All lectures are 1 hour, and all problem sessions too.

1 Lecture 1: Overview, topology, sheaves, examples

1.1 Introduction

We want to show in the workshop how some basic modern tools from topology, such as sheaves and cohomology, shed new light on an old theorem of Gauss in number theory: in how many ways can an integer be written as the sum of three squares? We are also very enthousiastic about the Stacks Project, and so we want to make people familiar with it. It seems very useful to us, even for the general theory of sheaves on topological spaces; users should try not to be intimidated by it, there is a lot of elementary material and it is an excellent place to learn a lot more.

Here is our plan for the 4 lectures:

1. Topological spaces, continuous maps,
2. Sheaves of groups acting on sheaves of sets, torsors,
3. The case of a transitive action,
4. Application to Gauss's theorem on 3 squares.

1.2 Topological spaces, continuous maps

Topological spaces were probably invented for catching the essence of the notion of continuity of maps between metric spaces: no more ϵ and δ, just "inverse image of open is open". A subset U of a metric space (X, d) is called *open* if for all x in X there is a δ in $\mathbb{R}_{>0}$ such that $B(x, \delta) \subset U$, where $B(x, \delta) = \{y \in X : d(x, y) < \delta\}$ is the open ball (in X) of radius δ around x.

1.2.1 Definition. A topological space is a pair $(S, \mathrm{Open}(S))$ with S a set and $\mathrm{Open}(S)$ a set of subsets of S, called the open subsets of S, such that:

1. $\emptyset$ and S are elements of $\mathrm{Open}(S)$,
2. if I is a set, and, for every i in I, U_i is in $\mathrm{Open}(S)$, then the union $\bigcup_{i\in I} U_i$ is in $\mathrm{Open}(S)$,
3. if U and V are in $\mathrm{Open}(S)$, then so is $U \cap V$.

1.2.2 Definition. Let $(S, \mathrm{Open}(S))$ be a topological space. A subset Z of S is *closed* if its complement $S - Z$ is open.

1.2.3 Remark. Note that this does not mean that a subset is closed if it is not open. For example, the interval $[0, 1)$ in $\mathbb{R}$ (with the usual topology) is not open and also not closed.

1.2.4 Example.

1. Each metric space (X, d) has a topology defined by d. Note that for metric spaces (X, d) and (Y, d'), a map $f\colon X \to Y$ is continuous if and only if for all $U \subset Y$ open, $f^{-1}U$ is open in X.
2. There are also natural non-Hausdorff spaces, occurring in algebraic geometry. The Zariski topology on $\mathbb{C}^n$ has as *closed* sets the $Z(T)$, where $T \subset \mathbb{C}[x_1, \ldots, x_n]$ is a set of polynomials and $Z(T) = \{a \in \mathbb{C}^n : \text{for all } f \in T,\ f(a) = 0\}$. It is a non-trivial exercise that this indeed satisfies the axioms for a topological space. For $n > 0$, $\mathbb{C}^n$ with this topology is not Hausdorff. For $n = 1$ this is easy to see: each non-empty open set is $\mathbb{C}$ minus a finite set, hence two non-empty open sets are not disjoint.

3. Let A be a ring. Then we let $\mathrm{Spec}(A)$ be the set of prime ideals of A, that is, the ideals $x \subset A$ such that A/x is an integral domain. For x in $\mathrm{Spec}(A)$, we have the quotient A/x which is an integral domain, and we let $\kappa(x)$ be the field of fractions of A/x. For f in A, and x in $\mathrm{Spec}(A)$, we call the image of f in $\kappa(x)$ the *value* $f(x)$ of f at x, and so f gives a function on $\mathrm{Spec}(A)$ with values in fields, but the field depends on the point where one takes the value. Then we can define the *Zariski topology* on $\mathrm{Spec}(A)$: the closed subsets are the sets of the form $Z(T)$, for T a subset of A, and

$$Z(T) = \{x \in \mathrm{Spec}(A) : \forall f \in T, f(x) = 0 \text{ in } \kappa(x)\}.$$

We have to prove that this satisfies the axioms for closed subsets.... Make an exercise of this? One has to show that $Z(T_1 \cdot T_2) = Z(T_1) \cup Z(T_2)$, here $T_1 \cdot T_2$ is the set $\{f_1 f_2 : f_1 \in T_1,\ f_2 \in T_2\}$. One inclusion is clear. So now assume that x is in $Z(T_1 \cdot T_2)$ and that x is not in $Z(T_1)$. Then there is an f_1 in T_1 such that $f_1(x) \neq 0$. Now let f_2 be in T_2. As $f_1 f_2$ is in $T_1 \cdot T_2$, we have $(f_1 f_2)x = 0$. But $(f_1 f_2)x = (f_1 x)(f_2 x)$ in $\kappa(x)$, and $f_1 x \neq 0$, hence $f_2 x = 0$.

1.2.5 Definition. Let X and Y be topological spaces, and $f\colon X \to Y$ a map. Then f is *continuous* if for all open $U \subset Y$, $f^{-1}U$ is open in X.

This gives us the category Top of topological spaces. In particular, a morphism of topological spaces $f\colon X \to Y$ is an isomorphism if and only if there exists a morphism $g\colon Y \to X$ such that $gf = \mathrm{id}_X$ and $fg = \mathrm{id}_Y$. The condition on f to be an isomorphism is of course equivalent to f being bijective and f^{-1} being continuous.

1.3 Presheaves

1.3.1 Definition. Let S be a topological space. A presheaf of sets $\mathcal{F}$ on S consists of:

- for each U in $\mathrm{Open}(S)$, a set $\mathcal{F}(U)$,
- for each inclusion $i_{V,U} : V \to U$ with V and U in $\mathrm{Open}(S)$, a map $\mathcal{F}(i_{V,U})\colon \mathcal{F}(U) \to \mathcal{F}(V)$,

such that:

1. for each U in $\mathrm{Open}(S)$, $\mathcal{F}(i_{U,U}) = \mathrm{id}_{\mathcal{F}(U)}$,
2. for all inclusions $W \subset V \subset U$ with U, V, W in $\mathrm{Open}(S)$, $\mathcal{F}(i_{W,U}) = \mathcal{F}(i_{W,V}) \circ \mathcal{F}(i_{V,U})$.

1.3.2 Remark. Often one writes the maps $\mathcal{F}(i_{V,U})\colon \mathcal{F}(U) \to \mathcal{F}(V)$ as $s \mapsto s|_V$, and calls them "restriction maps".

1.3.3 Remark. Category theory provides a very efficient way to summarise the definition of a presheaf: it is a contravariant functor $\mathrm{Open}(S) \to \mathrm{Sets}$, where $\mathrm{Open}(S)$ is then the category whose objects are the open subsets of S and whose morphisms are the inclusion maps, and where Sets is the category of sets. For those who want to be precise and correct and express everything in terms of set theory, let us make precise that we decide that all our categories are "small", that is, their collection of objects is a set. Actually, they are all subsets of one fixed big enough set, called a "universe". According to Albert, a previous visitor, Klaus Denecke, has talked a lot about categories in a workshop in Yogya in 2009.

1.3.4 Example. Let X be a topological space, then the presheaf $C^0_{X,\mathbb{R}}$ of continuous real functions on X is defined as $C^0_{X,\mathbb{R}}(U) = \{f\colon U \to \mathbb{R} : f \text{ is continuous}\}$, with, for $V \subset U$ an inclusion of open sets, and for $f \in C^0_{X,\mathbb{R}}(U)$, $f|_V$ the restriction of f to V.

Similarly, for X a smooth real manifold, we have the sheaf $C^\infty_{X,\mathbb{R}}$ of smooth real functions given by $C^\infty_{X,\mathbb{R}}(U) = \{f\colon U \to \mathbb{R} : f \text{ is smooth}\}$, with the usual restriction maps.

We could also consider a complex analytic manifold and define its sheaf of complex analytic functions.

1.3.5 Example. Let X be a topological space and let A be a set. Then the constant presheaf on X with values in A is $U \mapsto A$ for all U, and with all restriction maps id_A.

1.3.6 Definition. Let X be a topological space, $\mathcal{F}$ and $\mathcal{G}$ be presheaves on X. A morphism of presheaves $\phi\colon \mathcal{F} \to \mathcal{G}$ consists of maps $\phi(U)\colon \mathcal{F}(U) \to \mathcal{G}(U)$, for all open $U \subset X$, such that for all inclusions $V \subset U$, we have $\mathcal{G}(i_{V,U}) \circ \phi(U) = \phi(V) \circ \mathcal{F}(i_{V,U})$, that is, the diagram

$$\begin{array}{ccc} \mathcal{F}(U) & \xrightarrow{\phi(U)} & \mathcal{G}(U) \\ {\scriptstyle \mathcal{F}(i_{V,U})}\downarrow & & \downarrow{\scriptstyle \mathcal{G}(i_{V,U})} \\ \mathcal{F}(V) & \xrightarrow{\phi(V)} & \mathcal{G}(V) \end{array}$$

is commutative. In yet other words, morphisms of presheaves are the same as morphisms of functors.

A set-theoretic description is: a map from $\mathrm{Open}(X)$ to the union, over all U, of the $\mathrm{Hom}(\mathcal{F}(U), \mathcal{G}(U))$, such that the necessary conditions hold.

1.3.7 Remark. Similarly, we can define presheaves of of groups, of rings, and so on.

1.4 Sheaves

Sheaves are presheaves with the property that their sets of sections are "determined locally". The following definition makes this precise.

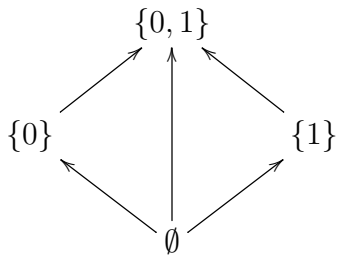
1.4.1 Definition. Let X be a topological space, and $\mathcal{F}$ a presheaf of sets on X. Then $\mathcal{F}$ is a sheaf of sets if for all open subsets $U \subset X$ and all open covers $(U_i)_{i\in I}$ (with I any set) of U, and for all collections $(s_i \in \mathcal{F}(U_i))_{i\in I}$ such that for all i and j in I, $s_i|_{U_i\cap U_j} = s_j|_{U_i\cap U_j}$, there is a unique $s \in \mathcal{F}(U)$ such that for all $i \in I$, $s_i = s|_{U_i}$.

For $\mathcal{F}$ and $\mathcal{G}$ sheaves of sets on X, a morphism from $\mathcal{F}$ to $\mathcal{G}$ is a morphism from $\mathcal{F}$ to $\mathcal{G}$ as presheaves.

