

Faculty of Physics

Physics and Geometry

Notes of lectures by dr. Ingmar Saberi

Academic year 2024/2025
Winter semester

Authors:

William Luciani
Simon Bukovšek

Last change: February 26, 2025

Contents

1 Observables and States **2**
1.1 An extended example: $\mathbb{R}^n$ as a space in different ways 10

2 Sheaves and Čech cohomology **14**
2.1 Ideas and examples 14
2.2 Definitions 18

3 Sheaf Cohomology **21**
3.1 Ideas and boring sheaves 21
3.2 Vector bundles as locally free sheaves 23
3.3 Derivations and one forms 27
3.4 Differential forms 31

4 From observables to symmetries **34**

5 Special lecture: Geometric quantization **38**
5.1 Quantizing Kähler manifolds 42

6 Poisson structure vs. Symplectic structure **44**

7 Origin of the symplectic structure **47**

8 Noether's theorem **50**

9 What is a directional derivative? **54**

10 Connections, curvature and compatibility **59**
10.1 The Fundamental Theorem of Riemannian Geometry 62

Index **64**

1 Observables and States

Roughly, there are three ways geometry shows up in physics:

- **Spacetime geometry**: general relativity, geodesics, ...
- **Phase space geometry**: Hamiltonian mechanics, Liouville theorem, ...
- **Geometry of gauge fields**: principal bundles, connections, monopoles, Dirac charge quantization, ...

The one we'll start with is the geometry of phase space. This is because of the following it's obviously abstract (not related to "real" space), there's no notion of distance, and we can only access it "indirectly" by measuring observables.
What structures are there on the set of all observables (Obs) in classical mechanics?

- I can add any two measurements: given $f, g \in \text{Obs}$, then we have $f+g \in \text{Obs}$.
- I can also multiply two measurements: $f, g \in \text{Obs} \Rightarrow f \cdot g \in \text{Obs}$ and in addition $f \cdot g = g \cdot f$.
- There are trivial measurements, 0 and 1, that are units for $+$ and $\cdot$.
- I can rescale measurements: given $f \in \text{Obs}$, we have $\alpha f \in \text{Obs}$, where α is some number.

The first three properties tell us that Obs is a commutative ring.

Definition 1.1. A *commutative ring* is a set R equipped with the following properties:

- an addition operation $+ : R \times R \to R$ that is associative and commutative;
- an element called 0 such that $a + 0 = a$ for all $a \in R$;
- a map called negative: $R \to R$, $a \mapsto -a$ such that $a + (-a) = 0$ for each $a \in R$;
- a multiplication operation $\cdot : R \times R$, $(a, b) \mapsto ab$ that is associative and commutative;
- an element called 1 such that $a \cdot 1 = a$ for all $a \in R$;
- a distribution law: $a(b + c) = ab + ac$ for all $a, b, c \in R$.

Examples of commutative rings are $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/n\mathbb{Z}$, fuctions of any set valued in any other ring, polynomials, smooth functions, ...

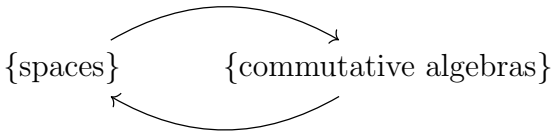
Definition 1.2. A map $\varphi : R \to S$ is a map of commutative rings when $\varphi(a+b) = \varphi(a) + \varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$.

The fourth property above tells us that there's even more structure: a *commutative algebra* over some collection of "numbers", which we call k (which itself is a commutative ring).

Now that we've passed from phase space to observables, some questions naturally arise:

- In what sense does the algebra Obs tell me the phase space?
- What is a "point" of Obs (or a "state")?
- Is the space of these states the same as the phase space?

The idea can be represented as follows



The arrow from left to right is what we've been considering up to now: we can get a commutative algebra from the phase space by considering observables, functions on phase space valued in k. An important question is whether we can go backwards, i.e. reconstruct the phase space by measuring observables. If we can go backward: what notion of "space" do we get? This is where we'll need a bit of topology.
In other words we'll try to understand the geometry of the phase space by studying algebraic properties of (or structures in) Obs.

To begin to answer these questions we first need to understand what a state should be. In a specific state, I should know a numerical value for every observable.
Attempt: A point or state is a map of commutative rings $\varphi : \text{Obs} \to k$. To decide whether this works, we need to think about k a bit more carefully. In particular, we insist that k is a field. This is related to being a complete measurement: knowing unambiguously if measurements are 0 or not.
Better attempt: A state is a map $\varphi : \text{Obs} \to k$, where k is a field.
Successful attempt:

Definition 1.3. A *state* is a map $\varphi : \mathrm{Obs} \to k$ (of commutative rings) valued in a field, up to the equivalence relation:

$$\varphi \sim \varphi' \iff \begin{array}{ccc} & \overset{\varphi}{\nearrow} & k \\ \mathrm{Obs} & & \downarrow \tau \\ & \underset{\varphi'}{\searrow} & k' \end{array}$$

Additionally, in the equivalence class there must be a surjective map.

This, maybe surprisingly, works! The collection of states defined this way usually goes under a different name:

Definition 1.4 (MaxSpec)**.** Let R be a commutative ring. The set of maps $R \to k$ into a field k as above is called the maximal spectrum of the ring R, $\mathrm{MaxSpec}\, R$ [a].

[a] $\mathrm{Spec}\, R$ has an analogous definition but without the surjectivity requirement for the maps.

So we'll indicate the collection of states by $\mathrm{MaxSpec}\, \mathrm{Obs}$.

Example 1. This example allows us to understand the additional condition of having a surjective map. Consider the ring $\mathbb{Z}$, we have the following maps into fields

$$\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$$
$$\mathbb{Z} \to \mathbb{Q}$$

which simply send $1 \in \mathbb{Z}$ to $1 \in \mathbb{Z}/p\mathbb{Z}, \mathbb{Q}$. In addition there are no maps $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p'\mathbb{Z}$ for two different primes p and p' and no maps $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Q}$ or $\mathbb{Q} \to \mathbb{Z}/p\mathbb{Z}$. However, the map into $\mathbb{Q}$ is not a surjective map and we don't want to include it since it's just the map from the ring into its field of fractions which is always present. It is often said that $\mathbb{Z}/p\mathbb{Z}$ are "points" of $\mathbb{Z}$ and $\mathbb{Q}$ is called a "generic point" and is something additional that we now don't want to consider[1]. We thus have

$$\mathrm{MaxSpec}(\mathbb{Z}) = \{\text{primes}\}\,.$$

End of lecture 1

[1] This is the difference between $\mathrm{Spec}(\mathbb{Z})$ and $\mathrm{MaxSpec}(\mathbb{Z})$, the presence of the generic point.

Recall: A *topology* on a set X is a collection Opens $\subset P(X)$ (the power set of X) satisfying:

- $\varnothing, X \in$ Opens,
- an arbitrary union of elements in Opens is again in Opens,
- a finite intersection of Opens is again in Opens.

A topology is *Hausdorff* when, for any x and y, $x \neq y$ there exist $U, V \in$ Opens with $x \in U$ and $y \in V$, so that $U \cap V = \varnothing$. It is also called a separation axiom. A function $f : X \to Y$ is continuous if $f^{-1}(V)$ is open for any open $V \subset Y$. If we pick Opens $(X) = \{\varnothing, X\}$ (the indiscrete topology) then every continuous map to $\mathbb{R}$ is constant.

Now, our definition of a state doesn't seem very practical: how can one check every field? It would be better if we had some intrinsic characterization of the set of states associated to Obs. In order to do this we have to think a bit more about maps of commutative rings in general. First we start with an analogy working with vector spaces. A map $f : V \to W$ of vector spaces has an image $\mathrm{im}(f) \subseteq W$ which is a vector subspace of W and allows us to define the *rank*

$$\mathrm{rk}(f) = \dim \mathrm{im}(f)$$

and we also have the rank–nullity theorem:

$$\mathrm{rk}(f) = \dim V - \dim \ker(f) = \dim(V/\ker f)$$

For any such map we have the following commutative diagram

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & W \\ \downarrow & & \uparrow \\ V/\ker f & \xrightarrow{\ \cong\ } & \mathrm{im}\, f \end{array}$$

which tells us that any map may be factorized into a quotient, an isomorphism and an injection.

If V is a vector space and K is a vector subspace, the *quotient* V/K is the set of equivalence classes of elements of V under the following relation: $v \sim w$ whenever $v - w \in K$. We write the equivalence classes in the form $v + K$. It turns out that V/K is a vector space:

- $0 + K$ is a class,
- $(v + K) + (w + K) = v + w + K$,

- $\alpha(v + K) = \alpha v + K$.

How does this play out for maps of (commutative) rings? First, consider a map $f : A \to B$. What structures are on $\ker f$ and $\operatorname{im} f$? Start with image of f. It is a subring of B:

- $0_B = f(0_A)$, $1_B = f(1_A)$,
- $a, a' \in A$: $f(a) + f(a') = f(a + a')$ and the same for multiplication.

What about $\ker f \subset A$? Recall that $\ker f = \{a \in A | f(a) = 0_B\}$. It holds:

- kernel is closed under sums: $a, a' \in A$, $f(a + a') = f(a) + f(a') = 0_B + 0_B = 0_B$;
- $0_A \in \ker f$;
- it is closed under arbitrary products with elements of A: $a \in \ker f$: $f(a \cdot a') = f(a)f(a') = 0_B f(a') = 0_B$ for any $a' \in A$.
- 1_A is typically not in $\ker A$. If it were, $\ker f = A$.

A subset satisfying the first three properties is called an *ideal* in A. Any kernel of a ring map is an ideal. Conversely: if $I \subset A$ is an ideal, A/I is a ring:

$$A/I = \{a \in A\}/\sim, \quad a \sim a' \text{ when } a - a' \in I.$$

We have

- $(a + I) + (a' + I) = a + a' + I$,
- $(a+I)(a'+I) = aa' + aI + a'I + II = aa' + I$, due to ideal property $I \cdot A \subset I$.

Just as for vector spaces, we can factor maps.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow & & \uparrow \\ A/\ker f & \xrightarrow{\cong} & \operatorname{im} f \end{array}$$

Idea: Replace an equivalence class of complete measurements by the ideal $\ker \varphi$. In order to make sense of this we first need to characterize fields better.

Claim 1.1. *A field is a commutative ring k with no proper nontrivial ideal.*

Proper ideal means it is not the whole ring and nontrivial means it's not the zero ideal.

Proof. Let I be a nontrivial ideal. Consider a nontrivial element $x \in I$, $x \neq 0$. Then we have $x^{-1}x \in I$ but then we have $1 \in I$ which tells us that $I = k$ and the ideal is not proper. □

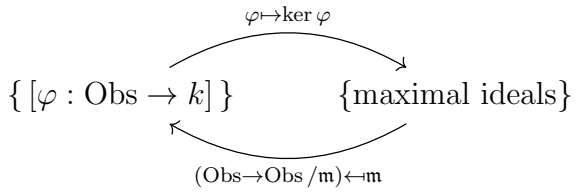
Intuition: An ideal is like a measurement. Which observables do I know to be zero?
A complete measurement is one after which there is nothing left to measure: *a maximal ideal.*
In particular, an ideal I is called maximal if there are no proper ideals J with $I \subset J \subset A$. If I is a maximal ideal, then the quotient A/I is actually a field.
We can now give a different definition of MaxSpec:

Definition 1.5 (MaxSpec, again)**.** The maximal spectrum of a commutative ring R is the collection of its maximal ideals.

This works because of the following. Consider the other definition, of a surjective map to a field $f : R \to k$, then $\operatorname{im} f = k$ and $R/\ker f \cong \operatorname{im} f$ is then also a field, so $\ker f$ is a maximal ideal[2].

End of lecture 2

We've given two definitions of MaxSpec, so we'd like to understand better why they're equivalent and how to translate between the two. We saw that we can get a maximal ideal from a map $\varphi : \mathrm{Obs} \to k$ by simply taking its kernel $\ker \varphi$. The only thing we should check is that two equivalent maps $\varphi \sim \varphi'$ have the same kernel, but that's immediately true because maps of fields $k \to k'$ are always injective.
We also have a natural way of getting a map to a field given a maximal ideal $\mathfrak{m}$: we can take the quotient map $\mathrm{Obs} \to \mathrm{Obs}/\mathfrak{m}$.
Schematically we have



So the two definitions coincide.

[2] An analgous definition works for Spec. If we consider maps to a field, then $\operatorname{im} f$ is a subring of the field k and since k has no zero divisors, also $\operatorname{im} f$ has none. Such a ring is called an integral domain and since $\operatorname{im} f \cong R/\ker f$ we have that $R/\ker f$ is also an integral domain. Such ideals, whose quotient is an integral domain, are called prime ideals. $\operatorname{Spec} R$ is thus the set of prime ideals of R.

Now that we have a notion of state, we can go back to another question we had: suppose I know Obs, the algebra of functions on a space X. Do I recover X as the set of "states"? And if X has some additional structure, do I recover that as well? The following proposition is a start towards answering these questions.

Proposition 1.2. *Let X be a topological space that is compact and Hausdorff and let $C^0(X, \mathbb{R})$ be the algebra of continuous real valued functions on X. Then* $\mathrm{MaxSpec}(C^0(X, \mathbb{R}))$ *is homeomorphic to X.*

In order to talk about continuous maps and homeomorphisms we need to put a topology on $\mathrm{MaxSpec}\, C^0(X, \mathbb{R})$. Since we're trying to identify the space of maximal ideals with the space X itself, the idea is that the topology of $\mathrm{MaxSpec}\, C^0(X, \mathbb{R})$ should be defined by requiring or declaring all elements of Obs to be continuous functions on $\mathrm{MaxSpec}\, C^0(X, \mathbb{R})$ itself! Take the basis of open sets as follows: for any element $f \in \mathrm{Obs}$ we take

$$U_f = \{\text{maximal ideals } \mathfrak{m} \subset \mathrm{Obs} \mid f \notin \mathfrak{m}\}$$

i.e. "the set of states for which f doesn't vanish", to be open. This topology is called the *Zariski topology*. In a sense what we're doing is implementing the idea we have that "inequalities are open conditions". By choosing these as a basis for the topology we're simply taking the weakest topology in which the sets that we expect to be open are in fact open. We'll see below that this is precisely the topology that implements the idea above.

We can now prove the preposition.

Proof.

Step 1: find a map $\varphi : X \to \mathrm{MaxSpec}(C^0(X))$. We take the map

$$x \mapsto ev_x = (f \mapsto f(x))$$

where we're thinking of ev_x as a map of commutative rings $C^0(X) \to \mathbb{R}$ which (by taking its equivalence class) gives an element of $\mathrm{MaxSpec}(C^0(X))$. This map is called "evaluation at x". We may also consider the other definition of MaxSpec and take the maximal ideal $\ker ev_x$ (from here on we'll use the two interchangeably).

Step 2: φ is surjective. Given $\mathfrak{m} \in \mathrm{MaxSpec}(C^0(X))$ we define

$$V_{\mathfrak{m}} = \{x \in X \mid f(x) = 0,\ \forall f \in \mathfrak{m}\}$$

which we call the "vanishing locus of $\mathfrak{m}$".
$V_{\mathfrak{m}}$ can't be empty for any proper maximal ideal. In fact suppose that at every point x we had a function $f_x \in \mathfrak{m}$ which at that point is nonzero, i.e. $f_x(x) \neq 0$.

Then we can take $f_x(x) = 1$ and $f_x \geq 0$ everywhere (e.g. by squaring it). Now, f_x is a continuous function, so there's a neighborhood $U_x \ni x$ such that $f_x|_{U_x} > 0$. Now the collection of U_x obviously cover X and by compactness we can find a finite subcover, labelled by a finite collection of points $I \subset X$, where I is a finite set. Then we can take the function

$$f = \sum_{x \in I} f_x$$

which by definition of the f_x is positive everywhere. Therefore we also have the continuous function $1/f$. But, since the f_x are in $\mathfrak{m}$, then so is f and so is $f \cdot 1/f = 1$. But if $1 \in \mathfrak{m}$ then we don't have a proper ideal.
So there exists $x \in V_{\mathfrak{m}}$, i.e. there is a point such that all functions in $\mathfrak{m}$ vanish at that point. Then $\mathfrak{m} \subset \ker ev_x$, but $\mathfrak{m}$ is maximal, so actually

$$\mathfrak{m} = \ker ev_x = \varphi(x)$$

i.e. φ is surjective.

Step 3: φ is injective. Intuitively, this means "any two states are distinguishable by at least one measurement". By definition injective means

$$\forall x, y \in X,\ \ker ev_x \neq \ker ev_y$$

or, equivalently,

$$\forall x, y \in X,\ \exists f \in C^0(X) : f(x) = 0, f(y) \neq 0$$

or, equivalently, any two points can be "separated" by a continuous function. Intuitively this is clear, but for arbitrarily weird topological spaces it may not be true. However, since X is Hausdorff we have this results thanks to Urysohn's lemma.

Step 4: φ is continuous. This is almost by design! We took the basis of opens to be

$$U_f = \{\mathfrak{m} \mid f \notin \mathfrak{m}\}$$

but we understood any $\mathfrak{m}$ is of the form $\ker ev_x$ for some $x \in X$. So U_f is the collection of $\ker ev_x$ such that $f \notin \ker ev_x$, i.e. the collection of all of the $\ker ev_x$ for which $f(x) \neq 0$. If we take the preimage of this open set then we get the set

$$\{x \in X \mid f(x) \neq 0\} = f^{-1}(\mathbb{R} \setminus \{0\})$$

which is open in X since f is continuous. □

This is great! Instead of considering the topological space X we can consider something entirely algebraic: the algebra of continuous functions on it $C^0(X)$. To go back we then simply take $\mathrm{MaxSpec}(C^0(X))$ with the topology above. In other words: the algebra of observables contains all the information about the topology of the phase space. However, this doesn't entirely answer our questions: the phase space is more than a topological space (smooth structure, symplectic structure...) and for now we don't see that additional structure at the level of the algebra.

1.1 An extended example: $\mathbb{R}^n$ as a space in different ways

Consider now the space $\mathbb{R}^n$. This will allow us to start thinking of the smooth structure since $\mathbb{R}^n$ is the simplest possible manifold.

There are lots of different types of maps on $\mathbb{R}^n$ that one can consider. Firstly we have concrete maps $\mathbb{R}^n \to \mathbb{R}$ of different types:

- arbitrary maps between the *sets* $\mathbb{R}^n$ and $\mathbb{R}$;
- continuous maps, $C^0(\mathbb{R})$;
- differentiable maps with continuous derivatives, $C^1(\mathbb{R})$;
- ...
- smooth maps, $C^\infty(\mathbb{R})$;
- analytic maps, functions that are represented by their Taylor series (like $\cos, \sin, \dots$);
- polynomial functions, *finite* Taylor series;
- constant functions.

