Notes on Geometry

(Under heavy construction!!)

November 13, 2024

Contents

Part I

The Algebraic Viewpoint

Chapter 1

Foundational Algebraic Geometry

Contents

1.1 A guiding example

Let X be a compact Hausdorff topological space. In this section we would like to portray the main point of scheme theory in the case of space X, that is, one can study the geometry over "base" space completely by studying the algebra of ring of suitable functions over it. In particular, we would like to establish the following result.

Proposition 1.1.0.1. *Let X be a compact Hausdorff topological space. Denote R to be the ring of continuous real-valued functions on X under pointwise addition and multiplication and denote* $\mathrm{mSpec}\,(R)$ *to be the set of maximal ideals of R. Then,*

1. *We have a set bijection:*
$$\mathrm{mSpec}\,(R) \cong X.$$
2. *We have that* $\mathrm{mSpec}\,(R)$ *and X are isomorphic as topological spaces:*
$$\mathrm{mSpec}\,(R) \cong X$$
where $\mathrm{mSpec}\,(R)$ *is given its Zariski topology.*

Proof. 1. Let $x \in X$ be an arbitrary point. Denote $\mathfrak{m}_x := \{f \in R \mid f(x) = 0\}$ to be the vanishing ideal of point x. This ideal is maximal because the quotient $R/\mathfrak{m}_x \cong \mathbb{R}$ via the map $f + \mathfrak{m}_x \mapsto f(x)$. Indeed, it is a valid ring homomorphism and is surjective by virtue of the continuous map constant at a point in $\mathbb{R}$. Moreover, if $f(x) = g(x)$ for $f, g \in R$, then $f - g \in \mathfrak{m}_x$ and hence $f + \mathfrak{m}_x = g + \mathfrak{m}_x$, so it is injective as well. Now consider the function:

$$\begin{aligned} \varphi : X &\to \mathrm{mSpec}\,(R) \\ x &\mapsto \mathfrak{m}_x. \end{aligned}$$

We claim that φ is bijective. To see injectivity, suppose $\mathfrak{m}_x = \mathfrak{m}_y$ for $x, y \in X$. Then, we have that $R/\mathfrak{m}_x = R/\mathfrak{m}_y \cong \mathbb{R}$. This tells us that for each $f \in R$, $f(x) = f(y) \in \mathbb{R}$. Now assume that $x \neq y$. Since X is T_1, therefore $\{x\}, \{y\}$ are two disjoint closed subspaces of X. Then, by Urysohn's lemma (we have that X is compact Hausdorff), we get that there exists a continuous $\mathbb{R}$-valued function $f : X \to \mathbb{R}$ such that $f(x) = 0$ and $f(y) = 1$, a contradiction. Hence $x = y$.

Pick any maximal ideal $\mathfrak{m} \in \mathrm{mSpec}\,(R)$. We show that it is kernel of evaluation at some point. If not, then for all $x \in M$, there exists $f_x \in \mathfrak{m}$ such that $f_x(x) \neq 0$. As $f_x : M \to \mathbb{R}$ is continuous, therefore there exists an open $x \in U \subseteq M$ such that $f_x(y) \neq 0$ for all $y \in U_x$. We have thus obtained a cover of M by $\{U_x\}$. By shrinking each U_x if necessary, we may assume that $U_x \subseteq C_x \subseteq V_x$ where C_x is a compact set of M and V_x is open in M. It follows by compactness that there is a finite cover $M = \bigcup_{i=1}^n U_{x_i}$. As M is compact Hausdorff, therefore there exists smooth bump functions on each open U_{x_i}. Thus we have maps $\rho_i : M \to \mathbb{R}$ such that $\rho_i = 1$ on U_{x_i}. Consider then the map $g = \sum_{i=1}^n \rho_i f_{x_i}^2$. This is a global smooth map $g : M \to \mathbb{R}$ such that $g(x) = \sum_{i=1}^n \rho_i f_{x_i}^2(x) \neq 0$ as for any $x \in X$, there are finitely many U_{x_i} containing x on which atleast one of f_{x_i} is non-zero and ρ_i is 1. Hence g is invertible. As $f_{x_i}^2 \in \mathfrak{m}$, therefore $g \in \mathfrak{m}$ and hence $\mathfrak{m} = R$, a contradiction. Thus α is surjective.

2. Let us first establish that φ as in item 1 above is continuous. Indeed, let $I \leq R$ be an ideal

and $V(I) = \{\mathfrak{m} \in \text{mSpec}\,(R) \mid \mathfrak{m} \supseteq I\}$. A closed set of mSpec (R) looks exactly like above. We wish to show that $\varphi^{-1}(V(I))$ is closed in X. It is immediate to observe by item 1 that

$$\varphi^{-1}(V(I)) = \bigcap_{f\in I}\{x \in X \mid f(x) = 0\}.$$

Since $f : X \to \mathbb{R}$ is continuous, so it follows that $\varphi^{-1}(V(I))$ is closed. This shows the continuity of $\varphi : X \to \text{mSpec}\,(R)$. As X is compact and φ a bijective homeomorphism, it is thus sufficient to show that mSpec (R) is Hausdorff.

Fix two points $\mathfrak{m}_x \neq \mathfrak{m}_y$ in mSpec (R) for $x \neq y \in X$. Fix two opens U, V of X such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$. Let $C = X \setminus U$ and $D = X \setminus V$. Note that $C \cup D = X$. Now applying Urysohn's lemma on C, D yields $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ such that $f(C) = 0$, $f(D) = 1$ and $g(D) = 0$, $g(C) = 1$. Consequently, $fg = 0$ over X. Now consider the basic opens $D(f), D(g) \subseteq \text{mSpec}\,(R)$. As $f(x) \neq 0$ since $x \in D$, therefore $D(f) \ni x$. Similarly, $D(g) \ni y$. Since $D(f) \cap D(g) = D(fg) = D(0) = \emptyset$, therefore x and y can be separated, as required. □

Remark 1.1.0.2. An important corollary of the above result is that we can actually distinguish between the points of X by looking at maximal ideals of R; for $x, y \in X$, $x \neq y$ if and only if $\mathfrak{m}_x \neq \mathfrak{m}_y$. This is interesting because a fundamental goal of algebraic geometry is to study geometric properties of varieties over an algebraically closed field k and dominant maps between them. A fundamental equivalence tells that this is equivalent to studying the ring of regular functions over such a variety. Moreover, this ring recovers the important topology on the variety (there can be atleast two topologies on the variety if we are in, say $\mathbb{C}$). Hence one motivation to undergo this switch of viewpoint, where we try to do everything algebraically is that 1) we can completely recover the points of the variety and the relevant topology on it and that 2) we have a broad generalization of algebro-geometric techniques and constructions to an arbitrary commutative and unital ring R.

Caution 1.1.0.3. While in the sequel we will encounter spaces which are compact, it would rarely (unless you are interested in Boolean rings) be the case that the spaces will be Hausdorff. However, if one notices the way Hausdorff property is used in the above result, then one can see that if we somehow makes sure that the space X constructed out of a ring R is such that every point of X can be "distinguished" by functions on X in R, then you don't need Hausdorff property. This is precisely what will happen.

1.2 Affine schemes and basic properties

Let us first swiftly give an account of basic global constructions in scheme theory. The foundational philosophy of scheme theory is to handle a space completely by the ring of globally defined *nice* functions on it. This is taken to an unprecedented extreme by the definition of an affine scheme, which tells us that one can even do geometry on the base space by the knowledge of globally defined functions on the base space alone; you can indeed *reconstruct* the base space! So, we begin with a general ring R and construct a topological space Spec (R). The way we will define its points is by thinking of each point of this base space Spec (R) as that subset of R, each of whose function becomes zero at a common point. One then sees that these are exactly the prime ideals of

R. Hence, the base space Spec (R) is:

$$\text{Spec}\,(R) := \{\mathfrak{p} \subset R \mid \mathfrak{p} \text{ is a prime ideal of } R\}.$$

Next thing we wish to do is to actually get a *space* structure on this constructed base space, that is, a topology on Spec (R). This is, again, given with the help of the ring R. In particular, we give a topology on Spec (R) where every closed set is given by the zero locus of collections of functions $S \subseteq R$, that is, $V(S) := \{\mathfrak{p} \in \text{Spec}\,(R) \mid \mathfrak{p} \supseteq S\} = \{x \in \text{Spec}\,(R) \mid f(x) = 0\ \forall f \in S\}$ where the last equation tells one how to think about the definition of $V(S)$. This is known as Zariski topology on Spec (R) and is defined by the following:

$$A \subseteq \text{Spec}\,(R) \text{ is closed } \iff A = V(S) \text{ for some } S \subseteq R.$$

After defining the topology on Spec (R), one is interested in interested in understanding the set of all *germs* of functions at a point $\mathfrak{p} \in$ Spec (R). What are germs of functions at a point? Well, heuristically, they are all possible ways a function can *look* different at the given point. So for this, we have to atleast gather all those functions in R which takes different values at point $\mathfrak{p} \in$ Spec (R). Clearly this is given by the quotient domain $R/\mathfrak{p}$. Now from this, we construct the *residue field* of Spec (R) at point $\mathfrak{p}$, denoted $\kappa(\mathfrak{p}) := (R/\mathfrak{p})_{\langle 0 \rangle}$, that is, the fraction field of domain $R/\mathfrak{p}$. What does this $\kappa(\mathfrak{p})$ denotes geometrically? Well, it denotes the field of all different values a function can take at point $\mathfrak{p} \in$ Spec (R). Now, if that is the case, then one sees that if one takes any function $f \in R$, then "evaluating" f at $\mathfrak{p}$ should yield a point $f(\mathfrak{p})$ in $\kappa(\mathfrak{p})$. Indeed, we have the natural quotient maps:

$$R \to R/\mathfrak{p} \to \kappa(\mathfrak{p}).$$

So one should see

$\kappa(\mathfrak{p})$ as the field of possible values that a function $f \in R$ can take at point $\mathfrak{p}$.

However, we have not yet made the set of germs at a point $\mathfrak{p}$. The relation between two functions of having equal germs on R at a point $\mathfrak{p}$ is given by the heuristic that $f, g \in R$ should become equal in some open neighborhood around $\mathfrak{p}$. Since we have a topology on Spec (R), so one can actually do this formally. One will then see this that the set of all germs at point $\mathfrak{p}$ are actually all rational functions of R definable at $\mathfrak{p}$, that is, heuristically, f/g with $g(\mathfrak{p}) \neq 0$ for $f, g \in R$. This in our language turns out to be all the symbols of the form f/g with $g \notin \mathfrak{p}$. This is exactly the local ring $R_{\mathfrak{p}}$, the localization of the ring R (seen as ring of functions over Spec (R)) at the point $\mathfrak{p} \in$ Spec (R). So

germs of functions of R at $\mathfrak{p}$ is $R_{\mathfrak{p}}$.

We will expand more on this when we will talk about the structure sheaf of Spec (R).