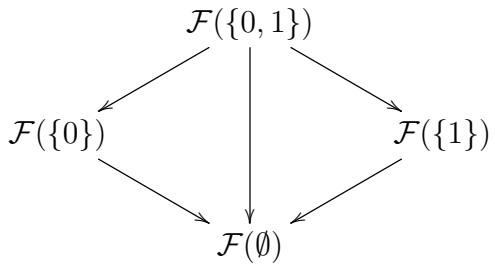
Similarly, one has sheaves of groups, rings etc.

1.4.2 Example. For X a smooth manifold, the presheaf $C^\infty_{X,\mathbb{R}}$ is a sheaf, because the condition for a function $f\colon U \to \mathbb{R}$ to be smooth is a *local* condition.

1.4.3 Example. Let $X = \{0, 1\}$ with the discrete topology. Then we have 4 open subsets: $\emptyset$, $\{0\}$, $\{1\}$, $\{0, 1\}$, with their inclusions. The diagram is a very familiar partially ordered set with sup and inf:



It is then clear what a presheaf of sets on X is: sets $\mathcal{F}(\{0,1\})$, $\mathcal{F}(\{0\})$, $\mathcal{F}(\{1\})$, $\mathcal{F}(\emptyset)$, with maps "in the other direction":



Now what is the condition that $\mathcal{F}$ is a sheaf? Here is the answer. First of all $\mathcal{F}(\emptyset)$ is a set with one element (we do not know which one point set that is, but it does not matter); the proof of this is an exercise in logic/set theory. Note that each one point set is a final object in Set, the category of sets. Considering the covers of $\{0\}$ leads to the conclusion that there is no condition on $\mathcal{F}(\{0\})$, and by symmetry, also no condition on $\mathcal{F}(\{1\})$. But $\{0,1\}$ is covered by $U_0 = \{0\}$ and $U_1 = \{1\}$, and that gives that the condition that the map $\mathcal{F}(\{0,1\}) \to \mathcal{F}(\{0\}) \times \mathcal{F}(\{1\})$ is bijective.

To summarise: a presheaf of sets $\mathcal{F}$ on $\{0,1\}$ with the discrete topology is a sheaf if and only if ($\mathcal{F}(\emptyset)$ is a one point set, and $\mathcal{F}(\{0,1\}) \to \mathcal{F}(\{0\}) \times \mathcal{F}(\{1\})$ is bijective).

Note that we have now seen that there are presheaves that are not sheaves. In particular, for A with at least two elements, the constant presheaf with values in A on $\{0,1\}$ with the discrete topology is not a sheaf, for two reasons.

1.4.4 Example. Let X be a topological space, and let A be a set. Then we define the *constant sheaf* A_X with values in A on X by $A_X(U) = \{f \colon U \to A : f \text{ is locally constant}\}$, with the restriction maps. This presheaf is a sheaf.

1.4.5 Remark. The chapter "Sheaves on Spaces" of the Stacks Project is a highly recommended reference for all the material of this lecture. It has all the necessary details, and is very accessible. One important notion that we left out in this lecture is that of the stalk of a presheaf at a point. If we need it at some point, then it will come up naturally.

2 Problem Session 1

1. Is every discrete topological space a metric space?

2. Let $A = \mathbb{C}[t]$. Show that $\mathrm{Spec}(A)$ has one point that is dense, and that all other points are closed. Give a bijection between the set of closed points and $\mathbb{C}$.

3. Show that for $\mathcal{F}$ a sheaf of sets on a topological space X, $\mathcal{F}(\emptyset)$ is a set with exactly one element. This *is* an important exercise, because it is about how things are really defined, and so one *has* to understand it (Hartshorne would not agree with this, he *defines* $\mathcal{F}(\emptyset)$ to be what he wants it to be; and he does not even *define* sheaves of sets).

4. Let X be the topological space $\{0,1\}$ with open sets $\emptyset$, $\{0\}$ and $\{0,1\}$. Describe the sheaves of sets on X. (This is more interesting than the preceding exercise.)

5. Take a look at $\mathrm{Spec}(\mathbb{Z})$. As a set it is the set of prime ideals of the ring $\mathbb{Z}$, hence all maximal ideals ($p\mathbb{Z}$ for p a prime number) and the zero ideal. What are the residue fields $\kappa(x)$? Show that the closed subsets are: $\mathrm{Spec}(\mathbb{Z})$ itself, and the finite sets of maximal ideals.

6. (In case you have time.) Let A and B be rings, and $\phi\colon A \to B$ a morphism of rings. Show that for x in $\mathrm{Spec}(B)$, $\phi^{-1}x$, the inverse image of x in A, is a prime ideal of A, and that the map $\phi^{-1}\colon \mathrm{Spec}(B) \to \mathrm{Spec}(A)$, $x \mapsto \phi^{-1}x$ is continuous (hint: give a formula for $(\phi^{-1})^{-1}Z(a)$ for any $a \in A$).

3 Lecture 2: Sheaves of groups acting on sheaves of sets, torsors

3.1 Sheaves of groups

Recall that a presheaf of groups $\mathcal{G}$ on a topological space S consists of groups $\mathcal{G}(U)$, for all $U \subset S$ open, and for each inclusion $V \subset U$ of open subsets of S, a morphism of groups $\rho_{U,V}\colon \mathcal{G}(U) \to \mathcal{G}(V)$, such that for all U, $\rho_{U,U} = \mathrm{id}$, and for all $W \subset V \subset U$, $\rho_{U,W} = \rho_{V,W} \circ \rho_{U,V}$.

A sheaf of groups on S is a presheaf of groups that, as a presheaf of sets, is a sheaf.

Let us give an interesting example. For a ring A, and $n \in \mathbb{Z}_{\geq 0}$, we let $\mathrm{GL}_n(A)$ be the group of invertible n by n matrices with coefficients in A. It is the automorphism group of the free A-module A^n of rank n. For $\phi\colon A \to B$, we get a morphism of groups $\mathrm{GL}_n(A) \to \mathrm{GL}_n(B)$ by applying ϕ to all the coefficients. This makes GL_n into a functor from rings to groups.

3.1.1 Example. Let S be a smooth manifold. Then we have the sheaf of $\mathbb{R}$-algebras $C^\infty_{S,\mathbb{R}}$ on S. This gives us, for $n \in \mathbb{Z}_{\geq 0}$, a presheaf $U \mapsto \mathrm{GL}_n(C^\infty_{S,\mathbb{R}}(U))$, with the obvious restriction maps. This is a sheaf of groups. We have composed the contravariant functor $C^\infty_{S,\mathbb{R}}$ with the covariant functor GL_n. We will denote this sheaf of groups by $\mathrm{GL}_n(C^\infty_{S,\mathbb{R}})$

More generally, if S is a topological space and $\mathcal{O}$ is a sheaf of rings (rings are always assumed to be commutative) on S, then we get the sheaves of groups $\mathrm{GL}_n(\mathcal{O})$ on S.

In the case $n = 1$, we use the notation $A^\times = \mathrm{GL}_1(A)$ and $\mathcal{O}^\times = \mathrm{GL}_1(\mathcal{O})$.

3.2 Actions, quotients

For a group G and a set X, we know what left and right actions of G on X are. We generalise this to sheaves. (Indeed, it is a generalisation, because sheaves on $\{0\}$ are sets, and presheaves on $\emptyset$ are sets; isomorphisms of categories, if we want to be precise.)

3.2.1 Definition. Let S be a topological space, $\mathcal{G}$ a presheaf of groups on S, and $\mathcal{X}$ a presheaf of sets on S. An *action* of $\mathcal{G}$ on $\mathcal{X}$ consists of an action of the group $\mathcal{G}(U)$ on the set $\mathcal{X}(U)$, for all $U \subset S$ open, such that for all inclusions $V \subset U$, for all $g \in \mathcal{G}(U)$ and $x \in \mathcal{X}(U)$, $(gx)|_V = (g|_V)(x|_V)$.

Equivalently, an action of $\mathcal{G}$ on $\mathcal{X}$ is a morphism of presheaves $\mathcal{G} \times \mathcal{X} \to \mathcal{X}$ such that for each $U \subset S$ open, the map $\mathcal{G}(U) \times \mathcal{X}(U) = (\mathcal{G} \times \mathcal{X})(U) \to \mathcal{X}(U)$ is an action of $\mathcal{G}(U)$ on $\mathcal{X}(U)$. (But what are products of presheaves? Is a product of sheaves a sheaf?)

If $\mathcal{G}$ and $\mathcal{X}$ are sheaves, then an action of $\mathcal{G}$ on $\mathcal{X}$ is an action of presheaves.

What we have defined are left-actions. Of course, there are also right-actions.

We want to take the quotient of a sheaf of sets by the action of a sheaf of groups. Here, it makes a difference if we do this for presheaves, or for sheaves. First of all, we must make clear what we mean by a quotient. As usual, it is best to do this with a universal property.

For G a group acting from the right on a set X, the quotient map $q\colon X \to X/G$ has the property that for any set Y, and for any map $f\colon X \to Y$ such that for all x in X and all g in G one has $f(xg) = f(x)$, there is a unique map $\overline{f}\colon X/G \to Y$ such that $f = \overline{f} \circ q$.

To define the quotient for an action of sheaves or presheaves, we use this universal property, but then in the category of sheaves or presheaves.

3.2.2 Definition. Let S be a topological space, $\mathcal{X}$ a presheaf of sets on S with a right-action by a presheaf of groups $\mathcal{G}$ on S. A morphism of presheaves $q\colon \mathcal{X} \to \mathcal{Y}$ is called a *quotient* for the $\mathcal{G}$-action on $\mathcal{X}$ if for every morphism of presheaves $f\colon \mathcal{X} \to \mathcal{Z}$ such that for all $U \subset S$ open, all $g \in \mathcal{G}(U)$, all $x \in \mathcal{X}(U)$ we have $(fU)(xg) = (fU)(x)$, there is a unique morphism of presheaves $\overline{f}\colon \mathcal{Y} \to \mathcal{Z}$ such that $f = \overline{f} \circ q$.

Of course, *if* such a quotient exists, it is unique up to unique isomorphism, and therefore, in vague terms, "well-defined".

3.2.3 Proposition. *Let S be a topological space, $\mathcal{X}$ a presheaf of sets on S with a right-action by a presheaf of groups $\mathcal{G}$ on S. Then $U \mapsto \mathcal{X}(U)/\mathcal{G}(U)$, with the obvious restriction maps, is a quotient. Notation: $(\mathcal{X}/\mathcal{G})_p$ (with the subscript p of "presheaf").*

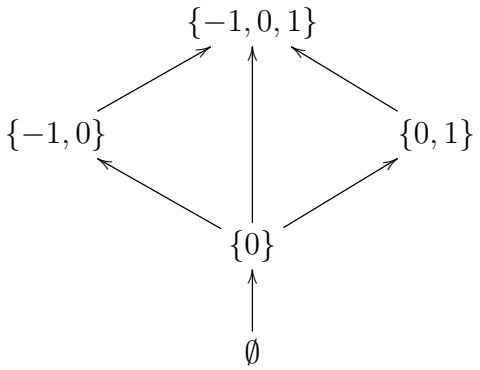
Proof. Straightforward. □

But in the category of sheaves the situation is more complicated.