From the previous discussion, the point is that we'd like to have an algebra associated to $\mathbb{R}^n$ and while the algebras above are nice in the sense that they're collections of maps out of $\mathbb{R}^n$, we can even take abstract algebras such as formal power series, which we just treat as formal expressions, not worrying about convergence. We can also take power series modulo n-th order, which is what one typically does in perturbation theory in QFT, when calculating quantities up to order n in some "small parameter".
The relation between all these maps is shown in Figure 1.

Now, depending on what algebra we take on $\mathbb{R}^n$ we get a different induced topology on $\mathbb{R}^n$. Let's start by considering the extremes: if we consider maps of sets $\mathbb{R}^n \to \mathbb{R}$ then we can just take a function which is nonzero at a single point, so that single point (and any other point) is an open set in $\mathbb{R}^n$. Since any single

point forms an open set, any arbitrary subset of $\mathbb{R}^n$ is also open. This is called the *discrete topology* and it doesn't give us much information because every possible subset is open.
Constant functions instead are the opposite. An arbitrary function is either zero everywhere or zero nowhere, meaning that the only open sets in $\mathbb{R}^n$ are the empty set and $\mathbb{R}^n$ itself. This is called the *indiscrete topology*.
These are the two extremes, because one says all subsets are open, while the other says no subsets are open (except for the two that must be open). All of the others will be somewhere in the middle.

Continuous functions are what we studied before: the induced topology is simply the usual one on $\mathbb{R}^n$ (*standard topology*). What about the more regular functions? They actually induce the same topology on $\mathbb{R}^n$! To see this first note that, by the argument above, the topology may only be coarser. So if we were able to find a basis for the smooth topology which is also a basis for the normal topology then we'd be done.
To prove this consider the case $n = 1$. In this case it's enough to find a smooth function $f : \mathbb{R} \to \mathbb{R}$ for every $(a, b) \subset \mathbb{R}$ such that

$$\{x \in \mathbb{R} \mid f(x) \neq 0\} = (a, b)$$

This is possible thanks to bump functions! The starting point is that the function

$$f(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

is actually smooth. This is true because

$$\lim_{x \to 0^+} \frac{d^n}{dx^n} e^{-1/x} = 0, \quad \forall n$$

since taking derivatives simply gives higher order polynomials in the denominator which diverge slower than the exponential.
Now the bump function $f(x - a)f(b - x)$ does the trick! For $n > 1$ it's enough to note that a basis is just given by products of open intervals, so the proof goes through the same way. So we've proven that the smooth topology is the same as the normal one! This actually shouldn't be too hard to believe since one usually defines manifolds (which have a *smooth* structure) to be *topological spaces* with additional properties.

Now, going to analytic or polynomial functions this reasoning does *not* work, since there are no bump functions. So the "analytic" topology and the "polynomial" (Zariski) topology are actually different from the normal one. The induced topologies from the formal power series and power series modulo n-th order are even weirder. The induced topologies on $\mathbb{R}^n$ are also indicated in Figure 1.

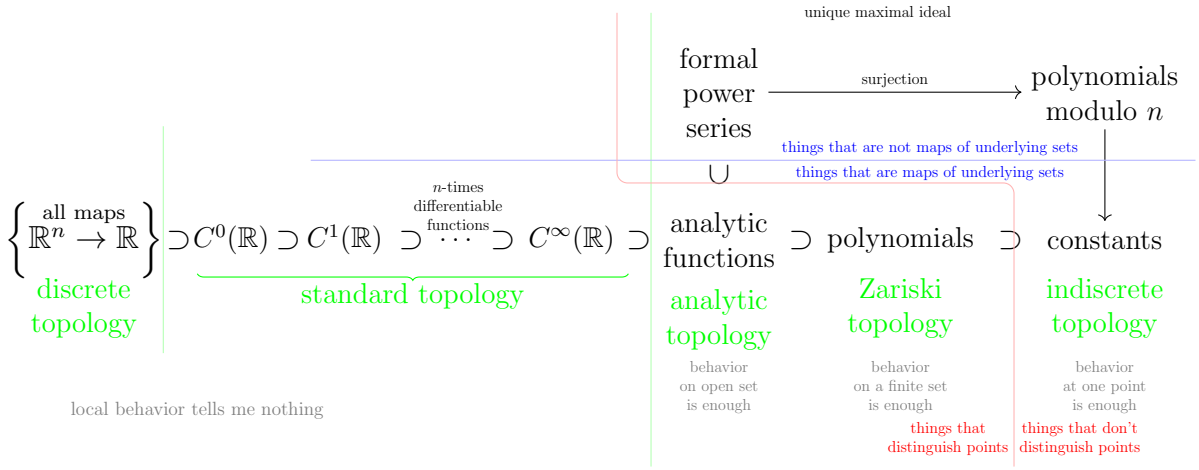


Figure 1: Many types of maps on $\mathbb{R}^n$, the relations among them and their induced topologies on $\mathbb{R}^n$.

Passing from smooth functions to analytic ones there's also something else interesting that happens. For smooth functions one can say that "information in some region is purely local", meaning that if I know a function in some open set, I know nothing of what that function is outside of my open set, since there are infinitely many ways to extend it. Instead for analytic functions one may say that "local information gives global constraints", since if I know the function on some open set, its global behavior is also determined.

End of lecture 3

The point of the discussion above is that the manner in which X is a space is characterized by which algebra of functions on X we choose to consider. From the example we saw how that was in fact a choice and different choices leas to different spaces. For differential geometry, for example we're interested in taking smooth functions.

In Figure 1 one can group some of the algebras together according to the properties they share. For example, all the algebras below the blue line differ from those above it by the fact that the ones below are actual maps of sets, while the ones above are just abstract algebras. The algebras between the green lines (i.e. from continuous functions up to smooth functions) also belong together since they all induce the same topology on $\mathbb{R}^n$. In addition, algebras to the right of the red line (formal power series, power series modulo n-th order and constant functions) also go together since they share the property of "not distinguishing points of $\mathbb{R}^n$". These all have the property that MaxSpec only has one point, a unique maximal ideal! For constant functions, this is because the algebra of constant functions is

isomorphic to the algebra $\mathbb{R}$, which, being a field, has the unique maximal ideal given by $\{0\}$. For formal power series, we can give a short proof.

Claim 1.3. *The ring $\mathbb{C}[|x|]$ of formal power series in a single variable has a unique maximal ideal.*

We start by proving a useful lemma:

Lemma 1.4. *$f = \sum_{k\geq 0} f_k t^k$, $f_k \in \mathbb{C}$ is invertible if and only if $f_0 \neq 0$.*

Proof. This is true because we can construct the inverse as follows. We want to find $g = \sum_{k\geq 0} g_k t^k$ such that $fg = 1$. But this just gives an infinite number of equations:

$$\begin{aligned} f_0 g_0 &= 1 \\ f_0 g_1 + f_1 g_0 &= 0 \\ f_0 g_2 + f_1 g_1 + f_2 g_0 &= 0 \\ &\dots \end{aligned}$$

To solve the first one we have to take $g_0 = f_0^{-1}$ which is only possible if $f_0 \neq 0$. Then we have

$$\begin{aligned} g_1 &= -(f_1 g_0) f_0^{-1} = 0 \\ g_2 &= -(f_1 g_1 + f_2 g_0) f_0^{-1} = 0 \\ &\dots \end{aligned}$$

so we can solve all the equations provided that $f_0 \neq 0$. □

A typical example is that in formal power series we have

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \dots$$

Normally this is called the geometric series and there's a problem of convergence. But here with formal power series we never think about plugging in a number and that equation holds.

Proof of the claim. By the lemma, no proper ideal I can contain an element with nonzero constant term. So there exists $k > 0$ such that all elements in I take the form $t^k f$ for some power series $f = \sum_{i\geq 0} f_i t^i$. In particular we may take k to be the smallest integer with that property. So the ideal actually takes the form $I = \langle t^k \rangle$. This is maximal if and only if $k = 1$ since $\langle t^k \rangle \subseteq \langle t^\ell \rangle$ whenever $\ell \leq k$. So the unique maximal ideal is $\langle t \rangle$. □

This is an example of a *local* ring, i.e. a ring with a unique maximal ideal.

2 Sheaves and Čech cohomology

2.1 Ideas and examples

We now restrict in the above diagram to examples that are actual maps of the underlying sets but for which local information constrains global behavior, i.e. analytic, polynomial and constant functions. Firstly, note that there are importnat differences between these 3:

- for analytic functions: knowing the behavior on one open set is enough to know the global behavior, since from the neighborhood I can find the Taylor series and the function is globally represented by its Taylor series;
- for polynomial functions: the behavior on a finite set of points is enough;
- for constant functions: the behavior at a single point is enough.

This leads to issues in gluing together local data, as we'll see below. Recall that these are radically different from smooth functions (and C^k functions) since their local behavior tells me nothing about their global behavior.

Consider a collection of observers who make local measurements in some regions of spacetime. We'd like to be able to glue those together in a consistent way in order to have a measurement over some larger region of spacetime that contains them.

Before we actually get to that, we recall a notion which will be useful.

Definition 2.1. An *exact sequence* of vector spaces is a sequence of maps between vector spaces which can be represented as follows

$$0 \longrightarrow A \xrightarrow{t} B \xrightarrow{p} C \longrightarrow 0 \tag{1}$$

such that t is injective, p is surjective and $\ker p = \operatorname{im} t$.

In general the notion of exactness is that the kernel of the outgoing map is the same as the image of the incoming map; applying this to A and C above we see that t must be injective and p surjective. This may seem abstract and not so useful but it represents something that we often think about: we can take A to be a vector subspace of B and C to be the quotient $C = B/A$, with maps t and p respectively being inclusion and quotient map.

We recalled this notion because we actually have something similar when talking about topological spaces. Suppose there are two observers who make local measurements in some overlapping regions U_1 and U_2 of the spacetime X. We now

want to combine their local measurements to a measurement on the larger region $U_1 \cup U_2$. We can represents these sets in the following diagram, where all maps are the obvious inclusions of open subsets of X.

$$\begin{array}{ccccc} & & U_1 & & \\ & \overset{i_1}{\nearrow} & & \overset{j_1}{\searrow} & \\ U_1 \cap U_2 & & & & U_1 \cup U_2 \,. \\ & \underset{i_2}{\searrow} & & \underset{j_2}{\nearrow} & \\ & & U_2 & & \end{array} \tag{2}$$

This diagram is kind of like an exact sequence since the maps to the right are surjective, the maps on the left are injective and we can also think of the object on the right as a quotient:

$$U_1 \cup U_2 = \frac{U_1 \sqcup U_2}{i_1(x) \sim i_2(x)}\,, \quad \forall x \in U_1 \cap U_2\,.$$

So the diagram above is "morally akin" to an exact sequence.

Now, going back to measurements, for any region U we have the space of measurements that can I can make on that region:

$$\mathscr{E}(U) = \text{functions (of whatever type I'm measuring) on } U$$

and for any inclusion $V \overset{i}{\subset} U$ we get a restriction map $i^* : \mathscr{E}(U) \to \mathscr{E}(V)$, because I can always think of a measurement in some large region U as also being a measurement on some smaller region V in U. As usual $\mathscr{E}(U)$ is actually an algebra and in particular also a vector space.

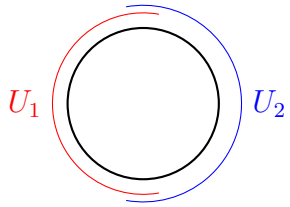
Now that we kind of have two notions of exact sequences, one for topological spaces as in (2) and one for vector spaces as in (1), it makes sense to ask the following question: if I take the space of measurements of the sets in (2), do they fit in an exact sequence of vector spaces? i.e. is the diagram below an exact sequence?

$$0 \longrightarrow \mathscr{E}(U_1 \cup U_2) \begin{array}{ccc} & \overset{j_1^*}{\nearrow} \mathscr{E}(U_1) & \\ & \bigoplus & \\ & \underset{j_2^*}{\searrow} \mathscr{E}(U_2) & \end{array} \begin{array}{ccc} & \overset{i_1^*}{\searrow} & \\ & & \\ & \underset{-i_2^*}{\nearrow} & \end{array} \mathscr{E}(U_1 \cap U_2) \longrightarrow 0\,.$$

Firstly, let's interpret the objects a bit. On the left we have $\mathscr{E}(U_1 \cup U_2)$, measurements on the larger region, which we can think of as "global data". In the middle we have a pair consisting of a measurement in U_1 and one in U_2 (that's what the

direct sum means) and we think of this as the space of "local measurements". Instead on the right side we have functions on the intersection which we think of the "mismatch of the local measurements on the overlap". That's because of the maps indicated on the right: for measurements on U_1 I simply restrict to $U_1 \cap U_2$ with the map i_1^*, while for measurements on U_2 I restrict and take the opposite with the map $-i_2^*$. This way we obtain in $\mathscr{E}(U_1 \cap U_2)$ the difference between the two local measurements.
Let us now consider exactness. The first map should be injective, i.e. a function should determined by its restriction to the covering regions, which makes sense. Exactness in the middle says that two local functions that agree on the overlap actually glue together to a function on the larger region, which also makes sense. Together these are saying: a global function is the same thing as a pair of local measurements which agree on their overlap. What about the final map being surjective? We should have that every function on the overlap is the mismatch of some pair of functions f_1 and f_2 on the two regions, which is not a very intuitive condition. In fact this is wrong! We can see what goes wrong in an example.

Example 2. Consider $X = S^1$ and take the cover by two open sets U_1 and U_2 as follows and take constant functions as measurements.



We want to see if the sequence above is exact on the right, i.e. if the map into $\mathscr{E}(U_1 \cap U_2)$ is surjective. But since we're considering constant functions, all of these vector spaces are very simple: $\mathscr{E}(U_1)$, $\mathscr{E}(U_2)$ and $\mathscr{E}(U_1 \cup U_2)$ are all simply $\mathbb{R}$. $\mathscr{E}(U_1 \cap U_2)$ instead consists of two numbers, the value on U_1 and the one on U_2, so $\mathscr{E}(U_1 \cap U_2) = \mathbb{R} \oplus \mathbb{R}$. So the diagram above becomes

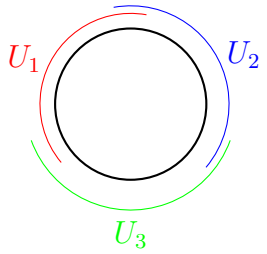
$$0 \longrightarrow \mathbb{R} \xrightarrow[j_2^*]{j_1^*} \mathbb{R} \oplus \mathbb{R} \xrightarrow[-i_2^*]{i_1^*} \mathbb{R} \oplus \mathbb{R} \longrightarrow 0\,.$$

Now, take an element in $\mathscr{E}(U_1) \oplus \mathscr{E}(U_2) = \mathbb{R} \oplus \mathbb{R}$, i.e. a pair of two real numbers (a, b). This element is mapped to $(a - b, b - a)$ in $\mathscr{E}(U_1 \cap U_2) = \mathbb{R} \oplus \mathbb{R}$. The map then is clearly not surjective since not all elements in $\mathbb{R} \oplus \mathbb{R}$ are of the form $(c, -c)$.

Note that the kernel of the map consists of pairs $(a, a) \in \mathscr{E}(U_1) \oplus \mathscr{E}(U_2)$, which make up a global measurement $a \in \mathscr{E}(U_1 \cup U_2)$, just like we noted above.

In a sense, this happened because my system of observers wasn't local enough: the overlap of the two regions consisted of two disjoint arcs which doesn't seem to be a region over which we could have an observer. We can then add the following requirement: all regions U_i should be nice (contractible) and all overlaps $U_{ij} = U_i \cap U_j$ should be as well, all triple overlaps as well, and so on... This is sometimes called a *good cover*.

Example 3. We can now do the circle again, but with three observers: take the cover by three open sets U_1, U_2 and U_3 as shown below.



These are in fact a good cover since their overlaps U_{12}, U_{23} and U_{13} are all arcs, which are in fact contractible. Instead for triple overlaps there's nothing to check since there is no triple overlap. The diagram above now becomes

$$\begin{array}{ccc} \mathscr{E}(U_1) & \longrightarrow & \mathscr{E}(U_{12}) \\ \mathscr{E}(U_2) & \longrightarrow & \mathscr{E}(U_{23}) \\ \mathscr{E}(U_3) & \longrightarrow & \mathscr{E}(U_{13}) \end{array}$$

and we can think of this collection of maps as a map $\mathscr{E}(U_1) \oplus \mathscr{E}(U_2) \oplus \mathscr{E}(U_3) \to \mathscr{E}(U_{12}) \oplus E(U_{23}) \oplus E(U_{13})$, i.e. a map $\mathbb{R}^3 \to \mathbb{R}^3$ which we can represent as the matrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

This construction is called Čech complex which we may write as

$$\mathscr{E}(U_1) \oplus \mathscr{E}(U_2) \oplus \mathscr{E}(U_3) \xrightarrow{A} \mathscr{E}(U_{12}) \oplus E(U_{23}) \oplus E(U_{13}) \longrightarrow 0 \longrightarrow \dots$$

where we have nothing from the degree 2 onwards since there are no triple intersections. As usual when given have a complex, one can take its cohomology:

$$H^k = \frac{\ker d_k}{\operatorname{im} d_{k-1}},$$

which for the complex above means

$$H^0 = \ker A\,,$$
$$H^1 = \mathscr{E}(U_{12}) \oplus E(U_{23}) \oplus E(U_{13})/\operatorname{im} A\,.$$

Now, $\ker A = (a, a, a)$ for a any real number. This has the same meaning as in the previous case: local measurements which agree on their intersections give a global measurement (a on $\mathscr{E}(X) = \mathscr{E}(U_1 \cup U_2 \cup U_3)$). However now there's something more: we can also calculate cohomology in degree 1:

$$H^1 = \frac{\mathbb{R}^3}{(a, b, c) \text{ with } a + b + c = 0}$$

These two vector spaces give the *Čech cohomology* of (the sheaf of locally constant functions on) the circle:

$$H^0 = \mathbb{R}\,,$$
$$H^1 = \mathbb{R}\,.$$

Note that there are many other ways of finding the cohomology of a circle but this may be the way which is more intuitive from a physics perspective.

End of lecture 4

2.2 Definitions

We now want to give the definition of a sheaf which is what we've secretly been talking about. A sheaf assigns some kind of locally defined data to any open set of my space and the reason this is useful is because this is what we have in field theory, which will be our motivating example.