Let us now see a basic but important dictionary between the topology of space Spec (R) and the algebra of ideals of R:

Lemma 1.2.0.1. *Let R be a ring. We then have the following:*

1. *If $\mathfrak{a}$, $\mathfrak{b}$ are two ideals of R, then $V(\mathfrak{a}\mathfrak{b}) = V(\mathfrak{a}) \cup V(\mathfrak{b})$.*

2. *If* $\{\mathfrak{a}_n\}$ *is a collection of ideals of* R*, then* $V\left(\sum_n \mathfrak{a}_n\right) = \bigcap_n V(\mathfrak{a}_n)$*.*
3. *If* $\mathfrak{a}, \mathfrak{b}$ *are two ideals of* R*, then* $V(\mathfrak{a}) \subseteq V(\mathfrak{b})$ *if and only if* $\sqrt{a} \supseteq \sqrt{b}$*.*

Proof. 1. First, let us see that $V(\mathfrak{ab}) \subseteq V(\mathfrak{a}) \cup V(\mathfrak{b})$. Take any $\mathfrak{p} \supseteq \mathfrak{ab}$. Suppose $p \notin V(\mathfrak{a})$ and $\mathfrak{p} \notin V(\mathfrak{b})$. Then there exists $f \in \mathfrak{a}$, $g \in \mathfrak{b}$ such that $fg \in \mathfrak{ab} \subseteq \mathfrak{p}$. Thus, $f \in \mathfrak{p}$ or $g \in \mathfrak{p}$, a contradiction in both cases. Second, it is easy to see that $V(\mathfrak{a}) \cup V(\mathfrak{b}) \subseteq V(\mathfrak{ab})$ as if either $p \supseteq \mathfrak{a}$ or $\mathfrak{b} \subseteq \mathfrak{p}$, then since $\mathfrak{ab} \subseteq \mathfrak{a} \cap \mathfrak{b} \subseteq \mathfrak{a}$, therefore $\mathfrak{ab} \subseteq \mathfrak{p}$.
2. Let $\mathfrak{p} \supseteq \sum_n \mathfrak{a}_n$. Since ideals are abelian groups so the sum contains each $\mathfrak{a}_n$, hence $\mathfrak{p} \supseteq \mathfrak{a}_n$ for each n, and so $\mathfrak{p} \in \bigcap_n V(\mathfrak{a}_n)$. Conversely, if $\mathfrak{p} \supseteq \mathfrak{a}_n$ for each n, then $\mathfrak{p} = \sum_n \mathfrak{p} \supseteq \sum_n \mathfrak{a}_n$.
3. (L $\implies$ R) Since each prime ideal containing $\mathfrak{a}$ also contains $\mathfrak{b}$, therefore the intersection of all prime ideals containing $\mathfrak{a}$ will contain the intersection of all prime ideals containing $\mathfrak{b}$.
(R $\implies$ L) Take any prime ideal $\mathfrak{p} \supseteq \mathfrak{a}$. Since $\sqrt{a} \supseteq \sqrt{b}$, therefore $\mathfrak{p} \supseteq \mathfrak{b}$. □

1.2.1 Topological properties of Spec (R)

Let us begin by an algebraic characterization of irreducible closed subspaces of Spec (R).

Lemma 1.2.1.1. *Let* R *be a ring and* $X \hookrightarrow \operatorname{Spec}(R)$ *be a closed subspace. Then the following are equivalent:*

1. X *is irreducible.*
2. *There is a unique point* $\mathfrak{p} \in \operatorname{Spec}(R)$ *such that* $X = V(\mathfrak{p})$*.*

One calls the point $\mathfrak{p}$ *the generic point of the irreducible closed subspace* X[1]*.*

Proof. (1. $\Rightarrow$ 2.) Since X is closed therefore $X = V(\mathfrak{a})$ for some ideal $\mathfrak{a}$ of R. If we assume that $X \neq V(\mathfrak{p})$ for each prime $\mathfrak{p} \subseteq R$, then this holds true for points $\mathfrak{p} \in X$ as well. Hence take $\mathfrak{p} \in X$ and consider the proper closed subset $V(\mathfrak{p}) \subsetneq X$. Let $\mathfrak{q} \notin V(\mathfrak{p})$. Then, $V(\mathfrak{q}) \subsetneq X$ as well. Hence we get that $V(\mathfrak{p}) \cup V(\mathfrak{q}) = V(\mathfrak{a})$, which stands in contradiction to the fact that X is irreducible. Hence there exists a prime $\mathfrak{p} \in \operatorname{Spec}(R)$ such that $X = V(\mathfrak{p})$. Uniqueness is quite clear.
(2. $\Rightarrow$ 1.) Suppose $Y = V(\mathfrak{a})$ and $Z = V(\mathfrak{b})$ are two closed subspaces of $X = V(\mathfrak{p})$ such that $X = Y \cup Z = V(\mathfrak{a}) \cup V(\mathfrak{b}) = V(\mathfrak{a} \cap \mathfrak{b}) = V(\mathfrak{ab})$ (Lemma 1.2.0.1). Assume that Y, Z are proper inside X. Then, there are two points $\mathfrak{q}_1 \in Y \setminus Z$ and $\mathfrak{q}_2 \in Z \setminus Y$. Algebraically, this is equivalent to saying that $\mathfrak{q}_1 \supseteq \mathfrak{a}$, $\mathfrak{q}_1 \not\supseteq \mathfrak{b}$ and $\mathfrak{q}_2 \supseteq \mathfrak{b}$, $\mathfrak{q}_2 \not\supseteq \mathfrak{a}$. It follows that $\mathfrak{q}_1 \cap \mathfrak{q}_2$ is also a prime ideal which contains $\mathfrak{a} \cap \mathfrak{b} = \mathfrak{ab}$. Since $X = V(\mathfrak{ab}) = V(\mathfrak{p}) \ni \mathfrak{p}$, hence it follows that $\mathfrak{q}_1 \cap \mathfrak{q}_2 \supseteq \mathfrak{p}$ as it already contains $\mathfrak{ab}$. Thus $\mathfrak{q}_1 \cap \mathfrak{q}_2 \in V(\mathfrak{a}) \cap V(\mathfrak{b}) \subseteq V(\mathfrak{p})$. Since $V(\mathfrak{a}) \cap V(\mathfrak{b}) = V(\mathfrak{a} + \mathfrak{b})$, hence it follows that $\mathfrak{q}_1 \cap \mathfrak{q}_2 \supseteq \mathfrak{a}, \mathfrak{b}$, which implies in particular that $\mathfrak{q}_1 \supseteq \mathfrak{a}, \mathfrak{b}$, a contradiction. □

Remark 1.2.1.2. The main idea of the above proof has been to first translate the topological condition to algebraic, and then using the critical observation that the closed subspace $V(\mathfrak{p})$ contains point $\mathfrak{p}$ itself.

A simple corollary of above gives all closed points of an affine scheme.

Lemma 1.2.1.3. *Let* R *be a ring. Then*

$$\{\textit{Closed points of } \operatorname{Spec}(R)\} \cong \{\textit{Maximal ideals of } R.\}$$

Proof. Follows immediately from Lemma 1.2.1.1. □

[1] Such spaces where every irreducible closed set has a unique generic point are called sober spaces.

Let us next observe a simple but important observation about topology of $\mathrm{Spec}\,(R)$.

Lemma 1.2.1.4. *Let R be a ring. For $f \in R$, define $\mathrm{Spec}\,(R)_f := \{\mathfrak{p} \in \mathrm{Spec}\,(R) \mid f \notin \mathfrak{p}\}$. Then,*

1. $\mathrm{Spec}\,(R)_f \hookrightarrow \mathrm{Spec}\,(R)$ *is an open set and such open sets form a basis of the Zariski topology on* $\mathrm{Spec}\,(R)$.
2. $\mathrm{Spec}\,(R)_f \hookrightarrow \mathrm{Spec}\,(R)_g$ *if and only if $f \in \sqrt{Rg}$.*

Proof. 1. Clearly $\mathrm{Spec}\,(R)_f = \mathrm{Spec}\,(R) \setminus V(f)$ where we know that $V(f) = \{\mathfrak{p} \in \mathrm{Spec}\,(R) \mid f \in \mathfrak{p}\}$. Hence X_f is open. It is also clear that if $U \subseteq \mathrm{Spec}\,(R)$ is open, then $\mathrm{Spec}\,(R) \setminus U = V(\mathfrak{a})$ is closed and hence $U = \bigcup_{f\in\mathfrak{a}} \mathrm{Spec}\,(R)_f$. Further, $\mathrm{Spec}\,(R) = \mathrm{Spec}\,(R)_1$ and $\emptyset = \mathrm{Spec}\,(R)_0$.

2. This follows from the following equivalences. Let $\mathrm{Spec}\,(R)_f \hookrightarrow \mathrm{Spec}\,(R)_g$, then we get the following (we implicitly use Hilbert Nullstellensatz)

$$\mathrm{Spec}\,(R)_f \hookrightarrow \mathrm{Spec}\,(R)_g \iff f(\mathfrak{p}) \neq 0 \implies g(\mathfrak{p}) \neq 0 \iff g(\mathfrak{p}) = 0 \implies f(\mathfrak{p}) = 0 \iff V(g) \subseteq V(f)$$
$$\iff \sqrt{Rg} \supseteq \sqrt{Rf} \supseteq Rf \iff f \in \sqrt{Rg}.$$

This completes the proof. □

Next we observe the equivalent formulation of partitions of unity in the context of algebra.

Lemma 1.2.1.5. *Let R be a ring. Then,*

1. *If $U \hookrightarrow \mathrm{Spec}\,(R)$ is any open set given by $U = \bigcup_{f\in S} \mathrm{Spec}\,(R)_f$ for some subset $S \subseteq R$, then*
$$\mathrm{Spec}\,(R) \setminus U = V\left(\sum_{f\in S} Rf\right).$$
2. $\mathrm{Spec}\,(R) = \bigcup_{f\in S} \mathrm{Spec}\,(R)_f$ *for some $S \subseteq R$ if and only if the ideal of R generated by S is the whole of R.*

Proof. 1. Let $U \hookrightarrow \mathrm{Spec}\,(R)$ be an open set. Then, $\mathfrak{p} \in \mathrm{Spec}\,(R) \setminus U \iff \mathfrak{p} \notin U \iff \forall f \in S,\ \mathfrak{p} \notin \mathrm{Spec}\,(R)_f \iff \forall f \in S,\ f \in \mathfrak{p} \iff \mathfrak{p} \supseteq S \iff \mathfrak{p} \in V(S)$.
2. Follows from 1. □

We next have an interesting observation that $\mathrm{Spec}\,(R)$ are always quasicompact[2].

Lemma 1.2.1.6. *Let R be a ring. Then $\mathrm{Spec}\,(R)$ is quasicompact.*

Proof. Take any arbitrary basic open cover $\bigcup_{f\in S} \mathrm{Spec}\,(R)_f$ for some $S \subseteq R$. Then by Lemma 1.2.1.5, 2, we get that $\sum_{f\in S} Rf \ni 1$ and hence there are $f_1, \dots, f_n \in S$ such that $g_1 f_1 + \dots g_n f_n = 1$ for some $g_i \in R$. Hence $\mathrm{Spec}\,(R) \setminus \bigcup_{i=1}^n = V(f_1, \dots, f_n) = V(R) = \emptyset$. □

Next, we see the topological effects on space $\mathrm{Spec}\,(R)$ of Noetherian hypothesis on ring R. In particular, we see that the space $\mathrm{Spec}\,(R)$ itself becomes *noetherian topological space*, that is, it's closed sets satisfies descending chain condition.

[2]it is customary in algebraic geometry to call the topological compactness as quasi-compactness; compactness in algebraic geometry historically means Hausdorff *and* topological compactness.

Lemma 1.2.1.7. *Let R be a ring. If R is noetherian, then* Spec (R) *is noetherian.*

Proof. Use $V(-)$ and $I(-)$, where $I(Y) = \{f \in R \mid f \in \mathfrak{p} \forall \mathfrak{p} \in Y\}$. Rest is trivial. □

We next discuss few things about the irreducible subsets of a closed set of Spec (R). Let $F \hookrightarrow$ Spec (R) be a closed subset. Then we can contemplate irreducible subsets of F. Clearly, each irreducible subset has to be in a maximal irreducible subset, which are called *irreducible components* of Spec (R). We have few basic observations about irreducible components.