3.2.4 Definition. Let S be a topological space, $\mathcal{X}$ a sheaf of sets on S with a right-action by a sheaf of groups $\mathcal{G}$ on S. A morphism of sheaves of sets $q\colon \mathcal{X} \to \mathcal{Y}$ is called a *quotient* for the $\mathcal{G}$-action on $\mathcal{X}$ if for every morphism of sheaves of sets $f\colon \mathcal{X} \to \mathcal{Z}$ such that for all $U \subset S$ open, all $g \in \mathcal{G}(U)$, all $x \in \mathcal{X}(U)$ we have $(fU)(xg) = (fU)(x)$, there is a unique morphism of sheaves $\overline{f}\colon \mathcal{Y} \to \mathcal{Z}$ such that $f = \overline{f} \circ q$.

Of course, we have the presheaf $(\mathcal{X}/\mathcal{G})_p$, and *if* it is a sheaf, then it has the universal property of the desired quotient. But the problem is that this presheaf is *not* always a sheaf. As this is very important, we give an example.

3.2.5 Example. Let $S = \{-1, 0, 1\}$ with opens $\emptyset$, $\{0\}$, $\{-1, 0\}$, $\{0, 1\}$ and $\{-1, 0, 1\}$. Here is the diagrams of open sets:



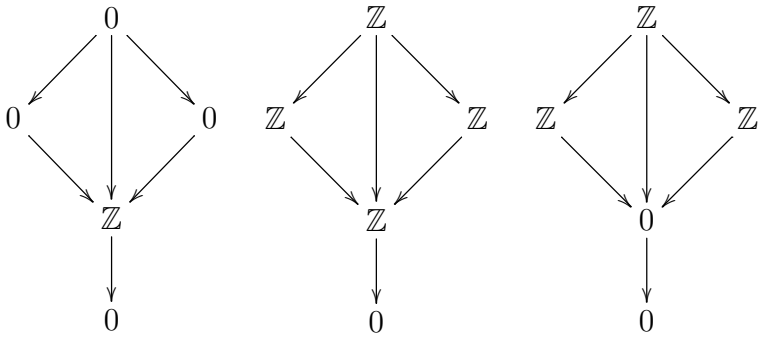
Then a presheaf $\mathcal{F}$ is a sheaf if and only if $\mathcal{F}(\emptyset)$ is a one point set and

$$\mathcal{F}(S) \to \{(a, b) \in \mathcal{F}(\{-1, 0\}) \times \mathcal{F}(\{0, 1\}) : a = b \text{ in } \mathcal{F}(\{0\})\}$$

is a bijection.

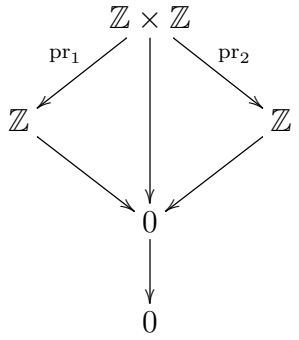
Let now $\mathcal{X}$ be the constant sheaf $\mathbb{Z}_S$; it is in fact a sheaf of groups. And we let $\mathcal{G}$ be the subsheaf of groups with $\mathcal{G}(S) = \{0\}$, $\mathcal{G}(\{-1, 0\}) = 0$, $\mathcal{G}(\{0, 1\}) = 0$ and $\mathcal{G}(\{0\}) = \mathbb{Z}$, and we

let $\mathcal{G}$ act on $\mathcal{X}$ "by translations". Here we "draw" $\mathcal{X}$, $\mathcal{G}$ and the presheaf quotient $(\mathcal{X}/\mathcal{G})_p$:



Then $\mathcal{X}$ and $\mathcal{G}$ are both sheaves of groups (addition), and $\mathcal{G}$ is a sheaf of subgroups $\mathcal{X}$. The presheaf quotient $(\mathcal{X}/\mathcal{G})_p$ is not a sheaf, because $(\mathcal{X}/\mathcal{G})_p(S) \to (\mathcal{X}/\mathcal{G})_p(\{-1,0\}) \times (\mathcal{X}/\mathcal{G})_p(\{0,1\})$ does not have the right image; we have the diagonal map $\mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ and it should be a bijection. In other words: not all compatible systems of local sections are given by a global section.

But what we can do is to replace $(\mathcal{X}/\mathcal{G})_p(S)$ by what it should be, and see if that sheaf

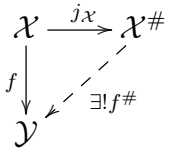


has the required property as in the definition. This works, one uses the sheaf property of $\mathcal{Y}$! (work it out in a diagram in the lecture, it is too much to type here).

There is standard notation for the sheaves in this example: $j_!\mathbb{Z}_U \hookrightarrow \mathbb{Z}_S \twoheadrightarrow i_*\mathbb{Z}_Z$, where U is the open subset $\{0\}$, $Z = \{-1, 1\}$ its complement, which is closed, and j and i the inclusion maps. The lower star i_* is the direct image which exists for arbitrary continuous maps, and the lower shriek $j_!$ is the extension by zero for the open immersion j.

There is a general procedure to make a sheaf from a presheaf.

3.2.6 Theorem. *Let S be a topological space, and let $\mathcal{X}$ be a presheaf of sets on S. Then there is a sheaf $\mathcal{X}^\#$ and a morphism of presheaves $j_\mathcal{X} \colon \mathcal{X} \to \mathcal{X}^\#$ such that for every morphism of presheaves $f \colon \mathcal{X} \to \mathcal{Y}$ with $\mathcal{Y}$ a sheaf, there is a unique $f^\# \colon \mathcal{X}^\# \to \mathcal{Y}$ such that $f = f^\# \circ j_\mathcal{F}$. In a diagram:*



Proof. See the section on sheafification in the chapter on sheaves on spaces in the Stacks Project. Their the usual construction using stalks at points is given. Another possibility is to do the operation suggested by our example, that is, replace, for every U, $\mathcal{X}(U)$ "by what it should be", *twice*, because there are two problems to solve: injectivity and surjectivity. This last procedure is what one does in the more general situation of presheaves on sites. □

3.2.7 Theorem. *Let S be a topological space, $\mathcal{X}$ a sheaf of sets on S with a right-action by a sheaf of groups $\mathcal{G}$ on S. Then $\mathcal{X} \to (\mathcal{X}/\mathcal{G})_p \to ((\mathcal{X}/\mathcal{G})_p)^\#$ is a quotient for the action by $\mathcal{G}$ on $\mathcal{X}$. Notation: $\mathcal{X}/\mathcal{G}$.*

Proof. Apply the previous theorem and the previous proposition. □

3.3 Torsors

These are the generalisation of non-empty sets with a free and transitive group action. Let G be a group and X a set with a G-action. For x in X, the stabiliser in G of x is the subset $G_x := \{g \in G : gx = x\}$ of elements that fix x; it is a subgroup of G. For x in X, the orbit of x under G is the set $G{\cdot}x := \{y \in X : \text{there exists } g \in G \text{ such that } y = gx\} = \{gx : g \in G\}$. The action of G on X is *free* if for all x in X we have $G_x = \{1\}$. The action is *transitive* if for all x and y in X there is a g in G such that $y = gx$.

A useful observation is that when X and Y non-empty right G-sets that are free and transitive, any G-equivariant map $f\colon X \to Y$ (meaning $f(xg) = (fx)g$) is bijective.

Non-empty sets with a free and transitive group action occur frequently, and are often used to "identify the set with the group". Think of affine geometry, for example: the line in $\mathbb{R}^2$ given by the equation $x + y = 1$ is acted upon freely and transitively by the sub $\mathbb{R}$-vectorspace of $\mathbb{R}^2$ given by the equation $x + y = 0$, via addition in $\mathbb{R}^2$. Choosing an element in the first line identifies it with the second one, but there is no natural choice.

We define the same properties in the context of sheaves. This is done by inserting "locally" at the right places. If one looks well at Example 3.2.5, one sees that this has to be done in the definition of "transitive".

3.3.1 Definition. Let S be a topological space, $\mathcal{G}$ a sheaf of groups, acting on a sheaf of sets $\mathcal{X}$.

- For x in $X(S)$, the stabiliser $\mathcal{G}_x$ of x in $\mathcal{G}$ is the sheaf of subgroups given by $\mathcal{G}_x(U) = \mathcal{G}(U)|_{x|_U}$ (it is indeed a sheaf).
- The action of $\mathcal{G}$ on $\mathcal{X}$ is *free* if for all $U \subset S$ open, $\mathcal{G}(U)$ acts freely on $\mathcal{X}(U)$.
- The action of $\mathcal{G}$ on $\mathcal{X}$ is *transitive* if for $U \subset S$ open, for all x and y in $\mathcal{X}(U)$, there exists an open cover $(U_i)_{i\in I}$ of U, and $(g_i \in \mathcal{G}(U_i))_{i\in I}$, such that for all $i \in I$, $g_i{\cdot}x|_{U_i} = y|_{U_i}$.

3.3.2 Lemma. *Let S be a topological space, $\mathcal{G}$ a sheaf of groups, and $\mathcal{X}$ and $\mathcal{Y}$ right G-torsors. Then every morphism $f\colon \mathcal{X} \to \mathcal{Y}$ of $\mathcal{G}$-torsors is an isomorphism.*

Proof. This will be an exercise. □

3.3.3 Definition. Let S be a topological space, $\mathcal{G}$ a sheaf of groups acting from the right on a sheaf of sets $\mathcal{X}$. Then $\mathcal{X}$ is called a *right-$\mathcal{G}$-torsor* if the following conditions hold:

1. the action of $\mathcal{G}$ on $\mathcal{X}$ is free and transitive,
2. locally $\mathcal{X}$ has sections: there is an open cover $(U_i)_{i\in I}$ of S, such that for each $i \in I$, $\mathcal{X}(U_i) \neq \emptyset$.

Let us give a very useful example of how torsors can arise: vector bundles on manifolds. Let S be a smooth manifold, and $p\colon E \to S$ (with the necessary additional data) a vector bundle, of rank n, say. Then, locally on S, E is isomorphic to the trivial vector bundle $\mathrm{pr}\colon \mathbb{R}^n \times S \to S$. This suggests to define a sheaf $\mathbf{Isom}_S(\mathbb{R}^n \times S, E)$, sending U to the set of isomorphisms of vector bundles on U from $\mathbb{R}^n \times U$ to $E|_U$, $\mathrm{Isom}_U(\mathbb{R}^n \times U, E|_U)$. There are natural restriction

maps, and in fact this is a sheaf. We also have the sheaf of groups $\mathbf{Aut}_S(\mathbb{R}^n \times S)$, which is the same as $\mathrm{GL}_n(C^\infty_{S,\mathbb{R}})$ which we have already seen. By construction, $\mathrm{GL}_n(C^\infty_{S,\mathbb{R}})$ acts on the right on $\mathbf{Isom}_S(\mathbb{R}^n \times S, E)$, by composition, and it makes $\mathbf{Isom}_S(\mathbb{R}^n \times S, E)$ into a right-$\mathrm{GL}_n(C^\infty_{S,\mathbb{R}})$-torsor.