To do this, first pick a suitable "category" $\mathscr{C}$ of objects of the type of $\mathscr{E}(U)$, like rings, groups, vector space, spaces... Up to now we've mostly been considering vector spaces, but the space of fields isn't necessarily a vector space. For example if we have a σ-model, then the fields are maps from my space into some fixed manifold so we don't know what it should mean to "add" two fields. String theory is an example of a field theory of this type.

Definition 2.2. A *presheaf* $\mathscr{E}$ on a topological space X, valued in the category of $\mathscr{C}$, is an assignment:

- to each U in $\mathrm{Opens}(X)$, an object $\mathscr{E}(U)$ of $\mathscr{C}$;
- to each inclusion $i : V \subseteq U$, a "restriction" map $i^* : \mathscr{E}(U) \to \mathscr{E}(V)$;
- consistent with composition: $W \overset{i}{\subseteq} V \overset{j}{\subseteq} U$, $(ij)^* = j^* i^*$.

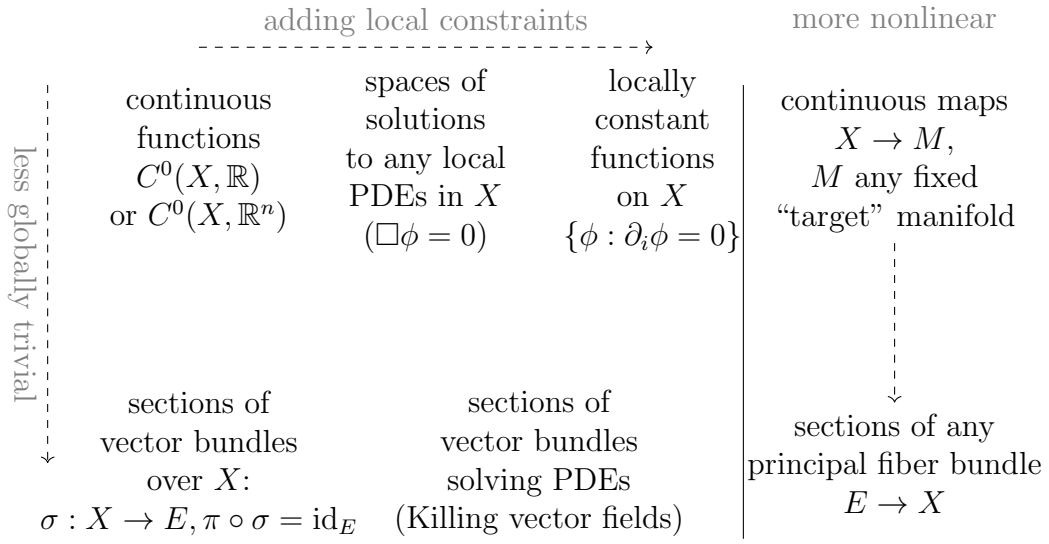
Definition 2.3. Let ε be a presheaf on X. $\mathscr{E}$ is a *sheaf* if the sequence

$$0 \longrightarrow \mathscr{E}(U) \xrightarrow[i_2^*]{i_1^*} \mathscr{E}(U_1) \oplus \mathscr{E}(U_2) \xrightarrow[j_2^*]{i_1^*} \mathscr{E}(U_{12}) \tag{3}$$

is "exact" in $\mathscr{C}$.

We call the elements of $\mathscr{E}(U)$ "sections" of $\mathscr{E}$ over U (if it makes sense in the category $\mathscr{C}$). Exactness in the first spot means that restricting from U to the cover $U_1 \sqcup U_2$ is injective. Exactness in the middle means that any pair of sections (σ_1, σ_2) over $U_1 \sqcup U_2$ that agree on $U_1 \cap U_2$ glue to a section on U. These conditions together mean that a section on U is the same as a collection of sections on a cover of U that agree on every overlap. This is a locality condition (often called "gluing").

Example 4. There are lots of examples! The simplest is what we've extensively considered: continuous functions into $\mathbb{R}$. We can then get more examples in different ways: we can consider adding local constraints (like satisfying some PDE) or we can have something "less globally trivial" by considering sections of vector bundles or we can even consider more nonlinear maps ("σ-models"), by mapping into something that's not a vector space. Some examples are in the table below, where we can imagine "more nonlinear" to be an additional dimension in which we can extend.



Recall that a vector bundle on X is a space E with a projection map $\pi : E \to X$ such that the fibers of this map $\pi^{-1}(p), \forall p$ are vector spaces $\mathbb{R}^n$ and locally we have $\pi^{-1}(U) \cong U \times \mathbb{R}^n$. A section of a vector bundle is then a map $\sigma : X \to E$ satisfying $\pi \circ \sigma = \mathrm{id}_X$. This means that at every point p we pick a vector in E and in particular the second condition says that the vector is in the fiber over p. (Almost) all fields that one considers in field theory are of one of the types in the table above.

A (potentially) less intuitive example.

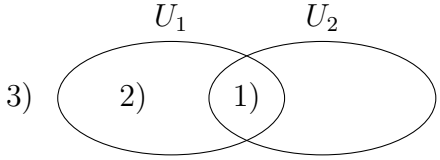
Example 5 (Skyscraper sheaf)**.** Pick one point $x \in X$ and one object $A \in \mathscr{C}$. For $U \in \mathrm{Opens}(X)$, define

$$\mathscr{E}(U) = \begin{cases} A, & x \in U; \\ 0, & x \notin U. \end{cases}$$

We also need to give a function for any inclusion of subsets. For $V \overset{i}{\subseteq} U$ we have three options:

1. $x \in V$, $\quad i^*$ is the identity $A \to A$.
2. $x \in U \setminus V$, $\quad i^*$ is the zero map $A \to 0$.
3. $x \notin U$, $\quad i^*$ is again the zero map $0 \to 0$.

Does this glue well? We have to see if the sequence (3) is exact and we again have three cases to consider which we can represent as:



1) If $x \in U_{12}$, then we have the exact sequence $0 \to A \to A \oplus A \to A$.

2) If $x \in U_2 \setminus U_1$, we have $0 \to A \xrightarrow{\text{id}} A \oplus 0 \to 0$, which is also exact.

3) If $x \notin U$, then we have the trivial exact sequence $0 \to 0 \to 0 \oplus 0 \to 0$.

We can view this sheaf as the space of sections of a "not quite vector bundle", which over the special point x has the fiber A, but over all other points has fiber given by the trivial vector space.

One can also make the same construction as above but over a line (or some higher dimensional subspace) in X, so in a sense sheaves allow for "defects" of this kind. Every good ("coherent") sheaf looks more or less like this. Vector bundles instead do *not* allow for defects because we precisely define them to look the same everywhere.

3 Sheaf Cohomology

3.1 Ideas and boring sheaves

Lesson from last time: A sheaf $\mathscr{E}$ on X contains more global information than the space of "global sections" $\mathscr{E}(X)$! We took the example of the locally constant sheaf $\mathscr{E} = \underline{\mathbb{R}}$ over the circle $X = S^1$ and we recovered the entire cohomology of S^1! Which is more information than just $\mathscr{E}(X)$ which is simply $\mathbb{R}$, since $\mathscr{E}(U) = \mathbb{R}$ for any connected X. Note that we had the information of $\mathscr{E}(X)$ since it was the 0-th cohomology H^0.

How did we do this?

- Cover X by contractible open sets U_i, with $U_{ij} = U_i \cap U_j$ being contractible, U_{ijk} contractible and so on.
- Consider the sequence of maps of spaces

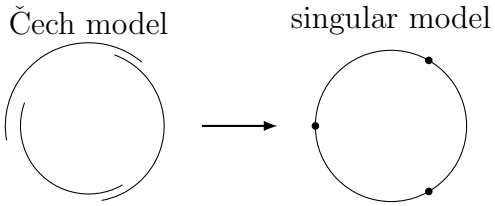
$$X \;\substack{\leftarrow\\ \leftarrow\\ \leftarrow}\; \bigsqcup U_i \;\substack{\leftarrow\\ \leftarrow\\ \leftarrow}\; \bigsqcup U_{ij} \;\substack{\leftarrow\\ \leftarrow\\ \leftarrow}\; \bigsqcup U_{ijk} \;\substack{\leftarrow\\ \leftarrow\\ \leftarrow}\; \dots$$

The sequence has two key properties:

1. it's "exact" (in whatever sense),
2. all terms except X are "boring".

In a sense what we're doing is that instead of evaluating $\mathscr{E}$ on X (which would correspond to an observer over the whole universe, which doesn't seem very physical), we're evaluating it on the whole observer system. We're then "gluing" together (taking cohomology) and in doing so we get $\mathscr{E}(X)$ but also much more (higher cohomology). This defines the cohomology of the sheaf $\mathscr{E}$ (this big diagram is the "Čech complex").

For $\mathscr{E} = \underline{\mathbb{R}}$ or $\underline{\mathbb{Z}}$ we recover the ordinary (real or integral) cohomology of the space X. Typically, one first sees real cohomology using differential forms, while integral cohomology was first obtained by studying triangulations of the space (singular cohomology). Conceptually what we're doing is not so different, but instead of triangulating our space we're covering it with open sets:



Q: We discovered cohomology by replacing the "interesting" object X by a collection of boring objects $(U_i, U_{ij}, \dots)$ glued together. Can we do something similar that replaces $\underline{\mathbb{R}}$ instead? I.e. if I want to consider a complicated sheaf I may want to replace it with a (collection of) simpler sheaves.

A: If you could do this, you reinvent the theory of differential forms.

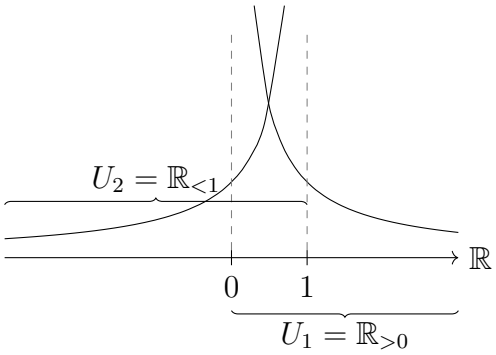
End of lecture 5

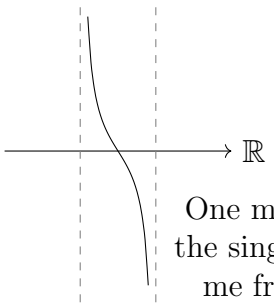
Last time we ended with a big idea: we should replace complicated objects by models for gluing them together out of boring objects. We saw a concrete example of this when considering the global sections: we covered our manifold M by open sets U_i which formed a good open cover; we could then think of M as the gluing of all of the U_i along their intersections. However their intersections may also have intersections which we should glue together and this continues with higher and higher intersections. We can then consider this gluing instead of M itself. Applying a sheaf $\mathscr{E}$ to this replacement then discovers not only the global sections, but also more information, given by the sheaf cohomology.

$$M \;\substack{\longleftarrow\\ \longleftarrow}\; \bigsqcup U_i \;\substack{\longleftarrow\\ \longleftarrow\\ \longleftarrow}\; \bigsqcup U_{ij} \;\substack{\longleftarrow\\ \longleftarrow\\ \longleftarrow\\ \longleftarrow}\; \bigsqcup U_{ijk} \;\substack{\longleftarrow\\ \longleftarrow\\ \longleftarrow\\ \longleftarrow\\ \longleftarrow}\; \dots$$

Q: This procedure took two inputs, $\mathscr{E}$ and M, and we replaced M with boring spaces. Could we instead replace $\mathscr{E}$ by boring sheaves? In order to try this, what is boring sheaf? The topological answer is: "one that has no higher cohomology in *any* space" – there is no extra information other than global sections (such a sheaf is called *acyclic*). But what guarantees that a sheaf is boring? One possibility would be the existence of partitions of unity. The following is an intuitive example of why this should be the case.

Example 6. The function $1/x$ on the positive real axis is smooth. It doesn't extend to a smooth function on $\mathbb{R}$. Suppose I want to "patch together" $1/x$ on $U_1 = \mathbb{R}_{>0}$ and $1/(1-x)$ on $U_2 = \mathbb{R}_{<1}$.





One might worry that the singularity prevents me from extending.

$$\begin{array}{ccccc} & & \mathcal{C}^\infty(U_1) & & \\ & \nearrow & & \searrow & \\ \mathcal{C}^\infty(\mathbb{R}) & & \bigoplus & & \mathcal{C}^\infty(U_1 \cap U_2) \\ & \searrow & & \nearrow & \\ & & \mathcal{C}^\infty(U_2) & & \end{array}$$

In particular, smooth functions are fine, hence acyclic.

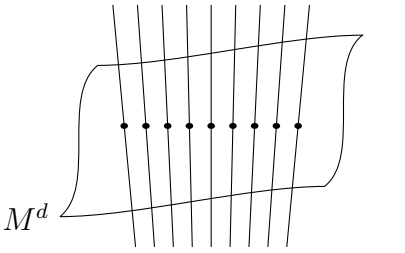
3.2 Vector bundles as locally free sheaves

Q: Can we "glue together" (e.g.) locally constant functions out of smooth functions or tuples of smooth functions?
Idea: A locally constant function is a smooth function with an additional constraint: all derivatives must vanish.
Q: What are "all possible derivatives"? Relatedly: how should I think of "tuples of smooth functions"? (Sections of) vector bundles!
The path we'll take to answer these questions is the following: we start with vector bundles, then view them as locally free sheaves and then considering vector fields (=derivations) will lead us to differential forms.

We start by recalling the notion of vector bundle. A vector bundle on M was a "family of n dimensional vector spaces, continuously parametrized by M", i.e. for every point $p \in M$ we have an n dimensional vector space E_p and these are "continuously parametrized" in the sense that we have a "total space" E, which contains both M and all these vector spaces, and has a continuous map $\pi : E \to M$ such that $\pi^{-1}(p) = E_p$. The simplest example is the "trivial bundle" $M \times \mathbb{R}^n$ where the map π simply forgets the second factor. For vector bundles we require that it is *locally trivial*, i.e. it locally looks like that, in the sense that for any point $p \in M$ there is an open neighborhood $p \ni U \subset M$ and an isomorphism $\pi^{-1}(U) \xrightarrow{\cong} U \times \mathbb{R}^n$ which fits into the diagram below

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\cong} & U \times \mathbb{R}^n \\ \downarrow & & \downarrow{\scriptstyle p_i} \\ U & \xrightarrow{\mathrm{id}} & U \end{array}$$

and where $\pi^{-1}(x) = E_x \xrightarrow{\sim} \mathbb{R}^n$ is a map of vector spaces. In other words the map $\pi^{-1}(U) \to U \times \mathbb{R}^n$ is "fiberwise linear". This is usually pictorially represented as follows.



A section of M is then a continuous map $\pi : E \to M$ such that $\pi \circ \sigma = \mathrm{id}_M$. We think of this as as a "continuous family of vectors" (a vector field). Sections of vector bundles clearly form a sheaf because I can restrict sections and they still maintain the property $\pi \circ \sigma = \mathrm{id}$:

$$\begin{array}{ccccc} \pi^{-1}(v) & \subseteq & \pi^{-1}(u) & \subseteq & E \\ {\scriptstyle\pi}\downarrow\uparrow{\scriptstyle\sigma|_V} & & {\scriptstyle\pi}\downarrow\uparrow{\scriptstyle\sigma|_U} & & {\scriptstyle\pi}\downarrow\uparrow{\scriptstyle\sigma} \\ V & \subseteq & U & \subseteq & M \end{array}$$

Note that when talking about sheaves we were a bit vague about what kinds of objects we get over every open set. In particular, when taking functions we actually got a sheaf of rings since I can add and multiply functions together. For sections of vector bundles, however, I don't have a way of multiplying them. The operations I have are the following:

- I can add two sections σ and σ' by adding vectors in the fiber: $(\sigma+\sigma')(p) = \sigma(p)+\sigma'(p)$. This addition is commutative and has an inverse because the fibers are vector spaces. This gives us a sheaf of abelian groups.
- I can also multiply a section σ by a function f by scalar multiplication in the fiber: $(f\sigma)(p) = f(p)\sigma(p)$ where on the right we have scalar multiplication in E_p. By restricting f to simply be a number we get a sheaf of vector spaces. By noting that we can take any function f we get a sheaf of *modules* over the ring of functions.

Let us then recall the definition of a module.

Definition 3.1. Given a commutative ring A, a module is an *abelian group* M with an operation $A \times M \to M$ which we call "scalar multiplication" which satisfies the following properties:

$$\begin{aligned} a(m+m') &= a\cdot m + a\cdot m' \\ (a+a')\cdot m &= a\cdot m + a'\cdot m \\ (aa')\cdot m &= a\cdot(a'\cdot m) \\ 1\cdot m &= m\,. \end{aligned}$$

Modules are useful because they're a generalization of vector spaces: if we take A above to actually be a field we recover the definition of a vector space. However, modules have a more interesting structure. In particular, there are not that many finite dimensional vector spaces: any vector space V of dimension n is isomorphic to $\mathbb{R}^n$ and the isomorphism corresponds to picking a basis of V. Of course there are many possible bases and correspondingly there are many possible isomorphisms with $\mathbb{R}^n$. We may then wonder if modules work the same way.
Q: Is it true that (finite dimensional) modules over A are isomorphic to $A^{\oplus n}$? The action of scalar multiplication and addition on such a module works as one might expect

$$b\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ba_1 \\ \vdots \\ ba_n \end{bmatrix}, \qquad \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} a_1' \\ \vdots \\ a_n' \end{bmatrix} = \begin{bmatrix} a_1+a_1' \\ \vdots \\ a_n+a_n' \end{bmatrix}.$$

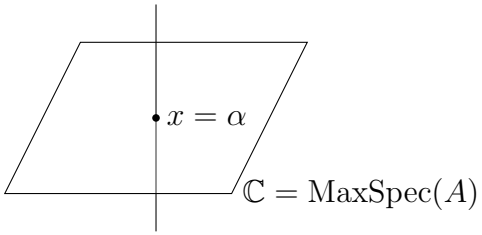
A: NO!

