 4.1.1. Let V, W be two vector spaces of the same field $\mathbb{K}$. We define the tensor product $V \otimes W$ to be the quotient space $F(V \times W)/\sim$. Here $F(V \times W)$ denotes the free vector space of $V \times W$, which consists of all possible finite expressions that are a sum of elements from $V \times W$ multiplied by some scalars in $\mathbb{K}$. The equivalence relation $\sim$ is the smallest equivalence relation that satisfies distributivity and scalar multiplication:

$$(v,w) + (v',w) \sim (v+v',w) \text{ and } (v,w) + (v,w') \sim (v,w+w') \qquad \text{(distributivity)}$$
$$c(v,w) \sim (cv,w) \sim (v,cw) \qquad \text{(scalar multiplication)}$$

for all $c \in \mathbb{K}$, $v, v' \in v$ and $w, w' \in W$. The equivalence class $[(v,w)]$ is denoted by $v \otimes w$. We define the k^{th} tensor power of V to be the tensor product of V with itself k times and denote this by T^kV. The tensor algebra $T(V)$ of V is defined as the algebra of all tensors on V with multiplication being the tensor product. This can be constructed by the following definition.

$$T(V) = \bigoplus_{k=0}^{\infty} T^k V. \tag{4.1}$$

Example 4.1.2. The tensor product of $\mathbb{C}^2$ with $\mathbb{C}^3$ consists is spanned by all kind of elements of the form (v,w), where $v \in \mathbb{C}^2$ and $w \in \mathbb{C}^2$. If we choose a basis e_i of $\mathbb{C}^2$ and f_i of $\mathbb{C}^3$ we can write that

$$v \otimes w = (v_1e_1 + v_2e_2) \otimes (w_1f_1 + w_2f_2 + w_3, f_3) = \sum_{i=1}^{2}\sum_{j=1}^{3} v_iw_je_i \otimes f_j \tag{4.2}$$

Thus all elements x in $\mathbb{C}^2 \otimes \mathbb{C}^3$ are sums of elements that can be written as linear combinations of $e_i \otimes f_j$, thus every x can also be written as a linear combination of $e_i \otimes f_j$. Therefore we see that all combinations of $e_i \otimes f_j$ form a basis of six elements. Moreover if we have a basis B_1 for V and a basis B_2 for W it can be shown that $B_1 \times B_2$ forms a basis of $V \otimes W$. Therefore we also see that $\mathbb{C}^n \otimes \mathbb{C}^k \cong \mathbb{C}^{nk}$. For the tensor algebra we see that our elements consist sums of scalars, vectors from V and various products of vectors. △

Definition 4.1.3. The exterior algebra $\Lambda(V)$ of a vector space V is the quotient algebra of the tensor algebra $T(V)$ by the ideal generated by elements of the form $x \otimes x$. The exterior product $\wedge$ of two elements in $\Lambda(V)$ is the induced multiplication from the tensor product in $T(V)$.

Example 4.1.4. For two elements $x, y \in V$ we know that $(x+y) \wedge (x+y)$ must be 0, since it is contained in the ideal where we have divided by. Thus we can calculate:

$$0 = (x+y) \wedge (x+y) = x \wedge x + x \wedge y + y \wedge x + y \wedge y = x \wedge y + y \wedge x. \tag{4.3}$$

Therefore we see the wedge product is antisymmetric. The elements in $\Lambda(V)$ that are the product of k vectors span the a space that we call the k^{th} exterior power. △

Definition 4.1.5. A vector bundle consists of a continuous surjection $\pi : E \to X$ called the bundle projection from the total space E to the base space X together with the structure of a vector space of dimension k on all fibers $\pi^{-1}(\{x\})$ for $x \in X$. It also has to satisfy the condition that for each $x \in X$ we can find an open U around x such that there is a homeomorphism from $\pi^{-1}(U)$ to $U \times \mathbb{K}^k$. Here $\mathbb{K}$ is the scalar field of our vector space structure for the fibers.

Example 4.1.6. Locally a vector bundle is just the Cartesian product of some vector space with your base space. But this local property does not always translate to the whole space, for instance we can see the Möbius strip as vector bundle over the circle, where our fibers have the structure of $\mathbb{R}$. If you look close by at a point on the circle it just looks like a piece of the plane, but because of the twist in the fibers the Möbius strip is a different space from $S^1 \times \mathbb{R}$. △

We will now use a definition from [4].

Definition 4.1.7. Suppose M is an n-dimensional smooth manifold and $x \in M$. A tangent vector at a point x can be defined as an equivalence class of smooth curves $\tilde{\gamma} : (-1, 1) \to M$ for which $\tilde{\gamma}(0) = x$. We consider two smooth curves $\tilde{\gamma}_1, \tilde{\gamma}_2$ equivalent if

$$\left.\frac{\mathrm{d}(\varphi\tilde{\gamma}_1(t))}{\mathrm{dt}}\right|_{t=0} = \left.\frac{\mathrm{d}(\varphi\tilde{\gamma}_2(t))}{\mathrm{dt}}\right|_{t=0}$$

in some coordinate system (U, φ).

A special kind of vector bundle is the tangent bundle. This is the vector bundle that assigns the tangent space as a fiber to each point in a manifold. The cotangent bundle is the bundle that assigns the dual of the tangent space to each point in the manifold.

Definition 4.1.8. Let M be a smooth manifold. A differential form of degree k is a smooth section of the k^{th} exterior power of the cotangent bundle of M. The set of all k-forms on M is a vector space denoted by $\Omega^k(M)$.

Example 4.1.9. In general we see that a k-form is a linear, alternating map that takes k vectors and returns a scalar. After all you need a vector for each linear map you get from the cotangent bundle. A 0-form is therefore takes no vectors and returns a scalar for each point in M, therefore a 0-form is a smooth function on M to $\mathbb{K}$. A 1-form is a smooth section of the cotangent bundle of M, therefore a 1-forms can be seen as a function that for each point in M assigns a linear map from the fiber at that point to $\mathbb{K}$. We can use the wedgeproduct we have in the the exterior algebra to multiply k-forms. The wedge-form of a p-form and a q-form is a $p + q$-form. △

At last for the regular n-forms, there is also some kind of derivative.

Definition 4.1.10. The exterior derivative d is the unique $\mathbb{K}$-linear mapping from k-forms to $k + 1$-forms that satisfies the following three conditions.

- $\mathrm{d}f = \sum_i \frac{\partial f}{\partial x_i}\mathrm{dx}_i$, so $\mathrm{d}f$ is the differential of f for a 0-form f.
- $\mathrm{d}(\mathrm{d}f) = 0$ for a 0-form f.
- $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^p(\alpha \wedge \mathrm{d}\beta)$, where α is a p-form

We can also formulate this in terms of local coordinates. We will write $\mathrm{dx}^I = \mathrm{dx}^{i_1} \wedge \cdots \wedge \mathrm{dx}^{i_k}$ for $I = (i_1, \ldots, i_k)$. Then we have for the exterior derivative d that

$$\mathrm{d}(g\ \mathrm{dx}^I) = \frac{\partial g}{\partial x^i}\mathrm{dx}^i \wedge \mathrm{dx}^I \tag{4.4}$$

Here we use the Einstein summation convention that if there is an index variable that appears twice in a single term, we take sum over all values.
Now that we have defined regular k-forms, we are ready to define vector-valued differential forms. We will state the definitions used in [19].

Definition 4.1.11. Let X be a manifold and F either a vector space or a manifold. Then F-valued differential forms $\Omega^r(U, F)$ on an open U of X are defined as $\Gamma^\infty(U, \Lambda^r(T^*X) \otimes F)$. Where $\Gamma^\infty(M, E)$ denotes the space of smooth sections from M to E.

If we have two vector spaces E, F with an action of E on F we also extend the definition of the wedge product to $\wedge : \Omega^p(U, E) \times \Omega^q(U, F) \to \Omega^{p+q}(U, F)$ by

$$a \wedge b = \frac{1}{p!q!} \sum_{P \in S_{p+q}} \mathrm{sign}(P) \alpha(X_{P(1),\ldots,X_{P(p)}}) \beta(X_{P(p+1)}, \ldots, X_{P(p+q)}). \tag{4.5}$$

Here S_n is the permutation group of order n and all X_i are vector fields on U. Moreover if $E = F$ and E is an algebra, then we define the commutator $[\alpha, \beta] = \alpha \wedge \beta - (-1)^{pq} \beta \wedge \alpha$.

Definition 4.1.12. The exterior derivative $\mathrm{d} : \Omega^k(U, F) \to \Omega^{k+1}(U, F)$ is defined as the linear operator with $\mathrm{d}\alpha = A \otimes \mathrm{d}\eta$ for $\alpha \in \Omega^k(U, F)$ with $\eta \in \Omega^k(U)$ and $A \in F$ such that $\alpha = A \otimes \eta$.

We will now see some remarks from [19] which we can use later to simplify some expressions.