Actually, as we are algebraists, we probably prefer to describe vector bundles in terms of their sheaves of local sections, which are then locally free $C^\infty_{S,\mathbb{R}}$-modules. Personally, I find the definition of a locally free sheaf of modules much simpler than that of a vector bundle, and the two notions are equivalent. So, we discuss sheaves of modules.

3.3.4 Definition. Let S be a topological space, and $\mathcal{O}$ a sheaf of rings on S. A *sheaf of $\mathcal{O}$-modules* is a sheaf $\mathcal{E}$ of abelian groups, together with, for all open $U \subset S$, a map $\mathcal{O}(U) \times \mathcal{E}(U) \to \mathcal{E}(U)$ that makes $\mathcal{E}(U)$ into an $\mathcal{O}(U)$-module, such that for all inclusions $V \subset U$ of opens of S, for all $f \in \mathcal{O}(U)$ and $e \in \mathcal{E}(U)$ we have $(fs)|_V = (f|_V)(e|_V)$.

A morphism of $\mathcal{O}$-modules $\phi\colon \mathcal{E} \to \mathcal{F}$ is a morphism of sheaves ϕ such that for all opens $U \subset S$, the morphism $\mathcal{E}(U) \to \mathcal{F}(U)$ is a morphism of $\mathcal{O}(U)$-modules.

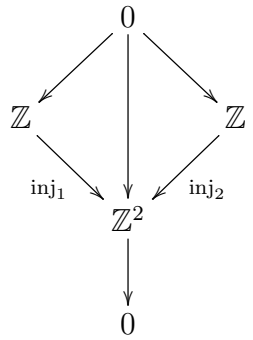
Let $n \in \mathbb{Z}_{\geq 0}$. A sheaf of $\mathcal{O}$-modules $\mathcal{E}$ is called locally free of rank n if there is an open cover $(U_i)_{i\in I}$ of S such that $\mathcal{E}|_{U_i}$ is isomorphic, as $\mathcal{O}|_{U_i}$-module, to the free $\mathcal{O}|_{U_i}$-module $\mathcal{O}|^n_{U_i}$. Concretely this means that we have $e_{i,1}, \ldots, e_{i,n}$ in $\mathcal{E}(U_i)$ such that for all open $V \subset U_i$ and all $e \in \mathcal{E}(V)$ there are unique $f_j \in \mathcal{O}(V)$, $1 \leq j \leq n$, such that $e = \sum_j f_j e_{i,j}|_V$.

3.3.5 Proposition. *Let S be a topological space, $\mathcal{O}$ a sheaf of rings on S, n in $\mathbb{Z}_{\geq 0}$ and $\mathcal{E}$ a locally free $\mathcal{O}$-module of rank n. Then the sheaf $\mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})$ is a right-$\mathrm{GL}_n(\mathcal{O})$-torsor.*

Proof. Straightforward. □

4 Problem Session 2

1. With S and $\mathcal{X}$ as in Example 3.2.5, take now $\mathcal{G}$ given by $\mathcal{G}(S) = \{0\}$, $\mathcal{G}(\{-1, 0\}) = \mathbb{Z}$, $\mathcal{G}(\{0, 1\}) = \mathbb{Z}$ and $\mathcal{G}(\{0\}) = \mathbb{Z}^2$. Here is the diagram:



where $\mathrm{inj}_1\colon \mathbb{Z} \to \mathbb{Z}^2$ is the map $x \mapsto (x, 0)$ and $\mathrm{inj}_2\colon \mathbb{Z} \to \mathbb{Z}^2$ the map $y \mapsto (0, y)$, they are the injections on 1st and 2nd coordinate. There is a morphism of sheaves from $\mathcal{G}$ to $\mathcal{X}$ given by:

$$\begin{aligned}
\mathcal{G}(S) = 0 \to \mathbb{Z} = \mathcal{X}(S) &\quad \text{is the zero map,}\\
\mathcal{G}(\{-1,0\}) = \mathbb{Z} \to \mathbb{Z} = \mathcal{X}(\{-1,0\}) &\quad \text{is the identity map,}\\
\mathcal{G}(\{0,1\}) = \mathbb{Z} \to \mathbb{Z} = \mathcal{X}(\{0,1\}) &\quad \text{is the identity map,}\\
\mathcal{G}(\{0\}) = \mathbb{Z}^2 \to \mathbb{Z} = \mathcal{X}(\{0\}) &\quad \text{is the } \textit{sum} \text{ map: } +\colon \mathbb{Z}^2 \to \mathbb{Z}.
\end{aligned}$$

Show that in this case $(\mathcal{X}/\mathcal{G})_p$ is not a sheaf, but this time the problem is that the map $(\mathcal{X}/\mathcal{G})_p(S) \to (\mathcal{X}/\mathcal{G})_p(\{-1, 0\}) \times (\mathcal{X}/\mathcal{G})_p(\{0, 1\})$ is not injective (non-zero global sections

are locally zero). Show that again replacing $(\mathcal{X}/\mathcal{G})_p(S)$ by what it should be gives a sheaf and that that sheaf has the universal property for the quotient in the category of sheaves.

Just for information, if we denote $U := \{-1, 0\}$, $V := \{0, 1\}$ and j_U and j_V their inclusions in S, then $\mathcal{G} = j_{U,!}\mathbb{Z}_U \oplus j_{V,!}\mathbb{Z}_V$.

2. Give the four missings proofs of the results in this lecture.

3. Browse the Stacks Project a bit: `stacks.math.columbia.edu`

5 Lecture 3: The case of a transitive action

We start by introducing certain operations with torsors.

5.1 Twisting with a torsor

First we discuss this for sets, not sheaves. Let G be a group, X a set with a right G-action, and Y a set with a left G-action. Then we define $X \otimes_G Y$ to be the quotient of $X \times Y$ by the right G-action $(x, y)\cdot g = (xg, g^{-1}y)$. This is the same as dividing $X \times Y$ by the equivalence relation

$$\{((xg, y), (x, gy)) : x \in X,\, y \in Y,\, g \in G\} \subset (X \times Y)^2.$$

We use the notation of the tensor product of modules over a ring: $M \otimes_A N$, because this construction has a similar universal property. For every set Z, for every map $f\colon X \times Y \to Z$ such that for all x, y, g one has $f(xg, y) = f(x, gy)$, there is a unique map $\overline{f}\colon X \otimes_G Y \to Z$ such that $\overline{f} \circ q = f$, where $q\colon X \times Y \to X \otimes_G Y$ is the quotient map.

Now for sheaves. For $\mathcal{G}$, $\mathcal{X}$, $\mathcal{Y}$ we get $\mathcal{X} \otimes_{\mathcal{G}} \mathcal{Y} := (\mathcal{X} \times \mathcal{Y})/\mathcal{G}$.

The construction of $\mathcal{X} \otimes_{\mathcal{G}} \mathcal{Y}$ is functorial in $\mathcal{X}$ and $\mathcal{Y}$: for $f\colon \mathcal{X} \to \mathcal{X}'$ and $g\colon \mathcal{Y} \to \mathcal{Y}'$, we get an induced morphism $f \otimes g\colon \mathcal{X} \otimes_{\mathcal{G}} \mathcal{Y} \to \mathcal{X}' \otimes_{\mathcal{G}} \mathcal{Y}'$.

Note that if we make $\mathcal{G}$ into a right $\mathcal{G}$-torsor by letting it act on itself by right multiplication, then $\mathcal{G} \times \mathcal{Y} \to \mathcal{Y}$, $(g, y) \mapsto gy$, induces an isomorphism $\mathcal{G} \otimes_{\mathcal{G}} \mathcal{Y} \to \mathcal{Y}$. Its inverse is given by $\mathcal{Y} \to \mathcal{G} \times \mathcal{Y}$, $y \mapsto (1_G, y)$. In particular, no sheafification is necessary for the quotient $q\colon \mathcal{G} \times \mathcal{Y} \to \mathcal{G} \otimes_{\mathcal{G}} \mathcal{Y}$.

If $\mathcal{X}$ is a right $\mathcal{G}$-torsor, then $\mathcal{X} \otimes_{\mathcal{G}} \mathcal{Y}$ is locally isomorphic to $\mathcal{Y}$, as sheaf of sets on S. Let us make this precise. For $U \subset S$ and $x_U \in \mathcal{X}(U)$, we have an isomorphism of right $\mathcal{G}|_U$-torsors: $i\colon \mathcal{G}|_U(V) \to \mathcal{X}|_U(V)$, $g \mapsto x_U|_V \cdot g$. Then $i \otimes \mathrm{id}_{\mathcal{Y}}$ is an isomorphism $(\mathcal{G} \otimes_{\mathcal{G}} \mathcal{Y})|_U \to (\mathcal{X} \otimes_{\mathcal{G}} \mathcal{Y})|_U$. And, we have seen that $(\mathcal{G} \otimes_{\mathcal{G}} \mathcal{Y})|_U$ is isomorphic to $\mathcal{Y}|_U$.

The next proposition shows that a locally free $\mathcal{O}$-module $\mathcal{E}$ on a topological space S can be recovered from the $\mathrm{GL}_n(\mathcal{O})$-torsor $\mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})$.

5.1.1 Proposition. *Let $\mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})$ be as in Prop. 3.3.5. Then the morphism of sheaves*

$$f(U)\colon \mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})(U) \times \mathcal{O}^n(U) \to \mathcal{E}(U), \quad (\phi, s) \mapsto (\phi(U))(s)$$

factors through $q\colon \mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E}) \times \mathcal{O}^n \to \mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E}) \otimes_{\mathrm{GL}_n(\mathcal{O})} \mathcal{O}^n$, and induces an isomorphism

$$\mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E}) \otimes_{\mathrm{GL}_n(\mathcal{O})} \mathcal{O}^n \to \mathcal{E}.$$

Proof. Let us show that f factors through q. For $\phi\colon \mathcal{O}^n|_U \to \mathcal{E}_U$ an isomorphism and s in $\mathcal{O}^n(U)$ and $g \in \mathrm{GL}_n(\mathcal{O}(U))$, we have to show that $(\phi \circ g, s)$ and $(\phi, g\cdot s)$ have the same image under $f(U)$. But that results from $f(\phi \circ g, s) = (\phi \circ g)s = \phi(g(s)) = f(\phi, g\cdot s)$.