Example 7. As a counterexample consider $A = \mathbb{C}[x]$ and $M = \mathbb{C}$. Make M into an A-module by defining $x\cdot m = \alpha m$ for some $\alpha \in \mathbb{C}$. This is enough because the properties above imply that $x^n \cdot m = \alpha^n m$ for any n, so I know how to multiply with anything in A. We'd like to view this as we were viewing vector bundles

above: the sheaf of sections was a sheaf of modules over the ring of functions on my manifold; now we take M as our module, so we want to view A as the ring of functions over some space. We saw how to do this: A is the same thing as functions on MaxSpec A! In addition MaxSpec$(A) = \mathbb{C}$, to see this recall that since we're working over $\mathbb{C}$ we have that any polynomial $f \in \mathbb{C}[x]$ may be factored as

$$f = \prod_i (x - \alpha_i)$$

for some collection of $\alpha_i \in \mathbb{C}$. Maximal ideals are generated by the irreducible polynomials which are simply $(x - \alpha_i)$ for any $\alpha_i \in \mathbb{C}$. So we indeed have $\mathbb{C} =$ MaxSpec(A), at least as sets (usual topology is different from the Zariski topology). Considering the module M again, we see that it's the module for which x takes the value α which we chose above, in other words we only have a fiber over the point $\alpha \in \mathbb{C}$. So we can view M as a skyscraper sheaf again!



Armed with the notion of modules, we can make precise what we meant by saying that the sheaf of sections of a vector bundle is a sheaf of modules over the ring of functions.

Definition 3.2. Let $\mathcal{O}$ be the sheaf of rings given by the functions on my space. A sheaf $\mathscr{E}$ of abelian groups is a sheaf of $\mathcal{O}$ modules if $\mathscr{E}(U)$ is a module over $\mathcal{O}(U)$ for every U.

The (counter)example above suggests that if we want to consider vector bundles we shouldn't consider arbitrary sheaves of $\mathcal{O}$-modules.

Definition 3.3. A sheaf of $\mathcal{O}$-modules is locally free if $\forall p \in M$, $\exists U \subset M$ containing p, such that $\mathscr{E}(U) \cong \mathcal{O}(U)^{\oplus n}$.

In other words, "there exist local vector fields that are a basis for the space of sections (over functions)". The following fact justifies why we should consider such sheaves.

Fact 3.1. *Finite dimensional vector bundles are equivalent to locally free sheaves of $\mathcal{O}$-modules of finite rank, i.e. there is a bijection*

$$\textit{f. d. vector bundles} \xleftrightarrow{1:1} \textit{locally free sheaves of } \mathcal{O}\textit{-modules of finite rank.}$$

3.3 Derivations and one forms

We can now pass to considering derivations. Recall that the Leibniz rule says that derivatives have the following property with respect to the product of functions:

$$\mathrm{D}(fg) = \mathrm{D}(f)g + f\mathrm{D}(g)\,. \tag{4}$$

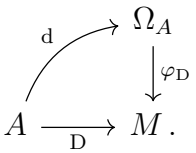
This is because it should capture the idea of a small variation of f, in the sense that if we consider the variation $f \to f + \epsilon \mathrm{D}f$ and we consider the variation of a product we have

$$(f + \varepsilon \mathrm{D}f)(g + \varepsilon \mathrm{D}g) = fg + \varepsilon(f\mathrm{D}g + (\mathrm{D}f)g) = fg + \varepsilon \mathrm{D}(fg) \tag{5}$$

because $\varepsilon^2 = 0$. So we can think of it as a condition of infinitesimal compatibility with the ring structure. In order for the formula to make sense, the source of the map D is some ring A and the target is some A module M since I have to know how to multiply $D(f)$ by g.

Definition 3.4 (Work over a field)**.** A derivation from A to M is a k linear map $\mathrm{D} : A \to M$ satisfying the Leibniz rule.

What's the universal way to get a derivation? We'd like for there to be an A module Ω_A such that any derivation D may be obtained as the composition $\varphi_{\mathrm{D}} \circ \mathrm{d}$ for some map of A modules φ_{D}.



In order to construct Ω_A we take the free A-module on symbols $\mathrm{d}a$, $\forall a \in A$ and divide out by the relations $\mathrm{d}(aa') = \mathrm{d}(a)a' + a\mathrm{d}a'$, $\forall a, a' \in A$. Module Ω_A is called "the module of *Kähler differentials*".

Example 8. Consider the case $A = \mathbb{R}[x_1, \ldots, x_n]$. We have $\mathrm{d}(1) = 0$ since

$$\mathrm{d}(f \cdot 1) = \mathrm{d}(f)1 + f\mathrm{d}(1) \quad \Longrightarrow \quad f\mathrm{d}(1) = 0 \quad \Longrightarrow \quad \mathrm{d}(1) = 0\,.$$

We can then take

$$\Omega_A = \bigoplus_{i=1}^{n} \mathbb{R}[x_1, \ldots, x_n] dx_i$$

with the map $\mathrm{d} : A \to \Omega_A$ acting as $x_i \mapsto \mathrm{d}x_i$. Thus, at least in the algebraic setting, it holds

$$\text{one-forms} \longleftrightarrow \text{Kähler differentials} \longleftrightarrow \begin{matrix}\text{universal target for 1st order} \\ \text{differential operators}\end{matrix}.$$

We now have another way of considering constant functions:

$$\ker(d : A \to \Omega_A) = \text{ constant functions}\,.$$

End of lecture 6

Recall from the last time:

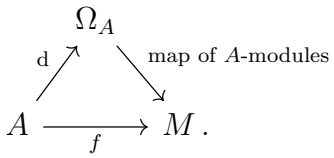
- Looking for a "resolution" of the constant sheaf $\underline{\mathbb{R}}$ by a chain complex of "fine sheaves" (such sheaves that admit a partition of unity, don't have any higher cohomology and are smooth functions or tuples of smooth functions).
- Geometrically they locally look like maps $U \to V$, where V is some (finite dimensional real) vector space.
- Vector bundle is equivalent to a locally free sheaf of $\mathcal{O}$-modules by taking its smooth sections.
- $\mathscr{E}$ is a sheaf of $\mathcal{O}$-modules if for $\mathcal{O}$ the sheaf of functions on my space and $\mathscr{E}$ a sheaf of vector spaces, $\mathscr{E}(U)$ is an $\mathcal{O}(U)$-module for each $U \subseteq X$ (where $\mathcal{O}(U)$ is an algebra).
- Not every module over A can be made isomorphic to $A^{\oplus n}$ by choosing a basis!

The point of the previous discussion is that a function is locally constant if all possible derivatives vanish. We defined the concept of derivation of an algebra A and we then found that there's a universal such pair: the module Ω_A of Kähler differentials

$$\Omega_A = \frac{\text{generated by symbols } \mathrm{d}a \text{ for } a \in A}{\text{all relations of the form } \mathrm{d}(ab) = \mathrm{d}a \cdot b + a \cdot \mathrm{d}b}\,. \tag{6}$$

The pair (Ω_A, d) is universal in the sense that any derivation map $A \to M$ factors uniquely through Ω_A via d, i.e. there is a unique map $\Omega_A \to M$ such that the

following diagram commutes



In particular this means that there's a bijection between the space of derivations $A \to M$ and the space of maps $\Omega_A \to M$:

$$\mathrm{Hom}_A(\Omega_A, M) \cong \mathrm{Der}(A, M)\,.$$

Example 9. Consider $A = \mathbb{C}[x_1, \ldots, x_n]$, we then may take

$$\Omega_A = \bigoplus_{i=1}^{n} \mathbb{C}[x_1, \ldots, x_n]\mathrm{d}x_i$$

which is actually free since it's isomorphic to $A^{\oplus n}$: it's just a collection of n polynomials labeled by $\mathrm{d}x_i$. So it's a free module of rank n.

However, for smooth functions, that is not exactly the case. To study this case, we first pose a question.
Q: What should a (smooth or polynomial) tangent vector field be?
A: It should be a derivation valued in (smooth or polynomial) functions! In the sense that a vector field gives me a directional derivative that takes in a function and gives another function:

$$\mathrm{Vect} = \mathrm{Der}(A, A)$$

where A is an algebra. As sheaves, the property above reads

$$\mathrm{Vect}(U) = \mathrm{Der}(\mathcal{O}(U), \mathcal{O}(U)) \quad \forall U \text{ open}$$

By the universal property

$$\mathrm{Vect} = \mathrm{Der}(A, A) = \mathrm{Hom}_A(\Omega_A, A) = \Omega_A^\vee,$$

where the $^\vee$ denotes the dual space.
Caution: this is the right definition of vector fields (whether smooth or polynomial). The right definition of one-forms is then the dual of vector fields:

$$\Omega^1 = \mathrm{Hom}_A(\mathrm{Vect}, A) = \mathrm{Vect}^\vee = (\Omega_A^\vee)^\vee.$$

For finite dimensional vector spaces we always have that a vector space and its double dual are canonically isomorphic, but this is not necessarily true for infinite

dimensional vector spaces. There is still a canonical map $\Omega_A \to \Omega^1 = (\Omega_A^\vee)^\vee$, but it's not injective in the smooth case (however, it is surjective). We can kind of understand this non injectivity by noting that the condition in (6) was in a sense a polynomial condition in the sense that it gives the properties we want for all polynomials by induction, but for example it doesn't imply

$$\mathrm{d}(e^x) = e^x \mathrm{d}x\,.$$

This suggests that the Kähler differentials are a little too coarse to consider smooth functions.

In one dimension we've already accomplished our goal of resolving the sheaf of locally constant functions $\underline{\mathbb{R}}$ by fine sheaves: consider the complex

$$\begin{array}{ccc} \underline{0} & & \underline{1} \\ \mathbb{R}[x] & \xrightarrow{x\mapsto \mathrm{d}x} & \mathbb{R}[x]dx \\ \| & & \| \\ A & \xrightarrow{\mathrm{d}} & \Omega_A\,, \end{array}$$

there is no cohomology in degree 1 (working locally on $\mathbb{R}$). Why is that? If $\alpha \in \mathbb{R}[x]\mathrm{d}x$ is a one-form, then

$$\alpha = \mathrm{d}f, \quad f : p \mapsto \int_0^p \alpha\,,$$

so it lies in the image of the differential and there is no cohomology. Note that this doesn't rely on specific properties of polynomials, the same proof could be done for smooth functions. So (in one dimension) we found a resolution of the sheaf of locally constant functions! In particular we're using sheaves that on boring spaces (here $\mathbb{R}$) have no cohomology.
These are "*differential forms*" in one dimension.

Example 10. Placing this on the circle recovers $H^\bullet(S^1) = \mathbb{R}$, as we already found using the Čech cohomology. On the circle we have

$$\begin{array}{ccc} \underline{0} & & \underline{1} \\ \{e^{in\theta}\}_{n\in\mathbb{Z}} & & \{e^{in\theta}\mathrm{d}\theta\}_{n\in\mathbb{Z}} \\ e^{in\theta} & \xmapsto{\mathrm{d}} & ine^{in\theta}\mathrm{d}\theta \end{array}$$

since we can think of functions on the circle as being periodic functions on the real line $\mathbb{R}$. The cohomology is given by the constant function 1 for $n = 0$ and by

the one form $1\mathrm{d}\theta$ for $n = 1$, since it's the only one form that we don't get as the differential of another function. So we have

$$H^\bullet = \begin{cases} \mathbb{R} \text{ or } \mathbb{C}, & n = 0, \\ \mathbb{R} \text{ or } \mathbb{C}, & n = 1, \end{cases}$$

depending on the choice of coefficients.

Note that this is the same result as we had previously found using Čech cohomology! Why did this work? Previously we had to take a good cover of our space, while now we just took directly global sections. This works because it precisely realizes the idea we had of taking boring sheaves (fine sheaves) to represent the sheaf of locally constant functions, instead of taking boring open sets to represent my space.

3.4 Differential forms

What about more than one dimension? Consider 2 dimensions. Let $A = C^\infty(\mathbb{R}^2)$ with coordinates x and y, so we're again considering a boring space so we don't want to have cohomology. The differential may then be written as $\mathrm{d} = \mathrm{d}x \cdot \partial x + \mathrm{d}y \cdot \partial y$ and we have

$$\begin{array}{ccc} \underline{0} & & \underline{1} \\ & \overset{\mathrm{d}x\cdot\partial x}{\nearrow} & A\mathrm{d}x \\ A & & \oplus \\ & \underset{dy\cdot\partial y}{\searrow} & A\mathrm{d}y \end{array}$$

but the cohomology here doesn't vanish yet. On polynomials, we have

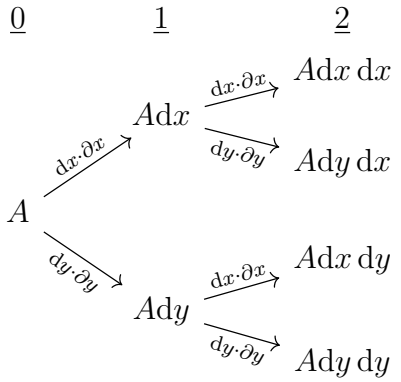
quadratic (0-forms)	linear (1-forms)
x^2	$x\mathrm{d}x$
xy	$x\mathrm{d}y + y\mathrm{d}x$
y^2	$y\mathrm{d}y$

But we see that $x\mathrm{d}y - y\mathrm{d}x$ is not in $\mathrm{im}(\mathrm{d})$! So in order to have zero cohomology we have to add something in degree 2! In particular I should have the map out of that which kills everything except for elements of the form $x\mathrm{d}y - y\mathrm{d}x$:

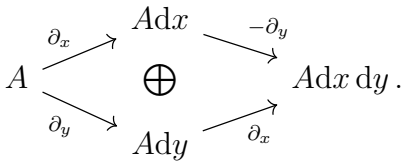
$$\begin{aligned} x\mathrm{d}x &\mapsto 0 \\ y\mathrm{d}y &\mapsto 0 \\ x\mathrm{d}y + y\mathrm{d}x &\mapsto 0\,. \end{aligned}$$

We therefore have an idea: Ω^1 represents the first order, while we now want to represent the second order, so we want to take things of the form $\mathrm{d}x\,\mathrm{d}x$ and so on,

with the map into it again being d. However, because of the properties above, we actually want that $\mathrm{d}x\,\mathrm{d}x = 0$, $\mathrm{d}y\,\mathrm{d}y = 0$ and $\mathrm{d}x\,\mathrm{d}y + \mathrm{d}y\,\mathrm{d}x = 0$. So in general we want the antisymmetric part! The complex



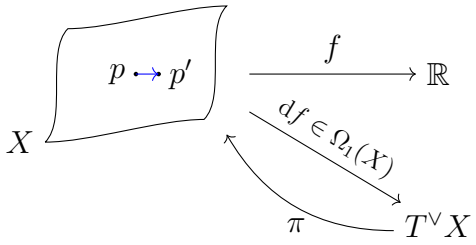
therefore simplifies to



which is the complex of differential forms in two dimensions $(\Omega^\bullet, d)$. So we've found a resolution for the sheaf of locally constant functions in two dimensions and we now have a hint of what to do to generalize to higher dimensions.

There's also a geometric reason for which the product of forms should be antisymmetric. Let $p, p' \in X$ be two points in a manifold X, so that $p' = p + \varepsilon v$. The derivation of a function $f : X \to \mathbb{R}$ is

$$\mathrm{d}f : V \mapsto \frac{f(p') - f(p)}{\varepsilon}.$$



Every 1-form $\mathrm{d}x$ represents an infinitesimal oriented 1-volume and a 2-form $\mathrm{d}x\,\mathrm{d}y$ is an infinitesimal oriented 2-volume. The reversal of order of $\mathrm{d}x$ and $\mathrm{d}y$ corresponds to the reversal in orientation of the infinitesimal volume.

We can also see this requirement another way.
Q: When is a pair of functions (g_1, g_2) such that $g_1 = \frac{\partial f}{\partial x}, g_2 = \frac{\partial f}{\partial y}$ for some function f?
A: If this is true, then

$$\frac{\partial g_1}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial g_2}{\partial x}.$$

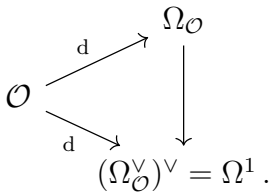
In fact, it's true if and only if that condition is satisfied. Then, if we consider a 2 form $g_1 dx + g_2 dy$, the condition $g_1 = \frac{\partial f}{\partial x}, g_2 = \frac{\partial f}{\partial y}$ says that it came from a 1 form and the next differential must kill it. We therefore define it in such a way that the pair of maps works as follows

$$(g_1, g_2) \mapsto \frac{\partial g_1}{\partial y} - \frac{\partial g_2}{\partial x}\,.$$

End of lecture 7

Last time we completed discussing "differential forms" (there $\mathcal{O} = C^\infty$ functions). We did this as follows

- construct one forms $\Omega^1 = \mathrm{Der}(\mathcal{O})^\vee = \mathrm{Hom}_{\mathcal{O}}(\mathrm{Der}(\mathcal{O}), \mathcal{O}) = (\Omega^\vee_{\mathcal{O}})^\vee$ because $\mathrm{Der}_{\mathcal{O}}(\mathcal{O}) = \mathrm{Hom}_{\mathcal{O}}(\Omega_{\mathcal{O}}, \mathcal{O}) = \Omega^\vee_{\mathcal{O}}$. Having a map from Kähler differentials to one forms then gives us a derivation on functions which is the map d that is usually introduced with differential forms. This map arises by simply composing the universal derivation $\mathcal{O} \to \Omega_{\mathcal{O}}$ and the map to the double dual $\Omega_{\mathcal{O}} \to (\Omega^\vee_{\mathcal{O}})^\vee$:

  The operator d stands for *de Rham*, or "exterior", derivative.

- define $\Omega^k = \bigwedge^k{}_{\mathcal{O}}(\Omega^1)$ where we extended d as a *graded derivation* satisfying $\mathrm{d}^2 = 0$. What we mean by graded derivation is that if we take

  $$\Omega^\bullet = \bigoplus_{k \geq 0} \Omega^k$$

  then we extend d to higher forms via the graded Leibniz rule

  $$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{|\alpha|} \alpha \wedge d\beta \tag{7}$$

  where the sign ensures that antisymmetry works out correctly (compare with Eq. (4)).

- We thus have the complex of differential forms

$$\begin{array}{ccccccccc} \Omega^0 & \xrightarrow{\mathrm{d}} & \Omega^1 & \xrightarrow{\mathrm{d}} & \Omega^2 & \xrightarrow{\mathrm{d}} & \dots & \xrightarrow{\mathrm{d}} & \Omega^n \\ \wr & & & & & & & & \\ \underline{\mathbb{R}} & & & & & & & & \end{array}$$

  where each of the terms is a fine sheaf, so we finally have a resolution for the sheaf of locally constant functions in the sense that forms are a "good" replacement for locally constant functions.