Lemma 1.2.1.8. *Let R be a ring and F be a closed subset of* Spec (R)*. Then,*

1. *Each irreducible component of F is closed.*
2. *If R is noetherian, then there are only finitely many irreducible components of* Spec (R)*.*
3. *We have that*

$$\{\textit{Irreducible components of } \operatorname{Spec}(R)\} = \{\textit{Closed sets } V(\mathfrak{p}),\ \mathfrak{p} \textit{ is minimal prime}\}.$$

Proof. Statement 1. follows from Lemma 1.2.1.1. Statement 2. follows from Lemma 1.2.1.7 and the fact that a noetherian topological space has only finitely many irreducible components. We now show statement 3. If Z is an irreducible component, then it is closed and $Z = V(\mathfrak{p})$ by Lemma 1.2.1.1. We claim that $\mathfrak{p}$ is a minimal prime. If not, then as every prime has a minimal prime, we will have $\mathfrak{p}' \subsetneq \mathfrak{p}$ such that $\mathfrak{p}'$ is minimal. Consequently, we get $V(\mathfrak{p}') \supsetneq V(\mathfrak{p})$. An another use of Lemma 1.2.1.1 yields that $V(\mathfrak{p}')$ is irreducible. But $V(\mathfrak{p})$ was irreducible component, giving a contradiction. We deduce that $\mathfrak{p}$ is a minimal prime, as required.

Conversely, if $\mathfrak{p}$ is minimal, then $V(\mathfrak{p})$ is an irreducible closed set which cannot be contained in a larger irreducible closed set as otherwise we will have $V(\mathfrak{p}') \supsetneq V(\mathfrak{p})$ and thus, $\sqrt{\mathfrak{p}'} \subsetneq \sqrt{\mathfrak{p}}$ (Lemma 1.2.0.1), but as the ideals are prime, so $\mathfrak{p}' \supsetneq \mathfrak{p}$, a contradiction to minimality. □

Note that we are already in a position to prove some algebraic statements using topological arguments, as the following lemma shows.

Lemma 1.2.1.9. *Let A be a ring and let $a_1, \dots, a_n \in A$ generate the unit ideal in A. Then for all $m > 0$, the collection $a_1^m, \dots, a_n^m \in A$ also generates the unit ideal in A.*

Proof. From Lemma 1.2.1.6, 2, it follows that $\{D(a_i)\}_{i=1,\dots,n}$ covers Spec (A). Since for any $a \in A$, the basic open $D(a) \subseteq$ Spec (A) is equal to $D(a^m)$ as a prime $\mathfrak{p}$ doesn't contain a if and only if it doesn't contain any of its power. Consequently, we get that $\{D(a_i^m)\}_{i=1,\dots,n}$ also forms a basic open cover of Spec (A). An application of Lemma 1.2.1.6, 2 again proves the result. □

1.2.2 The structure sheaf $\mathcal{O}_{\operatorname{Spec}(R)}$

The next important thing we want to consider on Spec (R) is a sheaf of suitable nice functions over it. This sheaf will be of utmost importance as it will not be treated as an additional structure, but will be an integral part (in-fact, the most important part) of the definition of an affine scheme.

The question now is, *what are* nice functions over Spec (R) whose sheaf we should take. We turn to classical algebraic varieties for that (one may skip the following if he/she find himself/herself to be brave enough to face the abstraction of the structure sheaf). See Section 1.5 for more details.

Example 1.2.2.1. (*Structure sheaf of an algebraic variety*) Let k be an algebraically closed field. An important aspect of varieties is their morphism. We will display this only in the affine case. Let X, Y be two affine varieties. To define a morphism between X and Y, we would first need to understand the notion of *regular functions* over any variety X. A function $\varphi : X \to k$ is said to be regular if it is locally rational. That is, for each $p \in X$, there exists an open set $U \ni p$ of X and there exists two polynomials $f, g \in k[x_1, \dots, x_n]$ such that $g(q) \neq 0 \forall q \in U$ and $\varphi|_U = f/g$. It then follows that a regular function is continuous when X and k are equipped with its Zariski topology (Lemma 3.1, [**??**] [Hartshorne]). We now define morphism of affine varieties.

A function $\varphi : X \to Y$ is said to be a morphism of varieties if

1. $\varphi : X \to Y$ is continuous,
2. for each open set $V \subseteq Y$ and a regular map $f : V \to k$, the map $f \circ \varphi$ as below

$$\begin{tikzcd} & V \arrow[dr, "f"] & \\ \varphi^{-1}(V) \arrow[ur, "\varphi"] \arrow[rr, "f\circ\varphi"'] & & k \end{tikzcd}$$

   is also a regular map.

Hence the main part of the data of a variety is the locally defined regular maps. This is what we will take as our motivation in defining the structure sheaf over Spec (R), as this example tells us to take care of these local functions to the base field. A question that may arise from this discussion is how are we going to define a regular map from an open set $U \hookrightarrow$ Spec (R) when we don't even have a field. The answer is, as we discussed previously, to work with residue field at a point instead.

We now start to define the structure sheaf of Spec (R). First, let us give the following lemma, which reduces the burden of construction only to basis elements of Spec (R).

Lemma 1.2.2.2. *Let X be a topological space and $\mathcal{B}$ be a basis. Let F be an assignment over sets of $\mathcal{B}$ which satisfies sheaf conditions for it. Then, F extends to a sheaf $\mathcal{F}$ over X.*

Proof. The main observation here is that we can find the stalk of $\mathcal{F}$ at each point x just by the knowledge of F, because of the basis $\mathcal{B}$. Take any point $x \in X$. We see that we can get the stalk $\mathcal{F}_x$ as follows:

$$\mathcal{F}_x := \varinjlim_{x \in B \in \mathcal{B}} F(B).$$

Once we have the stalks, we can define the sections of $\mathcal{F}$ quite easily as follows. Let $U \subseteq X$ be an open set. Then $\mathcal{F}(U)$ is defined to be the subset of $\prod_{x \in U} \mathcal{F}_x$ of those elements (s_x) where there exists a basic open cover $\{B_i\}$ of U and there exists elements $s_i \in F(B_i)$ such that $s_x = (s_i)_x$ for each $x \in B_i$. One can check that this satisfies the conditions of a sheaf. □

Construction 1.2.2.3. (*The* $\mathcal{O}_{\text{Spec}(R)}$) Let R be a ring. By virtue of Lemma 1.2.2.2, we will define $\mathcal{O}_{\text{Spec}(R)}$ only on basic open sets of the form Spec $(R)_f$. Let $X :=$ Spec (R). Motivated by Example 1.2.2.1, take a basic open set $X_f \hookrightarrow X$ for some $f \in R$ and then we wish to consider *rational functions* over X_f. This means those functions of the form g/h for $g, h \in R$ such that $h(\mathfrak{p}) \neq 0 \forall \mathfrak{p} \in$

X_f. This is equivalent to demanding that $h \notin \mathfrak{p} \forall \mathfrak{p} \in X_f$, that is, $X_h \supseteq X_f$. This is again equivalent to stating that $f \in \sqrt{Rh}$ by Lemma 1.2.1.4, 2. Hence $f^n = ah$ for some $n \in \mathbb{N}$ and $a \in R$. Thus, we see that the notion of rational functions over X_f is equivalent to all functions of the form g/f^n where $g \in R$ and $n \in \mathbb{N}$. Commutative algebra has an apt name for this, that is, the localization of R at f denoted by $R_f := \{a/f^n \mid a \in R,\ n \in \mathbb{N}\}$ which is again a ring by natural operation on fractions (see Special Topics, **??**). Thus, we should define the sections over X_f as:

$$\mathcal{O}_X(X_f) := R_f.$$

We would not verify the sheaf axioms here as it is a tedious but straightforward calculation. The sheaf $\mathcal{O}_X$ thus formed is called the structure sheaf on the space X. One should think of the sheaf $\mathcal{O}_X$ as natural as the ring R itself. In particular we will see in the next section that it indeed is the case.

Next, we would like to see the stalks of this sheaf $\mathcal{O}_X$. To understand this, we would have to understand the maps on sections induced by $X_f \hookrightarrow X_g$. As we saw earlier, this is equivalent to stating that $f^n = ag$ for some $n \in \mathbb{N}$ and $a \in R$. Hence, the induced map on sections are the restriction maps of the sheaf and is given by

$$\begin{aligned} \rho_{X_g,X_f} : R_g = \mathcal{O}_X(X_g) &\longrightarrow \mathcal{O}_X(X_f) = R_f \\ b/g^m &\longmapsto ba^m/a^m g^m = ba^m/f^{nm}. \end{aligned}$$

We are now ready to calculate the stalk. Take any point $x \in X$. The stalk becomes:

$$\begin{aligned} \mathcal{O}_{X,x} &:= \varinjlim_{x \in X_f} \mathcal{O}_X(X_f) \\ &= \varinjlim_{x \in X_f} R_f \\ &= \varinjlim_{f \notin x} R_f \\ &= R_x \end{aligned}$$

where the last equality follows from a small colimit calculation (which should really be thought of as a definition). Hence $\mathcal{O}_X$ is a sheaf whose stalks are local rings. So we have a complete description of the sheaf $\mathcal{O}_X$ when $X = \operatorname{Spec}(R)$.

We finally define an affine scheme.

Definition 1.2.2.4. (**Affine scheme**) Let R be a ring. Then the pair $(\operatorname{Spec}(R), \mathcal{O}_{\operatorname{Spec}(R)})$ is called an affine scheme.

Remark 1.2.2.5. (*Evaluation of functions*) Let $(\operatorname{Spec}(R), \mathcal{O}_{\operatorname{Spec}(R)})$ be an affine scheme. As noted earlier, we now see how all rational functions over $\operatorname{Spec}(R)$ are exactly the elements of R. In particular, since $\Gamma(\operatorname{Spec}(R), \mathcal{O}_{\operatorname{Spec}(R)}) = R_1 = R$. Hence if we interpret $\mathcal{O}_{\operatorname{Spec}(R)}$ as the sheaf of regular maps over $\operatorname{Spec}(R)$, then R itself appears as the globally defined regular maps.

Now take global map $f \in R$ and any point $\mathfrak{p} \in \operatorname{Spec}(R)$. We can "evaluate" f at $\mathfrak{p}$ via the following composite (note that $\kappa(\mathfrak{p}) = R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}} \cong (R_{\mathfrak{p}})_o$, the last one is the fraction field of $R_{\mathfrak{p}}$ obtained by localizing at 0 ideal):

$$\Gamma(\operatorname{Spec}(R), \mathcal{O}_{\operatorname{Spec}(R)}) \longrightarrow \mathcal{O}_{\operatorname{Spec}(R),\mathfrak{p}} \longrightarrow \kappa(\mathfrak{p})$$

where the first map on the left is the inclusion into the direct limit and the map on right is the natural quotient map. Algebraically, we have the following maps

$$R \longrightarrow R_{\mathfrak{p}} \longrightarrow R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$$

given by

$$f \longmapsto \frac{f}{1} \longmapsto \frac{f}{1} + \mathfrak{p}R_{\mathfrak{p}},$$

where $f/1 + \mathfrak{p}R_{\mathfrak{p}}$ denotes the class of all those functions in the stalk $\mathcal{O}_{\operatorname{Spec}(R),\mathfrak{p}} = R_{\mathfrak{p}}$ which takes same value at $\mathfrak{p}$ as f does.

For completeness' sake, we give a description of the section of the sheaf $\mathcal{O}_{\operatorname{Spec}(R)}$ on any open set $U \subseteq \operatorname{Spec}(R)$.

Lemma 1.2.2.6. *Let R be a ring and $(\operatorname{Spec}(R), \mathcal{O}_{\operatorname{Spec}(R)})$ the associated affine scheme. Let $U \subseteq \operatorname{Spec}(R) =: X$ be an open set. Then,*

$$\mathcal{O}_X(U) = \left\{ (s_{\mathfrak{p}}) \in \prod_{\mathfrak{p}\in U} R_{\mathfrak{p}} \mid \forall \mathfrak{p} \in U, \exists \textit{ basic open } X_g \ni \mathfrak{p} \;\&\; f/g^n \in R_g \textit{ s.t. } s_{\mathfrak{q}} = f/g^n \forall \mathfrak{q} \in X_g \right\}.$$

More concretely, we have

$$\mathcal{O}_X(U)=\left\{s:U\to\coprod_{\mathfrak{p}\in U} R_{\mathfrak{p}} \mid \forall\mathfrak{p}\in U,\ s(\mathfrak{p})\in R_{\mathfrak{p}} \;\&\; \exists \textit{ open } \mathfrak{p}\in V\subseteq U \;\&\; f,g\in R \textit{ s.t. } \forall\mathfrak{q}\in V,\ g\notin\mathfrak{q} \;\&\; s(\mathfrak{q})=f/g\right\}.$$

Proof. Follows from Lemma 1.2.2.2 and Construction 1.2.2.3. □

Ring morphisms and $\operatorname{Spec}(-)$

We now discuss some properties of ring morphisms and the associated map of affine schemes.