Now we must show that $\overline{f}\colon \mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E}) \otimes_{\mathrm{GL}_n(\mathcal{O})} \mathcal{O}^n \to \mathcal{E}$ is an isomorphism of sheaves. That is a local question, so we may assume that $\mathcal{E}$ is isomorphic to $\mathcal{O}^n$, and even that it *is* $\mathcal{O}^n$. But then $\mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})$ is $\mathrm{GL}_n(\mathcal{O})$, and the morphism f is the action, and we have seen above that this induces an isomorphism as desired. □

5.2 Functoriality of torsors

Let S be a topological space, let $\phi\colon \mathcal{H} \to \mathcal{G}$ be a morphism of sheaves of groups on S. Then, for each right $\mathcal{H}$-torsor $\mathcal{X}$, we obtain a right $\mathcal{G}$-torsor $\mathcal{X} \otimes_{\mathcal{H}} \mathcal{G}$, where we let $\mathcal{H}$ act from the left on $\mathcal{G}$ via left multiplication via ϕ: $h{\cdot}g := \phi(h)g$ (sections over some $U \subset S$), and where the right action of $\mathcal{G}$ on itself provides the right $\mathcal{G}$ action on $\mathcal{X} \otimes_{\mathcal{H}} \mathcal{G}$.

This construction is a functor from the category of right $\mathcal{H}$-torsors to that of right $\mathcal{G}$-torsors: $f\colon \mathcal{X} \to \mathcal{Y}$ induces $f \otimes \mathrm{id}_{\mathcal{G}}\colon \mathcal{X} \otimes_{\mathcal{H}} \mathcal{G} \to \mathcal{Y} \otimes_{\mathcal{H}} \mathcal{G}$.

5.3 The set of torsors up to isomorphism

5.3.1 Definition. Let S be a topological space, and $\mathcal{G}$ a sheaf of groups on S. Then we define $\mathrm{H}^1(S, \mathcal{G})$ to be the set of isomorphism classes of right $\mathcal{G}$-torsors on S. The isomorphism class of $\mathcal{X}$ will be denoted by $[\mathcal{X}] \in \mathrm{H}^1(S, \mathcal{G})$.

The set $\mathrm{H}^1(S, \mathcal{G})$ has a distinguished element: the isomorphism class of the trivial torsor $\mathcal{G}$ itself. Hence $\mathrm{H}^1(S, \mathcal{G})$ is actually a *pointed set*. It is called the *first cohomology set*. If $\mathcal{G}$ is commutative, then this set has an commutative group structure: $(\mathcal{T}_1, \mathcal{T}_2) \mapsto \mathcal{T}_1 \otimes_{\mathcal{G}} \mathcal{T}_2$ (there is no distinction between left and right, precisely because $\mathcal{G}$ is commutative). The inverse $\mathcal{T}^{-1}$ of $\mathcal{T}$ is $\mathcal{T}$ itself, but with $\mathcal{G}$ acting via $\mathcal{G} \to \mathcal{G}$, $g \mapsto g^{-1}$.

5.3.2 Example. Let S be a topological space and $\mathcal{O}$ a sheaf of rings on it. Then $\mathrm{H}^1(S, \mathrm{GL}_n(\mathcal{O}))$ is also the set of isomorphism classes of locally free $\mathcal{O}$-modules of rank n on S. This is an application of the constructions $\mathcal{E} \mapsto \mathbf{Isom}_S(\mathcal{O}^n, \mathcal{E})$ and $\mathcal{T} \mapsto \mathcal{T} \otimes_{\mathrm{GL}_n(\mathcal{O})} \mathcal{O}^n$ that give an equivalence of categories between the category of locally free $\mathcal{O}$-modules of rank n and with morphisms only isomorphisms, and the category of right $\mathrm{GL}_n(\mathcal{O})$-torsors.

5.4 A transitive action

The following theorem is the result from sheaf theory that will be applied to prove Gauss's theorem in the last lecture. We prefer to formulate one long statement.

5.4.1 Theorem. *Let S be a topological space, and $\mathcal{G}$ a sheaf of groups, $\mathcal{X}$ a sheaf of sets with a transitive left $\mathcal{G}$-action, and $x \in \mathcal{X}(S)$. We let $\mathcal{H} := \mathcal{G}_x$ the stabiliser of x in $\mathcal{G}$ (see Def. 3.3.1 if necessary), and let $i\colon \mathcal{H} \to \mathcal{G}$ denote the inclusion. Then $\mathcal{G}(S)$ acts on $\mathcal{X}(S)$, and we have maps*

$$\mathcal{X}(S) \xrightarrow{c} \mathrm{H}^1(S, \mathcal{H}) \xrightarrow{i} \mathrm{H}^1(S, \mathcal{G}) \tag{5.4.2}$$

where:

- *$c\colon \mathcal{X}(S) \to \mathrm{H}^1(S, \mathcal{H})$ sends $y \in \mathcal{X}(S)$ to the subsheaf $\mathcal{G}_{x,y}$ of $\mathcal{G}$ with (for all open $U \subset S$) $\mathcal{G}_{x,y}(U) = \{g \in \mathcal{G}(U) : g{\cdot}x|_U = y|_U\}$; $\mathcal{G}_{x,y}$ is a right $\mathcal{H}$-torsor;*
- *$i\colon \mathrm{H}^1(S, \mathcal{H}) \to \mathrm{H}^1(S, \mathcal{G})$ is the map that sends the isomorphism class of a right $\mathcal{H}$-torsor $\mathcal{X}$ to the isomorphism class of the right $\mathcal{G}$-torsor $\mathcal{X} \otimes_{\mathcal{H}} \mathcal{G}$, in other words, the map induced by $i\colon \mathcal{H} \to \mathcal{G}$.*

Then:

1. *for y_1 and y_2 in $\mathcal{X}(S)$, $c(y_1) = c(y_2)$ if and only if there exists $g \in \mathcal{G}(S)$ such that $y_2 = gy_1$;*
2. *for $\mathcal{T}$ a right $\mathcal{H}$-torsor, $\mathcal{T} \otimes_{\mathcal{H}} \mathcal{G}$ is trivial if and only if $[\mathcal{T}]$ is in the image of c.*

3. *if $\mathcal{H}$ is commutative, then for all y in $\mathcal{X}(S)$, $\mathcal{G}_y$ is naturally isomorphic to $\mathcal{H}$.*

4. *If $\mathcal{H}$ is commutative and $\mathcal{G}(S)/\mathcal{H}(S)$ is finite, then all fibres of c consist of $\#(\mathcal{G}(S)/\mathcal{H}(S))$ elements.*

Proof. Let us first show that for $y \in \mathcal{X}(S)$, the presheaf $\mathcal{G}_{x,y}$ is a sheaf. Let U be an open subset of S, and $(U_i)_{i\in I}$ an open cover of it, and, for $i \in I$, $g_i \in \mathcal{G}_{x,y}(U_i)$, such that for all $(i,j) \in I^2$, $g_i|_{U_{i,j}} = g_j|_{U_{i,j}}$ in $\mathcal{G}(U_{i,j})$. Note that the g_i are in $\mathcal{G}(U_i)$. As $\mathcal{G}$ is a sheaf, there is a unique $g \in \mathcal{G}(U)$ such that for all $i \in I$, $g_i = g|_{U_i}$. Then we have $g{\cdot}x|_U$ in $\mathcal{X}(U)$. Then for all i in I we have $(g{\cdot}x|_U)|_{U_i} = g|_{U_i}x|_{U_i} = g_ix|_{U_i} = y|_{U_i}$, hence, as $\mathcal{X}$ is a sheaf, $(g{\cdot}x|_U) = y|_U$, hence g is in $\mathcal{G}_{x,y}(U)$.

Let us now show that for $y \in \mathcal{X}(S)$, $\mathcal{G}_{x,y}$ is a right $\mathcal{H}$-torsor. First the right $\mathcal{H}$-action. For $U \subset S$ open, $h \in \mathcal{H}(U)$ and $g \in \mathcal{G}_{x,y}(U)$, we have gh in $\mathcal{G}(U)$ (h and g are both in $\mathcal{G}(U)$). By definition of $\mathcal{H}$, $hx|_U = x|_U$, and $gx|_U = y|_U$. Then $(gh)x|_U = g(hx|_U) = gx|_U = y|_U$. Hence indeed gh is in $\mathcal{G}_{x,y}(U)$. Let us show that for all U the action of $\mathcal{H}(U)$ on $\mathcal{G}_{x,y}(U)$ is free. Let g be in $\mathcal{G}_{x,y}(U)$ and h in $\mathcal{H}(U)$ such that $gh = g$. Then $h = g^{-1}gh = g^{-1}g = 1$ in $\mathcal{G}(U)$. So the action is free. Now we show that the action of $\mathcal{H}$ on $\mathcal{G}_{x,y}$ is transitive. Let U be open, g_1 and g_2 in $\mathcal{G}_{x,y}(U)$. Then $g_2 = g_1{\cdot}(g_1^{-1}g_2)$, and $h := g_1^{-1}g_2$ is in $\mathcal{H}(U)$ because $hx = g_1^{-1}g_2x = g_1^{-1}y = x$. Finally, we show that locally $\mathcal{G}_{x,y}$ has sections. But this is because $\mathcal{G}$ acts transitively on $\mathcal{X}$: there is a cover $(U_i)_{i\in I}$ and $g_i \in \mathcal{G}(U_i)$ such that $g_ix = y$ in $\mathcal{X}(U_i)$.

Let us prove (1). Let y_1 and y_2 be in $\mathcal{X}(S)$.

Suppose that g is in $\mathcal{G}(S)$ and that $gy_1 = y_2$. Then left multiplication by g in $\mathcal{G}$ gives us an isomorphism of right $\mathcal{H}$-torsors from $\mathcal{G}_{x,y_1}$ to $\mathcal{G}_{x,y_2}$

Suppose now that $c(y_1) = c(y_2)$. We have to show that there is a g in $\mathcal{G}(S)$ such that $gy_1 = y_2$. The assumption is that $\mathcal{G}_{x,y_1}$ and $\mathcal{G}_{x,y_2}$ are isomorphic. So let $\phi\colon \mathcal{G}_{x,y_1} \to \mathcal{G}_{x,y_2}$ be an isomorphism. Each point in S has an open neighborhood U such that there exists a t in $\mathcal{G}_{x,y_1}(U)$. For such a t, we have $\phi(t)$ in $\mathcal{G}_{x,y_2}(U)$, and hence $(\phi(t))t^{-1}$ in $\mathcal{G}(U)$ with $(\phi(t))t^{-1}{\cdot}y_1 = \phi(t)x = y_2$. We claim that this element $(\phi(t))t^{-1}$ does not depend on the choice of t. Any t' in $\mathcal{G}_{x,y_1}(U)$ is of the form th for a unique h in $\mathcal{H}(U)$. Then we have:

$$\phi(t')t'^{-1} = \phi(th)(th)^{-1} = \phi(t)hh^{-1}t^{-1} = \phi(t)t^{-1}\,.$$

So we let g_U be this element $\phi(t)t^{-1}$ of $\mathcal{G}(U)$. These g_U form a compatible collection of local sections of $\mathcal{G}$: for all U and V on which $\mathcal{G}_{x,y_1}$ has a section, g_U and g_V have the same restriction to $U \cap V$. As $\mathcal{G}$ is a sheaf, there is a unique g in $\mathcal{G}(S)$ such that for all U as above, $g_U = g|_U$. For each U we have $(gy_1)|_U = g|_Uy_1|_U = g_Uy_1|_U = y_2|_U$, hence (now using that $\mathcal{X}$ is a sheaf), $gy_1 = y_2$.