Remark 4.1.13 (Lindenhovius)**.** If we have some $\alpha \in \Omega^p(U, F)$ and $\beta \in \Omega^q(U, F)$ for some algebra F and write $\alpha = T_i \alpha^i$ and $\beta = Q_i \beta^i$ where α^i, β^i are all ordinary differential forms with $T_i, Q_i \in F$ then we have:

$$\begin{aligned} \alpha \wedge \beta &= T_i Q_j \alpha^i \wedge \beta^j \\ [\alpha, \beta] &= [T_i, Q_j] \alpha^i \wedge \beta^j \\ \mathrm{d}(\alpha \wedge \beta) &= \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta \\ \mathrm{d}[\alpha, \beta] &= [\mathrm{d}\alpha, \beta] + (-1)^p [\alpha, \mathrm{d}\beta] \end{aligned} \tag{4.6}$$

Definition 4.1.14. Let us work on a manifold with a tangent space of dimension m. We write $I = (\mu_1, \ldots, \mu_r)$ with $1 \leq \mu_1 < \cdots < \mu_r \leq m$ for the multi-index of length r. With $\bar{I} = (\bar{\mu}_1, \ldots, \bar{\mu}_{m-r})$ we denote the multi-index of length $m - r$ that is also increasing and complements I. Then we write $\mathrm{dx}^I = \mathrm{dx}^{\mu_1} \wedge \cdots \wedge \mathrm{dx}^{\mu_r}$. The hodge star operator $\star : \omega^k(U, F) \to \omega^{m-k}(U, F)$ is the unique linear mapping such that $\star \mathrm{dx}^I = (-1)^{\sigma(I)} \mathrm{dx}^{\bar{I}}$, where $(-1)^{\sigma(I)}$ is the sign of the permutation $\mu_1 \ldots \mu_r \bar{\mu}_1 \ldots \bar{\mu}_{m-r}$. If we have a Riemannian manifold we get that $\star^2 = (-1)^{r(m-r)} \mathrm{Id}$.

Remark 4.1.15. If we work over $\mathbb{R}^2$ with a trivial bundle of $\mathbb{R}^2$ attached to each point. We see that $\star 1 = (-1)^{\sigma(12)} \mathrm{dx}^{(1,2)} = \mathrm{dx}^1 \wedge \mathrm{dx}^2$. Moreover we see that $\star \mathrm{dx}^1 = \mathrm{dx}^2$, $\star \mathrm{dx}^2 = -\mathrm{dx}^1$ and $\star(\mathrm{dx}^1 \wedge \mathrm{dx}^2) = 1$. We see for $1, \mathrm{dx}^1 \wedge \mathrm{dx}^2$ that if we let $\star$ act twice we get the same element back, but for the 1-forms we get a minus sign. This agrees with the sign $(-1)^{r(m-r)}$.

Now we have defined vector-valued differentials forms, we have enough tools to continue and understand the Yang-Mills equations.

4.2 Connections and curvature

In physics, gauge theory is a type of field theory which studies invariants in space. More precisely, gauge theory is concerned with local transformations that do not change the Lagrangian. Mathematically this is formulated in terms of connections, which act on a vector bundle.

Definition 4.2.1. A connection A on a vector bundle $\pi : E \to M$ of rank n is a $\mathfrak{gl}(n)$-valued 1-form. For a connection $A \in \Omega(M, \mathfrak{gl}(n))$ we also get a corresponding differential operator on M:

$$\mathrm{d}_A := \mathrm{d} + \rho(A) \tag{4.7}$$

This operator is called the covariant exterior derivative. The adjoint operator of d_A is given by $\star \mathrm{d}_A \star$.

The ρ which we encounter in this equation appears because we need some representation for A. The connection A has some value in $\mathfrak{gl}(n)$ but it has act on some section $E \to M$. For this action there are all different kinds of possibilities how to act, ρ contains the information about how this action works.
Before we continue, we have some restrictions on our connection A. We will be working with Hermitian vector bundles, these give us an inner product on all fibers and we want our connection to be compatible with that inner product.

Definition 4.2.2. A Hermitian vector bundle $\pi : E \to M$ is a complex vector bundle over M together with a Hermitian inner product on each fiber.

For our connection to be compatible with the inner product it means that for all $\alpha_1, \alpha_2 \in \Omega^0(M, E)$ we want to have that:

$$\mathrm{d}(\alpha_1, \alpha_2) = (\mathrm{d}_A\alpha_1, \alpha_2) + (\alpha_1, \mathrm{d}_A\alpha_2) \tag{4.8}$$

Because of the Leibniz rule for derivatives we can see that this simplifies to $(\rho(A)\alpha_1, \alpha_2)+(\alpha_1, \rho(A)\alpha_2) = 0$ for all α_1, α_2. Therefore we see that A must be skew-Hermitian. Thus from now on we only look at connections A that are $\mathfrak{u}(n)$-valued 1-forms acting on E by ρ. With our connection A and its differential operator we can now also define the curvature 2-form.

Definition 4.2.3. The curvature F is defined as

$$F := \mathrm{d}_A A = \mathrm{d}A + A \wedge A = \mathrm{d}A + \frac{1}{2}[A, A] \tag{4.9}$$

For any connection there is already an identity that we can verify by calculation.

Theorem 4.2.4 (Bianchi Identity)**.** *For any connection A we have $d_A F = 0$.*

Proof. By definition we get:

$$\mathrm{d}_A F = (\mathrm{d} + \rho(A))(\mathrm{d}A + \frac{1}{2}[A, A]) = \mathrm{d}^2A + \frac{1}{2}\mathrm{d}[A, A] + \rho(A)(F) \tag{4.10}$$

We start by using $\mathrm{d}^2 = 0$, thus the first term vanishes. Then for the last term we have to know how $\rho(A)$ works on F. It works on F by the adjoint action. Thus we get:

$$\rho(A)(F) = [A, F] = [A, \mathrm{d}A + \frac{1}{2}[A, A]] = [A, \mathrm{d}A] + \frac{1}{2}[A, [A, A]] \tag{4.11}$$

Here we use the commutator for vector-valued differential forms. Now we write $A = \sum_\mu A_\mu(x)\mathrm{d}x^\mu$ in local coordinates and calculate:

$$[A, [A, A]] = \sum_{\mu,\nu,\rho} [A_\mu, [A_\nu, A_\rho]]\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \tag{4.12}$$

If we look at the coefficient in front of a term $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho$ there are two options. If μ, ν, ρ are all different then there are six terms in the sum contributing to this term and we see that this coefficient is equal to:

$$[A_\mu, [A_\nu, A_\rho]] + [A_\rho, [A_\mu, A_\nu]] + [A_\nu, [A_\rho, A_\mu]] - [A_\mu, [A_\rho, A_\nu]] - [A_\nu, [A_\mu, A_\rho]] - [A_\rho, [A_\nu, A_\mu]] \tag{4.13}$$

But because of the Jacobi-identity for matrices we know that both the positive terms and the negative terms sum to 0. The second option is that μ, ν, ρ are not all different, but then we see that $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\rho = 0$. So we see that all coefficients in $[A, [A, A]]$ are zero, thus $[A, [A, A]] = 0$. Now for the exterior derivative we see:

$$\begin{aligned} \mathrm{d}[A, A] + 2[A, \mathrm{d}A] &= [\mathrm{d}A, A] - [A, \mathrm{d}A] + 2[A, \mathrm{d}A] = [\mathrm{d}A, A] + [A, \mathrm{d}A] \\ &= \mathrm{d}A \wedge A - (-1)^{1\cdot 2} A \wedge \mathrm{d}A + A \wedge \mathrm{d}A - (-1)^{1\cdot 2}\mathrm{d}A \wedge A = 0 \end{aligned} \tag{4.14}$$

Here we have used the definition of the wedge product and the last equation from 4.6. So the two terms cancel and we are left with $\mathrm{d}_A F = 0$ □

4.3 Action and instantons

In this subsection we will see how we can define the action of a connection on the the manifold $M = \mathbb{R}^4$. As you see more often in physics, we are interested in connections A such that this action is minimal. We will then see what the topological charge is of a connection and moreover see why we actually should be working over S^4 instead of $\mathbb{R}^4$.

Definition 4.3.1. The action of our system is the functional $S_E[A]$ that depends on the connection A by the following equation:

$$S_E[A] = -\int_M \mathrm{Tr}(F \wedge \star F). \tag{4.15}$$

The gauge group acts on F by the adjoint operator, and moreover a change of basis also acts as a adjoint. Since the trace is cyclic we can pass both these operators to the same side where they cancel out. Thus the action is invariant under the Gauge group and under choice of coordinates. In the intersection of two opens the transition also acts on F by the adjoint operator, therefore this expression is also globally defined (2.2.7 in [19]).

We are interested in connections A such the action is minimal. We will use standard technique of setting the derivative equal to zero for this to happen. Assume we have some perturbation $A + t\alpha$ for small t, then we can calculate:

$$\begin{aligned} F[A + t\alpha] &= \mathrm{d}(A + t\alpha) + (A + t\alpha) \wedge (A + t\alpha) \\ &= \mathrm{d}A + \mathrm{d}t\alpha + A \wedge A + t\alpha \wedge A + tA \wedge \alpha + t^2\alpha \wedge \alpha \\ &= F[A] + t(\mathrm{d}\alpha + A \wedge \alpha + \alpha \wedge A) + t^2(\alpha \wedge \alpha) \\ &= F[A] + t(\mathrm{d}_A\alpha) + t^2(\alpha \wedge \alpha) \end{aligned} \tag{4.16}$$

Here we have used that the A and α are one-forms, thus the commutator is the sum of two wedge products. We also notice that $\int_M \mathrm{Tr}(A \wedge \star B)$ is a bilinear form that we will note by (A, B). We can now also calculate for the action that

$$S_E[A + t\alpha] = (F[A + t\alpha], F[A + t\alpha])^2 = (F[A], F[A])^2 + 2t(F[A], \mathrm{d}_A\alpha) + O(t^2) \tag{4.17}$$

Now we see that for A to minimize the action, we must have that the derivative to t is zero, therefore we should have $(F[A], \mathrm{d}_A\alpha) = 0$ for all α. Now we can use the fact that $\star\mathrm{d}_A\star$ is the adjoint operator of d_A to conclude that $\star\mathrm{d}_A \star F[A] = 0$, which implies that $\mathrm{d}_A \star F = 0$. Together with the Bianchi identity $\mathrm{d}_A F = 0$ these equations are called the Yang-Mills equations, and solutions to these equations are called Yang-Mills connections.