Lemma 1.2.2.7. [3] *Let A be a ring and $f \in A$. Then, $D(f) \subseteq \operatorname{Spec}(A)$ is empty if and only if f is nilpotent.*

Proof. Both sides follow immediately from the Lemma 16.1.2.9. □

We further obtain the following two results which corresponds to what happens on the level of sheaves.

Proposition 1.2.2.8. [4] *Let $X = \operatorname{Spec}(A)$ and $Y = \operatorname{Spec}(B)$ be two affine schemes and $\varphi : A \to B$ be a morphism of rings.*

[3]Exercise II.2.18, a of Hartshorne.

[4]Exercise II.2.18 b,c,d of Hartshorne.

1. *The ring map $\varphi : A \to B$ is injective if and only if the corresponding map of schemes $f : Y \to X$ yields injective map of structure sheaves, that is, $f^\flat : \mathcal{O}_X \to f_*\mathcal{O}_Y$ is injective.*
2. *If $\varphi : A \to B$ is injective, then $f : Y \to X$ is dominant*[5].
3. *The ring map $\varphi : A \to B$ is surjective if and only if the corresponding map of schemes $f : Y \to X$ is a closed immersion.*

Proof. 1. (L $\Rightarrow$ R) It suffices to show that $f^\flat$ is an injective map over basic opens of X. Pick any $g \in A$ and consider the basic open $D(g) \subseteq X$. We wish to show that the map

$$f^\flat_{D(g)} : \mathcal{O}_X(D(g)) \longrightarrow \mathcal{O}_Y(f^{-1}(D(g)))$$

is an injective homomorphism. Indeed, we first observe that $\mathcal{O}_X(D(g)) \cong A_g$ and $f^{-1}(D(g)) = D(\varphi(g))$, so that $\mathcal{O}_Y(D(\varphi(g))) \cong B_{\varphi(g)}$. It follows that the map $f^\flat_{D(g)} : A_g \to B_{\varphi(g)}$ is the localization map

$$\begin{aligned} \varphi_g : A_g &\longrightarrow B_{\varphi(g)} \\ \frac{a}{g^n} &\longmapsto \frac{\varphi(a)}{\varphi(g)^n}. \end{aligned}$$

We wish to show that the above map is injective. If $\varphi(a)/\varphi(g)^n = 0$, then for some $k \in \mathbb{N}$ we have $\varphi(g)^k\varphi(a) = 0$. It follows by injectivity of φ that $g^k a = 0$ in A. Consequently, we can write

$$\frac{a}{g^n} = \frac{ag^k}{g^{n+k}} = 0.$$

(R $\Rightarrow$ L) As a sheaf map is injective if and only if the kernel sheaf is zero (Theorem 20.3.0.7), where the latter is equivalent to the fact that every map on sections is injective. Consequently, over X, we get

$$f^\flat_X : \Gamma(\mathcal{O}_X, X) \longrightarrow \Gamma(\mathcal{O}_Y, Y)$$

Since $\Gamma(\mathcal{O}_X, X) \cong A$ and $\Gamma(\mathcal{O}_Y, Y) \cong B$, and the map $f^\flat_X : A \to B$ is just φ itself, therefore we are done.

2. We wish to show that for any basic non-empty open $D(g) \subseteq X$ for $g \in A$, the intersection $D(g) \cap f(Y)$ is non-empty. We have the following equalities:

$$\begin{aligned} D(g) \cap f(Y) &= \{\mathfrak{p} \in X \mid \mathfrak{p} \in f(Y) \ \&\ g \notin \mathfrak{p}\} \\ &= \{\varphi^{-1}(\mathfrak{q}) \in X \mid \mathfrak{q} \in Y,\ g \notin \varphi^{-1}(\mathfrak{q})\} \\ &= \{\varphi^{-1}(\mathfrak{q}) \in X \mid \mathfrak{q} \in Y,\ \varphi(g) \notin \mathfrak{q}\} \\ &= f(D(\varphi(g))). \end{aligned}$$

Conequently, $D(g) \cap f(Y)$ is non-empty if and only if $D(\varphi(g))$ is non-empty, which in turn implies by Lemma 1.2.2.7 that $D(g) \cap f(Y)$ is non-empty if and only if $\varphi(g)$ is not nilpotent. As g is not nilpotent because $D(g)$ is not empty, therefore $\varphi(g)$ is not nilpotent as φ is injective.

[5] that is, f has dense image.

3. (L $\Rightarrow$ R) Let $\varphi : A \to B$ be surjective and $I \leq A$ be the kernel. We wish to show that $f : Y \to X$ is a closed immersion. For that, we first need to show that f is a topological closed immersion, that is its image is closed and is homeomorphic to it. We claim that $f(Y) = V(I) \subseteq X$. Indeed, for any $\varphi^{-1}(\mathfrak{q}) \in f(Y)$, we have that $I \subseteq \varphi^{-1}(\mathfrak{q})$. Thus, $f(Y) \subseteq V(I)$. Conversely, for any $\mathfrak{p} \in V(I)$, as φ is surjective and $\mathfrak{p}$ contains I, therefore $\varphi(\mathfrak{p}) \in Y$ is a prime ideal such that $\varphi^{-1}(\varphi(\mathfrak{p})) = \mathfrak{p}$, so that $\mathfrak{q} = \varphi(\mathfrak{p}) \in Y$ is such that $f(\varphi(\mathfrak{p})) = \mathfrak{p}$, hence $\mathfrak{p} \in f(Y)$.

Next, we wish to show that f is homeomorphic to its image. It suffices to show that $f : Y \to f(Y)$ is a closed mapping. But this is immediate by the fact that a surjective map $\varphi : A \to B$ with kernel I induces an order preserving isomorphism of ideals of A containing I and ideals of B by mapping ideals of B to those of A containing I via φ^{-1}. Alternatively, one can see that $A/I \cong B$ and Spec $(A/I) \cong V(I) = f(Y)$, therefore application of Spec $(-)$ functor would do the job.

Next, we wish to show that $f^\flat : \mathcal{O}_X \to f_*\mathcal{O}_Y$ is surjective. We can check this on a basis of X. Let $D(g) \subseteq X$ for some $g \in A$. Indeed, for $t \in (f_*\mathcal{O}_Y)(D(g)) = \mathcal{O}_Y(D(\varphi(g)) \cong B_{\varphi(g)}$, we wish to find an open covering of $D(g)$ say U_i and $s_i \in \mathcal{O}_X(U_i)$ such that $f^\flat_{U_i}(s_i) = t|_{U_i}$ for each i. Indeed, the open set $D(g)$ as its own covering will suffice here as $\mathcal{O}_X(D(g)) \cong A_g$ and the map $f^\flat_{D(g)} = \varphi_g : A_g \to B_{\varphi(g)}$. As φ is surjective, therefore for $t = b/\varphi(g)^n \in B_{\varphi(g)}$, we obtain $a \in A$ such that $\varphi(a) = b$ and thus a/g^n is mapped by φ_g to $b/\varphi(g)^n$, as required.

(R $\Rightarrow$ L) Let $f : Y \to X$ be a closed immersion. We wish to show that $\varphi : A \to B$ is surjective. Pick $b \in B$. We wish to show that there exists $a \in A$ such that $\varphi(a) = b$. As the sheaf map $f^\flat : \mathcal{O}_X \to f_*\mathcal{O}_Y$ is surjective, therefore there exists a basic open covering (which will be finite by quasi-compactness of affine schemes, Lemma 1.2.1.6) namely $\{D(a_i)\}_{i=1,\dots,n}$ of X together with sections $s_i \in \mathcal{O}_X(D(a_i))$ such that $f^\flat_{D(a_i)}(s_i) \in \mathcal{O}_Y(f^{-1}(D(a_i)))$ is the restriction of $b \in \Gamma(\mathcal{O}_Y, Y)$ to $D(\varphi(a_i))$, namely $\rho_{X,D(\varphi(a_i))}(b)$. As we have $\mathcal{O}_X(D(a_i)) \cong A_{a_i}$, $\mathcal{O}_Y(f^{-1}(D(a_i))) = \mathcal{O}_Y(D(\varphi(a_i))) \cong B_{\varphi(a_i)}$ and that the restriction $\rho_{Y,D(\varphi(a_i))} : \Gamma(\mathcal{O}_X, X) \to \mathcal{O}_X(D(a_i))$ is just the natural localization map $A \to A_{a_i}$, therefore we may identify $s_i = \frac{c_i}{a_i^{k_i}} \in A_{a_i}$ and $\rho_{X,D(a_i)}(b) = \frac{b}{1} \in B_{\varphi(a_i)}$. Consequently, we have for each $i = 1, \dots, n$ the following equation in $B_{\varphi(a_i)}$

$$\frac{b}{1} = \frac{\varphi(c_i)}{\varphi(a_i)^{k_i}}.$$

It follows that we obtain an equation of the form

$$\varphi(a_i^{m_i})b = \varphi(c_i a^{l_i})$$

for some $m_i, l_i \geq 0$. Taking $M = \max_i m_i$, we obtain

$$\varphi(a_i^m)b = \varphi(d_i) \tag{$*$}$$

for some $d_i \in A$.

, 2, the collection $\{a_i\}_{i=1,\dots,n}$ generates the unit ideal in A. By Lemma 1.2.1.9, it follows that the collection $\{a_i^m\}_{i=1,\dots,n}$ also generates the unit ideal in A. Consequently, we have $r_1 a_1^m + \dots + r_n a_n^m = 1$ for some $r_i \in A$. Using this in $(*)$, we yield

$$b = \varphi\left(\sum_{i=1}^{n} r_i d_i\right),$$

as required[6]. □

1.2.3 $\mathcal{O}_{\mathrm{Spec}(R)}$-modules

As we pointed out in Construction 1.2.2.3, the structure sheaf $\mathcal{O}_{\mathrm{Spec}(R)}$ should really be thought of as natural as the ring R itself. This way of thought will be justified in this section, where we will see that, just like we can understand a ring by understanding the category of R-modules, we can understand the structure sheaf $\mathcal{O}_{\mathrm{Spec}(R)}$ by understanding the category of soon to be constructed $\mathcal{O}_{\mathrm{Spec}(R)}$-modules.

Let R be a ring and M be an R-module. Just like we underwent a "geometrification" to go from ring R (algebra) to the locally ringed space Spec (R) (geometry), we will also "geometrify" the notion of an R-module. This will yield us a sheaf $\widetilde{M}$ over Spec (R).

Definition 1.2.3.1. ($\widetilde{M}$) Let R be a ring and M be an R-module. The following presheaf on $X :=$ Spec (R) generated by the following definition on basic opens

$$X_f \longmapsto \widetilde{M}(X_f) := M_f = M \otimes_R R_f$$

and restrictions given by

$$(X_f \hookrightarrow X_g) \longmapsto M \otimes_R R_g \overset{\mathrm{id}\otimes\rho_{X_g,X_f}}{\to} M \otimes_R R_f$$

defines a unique sheaf on Spec (R) corresponding to R-module M denoted $\widetilde{M}$.

The above construction gives the sheaf $\widetilde{M}$ over R a structure of an $\mathcal{O}_{\mathrm{Spec}(R)}$-module, that is, a sheaf $\mathcal{F}$ of abelian groups where for each open $U \subseteq$ Spec (R) the group $\mathcal{F}(U)$ is a $\mathcal{O}_{\mathrm{Spec}(R)}(U)$-module. Since $\widetilde{M}(X_f) = M \otimes_R R_f$ is an $\mathcal{O}_X(X_f) = R_f$-module, therefore $\widetilde{M}$ are basic examples of $\mathcal{O}_{\mathrm{Spec}(R)}$-modules.