The proof of (2) is for the problem session. This proof is the heart of our workshop! Some beautiful things happen in it, and we hope the participants will discover them, with our help. After the problem session we will put on-line a version of these notes that contains a proof.

Let us prove (3). So now we assume that $\mathcal{H}$ is commutative. Let y be in $\mathcal{X}(S)$. Each point of S has an open neighborhood U such that there exists a g in $\mathcal{G}_{x,y}(U)$. Then, for each $V \subset U$, we have the map $c_g(V)\colon \mathcal{G}_x(V) \to \mathcal{G}_y(V)$ that sends h to ghg^{-1}. This map is an isomorphism of groups, and it is compatible with the restriction maps for $V' \subset V$, that is, c_g is an isomorphism of sheaves of groups from $\mathcal{G}_x|_U$ to $\mathcal{G}_y|_U$. We claim that c_g is in fact independent of the choice of g. Any g' in $\mathcal{G}_{x,y}(U)$ is of the form gh for a unique h in $\mathcal{H}(U)$. Then, for $V \subset U$:

$$c_{g'}(V)\colon k \mapsto g'kg'^{-1} = ghkh^{-1}g^{-1} = gkg^{-1} = (c_g(V))(k)\,.$$

Hence we can label the c_g as c_U, as they only depend on U. But then the $c_U\colon \mathcal{G}_x|_U \to \mathcal{G}_y|_U$ are a compatible collection of isomorphisms. Hence there is a unique isomorphism $c\colon \mathcal{G}_x \to \mathcal{G}_y$ such that for each U as above, $c_U = c|_U$.

Let us prove (4). By (1), the fibres of c are the orbits of $\mathcal{G}(S)$ acting on $\mathcal{X}(S)$. For y in $\mathcal{X}(S)$ the map $\mathcal{G}(S) \to \mathcal{X}(S)$, $g \mapsto gy$ factors through the quotient map $\mathcal{G}(S) \to \mathcal{G}(S)/\mathcal{G}_y(S)$, and gives a bijection from $\mathcal{G}(S)/\mathcal{G}_y(S)$ to the $\mathcal{G}(S)$-orbit of y. By (3), we have, for all y in$\mathcal{X}(S)$, an isomorphism $\mathcal{H} \to \mathcal{G}_y$, which shows that all $\mathcal{G}(S)$-orbits in $\mathcal{X}(S)$ have $\#(\mathcal{G}(S)/\mathcal{H}(S))$ elements. □

6 Problem Session 3

1. Prove (2) of Theorem 5.4.1. This is a very long exercise, but it is also the most important one. If you succeed, it is like getting the black belt in judo, but then for working with torsors. We give some hints to make it more likely that you succeed. But do not hesitate to ask us if you need help!

   - First do the easy half. Let y be in $\mathcal{X}(S)$. Then you have to prove that $\mathcal{G}_{x,y} \otimes_{\mathcal{H}} \mathcal{G}$ is trivial. This means that you have to show that $(\mathcal{G}_{x,y} \otimes_{\mathcal{H}} \mathcal{G})(S)$ is not empty. Your only hope for this is to construct an open cover $(U_i)_{i\in I}$ of S and a collection of compatible sections $f_i \in (\mathcal{G}_{x,y} \otimes_{\mathcal{H}} \mathcal{G})(U_i)$.
   - Recall that $\mathcal{G}_{x,y} \otimes_{\mathcal{H}} \mathcal{G}$ is the quotient of $\mathcal{G}_{x,y} \times \mathcal{G}$ by a suitable action of $\mathcal{H}$. This means that sections of $\mathcal{G}_{x,y} \otimes_{\mathcal{H}} \mathcal{G}$ locally come from sections of $\mathcal{G}_{x,y} \times \mathcal{G}$. Just be courageous now and *try something*.
   - Now the other implication. Let $\mathcal{T}$ be a right $\mathcal{H}$-torsor, and suppose that $\mathcal{T} \otimes_{\mathcal{H}} \mathcal{G}$ is trivial. Choose a global section f of it. Locally this f comes from $\mathcal{T} \times \mathcal{G}$, so choose something, and do something with it that gives a local section of $\mathcal{X}$.

2. Situation as in §5.3. Show that the formula given for the inverse in $\mathrm{H}^1(S, \mathcal{G})$ for $\mathcal{G}$ commutative actually *is* an inverse: $\mathcal{T}^{-1} \otimes_{\mathcal{G}} \mathcal{T}$ is isomorphic to $\mathcal{G}$, and give an isomorphism.

3. Situation as in Theorem 5.4.1. Let y_1 and y_2 be in $\mathcal{X}(S)$. Show that $\mathcal{G}_{y_2,x}$ is stable under the action of $\mathcal{H}$ on $\mathcal{G}$ by left translations, and that it is a left $\mathcal{H}$-torsor. And show that the morphism of sheaves $\mathcal{G}_{x,y_1}(U) \times \mathcal{G}_{y_2,x}(U) \to \mathcal{G}_{y_2,y_1}(U)$, sending (g_1, g_2) to $g_1 g_2$, induces an isomorphism of sheaves $\mathcal{G}_{x,y_1} \otimes_{\mathcal{H}} \mathcal{G}_{y_2,x} \to \mathcal{G}_{y_2,y_1}$. To do this, it helps to use the action by $\mathcal{G}_{y_2}$ on the right, or the $\mathcal{G}_{y_1}$ on the left, to see that the morphism is an isomorphism. This can be used to give another proof of one of the implications in (1) of Theorem 5.4.1.

7 Lecture 4: Application to Gauss's theorem on 3 squares

7.1 Picard groups, and the spectrum of a ring

In order to state Gauss's theorem, we need to introduce the notion of Picard group.

Let S be a topological space, and $\mathcal{O}$ a sheaf of rings on S. Then $\mathrm{Pic}(S, \mathcal{O})$ is the set of isomorphism classes of locally free $\mathcal{O}$-modules of rank one, also called invertible $\mathcal{O}$-modules. The class of an invertible $\mathcal{O}$-module $\mathcal{L}$ is denoted by $[\mathcal{L}]$. On this set $\mathrm{Pic}(S)$, there is a structure of commutative group: $[\mathcal{L}_1]\cdot[\mathcal{L}_2] = [\mathcal{L}_1 \otimes_{\mathcal{O}} \mathcal{L}_2]$, and $[\mathcal{L}]^{-1} = [\mathcal{L}^\vee]$, where $\mathcal{L}^\vee = \mathbf{Hom}_{\mathcal{O}}(\mathcal{L}, \mathcal{O})$ is the dual of $\mathcal{L}$. The equivalence of invertible $\mathcal{O}$-modules with $\mathcal{O}^\times$-torsors ($\mathcal{L}$ corresponds to $\mathbf{Isom}_{\mathcal{O}}(\mathcal{O}, \mathcal{L})$) shows that $\mathrm{Pic}(S, \mathcal{O}) = \mathrm{H}^1(S, \mathcal{O}^\times)$.

Now let A be a ring. Then we have the set $\mathrm{Spec}(A)$ of prime ideals of A, and the Zariski topology on $\mathrm{Spec}(A)$. There is a natural sheaf of rings $\mathcal{O}$ on $\mathrm{Spec}(A)$ that is analogous to the sheaf of regular functions on an algebraic variety, see for example Hartshorne's book "Algebraic

Geometry", Chapter II, section 2, or any other text in which affine schemes are defined. We do not want to go into this construction in detail, but instead give one property of $\mathcal{O}$ that determines it. For each f in A, we have the closed subset $Z(f)$ of $\mathrm{Spec}(A)$ (see Examples 1.2.4), and the complement $D(f) := \mathrm{Spec}(A) - Z(f)$ which is open in $\mathrm{Spec}(A)$. Then $\mathcal{O}(D(f)) = A_f$, the localisation of A with respect to the multiplicative system $\{f^n : n \in \mathbb{Z}_{\geq 0}\}$. We then have a ring morphism $\psi_f \colon A \to A_f$ that has the universal property that $\psi(f)$ is invertible, and for any ring morphism $\phi \colon A \to B$ such that $\phi(f)$ is invertible, there is a unique morphism $\phi' \colon A_f \to B$ such that $\phi = \phi' \circ \psi_f$. If f and g are in A and $D(g) \subset D(f)$, then there are $n \geq 0$ and $a \in A$ such that $g = f^n a$, and so there is a unique morphism $\psi_{f,g} \colon A_f \to A_g$ such that $\psi_g = \psi_{f,g} \circ \psi_f$. The restriction map $\mathcal{O}(D(f)) \to \mathcal{O}(D(g))$ is then $\psi_{f,g}$. This determines $\mathcal{O}$, because the $D(f)$ form a basis for the topology on $\mathrm{Spec}(A)$.

But then for a ring A we have the Picard group $\mathrm{Pic}(A) := \mathrm{Pic}(\mathrm{Spec}(A), \mathcal{O})$. This is the same as what number theorists call the class group of invertible ideals when A is an order in a number field (finite field extension of $\mathbb{Q}$). For such A, $\mathrm{Pic}(A)$ is a finite group, and it measures how much invertible ideals are not principal, and, if A is the ring of integers, $\mathrm{Pic}(A)$ is the obstruction to unique factorisation in A.

7.2 The theorem

For $d \in \mathbb{Z}$ not a square, $d \equiv 0, 1 (\mathrm{mod}\ 4)$, we let O_d be the subring of $\mathbb{C}$ generated by $u_d := (\sqrt{d} + d)/2$. It consists of the numbers $a + bu_d$ with a and b in $\mathbb{Z}$. It is free as $\mathbb{Z}$-module with basis $(1, u_d)$. The minimal polynomial of u_d is $f_d = x^2 - dx + (d^2 - d)/4$; the discriminant of f_d is d. The ring O_d is called the quadratic order of discriminant d. For such a d, we have the group $\mathrm{Pic}(O_d)$

Here is Gauss's theorem. It is article 291 (p. 336–339) of the english Springer edition of Disquisitiones Arithmeticae. Gauss wrote his book in Latin, finished in 1798 (published in 1801).