Now we can perform a similar calculation to another quantity called the topological charge.

Definition 4.3.2. Let M be some manifold with $\dim M = 4$. The topological charge k of a connection A is defined as the following integral:

$$k := -\frac{1}{8\pi^2} \int_M \mathrm{Tr}(F \wedge F) \tag{4.18}$$

When we perform the same variational argument we find that we encounter $\mathrm{d}_A F$ in the derivative instead of $\mathrm{d}_A \star F$. Because of the Bianchi identity that derivative is always zero for each connection, thus we see that k does not depend on A. It can be proven that if $M = S^4$, then k is an integer. This is because the expression for k can be recognized as some Chern class. Chern classes form some topological invariant for complex vector bundles. A proof for this can be found in [19].

The curvature 2-forms live in the space $\Omega^2(U, \mathfrak{u}(n))$ for U some open in a 4-dimensional manifold M. We can see that implies that $** : \Omega^2(U, \mathfrak{u}(n)) \to \Omega^2(U, \mathfrak{u}(n))$ is equal to the identity, so therefore the possible eigenvalues of $\star$ are ± 1. Therefore we can split our spaces in the direct sum of the eigen spaces of $\star$:

$$\Omega^2(U, \mathfrak{u}(n)) = \Omega^2_+ \oplus \Omega^2_- \tag{4.19}$$

We call these two spaces the self-dual and anti-self-dual spaces. For any curvature form F we can decompose F as $F_+ + F_-$, we can now calculate that:

$$\begin{aligned} -8\pi^2 k &= \int_M \mathrm{Tr}(F_+ + F_-) \wedge (F_+ + F_-) \\ &= \int_M \mathrm{Tr}(F_+ \wedge \star F_+) - \int_M \mathrm{Tr}(F_- \wedge \star F_-) \\ &= -\int_M |F_+|^2 + \int_M |F_-|^2 \end{aligned} \tag{4.20}$$

Here we have used that the eigenspaces are orthogonal. Moreover we can take the absolute values of both sides to find $8\pi^2|k| \leq |S_E[A]|$. Therefore we have a lower bound for the action, which we wanted to minimalize, that can only be attained if $F_+ = 0$ or $F_- = 0$. We also see that if one of those parts is 0, we actually get this minimum. Therefore we know that a solution on M must be self-dual or anti-self-dual. We can now also see that the minimal value of the action is equal for all connections, since in every minimum we have $8\pi^2 k$ as our value of $S_E(A)$. All connections that satisfy the anti-self-dual equations are called the Yang-Mills instantons. Let us consider an anti-self-dual instanton on $\mathbb{R}^4$. We can then write $F = \sum_{\mu,\nu} F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, where we sum over all $\mu < \nu$. We obtain that

$$\star F = F_{12}\mathrm{d}x^3 \wedge \mathrm{d}x^4 - F_{13}\mathrm{d}x^2 \wedge \mathrm{d}x^4 + F_{14}\mathrm{d}x^2 \wedge \mathrm{d}x^3 + F_{23}\mathrm{d}x^1 \wedge \mathrm{d}x^4 - F_{24}\mathrm{d}x^1 \wedge \mathrm{d}x^3 + F_{34}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \quad (4.21)$$

And we can compare this with $-F$ to get that

$$\begin{aligned} F_{12} + F_{34} &= 0 \\ F_{13} + F_{42} &= 0 \\ F_{14} + F_{23} &= 0 \end{aligned} \quad (4.22)$$

Here we have used that $F_{24} = -F_{42}$.
From the definition of F we can deduce that in coordinate language we get $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$.

Proposition 4.3.3. *We can write $F_{\mu\nu} = [D_\mu, D_\nu]$, where $D_\mu = (d_A)_\mu$.*

Proof. Let $\Phi \in \Omega^k(U, F)$ be some differential form our operators act on.

$$\begin{aligned} D_\mu(D_\nu\Phi) &= (\partial_\mu + A_\mu)((\partial_\nu + A_\nu)\Phi) = \partial_\mu(\partial_\nu\Phi) + \partial_\mu(A_\nu\Phi) + A_\mu(\partial_\nu\Phi) + A_\mu(A_\nu\Phi) \\ &= 0 + (\partial_\mu A_\nu\Phi + A_\nu\partial_\mu\Phi) + A_\mu\partial_\nu\Phi + A_\mu A_\nu\Phi \end{aligned} \quad (4.23)$$

$$\begin{aligned} [D_\mu, D_\nu]\Phi &= (\partial_\mu A_\nu\Phi + A_\nu\partial_\mu\Phi) + A_\mu\partial_\nu\Phi + A_\mu A_\nu\Phi - (\partial_\nu A_\mu\Phi + A_\mu\partial_\nu\Phi) - A_\nu\partial_\mu\Phi - A_\nu A_\mu\Phi \\ &= \partial_\mu A_\nu\Phi + A_\mu A_\nu\Phi - \partial_\nu A_\mu\Phi - A_\nu A_\mu\Phi = (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\Phi \end{aligned} \quad (4.24)$$

Here we have used the Leibniz rule for the derivative and the fact that $\mathrm{d}^2 = 0$. □

We can also look at what happens when the values of our connection 1-form are in an abelian group, instead of $\mathfrak{u}(n)$. Then our commutator in F vanishes and we get

$$||F||^2 = (F, F) = \pm(\star F, F) = \pm(\mathrm{d}A, \mathrm{d}A) = \pm(\star\mathrm{d}\star\star A, \mathrm{d}A) = \pm(\mathrm{d}^\dagger \star A, \mathrm{d}A) = \pm(\star A, \mathrm{d}^2 A) = 0. \quad (4.25)$$

Therefore we see that with an abelian structure group all instantons are trivial and equal to 0.
From now on we will focus on anti-self-dual instantons. We have two invariants for these instantons, namely n and k. An important remark is that it is not trivial that k is even an finite number, since we can possibly integrate over non-compact spaces such as $\mathbb{R}^4$. It is however shown by Uhlenbeck [20]that all instantons on $\mathbb{R}^4$ with a bounded norm, and thus bounded topological charge, may be obtained from an instanton on $S^4 = \mathbb{R}^4 \cup \{\infty\}$. We can now conclude that k should also be an integer for instantons on $\mathbb{R}^4$, since the topological charge is an integer for solutions on S^4. Such instantons with a finite topological charge are called framed instantons. We denote the moduli space of framed instantons that satisfy the anti-self-dual equation by $\mathcal{M}^{fr}_{ASD}$.

5 ADHM construction

In this section we will see how we can use some linear data to construct Yang-Mills instantons on $\mathbb{R}^4$. Moreover it is proven in [21] that this construction gives us all the solutions to the anti-self-dual equations. This idea is called the ADHM construction, named after Michael Atiyah, Vladimir Drinfeld, Nigel Hitchin and Yuri I. Manin.

5.1 ADHM data

If we take x_1, x_2, x_3, x_4 as our coordinates in $\mathbb{R}^4$, we can write this in $\mathbb{C}^2$ as $z_1 = x_2 + ix_1$, $z_2 = x_4 + x_3$. We can also take our real differential operators $(\mathrm{d}_A)_\mu = D_\mu$ and transform them to complex coordinates. We define our operators $\mathcal{D}_1, \mathcal{D}_2$ as follows:

$$\begin{aligned} \mathcal{D}_1 &= \frac{1}{2}(D_2 - iD_1) \\ \mathcal{D}_2 &= \frac{1}{2}(D_4 - iD_3) \end{aligned} \tag{5.1}$$

Theorem 5.1.1. *The anti-self-dual equations in terms of D_μ are equivalent to the following equations for $\mathcal{D}_1, \mathcal{D}_2$:*

$$\begin{aligned} [\mathcal{D}_1, \mathcal{D}_2] &= 0 \\ [\mathcal{D}_1, \mathcal{D}_1^\dagger] + [\mathcal{D}_2, \mathcal{D}_2^\dagger] &= 0 \end{aligned} \tag{5.2}$$

Proof. First of all we want to take the equations from 4.22 and use 4.3.3 to formulate them in terms of the D_μ operators. We then get that

$$\begin{aligned} [D_1, D_2] + [D_3, D_4] &= 0 \\ [D_1, D_3] + [D_4, D_2] &= 0 \\ [D_1, D_4] + [D_2, D_3] &= 0 \end{aligned} \tag{5.3}$$

We can fill in the definition of $\mathcal{D}_1, \mathcal{D}_2$ to get that

$$4[\mathcal{D}_1, \mathcal{D}_2] = [D_2 - iD_1, D_4 - iD_3] = [D_2, D_4] - [D_1, D_3] - i([D_1, D_4] + [D_2, D_3]) \tag{5.4}$$