A map $\eta : \mathcal{F} \to \mathcal{G}$ of $\mathcal{O}_{\mathrm{Spec}(R)}$-modules is just a sheaf morphism where for each inclusion $U \hookrightarrow V$ of Spec (R), we get that the following commutes

$$\begin{array}{ccc} \mathcal{F}(V) & \xrightarrow{\eta_V} & \mathcal{G}(V) \\ \downarrow & & \downarrow \\ \mathcal{F}(U) & \xrightarrow[\eta_U]{} & \mathcal{G}(U) \end{array}$$

where the top horizontal map is a $\mathcal{O}_{\mathrm{Spec}(R)}(V)$-module homomorphism, bottom horizontal is a $\mathcal{O}_{\mathrm{Spec}(R)}(U)$-module homomorphism and the verticals are the restriction map of sheaves $\mathcal{F}$ and $\mathcal{G}$, which are also module homomorphisms w.r.t. $\mathcal{O}_{\mathrm{Spec}(R)}(V) \to \mathcal{O}_{\mathrm{Spec}(R)}(U)$. The latter has the following meaning. If M is an R-module and N is an S-module, then a map $\phi : M \to N$ is a module homomorphism w.r.t $f : R \to S$ if $\phi(r \cdot m) = f(r) \cdot \phi(m)$.

[6] Note that in the whole proof, we didn't even required the fact that $f : Y \to X$ is also a topological closed immersion!

We thus get a functor

$$\begin{aligned}\widetilde{-} : \mathbf{Mod}(R) &\longrightarrow \mathbf{Mod}(\mathcal{O}_{\mathrm{Spec}(R)})\\ M &\longmapsto \widetilde{M}\\ f : M \to N &\longmapsto \tilde{f} : \widetilde{M} \to \widetilde{N}\end{aligned}$$

where $\tilde{f}_{X_f} : M_f \to N_f$ is given by localization. We may denote $\widetilde{\mathbf{Mod}}(\mathcal{O}_{\mathrm{Spec}(R)}) \hookrightarrow \mathbf{Mod}(\mathcal{O}_{\mathrm{Spec}(R)})$ to be the full subcategory of $\mathcal{O}_{\mathrm{Spec}(R)}$-modules of the form $\widetilde{M}$.

An explicit form of the sheaf $\widetilde{M}$ can be obtained by expanding the definition of the sheaf we obtain from it's definition on the basis.

Lemma 1.2.3.2. *Let M be an R-module and consider the associated $\mathcal{O}_{\mathrm{Spec}(R)}$-module $\widetilde{M}$. For any open $U \subseteq \mathrm{Spec}(R)$, we have*

$$\widetilde{M}(U) \cong \left\{ s : U \to \textstyle\coprod_{\mathfrak{p}\in U} M_{\mathfrak{p}} \mid \forall \mathfrak{p}\in U,\ s(\mathfrak{p})\in M_{\mathfrak{p}}\ \&\ \exists\ \textit{open}\ \mathfrak{p}\in V\subseteq U\ \&\ \exists m\in M, f\in R\ \textit{s.t.}\ \forall \mathfrak{q}\in V,\ f\notin \mathfrak{q}\ \&\ s(\mathfrak{q}) = m/f \right\}$$

Proof. Follows from Remark 20.2.0.4. □

We now collect properties of $\widetilde{M}$ below.

Proposition 1.2.3.3. *Let R be a ring and M, N, M_i be R-modules for $i \in I$,*

1. $\left(\widetilde{M}\right)_{\mathfrak{p}} \cong M_{\mathfrak{p}}$ *for all* $\mathfrak{p} \in \mathrm{Spec}(R)$,
2. $\widetilde{M}(\mathrm{Spec}(R)_f) \cong M_f$ *for all* $f \in R$,
3. $\Gamma(\widetilde{M}, \mathrm{Spec}(R)) \cong M$.

Proof. Statement 1 follows from the alternate definition given in Lemma 1.2.3.2. Indeed one considers the function

$$\begin{aligned}\varphi : \left(\widetilde{M}\right)_{\mathfrak{p}} &\longrightarrow M_{\mathfrak{p}}\\ (U, s)_{\mathfrak{p}} &\longmapsto s(\mathfrak{p}).\end{aligned}$$

One immediately sees this is R-linear. Injectivity and surjectivity is then also trivially checked by the above cited lemma.

Statements 3 follows from statement 2 by setting $f = 1$ and statement is just the Definition 1.2.3.1. □

We can also understand how $\mathcal{O}_{\mathrm{Spec}(R)}$-modules behave under morphism of affine schemes (see direct and inverse image of modules at Section 3.5)

Lemma 1.2.3.4. [7] *Let $f : \mathrm{Spec}(S) \to \mathrm{Spec}(R)$ be a morphism of affine schemes associated to map $\varphi : R \to S$ of rings. Then,*

1. *if N is an S-module, then $f_*\widetilde{N} \cong \widetilde{{}_RN}$ where ${}_RN$ is the R-module obtained by restriction of scalars by φ,*

[7]We will call it the *globalized* extension and restriction of scalars.

2. *if M is an R-module, then $f^*\widetilde{M} \cong \widetilde{(S \otimes_R M)}$ where $S \otimes_R M$ is the S-module obtained by extension of scalars by φ.*

Proof. The proof is routine with main observation being the facts that for $g \in R$, we have $({}_R N)_g \cong N_{\varphi(g)}$ and for $\mathfrak{q} \in \operatorname{Spec}(S)$, we get the natural isomorphism $(f^*\widetilde{M})_\mathfrak{q} \cong \widetilde{(S \otimes_R M)}_\mathfrak{q}$. □

Theorem 1.2.3.5. *Let R be a ring. There is an equivalence of categories between those of R-modules and $\mathcal{O}_{\operatorname{Spec}(R)}$-modules of the form $\widetilde{M}$:*

$$\mathbf{Mod}(R) \underset{\Gamma(X,-)}{\overset{\widetilde{(-)}}{\rightleftarrows}} \widetilde{\mathbf{Mod}}(\mathcal{O}_{\operatorname{Spec}(R)}) \quad (\equiv)$$

which moreover satisfies the following properties

1. *$\widetilde{(-)}$ is an exact functor; if $0 \to M' \to M \to M'' \to 0$ is exact, then $0 \to \widetilde{M'} \to \widetilde{M} \to \widetilde{M''} \to 0$ is exact,*
2. *$\widetilde{(-)}$ preserves tensor product; $\widetilde{M \otimes_R N} \cong \widetilde{M} \otimes_{\mathcal{O}_X} \widetilde{N}$,*
3. *$\widetilde{(-)}$ preserves coproducts; $\widetilde{\bigoplus_{i \in I} M_i} \cong \bigoplus_{i \in I} \widetilde{M_i}$.*

Proof. Let $X = \operatorname{Spec}(R)$. Consider the following map

$$\operatorname{Hom}_R(M, N) \to \operatorname{Hom}_{\mathcal{O}_X}\left(\widetilde{M}, \widetilde{N}\right)$$
$$f : M \to N \mapsto \widetilde{f} : \widetilde{M} \to \widetilde{N}$$
$$\eta_X : M \to N \leftarrow\!\shortmid \eta : \widetilde{M} \to \widetilde{N}$$

Now, beginning from η, we may show that $(\widetilde{\eta_X})_{X_g} = \eta_{X_g}$ for some basic open $X_g \hookrightarrow X$. The result follows from the fact that $\eta : \widetilde{M} \to \widetilde{N}$ is completely characterized by the map on global sections $\eta_X : M \to N$ from the following square

$$\begin{tikzcd} M_g \arrow[r, "\eta_{X_g}"] & N_g \\ M \arrow[u] \arrow[r, "\eta_X"'] & N \arrow[u] \end{tikzcd}$$

where the verticals are restrictions morphisms w.r.t $R \to R_g$ and the top horizontal is R_g-module homomorphism and bottom is R-module homomorphism.

For statement 1, by Theorem 20.3.0.8, the question is local in nature. We deduce the result then from Lemma 16.1.2.2.

For statement 2, we proceed as follows. To define an isomorphism

$$\varphi : \widetilde{M} \otimes_{\mathcal{O}_X} \widetilde{N} \to \widetilde{M \otimes_R N}$$

we need only define a map from the presheaf F given by $U \mapsto \widetilde{M}(U) \otimes_{\mathcal{O}_X(U)} \widetilde{N}(U)$ to $\widetilde{M \otimes_R N}$ such that on basic open sets, we have an isomorphism. Indeed, let $D(f) \subseteq \operatorname{Spec}(R)$ be an open set for some $f \in R$. We define

$$\varphi_U : M_f \otimes_{R_f} N_f \xrightarrow{\cong} (M \otimes_R N)_f$$

as the obvious natural isomorphism. One checks that this does define φ to be a sheaf map.

For statement 3, as $\widetilde{(-)}$ is a left adjoint, therefore it preserves all colimits. $\square$

Remark 1.2.3.6. We will later see that on affine schemes $\mathrm{Spec}\,(R)$, the category $\widetilde{\mathbf{Mod}}(\mathcal{O}_{\mathrm{Spec}(R)})$ is precisely the category of quasicoherent $\mathcal{O}_{\mathrm{Spec}(R)}$-modules, which is a class of modules of utmost importance in algebraic geometry.

1.3 Schemes and basic properties

We can now define scheme to be a locally ringed space (see Foundational Geometry, 3) with an affine open covering.

Definition 1.3.0.1. (**Schemes**) A locally ringed space $(X, \mathcal{O}_X)$ is a scheme if there exists an open affine cover $\{(\mathrm{Spec}\,(R_i), \mathcal{O}_{\mathrm{Spec}(R_i)})\}$ of $(X, \mathcal{O}_X)$ such that $\mathcal{O}_{X|\mathrm{Spec}(R_i)} \cong \mathcal{O}_{\mathrm{Spec}(R_i)}$.

As we go along in understanding schemes, it will be more and more apparent the need of sheaf language to talk about the "generalized functions" over the scheme X. Indeed, there is a fine interrelationship between the *space structure* of the scheme $(X, \mathcal{O}_X)$ (that is, the topological space X) and the *function structure* on the scheme (that is, the sheaf of functions $\mathcal{O}_X$). A big part of learning scheme theory is to understand and use this relationship between them.

We will now bring some global topological properties of schemes which reflect their affine origins. An analogue of Lemma 1.2.1.1 holds in the general case of schemes.

Lemma 1.3.0.2. [8] *Let X be a scheme. The following are equivalent.*

1. *$S \subseteq X$ is a closed irreducible subset.*
2. *There exists a point $x \in S$ such that $\overline{\{x\}} = S$.*

Proof. (1. $\Rightarrow$ 2.) Let U be an affine open in X intersecting S. Then $U \cap S$ is an open subset of S. As open subsets of irreducibles are dense, therefore $U \cap S$ is dense in S. Consequently, it suffices to show that there exists a point $x \in U \cap S$ such that $\overline{\{x\}} = U \cap S$. As open subsets of irreducibles are irreducible, therefore $U \cap S$ is irreducible. Replacing X by U, we may assume X is affine. The result then follows by Lemma 1.2.1.1.

(2. $\Rightarrow$ 1.) Since $x \in U$ for some open affine $U \subset X$, thus, $x \in U \cap S$. Since $U \cap S \subseteq U$ and U is open, therefore closure of $\{x\}$ in U is same as closure of $\{x\}$ in X. Now, $\overline{\{x\}} = S$ but $\overline{\{x\}} \subseteq U$. It thus follows that $S \subseteq U$ and hence S is in an open affine. The result follows by Lemma 1.2.1.1. $\square$

Every open subspace of a scheme is a scheme.