7.2.1 Theorem. (Gauss) *Let $n \in \mathbb{Z}_{\geq 1}$. Then:*

$$\#\{x \in \mathbb{Z}^3 : x_1^2 + x_2^2 + x_3^2 = n \text{ and } \gcd(x_1, x_2, x_3) = 1\} = \begin{cases} 0 & \text{if } n \equiv 0, 4, 7 (8), \\ 48 \cdot \dfrac{\#\mathrm{Pic}(O_{-n})}{\#(O_{-n}^\times)} & \text{if } n \equiv 3 (8), \\ 24 \cdot \dfrac{\#\mathrm{Pic}(O_{-4n})}{\#(O_{-4n}^\times)} & \text{if } n \equiv 1, 2 (4). \end{cases}$$

The first case in this theorem is very easy to prove. The squares in $\mathbb{Z}/8\mathbb{Z}$ are 0, 1 and 4. If (x_1, x_2, x_3) is in $\mathbb{Z}^3$ and $\gcd(x_1, x_2, x_3) = 1$, then at least one among the x_i is odd, hence $x_1^2 + x_2^2 + x_3^2$ cannot be 0, 4 and 7 in $\mathbb{Z}/8\mathbb{Z}$.

In the next sections, we will show how we apply Theorem 5.4.1 to prove the last two cases of Gauss's result, assuming that there is at least one solution.

7.3 How we apply the results on sheaves

We want to understand the set of primitive solutions in $\mathbb{Z}^3$ of the equation $x^2 + y^2 + z^2 = n$, where primitive means that $\gcd(x, y, z) = 1$. Theorem 7.2.1 says how many primitive solutions there are. This suffices for the problem of understanding all solutions, because if (x, y, z) is a solution and $d := \gcd(x, y, z) > 1$, then $(x/d, y/d, z/d)$ is a primitive solution of the equation $x^2 + y^2 + z^2 = n/d^2$. We will see that our method of using symmetries to study solutions is better suited for the problem of primitive solutions than all solutions.

The set of all solutions is a subset of the solutions (x, y, z) in $\mathbb{Q}^3$. The set of rational solutions is in fact very easy to parameterise, in the same way as one parameterises the rational points on a circle. If one has one rational point P on a circle C, then each line through P with rational direction D gives a second *rational* solution $Q = P + tD$ because the equation $\|P + tQ\|^2 = n$ is quadratic in t, with coefficients in $\mathbb{Q}$ and has already $t = 0$ as solution. Then the second solution is also in $\mathbb{Q}$.

Another way to get all rational solutions from P is to use symmetries in planes that are defined by an equation with rational coefficients. Let $Q \neq 0$ be in $\mathbb{Q}^3$. The symmetry about the plane orthogonal to Q is given by the formula:

$$s_Q \colon \mathbb{Q}^3 \to \mathbb{Q}^3, \quad R \mapsto R - 2\frac{\langle R, Q\rangle}{\langle Q, Q\rangle}Q,$$

where $\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3$ denotes the standard inner product. (Note that s_Q depends only on the direction of Q). This shows that if R and Q are in $\mathbb{Q}^3$, then so is $s_Q(R)$. And if $\|R\|^2 = n$, then $\|s_Q(R)\|^2 = n$ (symmetries are isometries, but you can check it by a calculation if you want). Finally, for R in $\mathbb{Q}^3$ with $R \neq P$ and $\|R\|^2 = n$, we can take $Q := R - P$ and then we have $s_Q(P) = R$ (elementary geometry, or a calculation but then you do not understand why this is so).

The idea for proving Theorem 7.2.1 is now to use these symmetries (rotations, actually, so, products of two symmetries), and do a serious administration concerning the denominators of the coordinates of the rational solutions to get information on the primitive integers solutions. The administration tool is sheaf theory on the topological space $\mathrm{Spec}(\mathbb{Z})$.

Let us define a sheaf $\mathcal{X}_n$ on $\mathrm{Spec}(\mathbb{Z})$. We define a, for each open $U = D(m)$ (with $m > 0$):

$$\mathcal{X}_n(U) := \{(x, y, z) \in \mathbb{Z}[1/m] : x^2 + y^2 + z^2 = n \text{ and } \gcd(x, y, z) = 1\}.$$

Then for $V \subset U$ we have $\mathcal{X}_n(U) \subset \mathcal{X}_n(V)$, these inclusions are our restriction maps, and make $\mathcal{X}_n$ into a presheaf. It is a sheaf (exercise 1 below).

We also want a sheaf of groups acting on $\mathcal{X}_n$. For this we take groups of rotations. For any ring A we define $\mathrm{SO}_3(A)$ as:

$$\mathrm{SO}_3(A) := \{g \in \mathrm{M}_3(A) : g^t \cdot g = 1_3 \text{ and } \det(g) = 1\},$$

where $\mathrm{M}_3(A)$ is the set of 3 by 3 matrices with coefficients in A. In other words, $\mathrm{SO}_3(A)$ is the group of automorphisms of the free A-module A^3 that fix the standard inner product and whose determinant is 1 (that is, they preserve the standard orientation).

For $U = D(m)$ $(m \neq 0)$, we define:

$$\mathcal{G}(U) := \mathrm{SO}_3(\mathcal{O}(U)) = \mathrm{SO}_3(\mathbb{Z}[1/m]).$$

Then also $\mathcal{G}$ is a sheaf, of groups, and it acts naturally on $\mathcal{X}_n$. The following (surprising) result then makes it possible to apply the sheaf theory from the previous lectures.

7.3.1 Theorem. *Let $n \in \mathbb{Z}_{>0}$. The action of $\mathcal{G}$ on $\mathcal{X}_n$ is transitive.*

Proof. Our proof will use symmetries as we described above. Transitivity of the action of $\mathcal{G}$ is a local property on $\mathrm{Spec}(\mathbb{Z})$, that is, a property at each maximal ideal of $\mathbb{Z}$.

For p a prime number we define $\mathbb{Z}_{(p)}$ as the subring of $\mathbb{Q}$, consisting of all x in $\mathbb{Q}$ for which there is an open U in $\mathrm{Spec}(\mathbb{Z})$ with $x \in \mathcal{O}(U)$. More explicitly, $\mathbb{Z}_{(p)}$ consists of the x in $\mathbb{Q}$ that are of the form a/b with a and b in $\mathbb{Z}$, with p not dividing b.

The local property at p that what we must show is that for each prime number p and all primitive P and Q in $\mathbb{Z}^3_{(p)}$ with $\langle P, P\rangle = n$, and $\langle Q, Q\rangle = n$, there exists a g in $\mathrm{SO}_3(\mathbb{Z}_{(p)})$ such that $gP = Q$.

If $P = Q$, then for $g := 1_3 \in \mathrm{SO}_3(\mathbb{Z}_{(p)})$ we have $g{\cdot}P = Q$. So assume that $P \neq Q$. First suppose that $p = 2$. Consider $v \in \mathbb{Z}^3$ that satisfies $\langle v, P\rangle = 0$ and v is primitive. The set $P^\perp = \{v \in \mathbb{Z}^3 : \langle v, P\rangle = 0\}$ is a free $\mathbb{Z}$-module of rank two, with the property that if v is in $\mathbb{Z}^3$ and $d \in \mathbb{Z}$ with $d \neq 0$ and $dv \in P^\perp$, then $v \in P^\perp$. Therefore we can take a primitive $v \in \mathbb{Z}^3$ such that $\langle v, P\rangle = 0$. At least one of the coordinates of $v \in \mathbb{Z}^3$ is an odd integer, so the residue of $\langle v, v\rangle$ modulo 4 is not equal to 0. Because of that, from the formula of s_v, we get $s_v \colon \mathbb{Z}^3_{(2)} \to \mathbb{Z}^3_{(2)}$. Also we get that $s_v(P) = P$. Now take w a primitive element in $\mathbb{Z}^3$ such that w is multiple of the vector $P - Q$ by some number in $\mathbb{Q}$. Then the symmetry $s_w \colon \mathbb{Z}^3_{(2)} \to \mathbb{Z}^3_{(2)}$ maps the point P to the point Q. So by construction $g := s_w \circ s_v \colon \mathbb{Z}^3_{(2)} \to \mathbb{Z}^3_{(2)}$ will be in $SO_3(\mathbb{Z}_{(2)})$ and $(s_w \circ s_v)(P) = s_w(P) = Q$.

Now let p be a prime number not equal to 2. We want to find v and w in $\mathbb{Z}^3_{(p)}$ such that s_v and s_w map $\mathbb{Z}^3_{(p)}$ to itself, and $(s_w \circ s_v)(P) = Q$. The idea is that for any v there is no choice for w: w must be a multiple of $s_v(P) - Q$. So, the conditions on $v \in \mathbb{Z}^3_{(p)}$ are: $\langle v, v\rangle$ is not divisible by p, and $w := s_v(P) - Q \in \mathbb{Z}^3_{(p)}$ has $\langle w, w\rangle$ not divisible by p, that is, its image in $\mathbb{F}_p$ is non-zero (here $\mathbb{F}_p$ denotes the finite field with p elements, $\mathbb{Z}/p\mathbb{Z}$).

Now the existence of a v as desired is a matter of showing that there exists an element v in the $\mathbb{F}_p$-vector space $\mathbb{F}_p^3$ (that has p^3 elements) such that $\langle v, v\rangle \neq 0$ and, with $w := s_v(P) - Q$, $\langle w, w\rangle \neq 0$.

Both conditions are *homogeneous* in v: they are satisfied by v if and only if they are satisfied by $\lambda{\cdot}v$ for all λ in $\mathbb{F}_p^\times$. So, it is better to study these conditions on $\mathbb{P}^2(\mathbb{F}_p) = (\mathbb{F}_p^3 - \{0\})/\mathbb{F}_p^\times$. The first condition, $\langle v, v\rangle \neq 0$ means that v does not lie on the conic C defined by the homogeneous equation $x_0^2 + x_1^2 + x_2^2 = 0$. A simple computation shows that the second condition is equivalent to:

$$\frac{\langle P, v\rangle\langle v, Q\rangle}{\langle v, v\rangle} \neq \frac{\langle P, Q\rangle - n}{2}.$$

Note that the left hand side of the last inequality defines a function

$$f \colon \mathbb{P}^2(\mathbb{F}_p) - C(\mathbb{F}_p) \to \mathbb{F}_p, \quad \overline{v} := \mathbb{F}_p^\times{\cdot}v \mapsto \frac{\langle P, v\rangle\langle v, Q\rangle}{\langle v, v\rangle}.$$

It suffices now to show that f is not constant. Now for $\overline{v}$ in $\mathbb{P}^2(\mathbb{F}_p) - C(\mathbb{F}_p)$ we have $f(\overline{v}) = 0$ if and only if $\overline{v}$ is on the (projective) line $P^\perp$ perpendicular to P (its equation is $P_1v_1 + P_2v_2 + P_3v_3 = 0$), or on the line $Q^\perp$ perpendicular to Q. Each of these has $p + 1$ $\mathbb{F}_p$-points, of which at most 2 are on C, hence there are $\overline{v}$ in $\mathbb{P}^2(\mathbb{F}_p) - C(\mathbb{F}_p)$ such that $f(\overline{v}) = 0$. We will now show that there is a $\overline{v}$ in $\mathbb{P}^2(\mathbb{F}_p) - C(\mathbb{F}_p)$ where $f(\overline{v})$ is not zero, by considering all $\overline{v}$ on a suitable line. The issue is that we want a proof that works for all $p \geq 3$, and not have to treat small primes differently. So, consider a line L that contains only one point R that lies on $P^\perp \cup Q^\perp$ (if $P^\perp \neq Q^\perp$ then this means that R is the intersection point of $P^\perp$ and $Q^\perp$). Then $L(\mathbb{F}_p)$ has $p + 1$ points, of which one is R and of which at most two are in $C(\mathbb{F}_p)$. Therefore there are at least $p + 1 - 3 = p - 2 > 0$ points of $L(\mathbb{F}_p)$ where f is defined and is non-zero. $\square$

7.4 Bilinear forms

Let us recall the definition of bilinear form on free modules.