We see that equating this to zero is equivalent to the second and third equation of 5.3.
Furthermore we want to know what $D_\mu^\dagger$ is. Component-wise we have $D_\mu = \partial_\mu + \rho(A_\mu)$. We now that A is $\mathfrak{u}(n)$-valued, thus A_μ is skew hermitian and thus $A_\mu^\dagger = -A_\mu$. Furthermore our inner product is the integral of the product of functions. With partial integration we can see that:

$$\int_{-\infty}^{\infty} (\partial_\mu f(x_\mu)) \cdot g(x_\mu) \mathrm{d}x_\mu + \int_{-\infty}^{\infty} f(x_\mu) \cdot (\partial_\mu g(x_\mu)) \mathrm{d}x_\mu = [f(x_\mu)g(x_\mu)]_{-\infty}^{\infty} = 0 \tag{5.5}$$

Thus in terms of inner products we see that $(\partial_\mu f, g) + (f, \partial_\mu g) = 0$, thus $\partial_\mu^\dagger = -\partial_\mu$. Therefore we get that $D_\mu^\dagger = -D_\mu$. Then we can also conclude for $\mathcal{D}_1$

$$\mathcal{D}_1^\dagger = \frac{1}{2}(D_2^\dagger - (iD_1)^\dagger) = \frac{1}{2}(D_2^\dagger + iD_1^\dagger) = \frac{1}{2}(-D_2 - iD_1) \tag{5.6}$$

Then we can calculate

$$4[\mathcal{D}_1, \mathcal{D}_1^\dagger] = [D_2 - iD_1, -D_2 - iD_1] = [D_2, -iD_1] + [-iD_1, -D_2] = 2i[D_1, D_2] \tag{5.7}$$

Analogously we see $4[\mathcal{D}_2, \mathcal{D}_2^\dagger] = 2i[D_3, D_4]$, thus we see that the second equation in 5.2 is equivalent to the first equation of 5.3. □

For the ADHM construction you start with some linear data that satisfies equations that look quite similar to these equations in $\mathcal{D}_1, \mathcal{D}_2$. However this data consists of just some matrices instead of linear operators, and are therefore a lot simpler to work with.

Definition 5.1.2 (ADHM Data)**.** An ADHM system on $\mathbb{C}^2$ is a set of linear data:

1. Vector spaces V, W over $\mathbb{C}$ with $\dim V = k$ and $\dim W = n$.
2. Complex $k \times k$ matrices B_1, B_2, a $k \times n$ matrix I and an $n \times k$ matrix J.

Remark 5.1.3. With this data we see that $B_1, B_2, B_1^\dagger, B_2^\dagger$ all take vectors from V and send them back to V, while $I^\dagger, J$ both take vectors from V and take them to W and finally $I, J^\dagger$ both take vectors from W and send them back to V.

Definition 5.1.4 (ADHM system)**.** A set of ADHM Data is an ADHM system $(U, V, W, B_1, B_2, I, J)$ if it satisfies the following two constraints:

1. The matrices B_1, B_2, I, J satisfy the ADHM equations:

$$\begin{aligned} [B_1, B_2] + IJ &= 0 \\ [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J &= 0 \end{aligned} \tag{5.8}$$

2. For all $x, y \in \mathbb{C}^2$ with $x = (z_1, z_2), y = (w_1, w_2)$ and z_1, z_2, w_1, w_2 not all 0 we have the maps $\alpha_{x,y} : V \to W \oplus V \oplus V$ and $\beta_{x,y} : W \oplus V \oplus V \to V$:

$$\begin{aligned} \alpha_{x,y} &= \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \\ \beta_{x,y} &= \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \end{pmatrix} \end{aligned} \tag{5.9}$$

   And we want $\alpha_{x,y}$ to always be injective, while $\beta_{x,y}$ should be surjective for all such x, y.

Lemma 5.1.5. *If (B_1, B_2, I, J) all satisfy the constraints for an ADHM system, then for any $g \in U(k)$ we also have that $(gB_1g^{-1}, gB_2g^{-1}, gI, Jg^{-1})$ satisfies the constraint*

Proof. Assume that (B_1, B_2, I, J) satisfies the ADHM constraint.

$$\begin{aligned} [gB_1g^{-1}, gB_2g^{-1}] + gIJg^{-1} = g[B_1, B_2]g^{-1} + gIJg^{-1} = g([B_1, B_2] + IJ)g^{-1} = 0 \\ [gB_1g^{-1}, (gB_1g^{-1})^\dagger] + [gB_2g^{-1}, (gB_2g^{-1})^\dagger] + gI(gI)^\dagger - (Jg^{-1})^\dagger Jg^{-1} = \\ g[B_1, B_1^\dagger]g^{-1}] + g[B_2, B_2^\dagger]g^{-1} + gII^\dagger g^{-1} - gJ^\dagger Jg^{-1} = \\ g([B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J)g^{-1} = 0 \end{aligned} \tag{5.10}$$

Here we have used $[gXg^{-1}, gYg^{-1}] = g[X, Y]g^{-1}$ and $(gXg^{-1})^\dagger = (g^{-1})^\dagger X^\dagger g^\dagger = gX^\dagger g^{-1}$ since g is unitary. We can also calculate that $\alpha_{x,y}$ and $\beta_{x,y}$ transform to:

$$\begin{aligned} \alpha_{x,y} &= \begin{pmatrix} g(w_2 J - w_1 I^\dagger)g^{-1} \\ g(-w_2 B_1 - w_1 B_2^\dagger - z_1)g^{-1} \\ g(w_2 B_2 - w_1 B_1^\dagger + z_2)g^{-1} \end{pmatrix} \\ \beta_{x,y} &= \begin{pmatrix} g(w_2 I + w_1 J^\dagger)g^{-1} & g(w_2 B_2 - w_1 B_1^\dagger + z_2)g^{-1} & g(w_2 B_1 + w_1 B_2^\dagger + z_1)g^{-1} \end{pmatrix} \end{aligned} \tag{5.11}$$

So we see injectivity and surjectivity are preserved since g is invertible. ∎

So we care about our ADHM systems modulo $U(V)$.

5.2 Making connections

We will now take some ADHM system and construct a Yang-Mills instanton from this data. This construction creates a bijection between the set of ADHM systems and the space of Yang-Mills instantons.

Theorem 5.2.1. *For an ADHM system the following diagram forms a short chain complex. Moreover the homology group of $W \oplus V \oplus V$ is a vector space with dimension n for all x, y.*

$$V \xrightarrow{\alpha_{x,y}} W \oplus V \oplus V \xrightarrow{\beta_{x,y}} V \tag{5.12}$$

Proof. For this to be a chain complex we only have to check that $\mathrm{Im}\alpha_{x,y} \subset \ker \beta_{x,y}$, or equivalently that $\beta_{x,y} \circ \alpha_{x,y} = 0$.

$$\begin{aligned}\beta_{x,y}\alpha_{x,y} = &\begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \end{pmatrix} \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \\ =&(w_2^2 IJ - w_2 w_1 II^\dagger + w_1 w_2 J^\dagger J - w_1^2 J^\dagger I^\dagger) + (-w_2^2 B_2 B_1 - w_1 w_2 B_2 B_2^\dagger - z_1 w_2 B_2 \\ &+ w_1 w_2 B_1^\dagger B_1 + w_1^2 B_1^\dagger B_2^\dagger + w_1 z_1 B_1^\dagger - w_2 z_2 B_1 - w_1 z_2 B_2^\dagger - z_1 z_2) + (w_2^2 B_1 B_2 \\ &- w_1 w_2 B_1 B_1^\dagger + w_2 z_2 B_1 + w_1 w_2 B_2^\dagger B_2 - w_1^2 B_2^\dagger B_1^\dagger + w_1 z_2 B_2^\dagger + w_2 z_1 B_2 - w_1 z_1 B_1^\dagger + z_1 z_2) \\ =&w_1^2 (J^\dagger I^\dagger + B_1^\dagger B_2^\dagger - B_2^\dagger B_1^\dagger) + w_1 w_2 (J^\dagger J - II^\dagger - B_2 B_2^\dagger + B_1^\dagger B_1 - B_1 B_1^\dagger + B_2^\dagger B_2) \\ &+ w_2^2 (IJ - B_2 B_1 + B_1 B_2) \\ =&w_1^2 ([B_1, B_2] + IJ)^\dagger - w_1 w_2 ([B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J) + w_2^2 ([B_1, B_2] + IJ)\end{aligned} \tag{5.13}$$

And in the last line we can recognize the ADHM equations, thus $\beta_{x,y} \circ \alpha_{x,y} = 0$. The homology group at $W \oplus V \oplus V$ is defined as $\ker \beta_{x,y} / \mathrm{Im}\alpha_{x,y}$. Since this is a vector space and we have an inner product on this space, we see that this is the same as $\ker \beta_{x,y} \cap (\mathrm{Im}\alpha_{x,y})^\perp$. Since we are working over finite-dimensional vector spaces we have $\mathrm{Im}\alpha_{x,y})^\perp = \ker \alpha_{x,y}^\dagger$. It now follows that

$$W \oplus V \oplus V = \ker(\alpha_{x,y}^\dagger) \oplus \ker(\alpha_{x,y}^\dagger)^\perp \subseteq \ker(\alpha_{x,y}^\dagger) \oplus \ker(\beta_{x,y}) \tag{5.14}$$