We can justify the sign that appears in (7) by considering again what brought us to the Leibniz rule in the first place, as in (5). Let $\alpha \in \Omega^k$ and $\beta \in \Omega^l$ and $\alpha \wedge \beta = (-)^{kl}\beta \wedge \alpha$. Derivations should preserve the product infinitesimally, so if we have φ of degree n and a parameter ε of degree $-n$, then

$$\begin{aligned}(\alpha + \varepsilon\varphi(\alpha)) \wedge (\beta + \varepsilon\varphi(\beta)) &= \alpha \wedge \beta + \varepsilon\varphi(\alpha) \wedge \beta + \alpha \wedge \varepsilon\varphi(\beta) \\ &= \alpha \wedge \beta + \varepsilon\varphi(\alpha) \wedge \beta + (-)^{kn}\varepsilon\alpha \wedge \varphi(\beta) \\ &= \alpha \wedge \beta + \varepsilon\varphi(\alpha \wedge \beta)\end{aligned}$$

4 From observables to symmetries

Today: A little bit of useful formal calculus (*Cartan's formula*). Recall we now have:

- a notion of observable (commutative algebra A),
- a notion of symmetry ($\mathrm{Der}(A)$).

Q: How can an observable generate a symmetry?
We expect this because in physics we always say that the Hamiltonian (an observable) generates time translations (a symmetry).

To start, let's do a quick recap of Cartan's formula and interior product. Let's fix some space with (smooth) functions $\mathcal{O}$ and let X be a vector field $X \in \mathrm{Der}(\mathcal{O})$. The field X acts on functions via the directional derivative. Being a derivation, it factors through Kähler differentials and in particular it also factors through Ω^1, i.e. there exists a map $i_X \in \mathrm{Hom}_{\mathcal{O}}(\Omega^1, \mathcal{O})$ so that the diagram below commutes.

$$\begin{array}{ccc} & & \Omega^1 \\ & \overset{\mathrm{d}}{\nearrow} & \downarrow{\scriptstyle i_X} \\ \mathcal{O} & \xrightarrow[f \mapsto X(f)]{} & \mathcal{O} \end{array}$$

Just like we did for d, we can extend i_X as a graded derivation to the whole complex $\Omega^\bullet$:

$$i_X(\alpha \wedge \beta) = i_X\alpha \wedge \beta + (-1)^{|\alpha|}\alpha \wedge i_X\beta\,.$$

Now i_X is a map of degree -1 on $\Omega^\bullet$ (while the map d was a degree $+1$ map). The way i_X acts is that if we have $\omega \in \Omega^k$ then $i_X(\omega) \in \Omega^{k-1}$ acts as

$$\omega(X, X_1, \ldots, X_{k-1}) =: i_X(\omega)(X_1, \ldots, X_{k-1})$$

for any collection of vector fields $(X_1, \ldots, X_{k-1})$. This map is sometimes called contraction or *inner product*. In our complex we thus have a pair of maps d and i_X going in opposite directions:

$$\Omega^0 \underset{i_x}{\overset{\mathrm{d}}{\rightleftarrows}} \Omega^1 \underset{i_x}{\overset{\mathrm{d}}{\rightleftarrows}} \Omega^2 \underset{i_x}{\overset{\mathrm{d}}{\rightleftarrows}} \ldots \underset{i_x}{\overset{\mathrm{d}}{\rightleftarrows}} \Omega^n\,.$$

Side remark: if I consider two derivations $X, Y \in \mathrm{Der}(\mathcal{O}) \subset \mathrm{Map}(\mathcal{O}, \mathcal{O})$, their product (composition) is not a derivation

$$\begin{aligned} XY(fg) &= X(Y(f)g + (-)^{|f||Y|}fY(g)) \\ &= XY(f)\cdot g + (-)^{|f|\,|Y|+|f|\,|X|}f\cdot XY(g) \\ &\quad + (-)^{|X|(|Y|+|f|)}Y(f)X(g) + (-)^{|f|\,|Y|}X(f)Y(g) \end{aligned}$$

because the two terms in the last line spoil the graded Leibniz rule. If we computed the same thing in the opposite order, we would get

$$\begin{aligned} YX(fg) &= Y(X(f)g + (-)^{|f||X|}fX(g)) \\ &= YX(f)\cdot g + (-)^{|f|\,|X|+|f|\,|Y|}f\cdot YX(g) \\ &\quad + (-)^{|Y|(|X|+|f|)}X(f)Y(g) + (-)^{|f|\,|X|}Y(f)X(g)\,. \end{aligned}$$

So if we take $[X, Y] = XY - (-)^{|X|\,|Y|}YX$, the terms that spoil the graded Leibniz rule exactly cancel and we see that it *is* a derivation:

$$[X, Y](fg) = [X, Y](f)\cdot g + (-)^{|f|\,|X|+|f|\,|Y|}f\cdot[X, Y](g)\,.$$

Upshot: Derivations form a (graded) *Lie algebra*! On derivations there's no product but there *is* a commutator! This is exactly the structure of a Lie algebra.

Definition 4.1. A (graded) Lie algebra is a (graded) vector space V with a map $[\cdot,\cdot] : V \otimes V \to V$ called commutator or Lie bracket. This map should be (graded) antisymmetric

$$[X,Y] = -(-)^{|Y||X|}[Y,X]$$

and satisfy the Jacobi identity:

$$[X,[Y,Z]] = [[X,Y],Z] + (-1)^{|Y||X|}[Y,[X,Z]]\,.$$

Note that these properties are automatic if we actually take commutators as Lie bracket maps. Lie algebras are very often used in physics because we're interested in studying infinitesimal symmetries. The last formula should look familiar: it looks like the property of being a derivation[3]! In fact it's exactly that for the map $[X,-] : V \to V$ which sends an element Y to its bracket $[X,Y]$ with some fixed vector X. In other words, for every $X \in V$ we get a derivation, so we have a map

$$\begin{aligned} V &\to \mathrm{Der}_{[\cdot,\cdot]}(V) \\ X &\mapsto [X,-]\,. \end{aligned}$$

In particular, going back to differential forms, we have two derivations d and i_X so we can take their commutator and get another derivation of Ω^0 which will be of degree 0. Let us see then see how

$$\mathscr{L}_X = \mathrm{d}i_X + i_X\mathrm{d} \tag{8}$$

acts on functions:

$$\mathscr{L}_X f = (\mathrm{d}i_X + i_X\mathrm{d})f = i_X\mathrm{d}f = X(f)$$

exactly the directional derivative along X! On exact one forms we have

$$\mathscr{L}_X\mathrm{d}g = (\mathrm{d}i_X + i_X\mathrm{d})\mathrm{d}g = \mathrm{d}(i_X\mathrm{d}g) = \mathrm{d}(X(g))$$

in other words $\mathscr{L}_X$ commutes with d. The object $\mathscr{L}_X$ is called *Lie derivative*. Equation (8) is sometimes called Cartan's magic formula. What's "magic" about is that its geometric definition seems rather involved, requiring to evaluate forms at different points and then taking a limit, while this expression is kind of the simplest degree 0 derivation we could write on differential forms.

We can put algebra and Lie algebras structures together:

[3] a derivation of a Lie algebra, i.e. a map $\phi : V \to V$ such that $\phi[Y,Z] = [\phi(Y),Z] + (-1)^{|Y||\phi|}[Y,\phi(Z)]$.

Definition 4.2. A *Poisson algebra* A is a vector space with

1. a commutative algebra structure $\cdot : A \otimes A \to A$;
2. a Lie algebra structure $\{\cdot,\cdot\} : A \otimes A \to A$;

such that the map $\{f, -\} \in \mathrm{Hom}(A, A)$ (the "adjoint" action by infinitesimal symmetries of A) is a derivation:

$$\{f, gh\} = \{f, g\}h + g\{f, h\}\,.$$

The final condition is the compatibility condition of the bracket with the product: the map $f \mapsto \{f, -\}$ should be a derivation of *both* structures.

In classical mechanics one defines a Poisson bracket as

$$\{f, g\} = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p^i} - \frac{\partial f}{\partial p^i}\frac{\partial g}{\partial q^i}$$

and this makes the observables $\mathcal{O}$ into a Poisson algebra. The output of the map $f \mapsto \{f, -\} \in \mathrm{Der}(\mathcal{O})$ is called the "Hamiltonian vector field" associated to f. Since $\mathrm{Vect} = \mathrm{Der}(\mathcal{O})$, we have $\{f, -\} \in \mathrm{Vect}$ and one usually writes the vector field as $X_f := \{f, -\}$. But now, by antisymmetry, the map from observables $\mathcal{O}$ to symmetries Vect given by $f \mapsto X_f$ is itself a derivation from $\mathcal{O}$ to Vect:

$$\{fg, -\} = f\{g, -\} + g\{f, -\}\,.$$

So it factors through one forms with some map π:

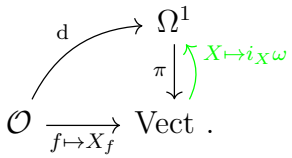
$$\begin{array}{ccc} & \overset{\mathrm{d}}{\nearrow} & \Omega^1 \\ & & \downarrow{\scriptstyle \pi} \\ \mathcal{O} & \xrightarrow[f \mapsto X_f]{} & \mathrm{Vect}\,. \end{array}$$

The map π is called *Poisson tensor* (or Poisson bivector). What kind of object is π? It's a map $\pi \in \mathrm{Hom}_{\mathcal{O}}(\Omega^1, \mathrm{Vect})$ so we could write it as an element of $\mathrm{Vect}^{\otimes 2}$ or rather of $\bigwedge^2{}_{\mathcal{O}}(\mathrm{Vect})$ since we ask for antisymmetry. Note that this is different from the symplectic form what one usually has in classical mechanics since a two form would be $\omega \in \Omega^2 = \bigwedge^2{}_{\mathcal{O}}(\Omega^1)$. We can see the π tensor by writing the Poisson bracket as

$$\{f, g\} = \pi^{ij}\partial_i f \partial_j g\,.$$

So in a sense if I have π and ω I have a way of raising indices and a way of lowering them. If I want these to be the same then ω should be non-degenerate and we

then have $\pi = \omega^{-1}$:



In this picture it's easy to think geometrically about conserved charges.

Example 11. Let $\mathfrak{g}$ be a Lie algebra. Then $\mathfrak{g}^\vee$ has a natural Poisson tensor. That's because functions on $\mathfrak{g}^\vee$ corresponds to elements of $\mathrm{Sym}^\bullet(\mathfrak{g})$ and we have a bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which extends uniquely as a derivation to a Poisson bracket on $\mathrm{Sym}^\bullet(\mathfrak{g})$. This example is interesting because it's clearly degenerate. For example if I take the element 0 then the map is just zero and in general bracketing with different elements acts differently, so the tensor varies in rank over $\mathfrak{g}^\vee$.

Let P be some phase space and A its Poisson algebra of functionals. $\mathfrak{g}$ acts on P by a map $\mathfrak{g} \to A$ of Lie algebras, or by a map $\mathrm{Sym}(\mathfrak{g}) \to A$ of Poisson algebras or by a map $P \to \mathfrak{g}^\vee$ of Poisson spaces.

End of lecture 8

5 Special lecture: Geometric quantization

Everything we've worked with has been classical so it's interesting to see how one would pass to the quantum version.
Idea of quantization: we'd like to have a map

$$\text{Classical} \longrightarrow \text{Quantum}\,.$$

It turns out that this construction will not be functorial. In addition, we also have a map going in the other direction, usually called "taking the classical limit". Of course it's not obvious how to go from one to the other since we have rather different structures:

	Classical $\overset{\text{Quantization}}{\underset{\text{Classical limit}}{\rightleftarrows}}$ **Quantum**		
Phase space	(M, ω)	$(\mathscr{H}, \langle\cdot,\cdot\rangle)$	Hilbert space
Observables	$f \in \mathcal{C}^\infty(M)$	$\hat{f}$	Hermitian operators
Poisson bracket	$\{\cdot,\cdot\}$	$[\cdot,\cdot]$	Commutator .

The starting point for a classical theory is phase space (M, ω), a space of observables $C^\infty(M)$ and a Poisson bracket $\{\cdot, \cdot\}$. On the other side, for a quantum theory we have a Hilbert space $(\mathscr{H}, \langle \cdot, \cdot \rangle)$, observables are Hermitian operators on the Hilbert space $\hat{f}$ and on this space of observables we have the bracket given by the commutator. There are a few options to obtain a quantum theory from a classical one, using different mathematical tools:

- Algebra: deformation quantization (Kontsevich),
- Analysis: stochastic quantization,
- Geometry: geometric quantization.

We'll focus on this last one.

Definition 5.1 (*Dirac's axioms*)**.** A *quantization* procedure assigns self adjoint operators $\hat{Q}_f$ on a Hilbert space $\mathscr{H}$ to smooth functions $f \in C^\infty(M)$, i.e. we have a map

$$\hat{Q} : C^\infty(M) \to \mathrm{End}(\mathscr{H}) \,.$$

This map must be

1. linear:
$$\hat{Q}_{\lambda f + \mu g} = \lambda \hat{Q}_f + \mu \hat{Q}_g \,,$$
2. normalized:
$$\hat{Q}_{\mathbb{1}} = \hat{\mathbb{1}} \,,$$
3. bracket preserving:
$$\hat{Q}_{\{f,g\}} = [\hat{Q}_f, \hat{Q}_g] \,,$$
4. minimal: a complete set of Poisson commuting algebras is quantized to a complete set of Lie commuting algebras.

A map Q satisfying 1-3 is called a *prequantization*. There is a no go theorem stating that not every symplectic manifold may be quantized. We're going to get this map Q geometrically.

We start with a recap on *symplectic geometry*. Given a symplectic (i.e. smooth, equipped with a closed non-degenerate differential 2-form) manifold (M, ω) we get a Lie group $\mathrm{Sympl}(M, \omega)$ which is the group of diffeomorphisms of M that preserve the *symplectic form* in the sense that $f^*\omega = \omega$. This group has Lie algebra given by symplectic vector fields $\chi_{\mathrm{Sympl}}(M, \omega)$ which are vector fields preserving the symplectic form in the sense that $\mathscr{L}_X \omega = 0$. The Hamiltonian vector field of a

smooth function f on M is the unique vector field X_f satisfying

$$\omega(X_f, \cdot) = df, \tag{9}$$

and we call $\chi_{\text{Ham}}(M, \omega)$ the collection of such vector fields.

Remark. Consider $X \in \chi_{\text{Sympl}}$, we then have

$$0 = \mathscr{L}_X \omega = i_X \mathrm{d}\omega + \mathrm{d} i_x \omega = \mathrm{d}(\omega(X, \cdot))$$

so the condition $0 = \mathscr{L}_X \omega$ is equivalent to the condition that $\omega(X, \cdot)$ is closed. Analogously, a Hamiltonian vector field $X_f \in \chi_{\text{Ham}}$ instead has the property that $\omega(X_f, \cdot) = \mathrm{d}f$ so $\omega(X_f, \cdot)$ is exact. The obstruction of a symplectic vector field into a Hamiltonian vector field is $H^1(M, \mathbb{R})$ in the sense that if $H^1(M, \mathbb{R}) = 0$ then any symplectic vector field is a Hamiltonian vector field, otherwise that is not the case.

Definition 5.2. The Poisson bracket of functions on a symplectic manifold (M, ω) is the map

$$\{\cdot, \cdot\} : \mathcal{C}^\infty(M) \times \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M),$$

$$f, g \mapsto \{f, g\} \coloneqq \omega(X_f, X_g).$$

Note that $\omega(X_f, X_g) = \mathrm{d}f(X_g) = X_g(f)$.

Theorem 5.1. *$(C^\infty(M), \{\cdot, \cdot\})$ is a Poisson algebra (and therefore also a Lie algebra) and if M is connected, then $g =$ constant if and only if $\{f, g\} = 0, \forall f \in C^\infty(M)$.*

Proposition 5.2. *The map*

$$\begin{aligned} (C^\infty(M), \{\cdot, \cdot\}) &\to (\chi_{\text{Ham}}(M, \omega), [\cdot, \cdot]) \\ f &\mapsto X_f \end{aligned}$$

is a Lie algebra homomorphism.

We can now pass to prequantization. Consider a $\mathbb{C}$-line bundle $L \to M$ equipped with a Hermitian metric $\langle \cdot, \cdot \rangle$. Then try to find a Hermitian connection on L, i.e. a connection ∇ satisfying

$$\mathrm{d}\langle s_1, s_2 \rangle = \langle \nabla s_1, s_2 \rangle + \langle s_1, \nabla s_2 \rangle \qquad \forall s_1, s_2 \in \Gamma(L)$$

such that the curvature of ∇ denoted R_∇ coincides with ω:

$$R_\nabla = -\frac{1}{i\hbar}\omega\,.$$

If such a connection exists, (M,ω) is called *prequantizable* and $(L,\langle\cdot,\cdot\rangle,\nabla)$ is called a prequantum line bundle. In fact we can get a prequantization as follows:

- Hilbert space: $\Gamma_{L^2}(L)$ (square integrable sections) with hermitian inner product
$$\langle s_1, s_2\rangle \coloneqq \frac{1}{(2\pi\hbar)^n}\int_M h(s_1,s_2)\frac{\omega^n}{n!}\,.$$
- Observables: let $a\in\mathcal{C}^\infty(M)$ and take the Hermitian vector field X_a:
$$\hat{Q}_a = -i\hbar\nabla_{X_0} + a$$
where a acts as a multiplication.
  - this operator is self-adjoint with respect to $\langle\cdot,\cdot\rangle$.
  - a theorem (Kostant-Souriau) says that $\hat{Q}$ satisfies Dirac conditions 1., 2. and 3. . Note that the third condition corresponds to
$$\hat{Q}_{\{a,b\}} = \frac{1}{i\hbar}[\hat{Q}_a,\hat{Q}_b].$$
where $[\cdot,\cdot]$ is the Lie bracket of vector fields.

However, this is not really what we want: this Hilbert space is too big! Sections of that type look like $s(x,p)$ whereas a wave function would be $\psi(x)$ or $\psi(p)$ so we have twice as many variables as we should have. We therefore should cut M to half of its dimension, but there's not a unique way of doing that! This is a choice called a *polarization*.

Definition 5.3. A polarization of (M,ω) is a complex subbundle $P\in TM\otimes\mathbb{C}$ such that

- P is involutive, i.e. $[P,P]\subset P$,
- $\forall m\in M$, P_m is Lagrangian.

However the fact that there's a choice involved is displeasing: we'd like for there to be a unique quantum theory corresponding to a classical theory, while here we have this arbitrary choice involved.

In conclusion, it's a conceptually nice way of quantizing a classical theory, but concretely it's not entirely satisfactory and often too impractical. If we instead upgrade our symplectic manifold to a Kähler manifold we actually get a natural polarization.