7.4.1 Definition. Let R be a ring (commutative, with identity, as always) and let M be a free R-module of finite rank. A bilinear form b on M is a function $b : M \times M \to R$ such that for every x, y, z in M and every r in R we have

$$b(x + y, z) = b(x, z) + b(y, z), \quad b(x, y + z) = b(x, y) + b(x, z), \quad b(rx, y) = rb(x, y) = b(x, ry).$$

A pair of (M, b) is called a *module with bilinear form*, for simplifying the notation sometime we just write it as M if the bilinear form b is clear in the context.

A bilinear form b is called *symmetric* if for every x, y in M we have $b(x, y) = b(y, x)$. A symmetric bilinear form b will be called *perfect* if the following strong non-degeneracy condition is satisfied: For each R-linear map $\phi\colon M \to R$ there should exist one and only one element x in M such that the homomorphism $y \mapsto b(x, y)$ from M to R is equal to ϕ. In other words, the map $M \to \mathrm{Hom}_R(M, R)$ sending x to $y \mapsto b(x, y)$ is a bijection. We use the notation $M^\vee$ for the dual $\mathrm{Hom}_R(M, R)$ of M.

Two elements x and y in M with a symmetric bilinear form b are called *orthogonal* if $b(x, y) = 0$. Note that for module with perfect symmetric bilinear M, an element x in M is orthogonal to every element y in M if and only if $x = 0$.

7.5 Minkowski's theorem, triviality of $\mathrm{H}^1(\mathrm{Spec}(\mathbb{Z}), \mathrm{SO}_3(\mathcal{O}))$

Now assume that the ring R is either $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$. Then a bilinear form b is called *positive definite* if for any x in M we have $b(x, x) \geq 0$ and $b(x, x) = 0$ if and only if $x = 0$.

Now let n be a positive integer and $\mathbb{R}^n$ be the $\mathbb{R}$-vector space consisting of all n-tuples $x = (x_1, \ldots, x_n)$ of real numbers. We can equip $\mathbb{R}^n$ with the *standard* symmetric bilinear form, i.e., the inner product: $x \cdot y = \sum_{i=1}^n x_i y_i$ for $x = (x_i)$ and $y = (y_i)$ in $\mathbb{R}^n$. The pair $\mathbb{R}^n$ with the inner product is called the *euclidean inner product space*. Note that for any n-dimensional vector space V over $\mathbb{R}$ with a positive definite symmetric bilinear form b will be isomorphic to the euclidean inner product space $\mathbb{R}^n$. This is because we can identify the orthonormal basis of V that is obtained from Gramm-Schmidt process to the standard basis of $\mathbb{R}^n$.

Let n be a positive integer and L be a free $\mathbb{Z}$-module of rank n with positive definite symmetric bilinear form b. We can embed L canonically to $V := L \otimes_{\mathbb{Z}} \mathbb{R}$ which is a n-dimensional $\mathbb{R}$ vector space with the bilinear form $b_{\mathbb{R}}$. It is still a positive definite symmetric bilinear form on V. By above we can assume from now that our L is a subgroup of $\mathbb{R}^n$. Suppose now $(x_1, \ldots, x_n)$ is a $\mathbb{Z}$-basis of L where x_i are vectors in $\mathbb{R}^n$. The volume of the quotient torus $\mathbb{R}^n/L$ is defined as the absolute value of the determinant of the matrix whose rows are $x_1, \ldots, x_n$. It is the same as the square root of the determinant of the Gramm matrix $B = (b(x_i, x_j))_{i,j}$.

Recall that a subset $X \subset \mathbb{R}^n$ is *convex* if x, y in X implies that $\lambda x + (1 - \lambda)y$ in X for all real numbers λ in the interval $0 \leq \lambda \leq 1$. A subset X of $\mathbb{R}^n$ is *symmetric about 0*, if x in X implies $-x$ in X. Now we can state the theorem of Minkowski's.

7.5.1 Theorem. (Minkowski) *Let X be a bounded, convex and symmetric about 0 subset of $\mathbb{R}^n$. If the volume of X is greater than 2^n times the volume of $\mathbb{R}^n/L$, then X contains a non-zero point of L.*

As a consequence of this one can show that $\mathrm{H}^1(\mathrm{Spec}(\mathbb{Z}), \mathrm{SO}_3) = \{1\}$. We cannot do this here because it uses something that we did not discuss: the equivalence between free $\mathbb{Z}$ modules of rank r and locally free $\mathcal{O}$-modules of rank r on $\mathrm{Spec}(\mathbb{Z})$.

7.6 The stabiliser in Gauss's theorem

An important step in proving Gauss's theorem using sheaves is to relate $\mathcal{H}$ of x in $\mathcal{X}_n(S)$ to the quadratic order O_d with $d = -n$ or $d = -4n$ depending on n mod 8, and $\mathrm{H}^1(S, \mathcal{H})$ to the Picard group of O_d. It turns out that there is an exact sequence:

$$0 \to \mathcal{O}^\times \to \mathcal{O}_d^\times \to \mathcal{H} \to 0.$$

Details will appear in Albert's thesis.

7.7 Existence of solutions

We have explained Gauss's result, except that we assumed the existence of a solution in the last two cases. What can be done about that?

The first thing that one can do is to follow Gauss's proof. It is a very beautiful argument, using his results on quadratic forms on $\mathbb{Z}^2$.

Another option is to use the Hasse principle (in fact a theorem of Hasse (for number fields) and Minkowski ($\mathbb{Q}$)) that says, in our case, that the equation $x_1^2 + x_2^2 + x_3^2 - nx_4^2 = 0$ has a non-zero solution if and only if it has solutions in all completions of $\mathbb{Q}$, that is, in $\mathbb{R}$ and in all p-adic fields $\mathbb{Q}_p$. Then, to deduce from a rational solution an integer solution one can use the sheaf approach that we have given above (see the "Open problems" below). A very funny argument is given in Cassels-Fröhlich's book "Algebraic Number Theory", Exercise 4.11 on page 359.

Yet another option is a suggestion by André Weil, in an article dedicated to Siegel. If one knows the formula for the number of solutions under the assumption that there are solutions, then one can prove it by relating it to the number of solutions of $x_1^2 + x_2^2 + x_3^2 + x_4^2 = n$ for which there is a famous fomula by Jacobi.

8 Problem Session 4

1. Let $U \subset \mathrm{Spec}(\mathbb{Z})$ be a non-empty open subset.

   (a) Prove that U is a principal open subset of $\mathrm{Spec}(\mathbb{Z})$, that is, prove that there exists an integer $n > 0$ such that $U = D(n)$. Hint: the complement of U is a finite set maximal ideals, say $p_1\mathbb{Z}, \ldots, p_r\mathbb{Z}$, with the p_i distinct prime numbers.

   (b) The ring $\mathcal{O}(U)$ is by definition equal to $\mathbb{Z}[1/n]$. Prove that this is the subring of $\mathbb{Q}$ consisting of the a/b with a and b in $\mathbb{Z}$, and b of the form $p_1^{e_1} \cdots p_r^{e_r}$.

   (c) Prove that $\mathcal{O}$ is a sheaf. Hint: think now of a rational number a/b ($b \neq 0$ and $\gcd(a, b) = 1$) as a function, that has poles at the primes dividing b, and that at any prime p not dividing b has value $\overline{a}/\overline{b}$ in the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, where $\overline{a}$ is the image of a in $\mathbb{F}_p$, etc. If you know algebraic curves, or Riemann surfaces, this should be familiar to you (rational or meromorphic functions, and regular or holomorphic functions). Then note that $\mathcal{O}(U)$ is the set of x in $\mathbb{Q}$ that do not have a pole at any p_i, $1 \leq i \leq r$. In particular, all $\mathcal{O}(U)$ for U non-empty are subrings of $\mathbb{Q}$; this makes this kind of sheaves easier to understand.

   (d) Prove that $\mathcal{X}_n$ defined as above is a sheaf. Hint: this should follow quite formally from the property that $\mathcal{O}$ on $\mathrm{Spec}(\mathbb{Z})$ is a sheaf.

2. Let R be a ring, M a free R-module, $(e_1, \ldots, e_n)$ a basis of M, and b: $M \times M \to R$ a symmetric bilinear form. Then we have the *Gramm matrix* $B = (b(e_i, e_j))_{i,j}$. Prove that the bilinear form b is perfect if and only if the matrix B is invertible (it has a 2-sided inverse). Hint: produce a basis for $M^\vee$, and think about the matrix of the linear map $M \to M^\vee$ given by b.

3. Let L be free $\mathbb{Z}$-module of rank 3 with positive definite perfect symmetric $\mathbb{Z}$-valued bilinear form b. Using Minkowski's theorem prove that L has an orthonormal $\mathbb{Z}$-basis.

9 Conclusions, Open problems

30 minutes.

I see several open problems, they are for Albert to work on, but if someone wants to work with him, that would be possible, I think, and nice.

1. Determine H as a groupscheme over $\mathbb{Z}$ (problem at 2).

2. Write down as explicit as possible the action of the Picard group on the set of solutions that the sheaf method gives (suggested by Zagier).

3. Try to get existence of a solution from the Hasse principle for $x^2 + y^2 + z^2 - nt^2 = 0$ and the fact that $\mathrm{H}^2(S, H) = 0$ (the dimension of S is one, and $2 > 1$). Sheaf methods say: if there are locally solutions, and this H^2 is zero, then there is a global solution (the technically advanced way to say this is that X with the G-action is an "H-gerbe").

4. Are there other examples where the Zariski topology works? Learn to work with Grothendieck topologies on categories, stronger topologies on schemes.