Because we also have inclusion the other way around we find that $\ker(\alpha_{x,y}^\dagger) \oplus \ker(\beta_{x,y}) = W \oplus V \oplus V$. Thus we see it has dimension $n + 2k$. Since $\alpha_{x,y}$ is injective we know that $\dim(\ker(\alpha_{x,y}^\dagger)) = n + k$, since $\alpha_{x,y}^\dagger$ is surjective. Moreover we also see that $\dim(\ker(\beta_{x,y})) = n + k$. With the dimension theorem we now get:

$$\dim(\ker(\alpha_{x,y}^\dagger) \cap \ker(\beta_{x,y})) = \dim(\ker(\alpha_{x,y}^\dagger)) + \dim(\ker(\beta_{x,y})) - \dim(\ker(\alpha_{x,y}^\dagger) \oplus \ker(\beta_{x,y})) = n \tag{5.15}$$

□

We can combine the given maps $\alpha_{x,y}$ and $\beta_{x,y}$ to define a new map $R_{x,y}$ as follows:

$$R_{x,y} := \begin{pmatrix} \beta_{x,y} \\ \alpha_{x,y}^\dagger \end{pmatrix} : W \oplus V \oplus V \to V \times V \tag{5.16}$$

The kernel of $R_{x,y}$ is exactly the homology group we saw before, therefore we see that $R_{x,y}$ has an n-dimensional kernel. Since it maps from an $n + 2k$-dimensional space to a $2k$-dimensional space this implies that $R_{x,y}$ is surjective. We can use the definitions of $\beta_{x,y}, \alpha_{x,y}^\dagger$ to write this more precisely:

$$R_{x,y} = \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \\ -\bar{w}_1 I + \bar{w}_2 J^\dagger & -\bar{w}_1 B_2 - \bar{w}_2 B_1^\dagger - \bar{z}_1 & -\bar{w}_1 B_1 + \bar{w}_2 B_2 + \bar{z}_2 \end{pmatrix} \tag{5.17}$$

The operator $R_{x,y}$ also gives rise to its adjoint operator which we denote by $\Delta_{x,y}$.

$$\Delta_{x,y} := R_{x,y}^\dagger = \begin{pmatrix} \beta_{x,y}^\dagger & \alpha_{x,y} \end{pmatrix} = \begin{pmatrix} \bar{w}_2 I^\dagger + \bar{w}_1 J & w_2 J - w_1 I^\dagger \\ \bar{w}_2 B_2^\dagger - \bar{w}_1 B_1 + \bar{z}_2 & -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ \bar{w}_2 B_1^\dagger + \bar{w}_1 B_2 + \bar{z}_1 & w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \tag{5.18}$$

The presence of x and y in this expression can be described more elegantly when we associate the complex pairs with their quaternionic operator. We begin by defining four matrices τ_μ by $\tau_\mu = i\sigma_\mu$ for the first three matrices (here σ_μ are the Pauli-matrices) and $\tau_4 = I_2$. These matrices give rise to a ring isomorphism from

$\mathbb{H}$ to the real span of τ_μ by $(i,j,k,1) \mapsto (\tau_1,\tau_2,\tau_3,\tau_4)$. We can then write for $x \in \mathbb{R}^4$ that $x = x^\mu \tau_\mu$ and similarly get $\bar{x} = x^\mu \tau_\mu^\dagger$. Now for $x = (z_1, z_2)$ we can write $z_1 = x_2 + ix_1$ and $z_2 = x_4 + x_3$ and use this isomorphism to get the following identification

$$q = (q_1, q_2) \leftrightarrow \begin{pmatrix} \bar{q}_2 & -q_1 \\ \bar{q}_1 & q_2 \end{pmatrix} \tag{5.19}$$

Now we will add to our identification that x should be seen as $x \otimes I_k$ and y as $y \otimes I_k$. Thus we get:

$$x = (z_1, z_2) \leftrightarrow \begin{pmatrix} \bar{z_2} I_k & -z_1 I_k \\ \bar{z_1} I_k & z_2 I_k \end{pmatrix} \tag{5.20}$$

We define two other matrices:

$$a = \begin{pmatrix} I^\dagger & J \\ B_2^\dagger & -B_1 \\ B_1^\dagger & B_2 \end{pmatrix}, b = \begin{pmatrix} 0 & 0 \\ I_k & 0 \\ 0 & I_k \end{pmatrix} \tag{5.21}$$

Now we can see that $\Delta_{x,y} = ay + bx$. With this new expression we can quickly calculate a new observation, namely that $\Delta^\dagger_{xq,yq} = \bar{q}\Delta^\dagger_{x,y}$. An important consequence of this observation is that $\ker \Delta^\dagger_{xq,yq} = \ker \Delta^\dagger_{x,y}$.

$$\Delta^\dagger_{xq,yq} = (awq + bzq)^\dagger = q^\dagger (aw + bz)^\dagger = q^\dagger \Delta^\dagger_{x,y} \tag{5.22}$$

We will now use these kernels to create a vector bundle. We write $E_{x,y} = \ker \Delta^\dagger_{x,y}$. Thus for each point in $\mathbb{C}^2 \backslash \{0\}$ we now have some vector space $E_{x,y}$. But we have also seen that this vector space does not change if we multiply by some quaternion. Therefore we can create a vector bundle E over $\mathbb{P}^1(\mathbb{H})$ by choosing some representative in the equivalence class, transforming it back to two complex numbers x, y and use $E_{x,y}$ as the vector space for this point. We have seen that this does not depend on which representative is chosen. Also we know that these vector spaces all have dimension n, thus they are all isomorphic. By construction we can also see that they change continuously and therefore form a vector bundle. As topological spaces we even have $\mathbb{P}^1(\mathbb{H}) \cong S^4$, thus we can also see this as a vector bundle over S^4.
Because our coordinates are projective over $\mathbb{H}$ we can from now on assume that $y = 1$, thus $(w_1, w_2) = (0, 1)$. When $y = 1$ we will not write the subscript anymore for our operators.
We will use the elements in $\ker \Delta^\dagger$ to construct our connection. First of all we can calculate $\Delta^\dagger \Delta : V \times V \to V \times V$.

$$\Delta^\dagger \Delta = \begin{pmatrix} \beta_{x,1} \\ \alpha^\dagger_{x,1} \end{pmatrix} \begin{pmatrix} \beta^\dagger_{x,1} & \alpha_{x,1} \end{pmatrix} = \begin{pmatrix} \beta_{x,1}\beta^\dagger_{x,1} & \beta_{x,1}\alpha_{x,1} \\ \alpha^\dagger_{x,1}\beta^\dagger_{x,1} & \alpha^\dagger_{x,1}\alpha_{x,1} \end{pmatrix} = \begin{pmatrix} \beta_{x,1}\beta^\dagger_{x,1} & 0 \\ 0 & \alpha^\dagger_{x,1}\alpha_{x,1} \end{pmatrix} \tag{5.23}$$

Here we have used that $\beta_{x,y}\alpha_{x,y} = 0$. Furthermore we see:

$$\beta_{x,1}\beta^\dagger_{x,1} - \alpha^\dagger_{x,1}\alpha_{x,1} = II^\dagger - J^\dagger J + [B_2, B_2^\dagger] + [B_1, B_1^\dagger] = 0 \tag{5.24}$$

Thus we see that $\beta_{x,1}\beta^\dagger_{x,1} = \alpha^\dagger_{x,1}\alpha_{x,1}$. Furthermore we know that $\text{rank}(\beta\beta^\dagger) = \text{rank}(\beta) = k$ since β is surjective, and since $\beta\beta^\dagger$ is a square $k \times k$ matrix we also see that it is invertible. So we can write $\beta\beta^\dagger = f^{-1}$ for some matrix f. Also we can clearly see that $\Delta^\dagger\Delta$ is Hermitian, thus f^{-1} is also Hermitian, and therefore f is also Hermitian.
We can also construct an orthonormal matrix M whose columns span $\ker \Delta^\dagger$. Since the domain of $\Delta^\dagger$ is $W \oplus V \oplus V$ and the dimension of its kernel is $n + 2k - 2k = n$ (due to surjectivity) we see that M is a $(n + 2k) \times n$ matrix.