Lemma 1.3.0.3. *Let X be a scheme and $U \subseteq X$ be an open subspace. Then $(U, \mathcal{O}_{X|U})$ is a scheme.*

Proof. Since for an affine scheme $\mathrm{Spec}\,(R)$, the basic open $\mathrm{Spec}\,(R)_f \cong \mathrm{Spec}\,(R_f)$ for $f \in R$, therefore for an open subspace $U \subseteq X$ and an affine open cover $\{U_i\}$ of X, $U_i \cap U$ is open in U_i and thus covered by affines of the form $\mathrm{Spec}\,(R_f)$. $\square$

[8] Exercise II.2.9 of Hartshorne.

Write **Sch** to be the category of schemes and $\mathbf{Sch}/S$ to be the category of schemes over S. Morphisms of schemes is merely the same concept as that of morphism of locally ringed spaces (see Foundational Geometry, Chapter 3).

Definition 1.3.0.4. (**Map of schemes**) Let X and Y be two schemes. A map of underlying locally ringed spaces $(f, f^\sharp) : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is called a map of schemes. In a more expanded form, $f : X \to Y$ is a continuous map and $f^\sharp : f^{-1}\mathcal{O}_Y \to \mathcal{O}_X$ is a map of sheaves such that the induced map (see Topics in Sheaf Theory, Chapter 20) on stalks for each $x \in X$

$$f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$$

is a map of local rings, i.e., $(f_x^\sharp)^{-1}(\mathfrak{m}_{X,x}) = \mathfrak{m}_{Y,f(x)}$.

An important theorem in global study of schemes is a complete characterization of schemes over $\operatorname{Spec}(R)$, which is of-course of paramount importance.

Theorem 1.3.0.5. *Let X be a scheme and R be a ring. Then, there's a natural bijection*

$$\operatorname{Hom}_{\mathbf{Sch}}(X, \operatorname{Spec}(R)) \cong \operatorname{Hom}_{\mathbf{Ring}}(R, \Gamma(X, \mathcal{O}_X)).$$

In other words, we have the following adjunction[9]

$$\begin{tikzcd} \mathbf{Sch} \arrow[r, shift left, "\Gamma(-)"] \arrow[r, phantom, "\perp"] & \mathbf{Ring}^{\mathrm{op}} \arrow[l, shift left, "\operatorname{Spec}(-)"] \end{tikzcd} .$$

Proof. The proof will be played out in two steps. In the first one we will show the candidates for the unit and counit of this adjunction. In the second play we will show that they indeed satisfy the required triangle identities.

Act 1 : *The units and counits.*

Let us first define the simpler one of them, the counit. For any $R \in \mathbf{Ring}$, we define a natural transformation $\epsilon : \mathrm{id}_{\mathbf{Ring}} \to \Gamma \circ \operatorname{Spec}()$ given by (note how we adjusted for the contravariant nature of $\operatorname{Spec}(-)$ and $\Gamma(-)$)

$$\begin{aligned} \epsilon_R : R &\longrightarrow \Gamma(\operatorname{Spec}(R)) \cong R \\ f &\longmapsto f. \end{aligned}$$

Thus, $\epsilon_R = \mathrm{id}_R$. Hence, $\epsilon = \mathrm{id}_{\mathbf{Ring}^{\mathrm{op}}}$.

Next, we define the more intricate part, which is the unit. Take any scheme $X \in \mathbf{Sch}$. We define $\eta : \mathrm{id}_{\mathbf{Sch}} \to \operatorname{Spec}(\Gamma)$ on X by

$$\begin{aligned} \eta_X : X &\longrightarrow \operatorname{Spec}(\Gamma(X)) \\ x &\longmapsto \mathfrak{p} = \eta_X(x) := \{f \in \Gamma(X) \mid f_x \in \mathfrak{m}_x\}. \end{aligned}$$

[9]This is also sometimes called the *algebra-geometry duality* or the *fundamental duality of algebraic geometry*.

Moreover, the map on structure sheaves is given by

$$(\eta_X)^\flat : \mathcal{O}_{\mathrm{Spec}(\Gamma(X))} \longrightarrow (\eta_X)_* \mathcal{O}_X$$

where as the map on global sections we keep it id and on a basic open $\mathrm{Spec}\,((\Gamma(X))_f)$ this is defined on sections by

$$(\eta_X)^\flat_{\mathrm{Spec}((\Gamma(X))_f)} : \Gamma(X)_f \cong \mathcal{O}_{\mathrm{Spec}(\Gamma(X))}((\mathrm{Spec}\,(\Gamma(X)))_f) \longrightarrow \mathcal{O}_X(\eta_X^{-1}((\mathrm{Spec}\,((\Gamma(X))))_f))$$

by the unique map that is obtained in the following diagram

$$\begin{tikzcd} \Gamma(X)_f \arrow[r, dashed] & \mathcal{O}_X(\eta_X^{-1}((\mathrm{Spec}\,(\Gamma(X)))_f)) \\ \Gamma(X) \arrow[u] \arrow[ur, "\rho"'] & \end{tikzcd},$$

where, indeed, $f \in \Gamma(X)$ is mapped to to an unit element in $\mathcal{O}_X(\eta_X^{-1}((\mathrm{Spec}\,(\Gamma(X)))_f$ because of the following simple lemma:

(*) For a locally ringed space $(X, \mathcal{O}_X)$ and an open subspace $U \subseteq X$, $f \in \mathcal{O}_X(U)$ is a unit if and only if $f_x \notin \mathfrak{m}_x \subset \mathcal{O}_{X,x}$ for all $x \in U$.

This construction has the following properties and we give the main idea which drives each one of them.

1. $\eta_X(x)$ *is a prime ideal of* $\Gamma(X)$: This follows from $\mathfrak{m}_x$ being a maximal (hence prime) ideal of $\mathcal{O}_{X,x}$.
2. η_X *is continuous* : Working with basis and reducing to assumption that $X = \mathrm{Spec}\,(S)$ is affine, we reduce to showing that $\{\mathfrak{p} \in \mathrm{Spec}\,(R) \mid f_\mathfrak{p} \notin \mathfrak{m}_\mathfrak{p}\}$ is open, which is true as it is equal to $(\mathrm{Spec}\,(S))_f$.
3. $\eta : \mathrm{id}_{\mathbf{Sch}} \to \mathrm{Spec}\,() \circ \Gamma$ *is a natural transformation* : We wish to show that commutativity of the natural square. For a map of schemes $f : X \to Y$, this reduces to showing that

$$\forall x \in X,\ \eta_Y(f(x)) = (f_Y^\flat)^{-1}(\eta_X(x)).$$

This further follows from the observation that for $g \in \Gamma(Y)$, $f_Y^\flat(g) \in \mathfrak{m}_x \iff f_x(g_{f(x)}) \in \mathfrak{m}_x$ and the latter is clearly true by the definition of maps of locally ringed spaces, where $f_x : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is the map on stalks.

Hence, we have obtained a map of schemes $(\eta_X, \eta_X^\flat) : X \to \mathrm{Spec}\,(\Gamma(X))$. This is our candidate for the unit of the adjunction.

Act 2 : *η and ϵ satisfies the triangle identities.*

It follows that we wish to show that the following two diagrams commute:

$$\begin{tikzcd} \Gamma(X) & \Gamma(\mathrm{Spec}\,(\Gamma(X))) \cong \Gamma(X) \arrow[l, "\Gamma\eta_X"'] \\ & \Gamma(X) \arrow[ul, "\mathrm{id}_{\Gamma(X)}"] \arrow[u, "\epsilon_{\Gamma(X)}"'] \end{tikzcd} \qquad \begin{tikzcd} \mathrm{Spec}\,(R) \arrow[r, "\eta_{\mathrm{Spec}(R)}"] \arrow[dr, "\mathrm{id}_{\mathrm{Spec}(R)}"'] & \mathrm{Spec}\,(\Gamma(\mathrm{Spec}\,(R))) \cong \mathrm{Spec}\,(R) \arrow[d, "\mathrm{Spec}(\epsilon_R)"] \\ & \mathrm{Spec}\,(R) \end{tikzcd}.$$

in **Ring** in **Sch**

This follows from a simple unraveling of the maps involved in the diagram as defined in Act 1. $\square$

Corollary 1.3.0.6. *The above adjunction restricts to the following equivalence of categories:*

$$\mathbf{AfSch} \underset{\mathrm{Spec}(-)}{\overset{\Gamma(-)}{\underset{\cong}{\rightleftarrows}}} \mathbf{Ring}^{\mathrm{op}} \ .$$

$\square$

Corollary 1.3.0.7. *Let X be a scheme over* $\mathrm{Spec}\,(R)$ *for a ring R. Then, for any open affine* $\mathrm{Spec}\,(S) \subseteq X$*, S is an R-algebra. Consequently, all stalks $\mathcal{O}_{X,p}$ are R-algebras.* $\square$

1.3.1 Basic properties

We can now observe some more basic properties.

Local rings at non-closed points

Let X be an arbitrary scheme and $p \in X$ be a non-closed point. One can show that the local ring $\mathcal{O}_{X,p}$ is obtained by localizing local rings at closed points. Indeed, we have the following simple observation in this direction.

Lemma 1.3.1.1. *Let X be a scheme and $p \in X$ be a non-closed point. Then, $\mathcal{O}_{X,p}$ is isomorphic to localization of a local ring $\mathcal{O}_{X,x}$ at a prime ideal, where $x \in X$ is a closed point.*

Proof. Let $p \in X$ be a non-closed point and $U = \mathrm{Spec}\,(A)$ be an open affine containing p. Consequently, p corresponds to a prime ideal $\mathfrak{p} \lneq A$ which is not maximal. Let $\mathfrak{m} \lneq A$ be a maximal ideal containing $\mathfrak{p}$ and let $m \in U$ be the corresponding closed point in X. As $\mathcal{O}_{X,p} \cong A_{\mathfrak{p}}$ and $\mathcal{O}_{X,m} \cong A_{\mathfrak{m}}$, and since $(A_{\mathfrak{m}})_{\mathfrak{p}_{\mathfrak{m}}} \cong A_{\mathfrak{p}}$, therefore we have that $\mathcal{O}_{X,p}$ is obtained by localizing $\mathcal{O}_{X,m}$ at a prime ideal, as required. $\square$

Using ideas similar to above, we can also prove the following simple result.

Lemma 1.3.1.2. *Let X be an integral scheme and $\eta \in X$ be a non-closed point. Then the fraction field of $\mathcal{O}_{X,\eta} \cong K(X)$ where $K(X)$ is the function field of X.* $\square$

Non-vanishing locus of a global section

We next see that how a global section of a scheme defines an open set which is the set of those points where that element, when treated as a function, is non-zero. One then finds what the ring of functions over this open set looks like. First, for any scheme X and any $f \in \Gamma(\mathcal{O}_X, X)$, define the *non-vanishing locus* of f by

$$X_f := \{x \in X \mid f \notin \mathfrak{m}_{X,x}\}.$$

We first have the following simple result about non-vanishing locus.