5.1 Quantizing Kähler manifolds

Definition 5.4. Let (M, ω, J), where J means almost complex structure, i.e. J is a bundle isomorphism $TM \to TM$ satisfying $J^2 = -\mathbb{1}$ on the fibers. Then if ω and J are such that

- (M, ω) is symplectic manifold,
- (M, J) is a complex manifold,

(M, ω, J) is called a *Kähler manifold*. On it there's a natural metric:

$$g(X, Y) \coloneqq \omega(JX, Y)$$

The idea is that a Kähler manifold has the structures of a symplectic structure, a complex structure and a metric and they're all compatible.

Example 12. The complex projective space

$$\mathbb{P}^n := \mathbb{C}^{n+1} \setminus \{0\} \quad /\mathbb{C}^\times$$

is Kähler. Therefore also any projective variety, i.e. a variety X with an inclusion $X \hookrightarrow \mathbb{P}^n$, is also Kähler.

In this context we have a natural polarization given by the antiholomorphic tangent bundle $P = T^{0,1}M = \{v \in TM | Jv = -iv\}$. P is involutive, it is Lagrangian and $P \cap \overline{P} = \emptyset$. Now let $\pi : L \to M$ be a holomorphic line bundle and fix a Hermitian metric $\langle \cdot, \cdot \rangle$ on L. Then there exists a unique connection on L which is compatible with both $\langle \cdot, \cdot \rangle$ and the complex structure. This is called Chern connection ∇ which may be thought of as a holomorphic version of Levi civita connection. It turns out that the curvature of the connection satisfies $[R_\nabla] \in H^{1,1}(M, \mathbb{C}) \cap H^2(M, \mathbb{Z})$ (respectively Dolbeault cohomology and characteristic class) i.e. we have a class $[R_\nabla]$ such that if I take its integral I get an integer and as a Dolbeault form it's of the form $(1, 1)$. This is called the first Chern class $c_1(L) \cong [R_\nabla]$.

Now the Hilbert space is given by sections of L

$$(\Gamma_{\text{hol}}(L), \langle \cdot, \cdot \rangle)$$

which is now a finite dimensional vector space. This fact shouldn't be too surprising because holomorphic functions on a compact Riemann surface (for example $\mathbb{P}^1 = \mathbb{C} \cup \{\infty\}$) are simply given by constants, i.e. the space of holomorphic functions on a compact Riemann surface is one dimensional. However in quantum mechanics

we usually have an infinite dimensional vector space! We can therefore take a collection of such vector spaces as our Hilbert space

$$\mathscr{H} = \bigoplus_{k=1}^{n} \mathscr{H}^{\otimes k}$$

and this in particular gives something sensible for Chern-Simons theory over a 3-manifold M_3 (the boundary $\partial M = \Sigma$ being a 2-manifold), which has a partition function of form:

$$Z = \int_{\mathcal{A}} \mathrm{D}A \exp\left(-ik \int_M A \wedge \mathrm{d}A + \frac{2}{3} A \wedge A \wedge A\right).$$

End of lecture 9

The idea of quantization can be summarized by the following diagram:

Classical state space			**Quatum state space**	
symplectic manifold	(M, ω)	$\xrightarrow[\text{quantization}]{}$	$\mathscr{H}$	Hilbert space
Poisson algebra	$(\mathcal{O}_M, \pi = \omega^{-1})$	$\longrightarrow$	$A_\hbar$	noncommutative algebra of linear operators on $\mathscr{H}$

$$A_{\hbar=0} = \mathcal{O}_M \quad \xleftarrow[\text{classical limit}]{\hbar \to 0} \quad A_\hbar$$

In addition, in $A_\hbar$ we have that $[X, Y] = XY - YX \in \hbar A_\hbar$ and

$$\lim_{\hbar \to 0} \frac{1}{\hbar}[X, Y] = \{X_{\mathcal{O}}, Y_{\mathcal{O}}\}.$$

so the Poisson bracket can be thought of as the first order term in $\hbar$ of the commutator.

A typical example of this procedure is the point particle with $M = \mathbb{R}^2_{x,p}$ and $\omega = \mathrm{d}p \wedge \mathrm{d}x$ from which we get the Hilbert space $\mathscr{H} = L^2(\mathbb{R}_x, \mathbb{C})$ (or $\mathscr{H} = L^2(\mathbb{R}_p, \mathbb{C})$) on which to the functions x and p we associate the following operators $x \mapsto x,\ p \mapsto \hbar\partial/\partial x$ (or $p \mapsto p\ x \mapsto -\hbar\partial/\partial p$).

We saw that in geometric quantization one start with a complex line bundle (a *line bundle* is a rank 1 vector bundle and a complex bundle is a vector bundle with complex vector spaces as fibers). The simplest example is the "trivial complex line bundle":

$$\begin{array}{c} M \times \mathbb{C} \\ \big\downarrow \\ M\,. \end{array}$$

The idea then was to take a subspace of $\mathcal{O}_M$ defined by a "polarization" (abstraction of "complete set of commuting observables") and to define $A_\hbar$ we take a choice of connection over M with curvature ω.

Maybe think of this as a motivation for the introduction of the notions of connection and curvature, which we'll get to.

6 Poisson structure vs. Symplectic structure

Reminder: we "discovered" the definition 4.2 of a Poisson algebra A which should describe both a space and the symmetries of that space simultaneously ("observables generate symmetries" and Noether's theorem talks about a map going the other way: given a symmetry I can find an observable that generates it). Recall that this meant A is at the same time a commutative algebra and a Lie algebra in a compatible way: for any $f \in A$, $\{f, -\}$ is a derivation of the product. In other words we get a map $A \to \operatorname{Der} A$ which acts as

$$f \mapsto \{f, -\} = X_f \tag{10}$$

where X_f is called the "Hamiltonian vector field". Applied to our algebra $A = \mathcal{O}_M$ we have a map $\mathcal{O}_M \to \operatorname{Vect}(M)$. By antisymmetry this map is also a derivation and as usual, this means that the map factors through one forms:

$$\begin{array}{ccc} & \overset{\mathrm{d}}{\nearrow} & \Omega^1 \\ & & \downarrow{\scriptstyle \pi} \\ \mathcal{O} & \xrightarrow[f \mapsto X_f]{} & \operatorname{Vect} . \end{array}$$

where π is a bundle map $T^*M \to TM$ or equivalently $\pi \in \Gamma(\bigwedge^2 TM)$.

Compare this with the notion of a symplectic form: $\omega \in \Gamma(\bigwedge^2 T^*M)$ or equivalently a map $\omega : TM \to T^*M$. If this is nondegenerate, i.e. the map $\omega : TM \xrightarrow{\sim} T^*M$ is an isomorphism, then we can take $\pi = \omega^{-1} : T^*M \xrightarrow{\sim} TM$. So if ω is invertible (or π is) then having a Poisson algebra is equivalent to having a symplectic form. Instead more in general we can have a Poisson algebra structure where π is not an isomorphism. Conversely, we can consider having a two form ω which *is* degenerate.

Definition 6.1. A *symplectic structure* on M is an element $\omega \in \Omega^2(M)$ satisfying:

(A) $\mathrm{d}\omega = 0$ ("closed"),

(B) ω is non-degenerate ($\omega : TM \to T^*M$ is an isomorphism).

A closed 2-form (not satisfying (B)) is called *presymplectic form*. Item (B) is violated by e.g. imposing constraints: restricting to subspaces $N \subset M$. Item (A) in this case is not violated. But what is (A) telling me? It turns out that if ω is nondegenerate we have

$$\mathrm{d}\omega = 0 \longleftrightarrow \text{Jacobi identity for } \pi = \omega^{-1}.$$

One way to see this is to use coordinates: write $\pi = \pi^{ij}\partial_i \wedge \partial_j$, then the Poisson bracket takes the form

$$\{f, g\} = \pi(\mathrm{d}f \wedge \mathrm{d}g) = \pi^{ij}\partial_i f \partial_j g$$

and taking the left hand side of the Jacobi identity gives

$$\begin{aligned}\{\{f, g\}, h\} + \text{cycl.} &= \pi^{ij}\partial_i(\pi^{kl}\partial_k f \partial_l g)\partial_j h + (j \to k \to l) \\ &= \pi^{ij}(\partial_i \pi^{kl})\partial_k f \partial_l g \partial_j h + (2^{\text{nd}} \text{ order correction}) + (j \to k \to l).\end{aligned}$$

By observing the blue term in the last line we conclude

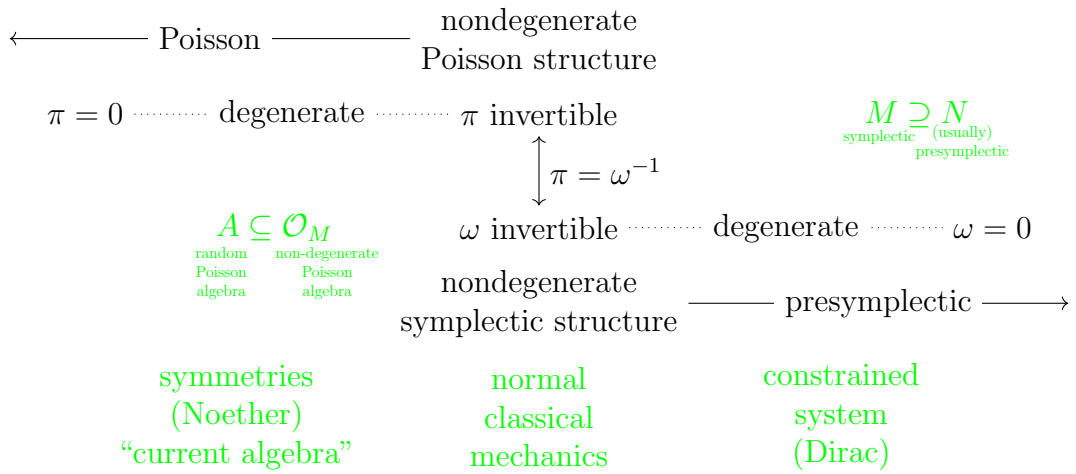
$$\pi^{ij}\partial_i \pi^{kl} + (j \to k \to l) = 0 \quad \in \bigwedge^3 TM \tag{11}$$

We can then use $\omega = \pi^{-1}$ to lower the 3 indices and get a 3-form:

$$\pi^{ji}\partial_i \pi^{kl}\omega_{jm}\omega_{kn}\omega_{lp} + + (j \to k \to l) \overset{(1)}{=} (\partial_m \pi^{kl})\omega_{kn}\omega_{lp} + (m \to n \to p) \overset{(2)}{=} \partial_m \omega_{np} = \mathrm{d}\omega,$$

where in step (1) we used $\pi^{ji}\omega_{jm} = -\delta^i_m$. While in step (2) we used $\mathrm{d}\omega = -\omega \cdot \mathrm{d}\pi \cdot \omega$ which holds since $\mathrm{d}(\pi\omega) = 0$ implies $(\mathrm{d}\pi)\omega + \pi\,\mathrm{d}\omega$. We then see that equation (11) is equivalent to the linear PDE $\mathrm{d}\omega = 0$.

If we consider the degenerate case we have the following diagram



So in general π and ω are inequivalent structures one can have, while in the special case in which they're degenerate, they're actually equivalent. The case ω degenerate arises because there are situations in which we're interested in subspaces $N \subset M$ of the phase space M and usually even though M is symplectic, N is only presymplectic.

A is some random Poisson algebra and $\mathcal{O}_M$ is some nondegenerate Poisson and $A \subseteq \mathcal{O}_M$. Symmetries (Noether), "current algebra". Recall that we mentioned that in some cases a group action G on M, which corresponds to a Lie map

$$\mathfrak{g} \longrightarrow \text{Vect}$$

and we have a map $\mathcal{O}_M \to \text{Vect}$ map sending $f \mapsto X_f$ and sometimes the map above lifts to a map to $\mathcal{O}_M$ in the sense that

$$\begin{array}{ccc} & & \mathcal{O}_M \\ & \nearrow & \downarrow \\ \mathfrak{g} & \longrightarrow & \text{Vect} \end{array}$$

commutes. In this case, we get a map $\text{Sym}(\mathfrak{g}) = \mathcal{O}_{\mathfrak{g}^\vee} \to \mathcal{O}_M$ by extending that map as a Poisson map.

What does "$\pi^{ji}\partial_i\pi^{kl} + \text{cycl.} = 0$" actually mean? We understand $\Omega^\bullet(M) = \Gamma(\bigwedge^\bullet T^*M)$ which has a universal differential d_{dR}. Instead, what does $\text{PV}^\bullet(M) = \Gamma(\bigwedge^\bullet TM)$ have? (where PV stands for poly-vector) We know the "first terms" of a (seemingly) universal structure: we have the Lie bracket of vector fields and of vector fields with functions:

$$[\text{PV}^1, \text{PV}^1] \to \text{PV}^1$$

$$\begin{aligned} [\text{PV}^1, \text{PV}^0] &\to \text{PV}^0 \\ [X, f] &\mapsto X(f) \end{aligned}$$

and we have a graded commutative structure $\wedge : \text{PV}^i \otimes \text{PV}^j \to \text{PV}^{i+j}$. We can then extend (uniquely) $[\cdot,\cdot]$ as a graded Lie bracket of degree -1 that makes $PV^\bullet$ into a graded Poisson algebra! So a coordinate free way of writing the expression $\pi^{ji}\partial_i\pi^{kl} + \text{cycl.}$ that we had before is that it is the map

$$[\text{PV}^2, \text{PV}^2] \to \text{PV}^3$$

inserting π in both slots. So we can state the condition that π correspond to a Poisson bracket as $[\pi, \pi] = 0$.

End of lecture 10

7 Origin of the symplectic structure

We started with π because it allows us to construct a vector field from an observable since we have this idea that observables should correspond to symmetries. In other words we have a map $C^\infty(M) \to \mathrm{Vect}(M) = \mathrm{Der}(C^\infty(M)) = \Gamma(TM)$ which sends $f \mapsto X_f = \pi(df, -)$. In addition to this, we get a Lie structure on $C^\infty(M)$ by taking the Poisson bracket $\{f, g\} = \pi(df \wedge dg)$ which is automatically a derivation $\{f, gh\} = g\{f, h\} + \{f, g\}h$, while the Jacobi identity for $\{\cdot, \cdot\}$ corresponds to $[\pi, \pi] = 0 \in PV^\bullet(M)$. In the nondegenerate case (π invertible), this (nonlinear) condition on π corresponds to the (linear) condition $d\omega = 0$ for ω. The case π nondegenerate is special because symmetries of the space exactly correspond to functions on the space.

Now, having done a course in classical mechanics we probably accepted the fact that this is what we should do, but it's not at all obvious what Newton's equations have to do with symplectic geometry. We now want to understand this connection. When we first see Newton's equations in the first year, we might meet something as

$$m\ddot{x} = F(x).$$

Maybe we want to only consider conservative forces, so we would write

$$m\ddot{x} = -\frac{\partial U}{\partial x}.$$

Later in the course, we figure that momentum is better than $\ddot{x}$, so we use the equation

$$\dot{p} = -\frac{\partial U}{\partial x}.$$

But then, when we see analytical mechanics for the first time, we encounter this novel function

$$H = \frac{p^2}{2m} + U(x)$$

and the equations of motion are suddenly nothing like what we knew before:

$$\dot{x} = \{H, x\}, \qquad \dot{p} = \{H, p\}.$$

Now, we want to enlighten these equations from the perspective of the mathematical tools that we developed during this course. Mathematically what's going on is that we're describing conservative motion on some manifold N and we find a nice way of describing it using the symplectic structure on T^*N. The question then arises: why is the symplectic structure on T^*N related to conservative motion on N? And does it only work for conservative motion?

The answer lies in the Lagrangian formulation. **Idea**: a set of PDEs is related to a Hamiltonian system if it is variational. More precisely, the PDEs we mean are the equations of motion of our system, by a Hamiltonian system we mean a symplectic manifold with a function called Hamiltonian which generates time evolution, and by variational system we mean that it comes from some principle of least action.

However, we have a lot to unpack to understand this idea. To start with, what is a Lagrangian? It's a functional on the space of possible histories of a system, valued in top forms on the spacetime. We can then take the integral of this top form to obtain the "action functional" $S : \{\text{histories}\} \to \mathbb{R}$

$$S = \int \mathscr{L},$$

and we then try to find histories for which $\delta S = 0$ ("S is extremal"). In particle mechanics in d dimensions histories correspond to trajectories of particles:

$$\mathscr{E} \coloneqq \{\text{histories}\} = C^\infty(\underbrace{\mathbb{R}}_{\text{time}} \to \underbrace{\mathbb{R}^d}_{\text{space}})$$

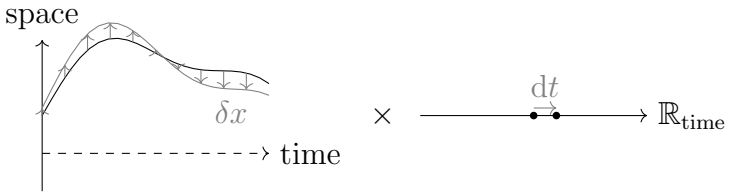
so we have functions of the form $x : t \mapsto x(t) \in \mathbb{R}^d$. Given such a trajectory x (and possibly a coordinate system), I can extend it to a map $J^k x \in C^\infty(\mathbb{R} \to \mathbb{R}^{d(k+1)})$ which acts as

$$t \mapsto (x(t), \dot{x}(t), \dots, x^{(k)}(t))$$

which is called the *k-jet* of x at t. Pick a polynomial $L : \mathbb{R}^{d(k+1)} \to \mathbb{R}$, we then obtain a Lagrangian since we can plug in histories to this polynomial and we write "$\mathscr{L}(x(t), \dot{x}(t))$" (composition of L and $J^1 x$) for the usual case $k = 1$ (or with higher derivatives in more generality). Now that we have some grasp on the concepts of Lagrangian and action, we can try to understand what δ in $\delta S = 0$ should be. Note that $\mathscr{L}$ above is actually a function on $\mathbb{R}_{\text{time}}$, equivalently we could think of it as a map:

$$\mathscr{L} : \mathscr{E} \times \mathbb{R}_{\text{time}} \to \mathbb{R}$$

by plugging in not only the k-jet $J^k x$, but also the time at which we evaluate it. We can now generalize to forms: think about $\Omega^{\bullet,\bullet}(\mathscr{E} \times \mathbb{R}_{\text{time}})$ which has two indices because I'm considering a product space. A differential form is then a combination of a "history variation" δx and a "time variation" $\mathrm{d}t$ which we respectively call vertical and horizontal.