Theorem 5.2.2. *Define Q as $\Delta f \Delta^\dagger$ and P as $MM^\dagger$. Here the f in the definition of Q acts on both components of elements in $V \times V$. Both Q and P map from $W \oplus V \oplus V$ to itself and are projection operators. Furthermore we have that $Q + P = 1$.*

Proof. To see that P, Q are projection operators we have to show that $P^2 = P$ and $Q^2 = Q$. Since M is orthonormal we have $M^\dagger M = \text{Id}$. Therefore it follows $P^2 = MM^\dagger MM^\dagger = M\text{Id}M^\dagger = P$. For Q we can calculate that

$$Q^2 = \Delta f \Delta^\dagger \Delta f \Delta^\dagger = \Delta f \begin{pmatrix} f^{-1} & 0 \\ 0 & f^{-1} \end{pmatrix} f \Delta^\dagger = \Delta f \Delta^\dagger = Q \tag{5.25}$$

We will define $S := 1 - P - Q$. We note that because of $\Delta^\dagger M = 0$ and $M^\dagger \Delta = 0$ it follows $PQ = QP = 0$. We now see that $QS = Q - PQ - Q^2 = Q - Q = 0$, and since f and Δ are of maximal rank it follows that $\Delta^\dagger S = 0$. But that implies that all rows of S are elements of $\ker \Delta^\dagger$, and thus are a linear combination of the rows in M. Therefore we see that $S = MB$ for some matrix B. But now we can calculate:

$$0 = P - P = PS = MM^\dagger MB = MB = S \tag{5.26}$$

And thus $S = 0$, and $Q + P = 1$. □

If we look at the trivial bundle $S^4 \times (W \oplus V \oplus V)$ and the vector bundle E, then $M^\dagger$ works as a mapping from $W \oplus V \oplus V$ to E_x for each $x \in S^4$. Notice that M depends on $\Delta_x^\dagger$, and is not constant for all x. Now from this bundle projection a connection A is induced on E by defining the covariant derivative d_A by (page 85 in [19]):

$$\mathrm{d}_A f = P\mathrm{d}f \tag{5.27}$$

where $f : U \subset M \to E$ is a section. Here we have to be precise about what is happening on the left side. Since f is a section on E and d is working on things in $W \oplus V \oplus V$, we need to embed elements from E_x to $W \oplus V \oplus V$. This is done by the map M so we actually get $P\mathrm{d}(Mf)$ for the right hand side. But since our result is now in $W \oplus V \oplus V$, we should also embed the left hand side from E_x to $W \oplus V \oplus V$, and thus the left hand side should actually be $M\mathrm{d}_A f$.
We can now calculate for a section s on E:

$$M\mathrm{d}_A s = P\mathrm{d}(Ms) = MM^\dagger(M\mathrm{d}s + (\mathrm{d}M)s) = M(\mathrm{d}s + (M^\dagger \mathrm{d}M)s) \tag{5.28}$$

And when we compare this to the definition of the covariant derivative we see that $A = M^\dagger \mathrm{d}M$. Thus we have now constructed a connection from our ADHM system. We can now also calculate that $A^\dagger = -A$, this follows from:

$$0 = \mathrm{d}(\mathrm{id}_{E_x}) = \mathrm{d}(M^\dagger M) = (\mathrm{d}M^\dagger)M + M^\dagger(\mathrm{d}M) = A^\dagger + A \tag{5.29}$$

Since A is anti-Hermitian, we see that it is $u(n)$-valued. To show that A is anti-self-dual we do some pre-computing.

$$\partial_\mu(M^\dagger \partial_\nu M) = (\partial_\mu M^\dagger)(\partial_\nu M) + M^\dagger(\partial_\mu \partial_\nu M) = (\partial_\mu M^\dagger)(\partial_\nu M) \tag{5.30}$$

$$0 = \partial_\mu(0) = \partial_\mu(\Delta_x^\dagger M) = (\partial_\mu \Delta_x^\dagger)M + \Delta_x^\dagger \partial_\mu M \tag{5.31}$$

$$\partial_\mu \Delta_x = \partial_\mu(a + b\tau_\nu x^\nu) = b\tau_\mu \tag{5.32}$$

Now we can calculate:

$$\begin{aligned}
\partial_\mu A_\nu + A_\mu A_\nu &= \partial_\mu(M^\dagger \partial_\nu M) + (M^\dagger \partial_\mu M)(M^\dagger \partial_\nu M) \\
&= (\partial_\mu M^\dagger)(\partial_\nu M) - (\partial_\mu M^\dagger)M(M^\dagger \partial_\nu M) && \text{(By 5.30)} \\
&= (\partial_\mu M^\dagger)(1 - MM^\dagger)(\partial_\nu M) \\
&= (\partial_\mu M^\dagger)Q(\partial_\nu M) \\
&= (\partial_\mu M^\dagger)\Delta_x f \Delta_x^\dagger(\partial_\nu M) \\
&= M^\dagger(\partial_\mu \Delta_x)f(\partial_\nu \Delta_x^\dagger)M && \text{(By 5.31)} \\
&= M^\dagger b\tau_\mu f \tau_\nu^\dagger b^\dagger M && \text{(By 5.32)}
\end{aligned}$$

Since x was identified with $x \oplus I_k$, we also get that τ_μ should be identified with $\tau_\mu \otimes I_k$. Now we see that:

$$\tau_\mu f \tau_\nu^\dagger = (\tau_\mu \otimes I_k)(I_2 \otimes f)(\tau_\nu^\dagger \otimes I_k) = \tau_\mu \tau_\nu^\dagger \otimes f \tag{5.33}$$

All these calculations are useful because we can now calculate $F_{\mu\nu}$.

$$\begin{aligned}
F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu \\
&= M^\dagger b\tau_\mu f \tau_\nu^\dagger b^\dagger M - M^\dagger b\tau_\nu f \tau_\mu^\dagger b^\dagger M \\
&= M^\dagger b((\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger) \otimes f)b^\dagger M
\end{aligned} \tag{5.34}$$

Now we can check by calculation that $\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger$ is anti-self-dual, and thus we see that $F_{\mu\nu}$ is also anti-self-dual. For this calculation we have to check that the following equation is satisfied.

$$\frac{1}{2}\epsilon_{\mu\nu\rho\phi}(\tau_\rho \tau_\phi^\dagger - \tau_\phi \tau_\rho^\dagger) = -(\tau_\mu \tau_\nu^\dagger - \tau_\nu \tau_\mu^\dagger) \tag{5.35}$$

Theorem 5.2.3. *The topological charge of the bundle E is $-k$.*

This theorem is proven in [19] and once again uses Chern-Theory to relate the integral to some Chern class. So from every $ADHM$ system we can create a Yang-Mills instanton over S^4, with a topological charge of $-k$ and a rank of n. Moreover we see that this must also correspond with a framed instanton on $\mathbb{R}^4$.

6 Geometrization

In this section we will look at properties of the space of instantons by looking at the space of ADHM systems. Moreover we will see the Hilbert scheme in another way and see how that connects a specific moduli space of instantons.

6.1 Moduli space of instantons

We will study the Yang-Mills instantons by studying the set of ADHM systems. Because of the bijective correspondence we found, we can reformulate properties back in the instanton space. We will start with an example.

Example 6.1.1. Assume we want to study all possible circles in $\mathbb{R}^2$. You could start with saying that three points in the plane often have a unique circle that crosses all three points. As long as those points are three different points and they are not all on a line there actually exists a unique circle through all three points. Therefore you could think of the set of all three points with those conditions as the set that parametrizes all circles. This is some set in $\mathbb{R}^6$ that is also a 6-dimensional manifold. But a lot of these triples correspond to the same circle, so you want to look at this set of triples modulo some relation to make sure that you actually get a one-to-one correspondence. One could also try to parametrize all circles by thinking of the set $\mathbb{R}^2 \times \mathbb{R}_{>0}$ by saying that each circle has a unique centre and some positive radius.
Another remark you could make is that two circles that have the same radius but different centres are essentially the same. Therefore you could also take the space of circles and look at it modulo translations. From that point of view the space of all circles is parametrized by $\mathbb{R}_{>0}$. Another question you could ask if you also want to include extreme cases. For instance you could ask if a circle with radius 0 exists. Or when you consider tangent circles and want to use Descartes' theorem it makes sense to define circles with infinite radius as lines. △

In this example the set $\mathbb{R}_{>0}$ is called a moduli space for the set of all circles. Generally a moduli space is a geometric space whose points represent objects of some fixed kind, or isomorphism classes of such objects. This idea helps us in the classification of those objects, since we can often use coordinates in our moduli space to describe all points. In our cases we will use the space of ADHM data as the moduli space for our Yang-Mills instantons.
We will now follow the notes from [18], but now lecture 2, to define the moduli space of instantons. All theorems that are not proven here can also be found there.

Definition 6.1.2. ADHM data consists of two vector spaces V, W with $\dim V = k$ and $\dim W = n$ and maps $B_1, B_2 \in \mathrm{End}(V), I \in \mathrm{Hom}(W, V), J \in \mathrm{Hom}(V, W)$. Where the all the maps satisfy $[B_1, B_2] + IJ = 0$ and $[B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J = 0$. We have that $U(V)$ acts on this space by $(gB_1g^{-1}, gB_2g^{-1}, gI, Jg^{-1})$.

Definition 6.1.3. The moduli space of framed genuine instantons $\mathcal{M}^{reg}(n, k)$ is given by taking the set of $ADHM$ data (B_1, B_2, I, J) with trivial stabilizer under the action of $U(V)$ and quotienting out $U(V)$.

There is also another construction which is equivalent to this space. Here we start with a smaller set and quotient by a larger set. This different notion will prove to be very useful, but will not be proven here.

Theorem 6.1.4. *If V, W are vector spaces with $\dim V = k$ and $\dim W = n$ and we have $B_1, B_2 \in End(V), I \in Hom(W, V), J \in Hom(V, W)$ satisfying only $[B_1, B_2] + IJ = 0$. Let GL_n act on this space as before. Then the space of all such data with trivial GL_n stabilizer, quotiented by GL_n is equivalent to $\mathcal{M}^{reg}(n, k)$.*

Definition 6.1.5. For ADHM data (B_1, B_2, I, J) there are two properties which we will name.

1. (Stability) There is no proper subspace $S \subset V$ such that $B_i(S) \subset S$ and $I(W) \subset S$.
2. (Co-stability) There is no proper subspace $S \subset V$ such that $B_i(S) \subset S$ and $S \subset \ker J$.