Lemma 1.3.1.3. *Let $f : X \to \mathrm{Spec}\,(B)$ be a scheme over a ring B and let $g \in B$. Let $\varphi : B \to \Gamma(\mathcal{O}_X, X)$ be the map induced on the global sections. Then,*

$$f^{-1}(D(g)) = X_{\varphi(g)}.$$

Proof. Observe that $x \in X_{\varphi(g)}$ if and only if $\varphi(g)_x \notin \mathfrak{m}_{X,x}$. As we have the following commutative square

$$\begin{tikzcd} B \arrow[r, "\varphi"] \arrow[d] & \Gamma(\mathcal{O}_X, X) \arrow[d] \\ \mathcal{O}_{\mathrm{Spec}(B),f(x)} \arrow[r, "f_x^\sharp"'] & \mathcal{O}_{X,x} \end{tikzcd}$$

where vertical arrows are image into the stalk, therefore we deduce that $\varphi(g)_x \notin \mathfrak{m}_{X,x}$ if and only if $f_x^\sharp(g_x) \notin \mathfrak{m}_{X,x}$. As $f_x^\sharp$ is a local ring homomorphism, therefore $f_x^\sharp(g_x) \notin \mathfrak{m}_{X,x}$ if and only if $g_x \notin \mathfrak{m}_{\mathrm{Spec}(B),f(x)} = f(x)B_{f(x)}$. As $B \to \mathcal{O}_{\mathrm{Spec}(B),f(x)}$ is just localization map $B \to B_{f(x)}$, therefore $g_x \notin f(x)B_{f(x)}$ if and only if $g \notin f(x)$, that is $f(x) \in D(g)$. This completes the proof. □

Proposition 1.3.1.4. [10] *Let X be a scheme and $f \in \Gamma(\mathcal{O}_X, X)$.*

1. *Let $U = \mathrm{Spec}\,(A)$ be an affine open subset of X and denote $\bar{f} = \rho_{X,U}(f)$. Then, $U \cap X_f = D(\bar{f})$. Consequently, $X_f \subseteq X$ is an open subscheme.*
2. *Let X be quasicompact and $a \in \Gamma(\mathcal{O}_X, X)$ such that $\rho_{X,X_f}(a) = 0$. Then, $f^n a = 0$ in $\Gamma(\mathcal{O}_X, X)$ for some $n > 0$.*
3. *Let X admit an affine open cover U_i such that $U_i \cap U_j$ is quasicompact. If $b \in \mathcal{O}_X(X_f)$, then there exists $a \in \Gamma(\mathcal{O}_X, X)$ and $n > 0$ such that $f^n b = \rho_{X,X_f}(a)$ in $\mathcal{O}_X(X_f)$.*
4. *There is an isomorphism of rings $\Gamma(\mathcal{O}_{X_f}, X_f) \cong (\Gamma(\mathcal{O}_X, X))_f$.*

Proof. 1. We wish to show that $\{x \in U \mid \bar{f}_x \notin \mathfrak{m}_{X,x}\} = \{x \in U \mid \bar{f} \notin x\}$, where $x \in U$ in latter is treated as a prime ideal of A. The side " $\subseteq$ " follows from the fact that for $x \in U$, we have $\mathcal{O}_{X,x} \cong A_x$, $\mathfrak{m}_{X,x} \cong xA_x$ and the fact that the map into stalks $\mathcal{O}_X(U) \to \mathcal{O}_{X,x}$ is given by the canonical map $A \to A_x$, $a \mapsto a/1$. One further would need the commutativity of the following diagram:

$$\begin{tikzcd} & \mathcal{O}_{X,x} \\ \mathcal{O}_X(U) \arrow[ur] & \Gamma(\mathcal{O}_X, X) \arrow[l] \arrow[u] \end{tikzcd} .$$

The side " $\supseteq$ " also follows from the commutativity of the above triangle together with the canonical isomorphisms of the local ring and its maximal ideal.

2. **TODO from notebook.** □

Locality of isomorphism on target

We now show a rather simple result on locality of isomorphism on target, but it is quite useful in scenarios where one understands the map well on individual opens of target but not on the global level.

[10] Exercise II.2.16 of Hartshorne.

Proposition 1.3.1.5. *Let $f : X \to Y$ be a map of schemes and $Y = \bigcup_{i \in I} U_i$ be an open cover of Y such that $f|_{f^{-1}(U_i)} : f^{-1}(U_i) \to U_i$ is an isomorphism. Then, f is an isomorphism.*

Proof. **TODO from notes.** □

Criterion for affineness

We now show a useful criterion for a scheme to be affine. This also portrays the power of previous result on locality of isomorphism.

Proposition 1.3.1.6. *Let X be a scheme and denote $A = \Gamma(\mathcal{O}_X, X)$. Then the following are equivalent:*

1. *X is affine,*
2. *there exists $f_1, \dots, f_r \in A$ such that X_{f_i} are open affine subsets of X and $\langle f_1, \dots, f_r \rangle = A$.*

Proof. **TODO from notes.** □

1.4 First notions on schemes

Having defined schemes, our next goal is to bring to light some of the obvious definitions that one can make on them. In some sense, having made the general definition of schemes, we are now trying to go back to try and find where does varieties lie in this big world of **Sch**. Indeed, we will see that the definitions introduced in the following few sections are bringing us ever closer to define varieties as certain type of schemes, which will thus enable us to bring to light the most important geometric notions on varieties.

1.4.1 Noetherian schemes

Definition 1.4.1.1. (**Noetherian schemes**) A scheme X is called *locally noetherian* if there exists an affine open cover $X = \cup_{i\in I} U_i$ where each $U_i = \operatorname{Spec}(A_i)$ where A_i is a noetherian ring. If moreover, X is quasicompact, then X is called *noetherian*.

Remark 1.4.1.2. Since $X = \operatorname{Spec}(A)$ is already quasi-compact (Lemma 1.2.1.6), therefore for affine schemes X, the notion of locally noetherian and noetherian are equal.

The only immediately important result about such schemes that one needs is that an affine scheme is noetherian if and only if the obvious thing happens.

Lemma 1.4.1.3. *Let $X = \operatorname{Spec}(A)$ be an affine scheme. Then, the following are equivalent:*

1. *X is a noetherian scheme.*
2. *A is a noetherian ring.*

Proof. (2. $\Rightarrow$ 1.) This follows from Remark 1.4.1.2 and the fact that localization of noetherian rings are noetherian (Proposition 16.3.0.7).
(1. $\Rightarrow$ 2.) Let X be noetherian. Then there is an affine open cover of X by spectra of noetherian rings. Pick any ideal $I \leq A$. We shall show it is finitely generated. There is a finite cover $\{\operatorname{Spec}(A_{f_i})\}_{i=1}^n$ of $\operatorname{Spec}(A)$ where A_{f_i} are noetherian and $f_i \in A$. Hence we have that the ideal IA_{f_i} of A_{f_i} is finitely generated for all $i = 1, \dots, n$. By Lemma 1.2.1.5, 2, we see that $f_1, \dots, f_n$ generate the whole ring A. The result then follows by Lemma 16.1.2.10. □

Example 1.4.1.4. By the Lemma 1.4.1.3, we observe that any of the variety over a field is a noetherian scheme (technically, we are identifying the affine variety with its associated scheme, see Section **??**, Schemes associated to varieties). So any of your favorite variety

$$\operatorname{Spec}\left(\frac{k[x,y,z]}{x^2+y^2-z^3-1}\right), k \text{ is algebraically closed}$$

gives a (*is* a) noetherian scheme.

Our next goal is to show that a noetherian scheme is a noetherian space.

Proposition 1.4.1.5. *If X is a noetherian scheme, then X is a noetherian space.*

Proof. As X has a finite open affine cover by spectra of noetherian rings and such spectra are noetherian schemes by Lemma 1.4.1.3, thus by the fact that finite union of noetherian spaces is noetherian we can complete the proof. □

Local rings of a locally noetherian scheme are noetherian.

Lemma 1.4.1.6. *If X is locally noetherian, then $\mathcal{O}_{X,x}$ is a noetherian ring.*

Proof. Since localization of a noetherian ring at a prime is again noetherian by Proposition 16.3.0.7, therefore $\mathcal{O}_{X,x}$ is noetherian. □

Being locally noetherian is a local property.

Proposition 1.4.1.7. *Let X be a locally noetherian scheme. If* $\operatorname{Spec}(A) \subseteq X$ *is an open affine, then* $\operatorname{Spec}(A)$ *is noetherian and thus A is a noetherian ring.*

Proof. Let $U_i = \operatorname{Spec}(A_i)$ be an open cover by noetherian affine schemes (A_i are noetherian). Then, a finitely many of U_i will cover $\operatorname{Spec}(A)$ by quasi-compactness of $\operatorname{Spec}(A)$, say $U_1, \dots, U_n$. Thus we obtain a finite basic open cover $D(f_i)$ of $\operatorname{Spec}(A)$ for $f_i \in A$ where each $D(f_i) \subseteq U_j$ for some j such that $D(f_i)$ is also basic in U_j (Lemma 1.4.4.3). As U_j is noetherian, therefore if we can show that $\mathcal{O}_{U_j}(D(f_i))$ is noetherian, then we would have shown that A_{f_i} is noetherian, which would complete the proof by Lemma 16.3.0.8. We thus reduce to assuming $X = \operatorname{Spec}(A)$ noetherian affine and to show that $U = D(f) \subseteq X$ is noetherian for $f \in A$.

In this case, as A is noetherian, therefore by Corollary 16.3.0.9, the ring A_f is noetherian, as required. □

Another important aspect of noetherian schemes is quasi-compactness of intersection of open affines.

Proposition 1.4.1.8. *Let X be a noetherian scheme and $U, V \subseteq X$ be two affine opens. Then $U \cap V$ is quasi-compact.*

Proof. **TODO** □

One can reduce a lot of arguments from non-noetherian to the noetherian case using the following.

Proposition 1.4.1.9. *Let X be a finitely presented scheme over A. Then there exists a noetherian ring $A_0 \hookrightarrow A$ and a finitely presented scheme X_0 over A_0 such that the base change $(X_0)_A$ is isomorphic to X.*

1.4.2 Reduced, integral schemes and function field

The following are the definitions required, which are clearly geometric in nature.

Definition 1.4.2.1. (**Reduced and integral schemes**) A scheme X is said to be *reduced* if local rings $\mathcal{O}_{X,x}$ for all $x \in X$ is a reduced ring; have no nilpotents. A scheme X is said to be *integral* if it is reduced and irreducible as a topological space.

The one basic result that must be seen about these two types of schemes is that they are characterized by algebraic properties of local sections. Thus being reduced or integral, while defined geometrically, is concretely controlled by the algebraic properties of the structure sheaf.

Lemma 1.4.2.2. *Let X be a scheme. Then,*

1. *X is reduced if and only if $\mathcal{O}_X(U)$ is a reduced ring for each open set $U \subseteq X$* [11].
2. *X is integral if and only if $\mathcal{O}_X(U)$ is an integral domain for each open set $U \subseteq X$.*

Proof. 1. (L $\Rightarrow$ R) Suppose for some open $U \subseteq X$ there exists a section $f \in \mathcal{O}_X(U)$ which is nilpotent. Using the homomorphism $\mathcal{O}_X(U) \to \mathcal{O}_{X,x}$ given by $s \mapsto s_x$, we see that $f_x \in \mathcal{O}_{X,x}$ is a nilpotent element.
(R $\Rightarrow$ L) Suppose X is not reduced. Hence for some germ $f_x \in \mathcal{O}_{X,x}$ at some point $x \in X$ is a nilpotent where $f \in \mathcal{O}_X(U)$ for some open $x \in U \subseteq X$. Since $f_x^n = 0$ for some $n \in \mathbb{N}$, we get that $f^n = 0$ for some open $W \subseteq U$. Thus $\rho_{U,W}(f) \in \mathcal{O}_X(W)$ is a nilpotent element[12].