The differential also splits into two: $\mathrm{d}+\delta$ respectively spacetime and "variational" differentials. Now $\mathscr{L} \in \Omega^{0,1}(\mathscr{E}\times\mathbb{R}_{\text{time}})$[4], in fact $\mathscr{L}$ sits in a subspace $\Omega^{\bullet,\bullet}_{\text{loc}}(\mathscr{E}\times\mathbb{R}_{\text{time}})$ of functionals that are "local" in time[5]. The action is then the horizontal integral: $S = \int \mathscr{L}$ where by $\int$ we mean the map $\Omega^1(\mathbb{R}_{\text{time}}) \to \mathbb{R}$ [6]. We then have $S \in \Omega^0(\mathscr{E})$. We now want to impose $\delta S = 0$, which we can write as

$$0 = \delta \int \mathscr{L} = \int \delta\mathscr{L} \quad \Longrightarrow \quad \delta\mathscr{L} = \mathrm{d}\gamma$$

and we want this to be true if the EOMs are satisfied. We can then write $\delta\mathscr{L} \in \Omega^{1,1}(\mathscr{E} \times \mathbb{R}_{\text{time}})$ as

$$\delta\mathscr{L} = (\mathrm{EOM}_x)\delta x + \mathrm{d}\gamma$$

for some $\gamma \in \Omega^{1,0}(\mathscr{E} \times \mathbb{R}_{\text{time}})$ and $\mathrm{EOM}_x\delta x \in \Omega^{1,1}(\mathscr{E} \times \mathbb{R}_{\text{time}})$. Consider the free particle:

$$\mathscr{L} = \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\mathrm{d}x \wedge \star\mathrm{d}x$$

then we have

$$\delta\mathscr{L} = m\mathrm{d}\delta x \wedge \star\mathrm{d}x = m\mathrm{d}(\delta x \wedge \star\mathrm{d}x) - m\delta x \wedge \underbrace{\mathrm{d}\star\mathrm{d}x}_{\mathrm{EOM}_x}$$

so the γ from above now has the form

$$\gamma = \delta x \wedge m\star\mathrm{d}x\,,$$

conceptually "$p_x\delta x$". It's the piece we used when we integrated by parts and it's what will give rise to the symplectic form. Now we can define $\omega \in \Omega^{2,0}(\mathscr{E} \times \mathbb{R})$ to be $\delta\gamma$ and we see that

$$d\omega = d\delta\gamma = \delta d\gamma = \delta(\delta\mathscr{L} - \mathrm{EOM}_x\delta x) = \delta(\mathrm{EOM}_x\delta x)$$

so $d\omega = 0$ when the equations of motion are satisfied!

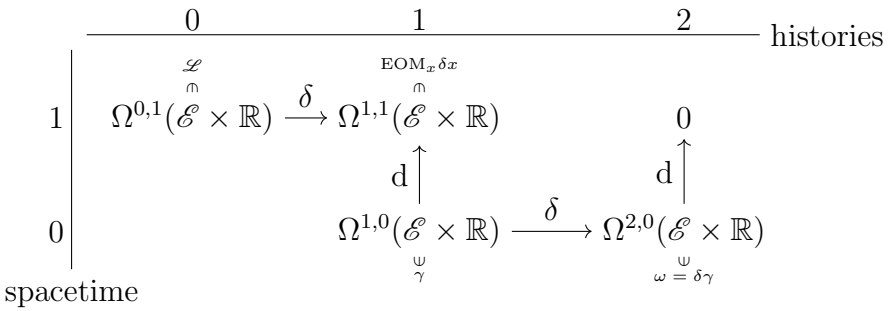
In order to understand all the moving pieces, it's useful to represent these

[4]note that the framework we're working in is basically field theory in 0+1 dimensions, if we worked in actual field theory then $\mathbb{R}_{\text{time}}$ would be upgraded to some spacetime manifold M_n and the Lagrangian would be $\mathscr{L} \in \Omega^{0,n}(\mathscr{E} \times M_n)$.

[5]or spacetime.

[6]or $\Omega^n(M_n) \to \mathbb{R}$.

(bigraded) differential forms as follows.



(12)

End of lecture 11

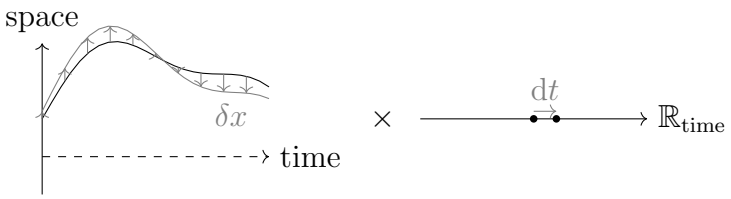
8 Noether's theorem

Remember from the last lecture, we were interested in the following question: which dynamical systems admit Hamiltonian formulations[7]? The answer was: the ones coming from (sufficiently nice) variational principles.

We started by thinking of classical mechanics. Here we have $M = \mathbb{R}_{\text{time}}$ being the space of parameters on which my theory depends, so for field theory we would replace it with some d-dimensional spacetime manifold. We then considered

$$\mathscr{E} = \{\text{space of "histories" of the system}\}$$

which for classical mechanics is simply the space of trajectories $C^\infty(\mathbb{R}_{\text{time}} \to \mathbb{R}^3) \ni x$ with x the map $t \mapsto x(t)$. There's some subspace $\mathscr{F} \subseteq \mathscr{E}$ consisting of histories that solve the equations of motion ("on shell"). We then wrote down the differential forms $(\Omega^{\bullet,\bullet}(\mathscr{E}\times M), \mathrm{d}_{\mathrm{dR}} = \delta + \mathrm{d}_{\mathrm{dR}})$ with a "vertical" differential (on $\mathscr{E}$) and a "horizontal" differential (on M). Note that the forms have double grading – one for $\mathscr{E}$ and one for M. We think of the product of the two spaces and the two differentials as follows.



[7]Phase space X with symplectic form $\omega \in \Omega^2_{\mathcal{O}}(X)$, $H \in \mathrm{Obs} = C^\infty(X)$, time evolution generated by $\{H, -\}$.

We also wrote down the functional that assigns numbers to histories: $S : \mathscr{E} \to \mathbb{R}$ ($S \in \Omega^0(\mathscr{E})$). The forms of interest were the "local" ones $\Omega^{\bullet,\bullet}_{\text{loc}} \subseteq \Omega^{\bullet,\bullet}$: at a point $t \in M$, such forms only depends on x and finitely many of its derivatives all evaluated at the point t. For the free particle it depends only on the first derivative:

$$\mathscr{L} = \frac{1}{2}|\dot{x}|^2 \, \mathrm{d}t.$$

The action should arise from integrating a local Lagrangian density:

$$S = \int_M \mathscr{L}, \qquad \int_M : \Omega^{0,d}_{\text{loc}}(\mathscr{E} \times M) \to \Omega^0(\mathscr{E}).$$

The least action principle then tells us that solutions to the equations of motion are the following subspace of $\mathscr{E}$:

$$\mathscr{F} = \{x \in \mathscr{E} | \delta S(x) = 0\}.$$

We can then write

$$\delta S = 0 \quad \iff \quad \int \delta \mathscr{L} = 0 \quad \iff \quad \delta \mathscr{L} = \mathrm{d}\gamma \, .$$

Note that we're thinking of the integral as a map

$$\int : \Omega^n_c \to \Omega^n_c / \mathrm{d}\Omega^{n-1}_c \cong \mathbb{R}$$

so its kernel is precisely made up of exact $n-1$ forms. The equation $\delta\mathscr{L} = \mathrm{d}\gamma$ holds on $\mathscr{F}$, while in general we have

$$\delta \mathscr{L} = \mathrm{EOM} + d\gamma \qquad \in \Omega^{1,d}_{\text{loc}}$$

where the equations of motion should only depend on the 0th Taylor coefficient of δx. We specify this condition to fix what the splitting of $\delta\mathscr{L}$ should be.

Example 13. Consider a free particle in $\mathbb{R}^3$

$$\mathscr{L} = \frac{1}{2}\langle \mathrm{d}x, \star \mathrm{d}x \rangle \, ,$$

which we write this way in order to highlight the structures present: for the Hodge star $\star$ we require a choice of volume form on M, while the angled brackets indicate an inner product on $\mathbb{R}^3$. Here we have

$$\mathrm{d}x = \frac{\partial x}{\partial t}\mathrm{d}t, \qquad \star \mathrm{d}x = \frac{\partial x}{\partial t} \, .$$

If we take the variation of the Lagrangian we get

$$\delta\mathscr{L} = \langle \mathrm{d}\delta x, \star \mathrm{d}x\rangle = \mathrm{d}\langle \delta x, \star \mathrm{d}x\rangle - \langle \delta x, \mathrm{d}\star \mathrm{d}x\rangle$$

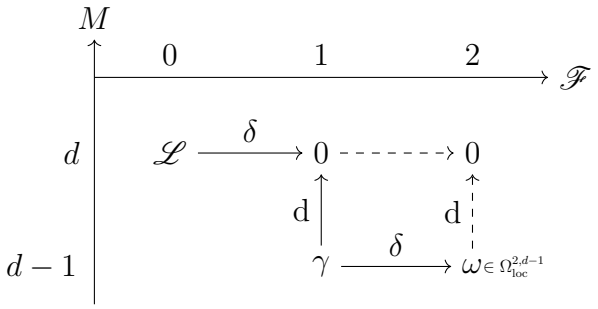
so we can identify $\gamma = \langle \delta x, \star \mathrm{d}x\rangle \in \Omega^{1,d-1}_{\mathrm{loc}}$ and $\langle \delta x, \mathrm{d}\star \mathrm{d}x\rangle = \mathrm{EOM} \in \Omega^{1,d}_{\mathrm{loc}}$ (and $d = 1$).

So we can define $\mathscr{F}$ as the "zero locus" of EOM. However EOM is a form so we can view $\mathscr{F}$ as the intersection of EOM with the identically 0 form which we view as $\mathscr{E}$ lying in $T^*\mathscr{E}$, i.e.

$$\mathscr{F} = \mathscr{E} \bigcap_{T^*\mathscr{E}} (\mathrm{EOM}).$$

If $\mathscr{E}$ were finite-dimensional, then a critical point x of S would satisfy $\partial_i S = 0$, $\forall i$ in every coordinate system at x.

Let us now restrict to $\mathscr{F}$, where interesting things happen. The bicomplex in (12) becomes:



By which we mean $\delta\mathscr{L} + d\gamma = 0, \quad \delta\gamma = \omega$ and $d\omega = 0$. We can then take N, a $d-1$ dimensional submanifold of M, and consider consider

$$\int_N \omega \qquad \in \Omega^2_{\mathrm{loc}}(\mathscr{F})$$

which is independent of (small perturbations) of N. We then have a closed two form on $\mathscr{F}$! So we've finally found the structure we were looking for: $\mathscr{F}$ is the phase space and $\int \omega$ is the symplectic form.

How do symmetries fit into the picture? Consider continuous symmetries (Lie groups) so that we can work infinitesimally (with the Lie algebra). A symmetry is then a vector field X on $\mathscr{E}$ such that $L_X(S) = 0$[8]. I can extend X to be a vector field on $\mathscr{E} \times M$ by taking it to be constant along M ($dX = 0$).

$$L_X(S) = 0 \iff \int L_X(\mathscr{L}) = 0 \iff L_X(\mathscr{L}) = \mathrm{d}\alpha_X$$

[8]here we indicate the Lie derivative by L since we reserve $\mathscr{L}$ for the Lagrangian.

for some $\alpha_X \in \Omega^{0,d-1}_{\text{loc}}(\mathscr{E} \times M)$. Now, we will use the Cartan's formula:

$$L_X(\mathscr{L}) = \delta i_X \mathscr{L} + i_X \delta \mathscr{L} = i_X \delta \mathscr{L}.$$

In particular, on shell (that is on $\mathscr{F} \times M$) it holds

$$i_X \delta \mathscr{L} = i_X \mathrm{d}\gamma = \mathrm{d}(i_X \gamma) = \mathrm{d}\alpha_X.$$

Hence, $\mathrm{d}(i_X\gamma - \alpha_X) = 0$ or

$$J_X = i_X \gamma - \alpha_X \qquad \in \Omega^{0,d-1}_{\text{loc}}$$

is a conserved current. Note that conservation is again only true on shell.

To a vector field X which is a symmetry of the system, there is associated a conserved current J_X. Note that the current is well defined on $\mathscr{E}$, but it's only closed on $\mathscr{F}$ because that's where the equation above is true.

Example 14. Consider, in the example above, the case where X is given by time translations? By this we mean that we define the action on x be a coordinate field on $\mathscr{E}$ to be

$$L_X(x) = X(x) \coloneqq i_{\partial_t} \mathrm{d}x = \star \mathrm{d}x$$

sometimes written as "$\delta_X x = \dot{x}$" and the action on the Lagrangian to be

$$L_X(\mathscr{L}) \coloneqq \mathrm{d} i_{\partial_t} \mathscr{L} = \mathrm{d} \star \mathscr{L}.$$

Where the second equalities hold since in one dimension $\star = i_{\partial_t}$. So $\alpha_X = \star\mathscr{L}$ and

$$i_X \gamma = \langle \star \mathrm{d}x, \star \mathrm{d}x \rangle$$

and

$$J_X = i_X \gamma - \alpha_X = \langle \star \mathrm{d}x, \star \mathrm{d}x \rangle - \star \mathscr{L} \qquad \in \Omega^{0,0}_{loc} \cong \Omega^0(\mathscr{F})$$

so it's a function on the phase space, an observable. In particular we have

$$J_X = |\dot{x}|^2 - \mathscr{L}(x, \dot{x})$$

which is exactly the Hamiltonian! So we found what we expected, that the conserved charge due to the symmetry of time translations is given by the Hamiltonian. If we had a potential in the Lagrangian, it wouldn't affect our derivation.

There's one thing left to check: the Hamiltonian (or more generally the conserved charge) should generate time translations (or its corresponsing symmetry)[9]. To check this we compute

$$L_X \gamma = \delta i_X \gamma + i_X \delta \gamma = \delta i_X \gamma + i_X \omega \,.$$

[9] in the sense of Equation (9) or (10): the symmetry should be given by a Hamiltonian vector field, whose corresponding function is the conserved charge.

On shell, we have $\delta\mathscr{L} + d\gamma = 0$, so in particular $L_X(\delta\mathscr{L} + d\gamma) = 0$ so we have

$$\begin{aligned} 0 &= \delta(L_X\mathscr{L}) + \mathrm{d}(L_X\gamma) \\ &= \delta(\mathrm{d}\alpha_X) + \mathrm{d}(L_X\gamma) \\ &= \mathrm{d}(L_x\gamma - \delta\alpha_X) \\ &= \mathrm{d}(i_X\omega + \delta(i_X\gamma - \alpha_X)) \\ &= \mathrm{d}(i_X\omega + \delta(J_X)) \end{aligned}$$

Now, because $\Omega^{\bullet,\bullet}$ has no cohomology in this degree we have

$$i_X\omega + \delta(J_X) = d\beta_X \qquad \in \Omega^{1,d-1}_{loc}$$

for some $\beta_X \in \Omega^{1,d-2}$. If we then integrate over some $d-1$ dimensional manifold N we find

$$i_X \int_N \omega + \delta \int_N J_X = 0$$

Here $\int_N \omega$ is the symplectic form on $\mathscr{F}$ and $\int_N J_X$ is the Noether charge in $\Omega^0(\mathscr{F})$. In words, this equation says: Noether charges generate the corresponding symmetries on phase space!

Example 15. Consider for example the case of a particle moving on a circle S^1. We therefore have phase space T^*S^1 with symplectic form $\omega = d\theta \wedge dp_\theta$. Here we have symplectic vector fields

$$\frac{\partial}{\partial\theta}, \qquad \frac{\partial}{\partial p_\theta}$$

since

$$i_{\partial_\theta}\omega = dp_\theta, \qquad i_{\partial_{p_\theta}}\omega = -d\theta\,.$$

However, only one of them is a Hamiltonian vector field since dp_θ is exact (being the differential of the function p_θ), while $d\theta$ is not, since there's not a function θ.

End of lecture 12

9 What is a directional derivative?

Recall: we had a basic example of a derivative, given by a vector field acting on a function, since we *defined* vector fields to be derivations of the algebra of smooth functions on X:

$$\mathrm{Vect}(X) = \mathrm{Der}(C^\infty(X)) = \mathrm{Der}(\Omega^0(X))\,.$$

The way this worked was that the directional derivative of a function along a vector field is obtained by taking its differential and then contracting the one form that we get with the vector field:

$$\begin{array}{ccc} & \overset{d}{\nearrow} & \Omega^1(X) \\ \Omega^0(X) & & \Big\downarrow i_X \\ & \underset{\mathscr{L}_X}{\searrow} & \Omega^0(X)\,. \end{array} \tag{13}$$

It had the following properties:

1. it satisfies the Leibniz rule (it's "infinitesimal").
2. It is "natural", i.e. it just depends on the manifold structure (not on some additional structure the manifold *may* have) and is therefore well defined on any arbitrary manifold. It can also be generalized to Lie derivatives for forms, vector fields, tensors / "natural objects".
3. It is linear over Ω^0 (it's "directional").

We'd like for the derivative to be something of the form

$$\text{“}\lim_{p'\to p}\frac{f(p')-f(p)}{p'-p}\text{”} \tag{14}$$

but, in a generic manifold, the difference of points in the denominator doesn't make sense. The Lie derivative allows to take the directional derivatives and so solves the problem of the denominator. However, the directional derivative just depends on the value of the vector field at the point where at which we take the derivative. So in particular if we rescale the vector field by some function, which at a point is just a number, we rescale the derivative by that same number, i.e.

$$\mathscr{L}_{fX}(g) = f\mathscr{L}_X(g)\,.$$

This is what we meant by linear over Ω^0 above. This can be seen by simply using Cartan's formula:

$$\mathscr{L}_{fX}(g) = i_{fX}dg = fi_X dg = f\mathscr{L}_X(g)\,.$$

Fact: This is the only case where all 3 properties simultaneously work! In fact consider some higher form $\alpha \in \Omega^k(X)$, then

$$\begin{aligned}\mathscr{L}_{fX}(\alpha) &= di_{fX}\alpha + i_{fX}d\alpha = d(fi_X\alpha) + fi_X d\alpha = df\wedge i_X\alpha + fdi_X\alpha + fi_X d\alpha \\ &= df\wedge i_X\alpha + f\mathscr{L}_X\alpha\end{aligned}$$

so $\mathscr{L}_X$ is not "directional" as above! This is instead a "numerator issue" for equation (14): we want to take the difference of forms at different points which live in different spaces $\Omega^k_{p'}(X)$ and $\Omega^k_p(X)$. We then get that α is identified at nearby points by acting on each tensor index with the Jacobian matrix. So the Lie derivative of forms is not directional, but it's for a good reason.