Proposition 6.1.6. *The ADHM data that was used for the construction of instantons is both stable and co-stable.*

Proof. Assume that S exists such that $S \subset \ker J$ and $B_i(S) \subset S$. Then we know that $J|_S = 0$, thus we see that B_1 and B_2 commute on S. Commuting matrices always share an eigenvector v. Therefore we can see that if we choose $(w_1, w_2) = (0, 1)$ and try to calculate $\alpha_{x,y}v$. Since $v \in S$ we have $Jv = 0$. Therefore we get:

$$\alpha_{x,y}v = \begin{pmatrix} Jv \\ -B_1v - z_1v \\ B_2v + z_2v \end{pmatrix} = \begin{pmatrix} 0 \\ -\lambda_1 v - z_1 v \\ \lambda_2 v + z_2 v \end{pmatrix} \tag{6.1}$$

Here λ_i are the eigenvalues of v for B_i. We can now choose $z_1 = -\lambda_1$ and $z_2 = -\lambda_2$ and see that we have found a non-zero vector that is in the kernel of $\alpha_{x,y}$, but this contradicts the assumption that $\alpha_{x,y}$ is injective for all x, y.
Thus we see that all ADHM data from our construction is co-stable. For stability, we will use a duality between ADHM data. It is easily verified that if (B_1, B_2, I, J) satisfies the ADHM equations and $\alpha_{x,y}, \beta_{x,y}$ are respectively injective and surjective, then $(B_2^\dagger, B_1^\dagger, J^\dagger, I^\dagger)$ also satisfies al those conditions. So we know that $(B_2^\dagger, B_1^\dagger, J^\dagger, I^\dagger)$ is co-stable. If there is a proper subspace $S \subset V$ that violates stability, then we also have that $B_i^\dagger(S^\perp) \subset S^\perp$, and $S^\perp \subset \ker I^\dagger$. Therefore this violates the co-stability of $(B_2^\dagger, B_1^\dagger, J^\dagger, I^\dagger)$. Thus our data is stable and co-stable. □

Theorem 6.1.7. *If (B_1, B_2, I, J) satisfies either the stability or the co-stability condition, then it has trivial stabilizer under GL_n.*

Proof. Assume that (B_1, B_2, I, J) has a non-trivial stabilizer. Then there must be some $g \in \mathrm{GL}_n$ such that $(B_1, B_2, I, J) = (gB_1g^{-1}, gB_2g^{-1}, gI, Jg^{-1})$. We can then take $S = \ker(g - \mathrm{id}_n)$ which violates stability. Since $(g - \mathrm{id}_n)Ix = gIx - Ix = 0$ and because g commutes with both B_i, we have $(g - \mathrm{id}_n)B_i = B_i(g - \mathrm{id}_n)$. Therefore we see that $B_i(S) \subset S$.
We can also see that $\mathrm{Im}(g - \mathrm{id}_n)$ violates co-stability. □

Moreover with these constraints we can reformulate our moduli space $\mathcal{M}^{reg}(n,k)$ to all ADHM data that satisfy both ADHM equations and are regular, still modulo $U(V)$. We can also use 6.1.4 to see that this space can also be seen as the space of $ADHM$ data that is regular and only satisfies the first ADHM equation, modulo GL_n.

Definition 6.1.8. The compactification $\hat{\mathcal{M}}(n,k)$ of $\mathcal{M}^{reg}(n,k)$ is defined as the set of ADHM data that satisfies the first ADHM equation, modulo GL_n. Furthermore we define $\tilde{\mathcal{M}}(n,k)$ as the set of ADHM data that satisfies the stability condition and also satisfies the first ADHM equation, again modulo GL_n.

Now it can be shown that the space $\hat{\mathcal{M}}(n,k)$ has singularities and $\tilde{\mathcal{M}}(n,k)$ is the minimal resolution of this space for all n, k.

6.2 Hilbert schemes in a different perspective

To connect our moduli spaces with the Hilbert scheme of points, we start by establishing a link from the Hilbert scheme to a definition in terms of matrices.

Definition 6.2.1. A cyclic vector v_0 for a set of $n \times n$ matrices is a vector such that the whole linear space is spanned by the set of vectors obtained by applying the matrices repeatedly on v_0.

Theorem 6.2.2. *Let A be the set of all pairs of $n \times n$ matrices (B_1, B_2) such that they commute and admit a cyclic vector. Here GL_n acts on this space by conjugation. Then $A/GL_n \cong \mathrm{Hilb}^n(\mathbb{C}^2)$.*

Proof. First of all we remark that if $g \in \mathrm{GL}_n$, then $gB_1g^{-1} \cdot gB_2g^{-1} = gB_1B_2g^{-1} = gB_2B_1g^{-1}$, and thus (gB_1g^{-1}, gB_2g^{-1}) is still a commuting pair of matrices. Moreover we see that a cyclic vector v_0 for (B_1, B_2) can be transformed into a cyclic vector for gB_1g^{-1}, gB_2g^{-1} by taking gv_0. Thus the action of GL_n is well-defined.
We will give two mappings which are each others inverses to show the correspondence.
First of all we construct a map from $\mathrm{Hilb}^n(\mathbb{C}^2)$ to A/GL_n. For some ideal I we know that $\mathbb{C}[x,y]/I$ is an n-dimensional vector space. Choose some basis B for this space and let M_x be the map that sends a

polynomial $P(x,y)$ to $xP(x,y)$. It is easily checked that this is a linear map. Now define M_y as the map that multiplies by y. Since multiplying by x and by y commute, we see that M_x and M_y will also commute. Moreover we see that $1 \in \mathbb{C}[x,y]/I$ corresponds to some vector that is cyclic, since all polynomials can be written as a sum of monomials, which are exactly the vectors obtained by applying M_x and M_y repeatedly to 1. If we choose a different basis, we will just get conjugation by some element in GL_n to our matrices, but that will give a pair of matrices in the same equivalence class.
Now for a map from A/GL_n back to $\mathrm{Hilb}^n(\mathbb{C}^2)$. For some $(M,N) \in A$ we define the ideal $I = \{P \in \mathbb{C}[x,y] | P(M,N) = 0\}$. We see that conjugation by some matrix in GL_n does not change these polynomials, since we have $P(gMg^{-1}, gNg^{-1}) = gP(M,N)g^{-1}$. Moreover this notion of polynomials is only well-defined because M, N commute, otherwise we would get that xy and yx are different polynomials. It is also clear by definition that I is some ideal, so we only have to show that its colength is n. We do know that (M,N) admit a cyclic vector v_0, thus we can choose n polynomials Q_i such that $Q_i(M,N)v_0$ form a basis of $\mathbb{C}^n$. We see that these also form a basis for $\mathbb{C}[x,y]/I$, since linear dependence in $\mathbb{C}[x,y]/I$ of these vectors is the same as linear dependence in $\mathbb{C}^n$ due to the fact that all polynomials in I are zero. Moreover we can choose our Q_i to be minimal in the sense that they must generate $\mathbb{C}[x,y]$.
It is easily checked that composing these maps gives the identity. □

Theorem 6.2.3. *We have an isomorphism* $\tilde{\mathcal{M}}(1,k) \cong \mathit{Hilb}^k(\mathbb{C}^2)$.

To prove this equivalence we first prove a lemma.

Lemma 6.2.4. *For all* $(B_1, B_2, I, J) \in \tilde{\mathcal{M}}(1,k)$ *we have* $J = 0$.

Proof. We see that $\mathbb{C}[B_1,B_2]I$ fulfils $B_i(S) = S$ and $I(W) \subset S$, thus because of stability we get that $\mathbb{C}[B_1,B_2]I = V$. Thus if we prove $JP(B_1,B_2)I$ for any polynomial P we can can conclude $J = 0$. Because of linearity it suffices to prove this for monomials. We also note that $JI : \mathbb{C} \to \mathbb{C}$, and is therefore some scalar multiplication. Thus we see that $JI = \mathrm{Tr}(IJ) = -\mathrm{Tr}([B_1,B_2]) = 0$, where we have used that $[B_1,B_2] + IJ = 0$ and the fact that the trace of a commutator is always zero. We will use induction on the degree of the monomial to prove $JP(B_1,B_2)I = 0$ for some monomial P. We will now prove that we can change the order of the matrices in P. We will use our induction hypothesis to see that for monomials P, Q in B_1, B_2 we have
$$JPB_2B_1QI = JP(IJ + B_1B_2)QI = JPB_1B_2QI. \tag{6.2}$$
Here we have used our induction hypothesis for P, since the total degree of the monomial is clearly larger than the degree of P. Furthermore we see that for any monomial R we have:
$$\mathrm{Tr}(JRI) = \mathrm{Tr}(IJR) = \mathrm{Tr}(B_2B_1R) - \mathrm{Tr}(B_1B_2R) \tag{6.3}$$
We see that if $R = B_i^k$ for some $k \in \mathbb{N}$ that we can use the cyclic property to show both traces are equal. For the other cases we will use a more combinatorial argument.
We can rewrite any monomial such that $R = B_1^n B_2^k$. We see that $\mathrm{Tr}(JRI)$ does not change under permutations of the order in R. We can now see
$$\begin{aligned}
\mathrm{Tr}(JRI) &= \mathrm{Tr}(B_2B_1B_1^nB_2^k) - \mathrm{Tr}(B_1B_2B_1^nB_2^k) = \mathrm{Tr}(B_2B_1B_1^nB_2^k) - \mathrm{Tr}(B_2B_1^nB_2^kB_1)\\
\mathrm{Tr}(JRI) &= \mathrm{Tr}(B_2B_1^nB_2^kB_1) - \mathrm{Tr}(B_1B_2B_1^{n-1}B_2^kB_1) = \mathrm{Tr}(B_2B_1^nB_2^kB_1) - \mathrm{Tr}(B_2B_1^{n-1}B_2^kB_1^2)\\
\mathrm{Tr}(JRI) &= \mathrm{Tr}(B_2B_1^{n-1}B_2^kB_1^2) - \mathrm{Tr}(B_1B_2B_1^{n-2}B_2^kB_1^2) = \mathrm{Tr}(B_2B_1^{n-1}B_2^kB_1^2) - \mathrm{Tr}(B_2B_1^{n-2}B_2^kB_1^3)
\end{aligned} \tag{6.4}$$
Here we have used 6.3 repeatedly and use the cyclic property to change the last trace such that it starts with B_2B_1. Then we choose a permutation of R such that our first term becomes the last term from the equation before. We can keep on repeating this process until we have a last term that we have already encountered as a first term. We can then sum up all traces to find that $l \cdot \mathrm{Tr}(JRI) = 0$ for some number l, thus $\mathrm{Tr}(JRI) = 0$.
We can conclude that $J\mathbb{C}[B_1,B_2]I = 0$, thus $J = 0$. □

Proof of 6.2.3. We saw that the elements of $\tilde{\mathcal{M}}(1,k)$ are quadruplets (B_1,B_2,I,J) such that $[B_1,B_2]+IJ = 0$ and the stability condition is satisfied, modulo GL_n.
We have seen in theorem 6.2.2 that we have to show that A/GL_n is isomorphic to $\bar{\mathcal{M}}(1,k)$.

We start with a map from A/GL_n. If we have some pair (B_1, B_2), we know that there must be some cyclic vector v_0. Since $n = 1$ we see that we need some map $I : \mathbb{C} \to V$. We will choose the map with $1 \mapsto v_0$. In total we will map our pair to $(B_1, B_2, v_0, 0)$. Because the image of I contains v_0, we see that the stability property must be satisfied.
Now if we have $(B_1, B_2, I, J) \in \bar{\mathcal{M}}(1, k)$ we have seen that $J = 0$, thus B_1 and B_2 commute. We can also look at the space $\mathbb{C}[B_1, B_2]I(W)$. Because of the stability property we see that this must be the whole of V, and thus $I(1)$ is a cyclic vector for B_1, B_2. $\square$

This equivalence does have some strange physical interpretation. It is known that there exist no $U(1)$ instantons, but we have a non-empty moduli space of these instantons. For instance in [22] it is discussed why these new instantons appear and how these should be interpreted.

7 Conclusion

Let us summarize what we have done in this thesis. We've seen the definition of the Euler characteristic for simplicial complexes and seen how it extends to compact complex manifolds. Moreover we have extended the definition for non-compact manifolds and varieties. Subsequently we have defined the Hilbert scheme of points over $\mathbb{C}^2$ and have seen how monomial ideals play an important role in this space. After equipping this set with a topology we have seen how we can use the flat limit to make a path between ideals. This all led to a calculation of the Euler characteristic of the Hilbert space of n points over $\mathbb{C}^2$, which is equal to the number of partitions of n as we saw in 3.3.2.

In chapter 4 we have introduced differential forms and seen how they are used in Yang-Mills theory. From the Yang-Mills functional the equations of motion can be derived by the principle of least action. We have then used the structure of $\mathbb{R}^4$ to formulate the self-dual and anti-self-dual equations. Thereafter we have defined Yang-Mills instantons and seen why we can view them over S^4 instead of $\mathbb{R}^4$. Furthermore we have constructed all Yang-Mills instantons over $\mathbb{R}^4$ with the ADHM construction. Finally we have studied the moduli space of all instantons. In 6.2.3 we have seen a connection between these moduli spaces and the Hilbert scheme of points.

For future research one could try to calculate the Euler characteristic for the Hilbert scheme of points over other surfaces. For instance in [23] a more general result for the Euler characteristic is proven. In [24] the Euler characteristic of moduli space of instantons plays an important role, which we have calculated in this thesis for a specific case. The ADHM construction can also be generalized to find the anti-self-dual instantons in more dimensions [25].

References

[1] Michael Friedman. *A History of Folding in Mathematics: Mathematizing the Margins*. Birkhäuser, 2018, p. 71. ISBN: 978-3-319-72486-7.

[2] James R. Munkres. *Elements of Algebraic Topology*. 2725 Sand Hill Road, Menlo Park, California 94025: Westview Press, 1984. ISBN: 978-0-2016-2728-2.

[3] Allen Hatcher. *Algebraic topology*. Cambridge University Press, 2002, pp. 128–130. ISBN: 0-521-79540-0.

[4] Viktor V. Prasolov. *Elements of Combinatorial and Differential Topology*. Vol. 74. The American Mathematical Society, 2006. ISBN: 0-8218-3809-1.

[5] W. Fulton. *Introduction to toric varieties*. Vol. 131. The William H. Roever Lectures in Geometry. Princeton,NJ: Princeton University Press, 1993, p. 141. ISBN: 0691033323.

[6] Jörg Schürmann. *Topology of Singular Spaces and Constructible Sheaves*. Vol. 63. Monografie Matematyczne. Birkhäuser Basel, 2003. ISBN: 978-3-7643-2189-5.

[7] Sylvain E. Cappell, Laurentiu Maxim, and Julius L. Shaneson. "Euler characteristics of algebraic varieties". In: (2006). arXiv: `math/0606654 [math.AT]`.

[8] Edwin H. Spanier. *Algebraic Topology*. Springer, New York, NY, 1966. ISBN: 978-1-4684-9322-1.

[9] Collectif. *Techniques de construction et théorèmes d'existence en géométrie algébrique IV : les schémas de Hilbert*. fr. Séminaire Bourbaki 6. talk:221. Société mathématique de France, 1961.

[10] John Fogarty. "Algebraic Families on an Algebraic Surface". In: *American Journal of Mathematics* 90.2 (1968), pp. 511–521. ISSN: 00029327, 10806377.

[11] Mark Haiman and Bernd Sturmfels. "Multigraded Hilbert Schemes". In: (2002). arXiv: `math/0201271 [math.AG]`.

[12] Ezra Miller and Bernd Sturmfels. *Combinatorial Commutative Algebra*. Springer Science+Bussiness Media, Inc., 2005. ISBN: 0-387-22356-8.

[13] JOHN W. MILNOR and JAMES D. STASHEFF. *Characteristic Classes. (AM-76)*. Princeton University Press, 1974. ISBN: 9780691081229. URL: `http://www.jstor.org/stable/j.ctt1b7x751`.

[14] Clifford H. Wagner. *Symmetric, cyclic, and permutation products of manifolds*. eng. Warszawa: Instytut Matematyczny Polskiej Akademi Nauk, 1980. URL: `http://eudml.org/doc/268344`.

[15] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Vol. 150. Graduate Texts in Mathematics. Spring-Verlag New York, 1995. ISBN: 978-0-387-94268-1.

[16] Heisuke Hironaka. "Resolution of Singularities of an Algebraic Variety Over a Field of Characteristic Zero: I". In: *Annals of Mathematics* 79.1 (1964), pp. 109–203. ISSN: 0003486X. URL: `http://www.jstor.org/stable/1970486`.

[17] A. Bialynicki-Birula. "Some Theorems on Actions of Algebraic Groups". In: *Annals of Mathematics* 98.3 (1973), pp. 480–497. ISSN: 0003486X. URL: `http://www.jstor.org/stable/1970915`.

[18] Alex Atanasov. *Instantons and the ADHM Construction*. lecture notes. 2016.

[19] A.J. Lindenhovius. *Instantons and the ADHM construction*. 2010.

[20] Karen K. Uhlenbeck. "Removable singularities in Yang-Mills fields". In: *Communications in Mathematical Physics* 83.1 (1982), pp. 11–29. DOI: `cmp/1103920742`. URL: `https://doi.org/`.

[21] M.F. Atiyah et al. "Construction of instantons". In: *Physics Letters A* 65.3 (1978), pp. 185–187. ISSN: 0375-9601. DOI: `https://doi.org/10.1016/0375-9601(78)90141-X`. URL: `https://www.sciencedirect.com/science/article/pii/037596017890141X`.

[22] H. W. Braden and N. A. Nekrasov. *Instantons, Hilbert Schemes and Integrability*. 2001. arXiv: `hep-th/0103204 [hep-th]`.

[23] L. Goettsche. "The Betti numbers of the Hilbert scheme of points on a smooth projective surface". In: *Mathematische Annalen* 286 (1990), pp. 193–207.

[24] Cumrun Vafa and Edward Witten. “A strong coupling test of S-duality”. In: *Nuclear Physics B* 431.1-2 (Dec. 1994), pp. 3–77. ISSN: 0550-3213. DOI: 10.1016/0550-3213(94)90097-3. URL: http://dx.doi.org/10.1016/0550-3213(94)90097-3.

[25] Koki Takesue. “ADHM construction of (anti-)self-dual instantons in 4n dimensions”. In: *Journal of High Energy Physics* 2017.7 (July 2017). ISSN: 1029-8479. DOI: 10.1007/jhep07(2017)110. URL: http://dx.doi.org/10.1007/JHEP07(2017)110.