We could instead give up on naturality and try to define some operation which satisfies the Leibniz rule and is linear over Ω^0. We then require some additional data, that of a connection[10]. This will allow us to construct a notion of directional derivative on sections of a vector bundle $\begin{matrix} E \\ \downarrow \\ X \end{matrix}$ which will not be natural: the *connection* is an additional datum.

We start by working in the setting of vector bundles, in which case the additional datum takes the name of "affine" or "Koszul" connection.

Definition 9.1. An affine or *Koszul connection* on a vector bundle $\begin{matrix} E \\ \downarrow \\ X \end{matrix}$ is a map

$$\begin{aligned} \nabla : \text{Vect} \times \mathscr{E} &\to \mathscr{E} \\ (x, s) &\mapsto \nabla_X(s) \end{aligned}$$

where $\mathscr{E} = \Gamma(E)$ (and of course Vect $= \Gamma(TX)$). This map satisfies:

1. $\nabla_X(fs) = X(f)s + f\nabla_X(s)$ (Leibniz);
3. $\nabla_{fX}(s) = f\nabla_X(s)$ (directional).

Since ∇ is Ω^0 linear in X, we can dualize it to get a map

$$\nabla : \mathscr{E} \to \mathscr{E} \otimes_{\Omega^0} \Omega^1 ,$$

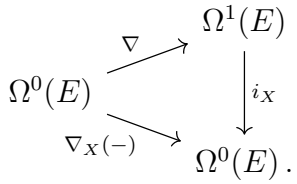
or equivalently

$$\nabla : \Omega^0(E) \to \Omega^1(E) .$$

So in a sense the connection replaces the de Rham operator for differential forms valued in E and allows us to define a directional derivative analogously to the Lie

[10] This is sometimes called a "Covariant Derivative" although that's not the best name because by covariant we usually mean that it doesn't depend on the coordinates used, but everything we've worked with doesn't depend on coordinates. In addition it may suggest that it doesn't depend on any additional structure on the manifold, but that's not true.

derivative in (13):



Now that we've introduced the datum of a connection, which is a choice, we may wonder what the possible choices are. In other words, how big is the space of possible connections? The following claim allows us to answer this question.

Claim 9.1. *Let ∇ and ∇' be two connections. Then $(\nabla - \nabla') : \mathrm{Vect} \times \mathscr{E} \to \mathscr{E}$ is a bundle map, i.e. $(\nabla - \nabla')$ is actually a map $TX \times E \to E$.*

Note that connections, individually, are *not* bundle maps, as they are not linear over functions (the same is true for the de Rham differential).

Proof.

$$(\nabla_X - \nabla'_X)(fs) = X(f)s + f\nabla_X(s) - X(f)s - f\nabla'_X(s) = f(\nabla_X - \nabla'_X)(s)$$

□

We then have

$$\nabla - \nabla' \in \mathrm{Hom}(T \otimes E \to E) \cong \mathrm{Hom}(T \to E^* \otimes E) \cong \Omega^1 \otimes \mathrm{End}(E)$$

If E has rank n, then $\mathrm{End}(E)$ is isomorphic to $\mathfrak{gl}(n)$ and then (by picking an isomorphism) we can think of $\nabla - \nabla'$ as a $\mathfrak{gl}(n)$ valued one form. This may be familiar from gauge theory! There one usually says that a gauge field is not really a Lie algebra valued one form, but if we choose a reference gauge field and take the difference of our gauge field and the reference one, then we get a Lie algebra valued one form. This is exactly what we've found. We then have that the space of connections in E is an *affine space*[11] over $\Omega^1(\mathrm{End}(E))$.

Here's a useful reformulation of this data.

Definition 9.2. Let E be a vector bundle on X, with sheaf of section $\mathscr{E}$ (so $\mathscr{E}$ is a locally free sheaf of Ω^0 modules). The *first jet bundle* $J^1(E)$ of E is defined as follows. As a vector space $\mathscr{J}^1(E) = \Omega^0(E) \oplus \Omega^1(E)$. As an Ω^0 module

$$f \cdot (s, \beta) = (f \cdot s, f \cdot \beta + df \otimes s) \tag{15}$$

where $f \in \Omega^0$ and $(s, \beta) \in \Omega^0(E) \oplus \Omega^1(E)$.

[11] a vector space without a fixed origin

The last term in the Ω^0 module structure is interesting. It tells us that a section of $J^1(E)$ should look something like the zeroth and first Taylor coefficient of an E valued function. Note that if $s = 0$ nothing interesting happens: a function acting on $(0, \beta)$ gives back something of the same form. We thus have $\Omega^1(E)$ sitting inside $J^1(E)$ and the quotient is $\Omega^0(E)$, i.e. we have the short exact sequence

$$0 \longrightarrow \Omega^1(E) \longrightarrow J^1(E) \longrightarrow \Omega^0(E) \longrightarrow 0\,.$$

This allows us to give another definition of a connection:

Definition 9.3. A connection in E is a splitting of the short exact sequence above.

Note that there's no obstruction to having a splitting in the smooth category, so connections always exist! This definition also gives you an obstruction to having a connection in a more general case. A splitting is a map $\Omega^0(E) \to J^1(E)$ such that its composition with the map $J^1(E) \to \Omega^0(E)$ is the identity on $\Omega^0(E)$. In particular, if we have a connection ∇ we can take the splitting

$$s \mapsto (s, \nabla s)$$

that is Ω^0 linear

$$\begin{array}{ccccc} s & \longmapsto & (s, \nabla s) & & \\ \downarrow & & & \searrow & \\ fs & \longmapsto & (fs, \nabla(fs)) & = & (fs, f\nabla s + \mathrm{d}f \otimes s) \end{array}$$

where equality is true exactly when ∇ satisfies the Leibniz rule. So this definition is equivalent to the previous one.

There's one thing we should check: is $\mathscr{J}^1(E)$ with the Ω^0 module structure in Equation (15) actually an Ω^0 module? i.e.

$$g(f(s, \beta)) \stackrel{?}{=} (gf)(s, \beta)\,.$$

For the left hand side we find

$$g(f \cdot s, f \cdot \beta + \mathrm{d}f \otimes s) = (gf \cdot s, gf \cdot \beta + g\,\mathrm{d}f \otimes s + \mathrm{d}g \otimes f \cdot s)$$

while for the right hand side

$$(gf \cdot s, gf \cdot \beta + \mathrm{d}(gf) \otimes s)$$

which are the same because d satisfies the Leibniz rule.

We now want to consideral a more general case, e.g. general fiber bundles.

Definition 9.4. A fiber bundle with fiber F over X is a space B with a map $\pi : B \to X$ with local identifications

$$\begin{array}{ccc} B \supseteq \pi^{-1}(U) & \xrightarrow{\sim} & U \times F \\ & {\scriptstyle\pi}\searrow \quad \swarrow{\scriptstyle \text{forget } F} & \\ & U & \end{array}$$

So it's simply a space that locally looks like a product of my space with some other space (which we *don't* require to be a vector space). So locally B looks like a collection of Fs parametrized by X. We usually represent such bundles as $\begin{array}{c} B \\ \downarrow{\scriptstyle\pi} \\ X \end{array}$ just like we did for vector bundles.

Let $V \subset TB$ be the vertical subbundle, i.e. the bundle of vectors that annihilate functions that are constant along the fibers. We then have the short exact sequence

$$0 \longrightarrow V \longrightarrow TB \xrightarrow{D\pi} \pi^*TX \longrightarrow 0$$

of vector bundles on B[12].

Inspired by (9.3) we can give the following definition.

Definition 9.5. An Ehresmann connection is a splitting of this sequence, equivalently choice of horizontal subbundle (a complementary bundle to V in TB).

End of lecture 13

10 Connections, curvature and compatibility

Last time we focused on the notion of a Koszul connection. We were considering a vector bundle $\begin{array}{c} E \\ \downarrow \\ X \end{array}$ and we where thinking of it in terms of its sheaf of sections $\mathscr{E} = \Gamma(E) = \Omega^0(E)$, which is a locally free sheaf of modules over functions on

[12]note that $D\pi$ may represent either the map $TB \to TX$ or the map $TB \to \pi^*TX$. Here we consider the second one since we want a sequence of bundles over B

X. Notationally, more in general, we have $\Omega^k(E) = \Gamma(\bigwedge^k T^*X \otimes E) = \Omega^k \otimes_{\Omega^0} \mathscr{E}$, where tensoring over Ω^0 means that

$$f\omega \otimes s = \omega \otimes fs, \qquad \omega \in \Omega^k, s \in \mathscr{E}, f \in \Omega^0 .$$

The connection then gave us a map $\nabla : \Omega^0(E) \to \Omega^1(E)$ which gave us a directional derivative (along some vector field $X \in \mathrm{Vect}$) by the the bottom line of the diagram

$$\begin{array}{ccc} & & \Omega^1(E) \\ & \overset{\nabla}{\nearrow} & \big\downarrow i_X \\ \Omega^0(E) & \xrightarrow[\nabla_X]{} & \Omega^0(E) . \end{array}$$

Additionally, ∇ satisfied the Leibniz rule:

$$\nabla(fs) = \mathrm{d}f \otimes s + f \otimes \nabla s \qquad (\text{in } \Omega^1(E)) .$$

Recall that for differential forms we started with $\mathrm{d} : \Omega^0 \to \Omega^1$ and we extended it uniquely to the complex of differential forms $\Omega^\bullet$ by requiring it to be a graded derivation, i.e. satisfying (7). Analogously, we can extend ∇ to an operator of degree $+1$ on $\Omega^\bullet(E)$, satisfying

$$\nabla(\omega \otimes s) = d\omega \otimes s + (-)^{|\omega|}\omega \wedge \nabla s$$

where $\mathrm{d}\omega \otimes s \in \Omega^{k+1}(E)$ and since $\nabla s \in \Omega^1(E)$ we have that also $\omega \wedge \nabla s \in \Omega^{k+1}(E)$. To think of the wedge in the last term, view ∇s as $\alpha \otimes t$ for $\alpha \in \Omega^1$ and $t \in \mathscr{E}$, and the wedge product above is defined as

$$\omega \wedge (\alpha \otimes t) := (\omega \wedge \alpha) \otimes t .$$

Note: the vector spaces $\Omega^k(E)$ are not an algebra since I have no multiplication on E. They are a module over $\Omega^\bullet$.

So we've found what looks like a generalization of the de Rham operator to forms valued in some vector bundle, but what is similar and what is different? Well, the (second) key property of the de Rham operator was $\mathrm{d}^2 = 0$[13], is that true now?

Instead of locally constant, we would now consider "covariantly constant", i.e. being in the kernel of ∇. So are we doing the same thing as in forms? Did we just find a "good replacement" for the sheaf of covariantly constant functions valued in E? No, this fails for ∇ for an interesting reason: *curvature*.

[13]That was important because $\Omega^0 \xrightarrow{\mathrm{d}} \Omega^1 \xrightarrow{\mathrm{d}} \Omega^2 \xrightarrow{\mathrm{d}} \dots$ was our "good replacement" for the sheaf of locally constant function on X.

Claim 10.1. *Let $s \in \Omega^0(E)$ be a section, so $\nabla^2(s) \in \Omega^2(E)$. The map $s \mapsto \nabla^2 s$ is linear over functions.*

This is in contrast with the map $s \mapsto \nabla s$, which is *not* linear over functions as it obeys Leibniz rule. How can that be?

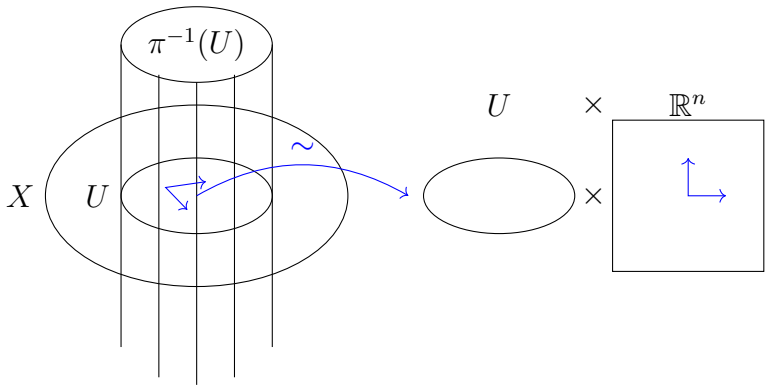
Proof. Let's check:

$$\nabla^2(f \cdot s) = \nabla(\mathrm{d}f \otimes s + f\nabla s) = \mathrm{d}^2 f \otimes s - \mathrm{d}f \otimes \nabla s + \mathrm{d}f \otimes \nabla s + f\nabla^2 s = f\nabla^2 s.$$

The firs term is zero, since $\mathrm{d}^2 f = 0$ and the middle two terms cancel each other. □

We thus have a vector bundle map $E \to \bigwedge^2 T^*X \otimes E$, i.e. a section of $\bigwedge^2 T^*X \otimes \mathrm{End}(E)$, i.e. an element of $\Omega^2(\mathrm{End}(E))$. This is a matrix valued 2-form $F_\nabla \in \Omega^2(\mathrm{End}(E))$ and is called the *curvature* of the connection. This should make some intuitive sense: we plug in an element of E and the antisymmetric combination of two vector fields and we get out another element of E, measuring the failure of ∇^2 to be zero.

Aside: recall that given two connections ∇ and ∇', we showed last time that $\nabla - \nabla' \in \Omega^1(\mathrm{End}(E))$. If we are working in a locally trivial neighbourhood, we have $\nabla =$ "d" $+ A$, where $A \in \Omega^1(\mathrm{End}(E))$. By "d", we mean that we use a local trivialization to write $s = \sum_i s_i \cdot e_i$, $s \in \Gamma(E, U)$ and then "d"$s = \sum_i ds_i e_i$. So we're taking basis vector fields from the trivialization to write a section as a collection of functions times those basis vectors, pictorially:



In other words, we're using a local trivialization to define a local connection. We then have

$$\nabla^2 = (\mathrm{d}+A)^2 = \mathrm{d}^2 + \mathrm{d}A + A\mathrm{d} + A\wedge A = 0 + (\mathrm{d}A) - A\mathrm{d} + A\mathrm{d} + A\wedge A = \mathrm{d}A + A\wedge A$$

so

$$F_\nabla = \mathrm{d}A + A \wedge A\,,$$

which is the formula one typically finds.

10.1 The Fundamental Theorem of Riemannian Geometry

One last thing that one can do cleanly by thinking like this is proving the fundamental theorem of Riemannian geometry. To do this, let X be a smooth manifold, ∇ a connection in TX, g a *metric*[14] on X. Via the Leibniz rule for tensor products, ∇ determines a connection in every tensor bundle:

$$\nabla(a \otimes b) = \nabla(a) \otimes b + a \otimes (\nabla b)$$

where a and b are tensor on which I already know how the connection acts. In particular ∇g is meaningful and allows to give the following definition.

Definition 10.1. A connection ∇ is called a *Levi-Civita connection* if:

1. $\nabla g = 0$, "metric compatible";
2. $\nabla_X Y - \nabla_Y X = [X, Y]$, "torsion free".

We can now state the theorem.

Theorem 10.2 (Fundamental theorem of Riemannian geometry)**.** *There exists a unique Levi-Civita connection on* (X, g).

Like other proofs we've seen, this relies on the fact that something built out of ∇ which in principle is a differential operator, is instead linear over functions.

Proof. Let X, Y, Z be vector fields and consider $X(g(Y, Z))$ which makes sense since $g(Y, Z)$ is a function, which may be acted on by a vector field X. We have

$$X(g(Y,Z)) = (\nabla_X g)(Y,Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

where we used the Leibniz rule and metric compatibility. Now, consider the combination of three terms

$$\begin{aligned} X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) = \quad & g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ & + g(\nabla_Y X, Z) + g(X, \nabla_Y Z) \\ & - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \end{aligned}$$

[14] a metric on X is a nondegenerate family of inner products at every point of X. In other words, it takes in two vector fields and gives a function, the inner product of the vector fields at every point. Therefore it is a nondegenerate section $g \in \mathrm{Sym}^2 T^*X$. It therefore gives us an isomorphism $\#_g : TX \to T^*X$ and the symmetry property reads $\#_g^\vee = +\#_g$.

where we can recognize $\nabla_X Z - \nabla_Z X$ and $\nabla_Y Z - \nabla_Z Y$ and use torsion freeness. Analogously,

$$\nabla_X Y + \nabla_Y X = 2\nabla_X Y - (\nabla_X Y - \nabla_Y X) = 2\nabla_X Y - [X, Y]$$

so we have

$$2g(\nabla_X Y, Z) + g(Y, [X, Z]) + g(X, [Y, Z]) - g([X, Y], Z)$$

where ∇ appears only once, so we can think of using this expression to *define* ∇. We thus have Koszul's formula:

$$\begin{aligned} g(\nabla_X Y, Z) = \frac{1}{2}\Big(& X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) \\ & - g(Y, [X, Z]) - g(X, [Y, Z]) + g([X, Y], Z)\Big) \end{aligned}$$

We can now check that the right hand side is linear over functions in Z: the error terms are

$$X(f)g(Y, Z) + Y(f)g(X, Z) - g(X, Z)Y(f) - g(Y, Z)X(f)$$

which simply cancel pairwise. Therefore $g(\nabla_X Y, Z)$ is some linear expression in Z and it *defines* a Levi-Civita connection starting from g. We've thus found the existence and uniqueness of the Levi-Civita connection. □

This can be generalized to show that there's a unique metric compatible connection with some fixed torsion. In particular we showed the case where the torsion is zero.

End of lecture 14

Index

A
abelian group 25
affine space 57

C
Cartan's formula 34
Čech cohomology 18
commutative algebra 3
commutative ring 2
connection 56
curvature 61

D
de Rham derivative 33
differential form 30
Dirac's axioms 39

E
exact sequence 14

F
first jet bundle 57

G
graded derivation 33

H
Hausdorff property 5

I
ideal 6
inner product 35

K
k-jet 48
Koszul connection 56
Kähler differential 27
Kähler manifold 42

L
Levi-Civita connection 62
Lie algebra 35
Lie derivative 36
line bundle 43

M
metric 62
module 25

P
Poisson algebra 37
Poisson tensor 37
polarization 41
prequantizability 41
prequantization 39
presheaf 19
presymplectic form 45

Q
quantization 39
quotient 5

R
rank 5

S
sheaf 19
    acyclic 23
state 4
symplectic form 39
symplectic geometry 39
symplectic structure 44

T
topology 5
    discrete 11
    indiscrete 11
    standard 11
    Zariski 8