2. (L $\Rightarrow$ R) Pick any open $U \subseteq X$. We wish to show that $\mathcal{O}_X(U)$ is an integral domain. In other words, we wish to show the proposition for the open subscheme $(U, \mathcal{O}_{X|U})$. Replacing X by U, we reduce to showing $\mathcal{O}_X(X)$ is an integral domain. So let $f, g \in \mathcal{O}_X(X)$ be such that $fg = 0$. We wish to show that either $f = 0$ or $g = 0$. Suppose neither f nor g is 0 but $fg = 0$. It follows from Lemma 1.2.0.1, 1, that $V(f)$ and $V(g)$ covers X and hence by irreducibility of X, either $V(f) = 0$ or $V(g) = 0$, that is, $f = 0$ or $g = 0$.
(R $\Rightarrow$ L) We first need to show that X is reduced. Indeed, by 1. it follows immediately as integral domains are reduced. We then wish to show that X is irreducible. Indeed, if there are two open subsets of X say $U_1, U_2 \subseteq X$ such that $U_1 \cap U_2 = \emptyset$, then we claim that $\mathcal{O}_X(U_1 \cup U_2) \cong \mathcal{O}_X(U_1) \times \mathcal{O}_X(U_2)$. Since both $\mathcal{O}_X(U_1), \mathcal{O}_X(U_2)$ have 0 and 1, thus $\mathcal{O}_X(U_1 \cup U_2)$ will have a zero-divisor, a contradiction. Indeed, consider the following homomorphism, denoting $U := U_1 \cup U_2$

$$\begin{aligned} \mathcal{O}_X(U) &\longrightarrow \mathcal{O}_X(U_1) \times \mathcal{O}_X(U_2) \\ s &\longmapsto (\rho_{U,U_1}(s), \rho_{U,U_2}(s)). \end{aligned}$$

This is injective by locality axiom and surjective by gluing axiom of sheaves. □

Corollary 1.4.2.3. *Let X be a scheme. If X is integral, then all local rings $\mathcal{O}_{X,x}$ are integral domains.*

Proof. Use Lemma 1.4.2.2, 2 together with the fact that localization of integral domains is an integral domain. □

Corollary 1.4.2.4. *Let $X = \operatorname{Spec}(A)$ be an affine scheme. Then X is integral if and only if A is an integral domain.*

Proof. Use Lemma 1.4.2.2, 2 on global sections together to get one side. For the "only if" side, stalks are reduced as they are integral (localizations of A) and X is irreducible as for any $V(\mathfrak{a}) \cup V(\mathfrak{b}) = X$, we have $V(\mathfrak{a}\mathfrak{b}) = X$ and thus $\mathfrak{a}\mathfrak{b} \subseteq \mathfrak{n}$ where $\mathfrak{n}$ is the intersection of all prime ideals, the nilradical (Lemma 16.1.2.9). Since A is integral, therefore $\mathfrak{o}$ is prime as well and hence $\mathfrak{n} = 0$, making $\mathfrak{a}\mathfrak{b} = 0$. Since A is integral, hence $\mathfrak{a} = \mathfrak{o}$ or $\mathfrak{b} = \mathfrak{o}$. □

Remark 1.4.2.5. (*Function field of an integral scheme*) Let X be an integral scheme. Since X is irreducible as a topological space, therefore there is a generic point η in X, i.e. a point whose closure

[11] Exercise II.2.3.a of Hartshorne.

[12] This is a very inefficient way of using the equality on stalks. Indeed, two germs are equal if and only if the representatives are equal on some common shrinking of their domains. This is how usually people work with stalks without being overly full of symbols.

is the whole of X (Lemma 1.3.0.2). Now let $\mathrm{Spec}\,(A) \subseteq X$ be an affine open such that $\eta \in \mathrm{Spec}\,(A)$. Thus, η is a generic point of $\mathrm{Spec}\,(A)$ as well. Hence η corresponds to the zero ideal of A, which is indeed an integral domain from Lemma 1.4.2.2, 2. Since $\mathcal{O}_{X,\eta} \cong \mathcal{O}_{\mathrm{Spec}(A),\eta} = A_{\mathfrak{o}}$, therefore $\mathcal{O}_{X,\eta}$ is a field, called the *function field of the integral scheme* X and is in particular given by field of fractions of any domain A such that open $\mathrm{Spec}\,(A)$ contains η. We denote the function field of X as $K(X)$[13].

Using the fact that the generic point of an integral scheme X will be in every non-empty open set, we can make some fascinating observations about the function field $K(X)$, which thus justifies its name.

Lemma 1.4.2.6. *Let X be an integral scheme with function field $K(X)$. Then for all $x \in X$, the local ring $\mathcal{O}_{X,x}$ is contained in $K(X)$.*

Proof. Let $x \in X$, $\eta \in X$ be the generic point and $U = \mathrm{Spec}\,(A)$ be an open affine in X. By Lemma 1.4.2.2, 2, A is a domain. Clearly, $\eta \in U$ and it corresponds to the zero ideal $\mathfrak{o} \lneq A$. Further we have $\mathcal{O}_{X,x} \cong A_{\mathfrak{p}}$, $\mathfrak{p} \in U$ is equal to the point $x \in U$. By definition $K(X) = A_{\mathfrak{o}}$. The result follows by observing that $A_{\mathfrak{p}} \subseteq A_{\mathfrak{o}}$. $\square$

The following lemma shows that restriction of functions in an integral scheme is injective.

Lemma 1.4.2.7. *Let X be an integral scheme and $U \hookrightarrow V$ be an inclusion of open sets. Then, the restriction maps $\rho : \mathcal{O}_X(V) \to \mathcal{O}_X(U)$ is an injective ring homomorphism.*

Proof. By Lemma 20.3.0.2, we need only show that for any $x \in V$ and any $s \in \mathcal{O}_X(V)$, we have $(V,s)_x = 0$ in $\mathcal{O}_{X,x}$. Let $W = \mathrm{Spec}\,(A)$ be an open affine containing x. As U is open in X and X is irreducible, therefore it is dense. Consequently, $U \cap W$ is an open non-empty set in X. We may write $\rho_{V,W}(s) = a \in A$. Let $D(f) \subseteq U \cap W$ be a basic open set of W. Since taking germs commutes with restrictions, therefore we have the restriction map $\mathcal{O}_X(W) \to \mathcal{O}_X(D(f))$ which is the localization map $A \to A_f$, which takes $a \mapsto \frac{a}{1}$. As s on U is 0, therefore, s is 0 on $W \cap U$ and thus on $D(f)$. Consequently, we have $\frac{a}{1} = 0$ in A_f. As A is a domain by Lemma 1.4.2.2, it follows that $a = 0$ in A. Thus, $\rho_{V,W}(s) = 0$, hence, $(V,s)_x = 0$ in $\mathcal{O}_{X,x}$, as required. $\square$

Example 1.4.2.8. ($\mathrm{Spec}\,(\mathbb{Z})$) Since $\mathbb{Z}$ is an integral domain, therefore by Corollary 1.4.2.4, $X = \mathrm{Spec}\,(\mathbb{Z})$ is an integral scheme. Clearly, X as a topological space consists of all prime numbers and a generic point given by the zero ideal $\mathfrak{o}$. Further, the topology is thus given by cofinite topology. At the level of stalks, we have that for a prime $\mathfrak{p} \in X$, $\mathcal{O}_{X,\mathfrak{p}} \cong \mathbb{Z}_{\mathfrak{p}}$ and we can describe $\mathbb{Z}_{\mathfrak{p}}$ as all those rationals whose denominator is not a multiple of prime p where $\mathfrak{p} = \langle p \rangle$ as $\mathbb{Z}$ is a PID (it's ED). Clearly, localizing X at the generic point $\mathfrak{o}$ would yield $\mathcal{O}_{X,\mathfrak{o}} \cong \mathbb{Q}$. More fascinatingly, for a prime $\mathfrak{p} = \langle p \rangle$ in X, the residue field at point $\mathfrak{p}$ is $\kappa(\mathfrak{p}) = \mathbb{Z}_{\mathfrak{p}}/\mathfrak{p}\mathbb{Z}_{\mathfrak{p}} \cong \mathbb{F}_p$, the finite field with p elements!

Now for any affine scheme $\mathrm{Spec}\,(A)$, consider a map $f : X \to \mathrm{Spec}\,(\mathbb{Z})$. By the fact that $\mathbb{Z}$ is initial in category of rings, therefore $\mathrm{Spec}\,(\mathbb{Z})$ is terminal in the category of affine schemes (Corollary 1.3.0.6). Since any scheme is locally affine, it further follows that $\mathrm{Spec}\,(\mathbb{Z})$ is terminal in the category of schemes.

We now introduce a concept which will be used while discussing divisors.

[13] Exercise II.3.6 of Hartshorne.

Definition 1.4.2.9. (**Center of a valuation**) Let X be an integral scheme with function field K and $v : K \to G$ be a valuation over K with valuation ring $R \subset K$. A center of v is defined to be a point $x \in X$ such that R dominates $\mathcal{O}_{X,x}$ in K (see Definition 16.10.1.5).

1.4.3 (Locally) finite type schemes over k

This section is the beginning of a theme which we would like to understand intimately, schemes over a field. This is because most of the schemes we will encounter in nature will be varieties whose coordinate rings would be algebras over a field. Here we first understand in scheme language the first thing about coordinate rings of varieties over k, the fact that they are finitely generated as an k-algebra. Indeed, this is what we seek from the following definition.

Definition 1.4.3.1. (**Finite and locally finite type schemes over a field**) Let k be a field and let $X \to \text{Spec}\,(k)$ be a scheme over k. Then X is said to be locally finite type if there exists an affine open covering $\{\text{Spec}\,(A_i)\}_{i\in I}$ of X such that each A_i is a finitely generated k-algebra. Moreover, X is said to be finite type if X is locally finite type and quasi-compact.

Example 1.4.3.2. Our $\mathfrak{h}$yperboloid of one sheet (introduced in Example 1.5.1.3) has the following coordinate ring:

$$\frac{k[x,y,z]}{I(V(p))}$$

where $p(x,y,z) = x^2 + y^2 - z^2 - 1$, where we have chosen $a = b = c = 1$ for simplicity. Let $\mathfrak{h} := I(V(p))$. Clearly $\text{Spec}\,(k[x,y,z]/\mathfrak{h})$ is a finite type k-scheme.

Great thing about the above definition is that it really doesn't depend on the affine open cover that is chosen.

Lemma 1.4.3.3. *Let k be a field and X be a k-scheme. Then the following are equivalent.*

1. *X is of locally finite type over k.*
2. *For all open affine $U \hookrightarrow X$, the ring $\mathcal{O}_X(U)$ is finitely generated k-algebra.*

Proof. (2. $\Rightarrow$ 1.) Immediate.
e it's proof,
(1. $\Rightarrow$ 2.) We shall use Lemma 16.1.2.11 for this. $\square$

1.4.4 Subschemes and immersions

These notions are important in what is to come next.

Definition 1.4.4.1. (**Open subscheme**) Let X be a scheme. An open set $U \subseteq X$ has a canonical scheme structure, given by $(U, \mathcal{O}_{X|U})$. We call $(U, \mathcal{O}_{X|U})$ an open subscheme of X.

Indeed, locally U will look affine via the open affine cover of X. We can relativize this notion to define open immersions.

Definition 1.4.4.2. (**Open immersion**) A map $f : X \to Y$ of schemes is said to be an open immersion if $f : X \to f(X)$ is a homeomorphism, $f(X) \subseteq Y$ is open and $f^\flat_{|f(X)} : \mathcal{O}_{Y|f(X)} \to (f_*\mathcal{O}_X)_{|f(X)}$ is an isomorphism.

We observe that for any point in an intersection of open subschemes is contained in some special open subscheme. This is a very important result as this will be used as a technical tool to allow passage from one open affine with certain properties to another open affine, all the time while handling only basic open sets.

Lemma 1.4.4.3. *Let $U = \mathrm{Spec}\,(A), V = \mathrm{Spec}\,(B) \hookrightarrow X$ be two affine open subsets. For each $x \in U \cap V$, there exists an affine open subset $x \in W \hookrightarrow U \cap V$ such that $W = \mathrm{Spec}\,(A_f)$ and $W = \mathrm{Spec}\,(B_g)$ for some $f \in A$ and $g \in B$. Moreover, under the isomorphism $A_f \cong B_g$, the element $f \in A_f$ maps to $g \in B_g$.*

Proof. By replacing B by B_g for some $g \in B$, we may assume that $x \in V \subseteq U$. Consequently, let $f \in A$ be such that $D_U(f) \subseteq V$ and contains x, where $D_U(f) = \{\mathfrak{p} \in U \mid f \notin \mathfrak{p}\}$. We thus have $x \in D_U(f) \subseteq V \subseteq U$. Consider the restriction $h = \rho_{U,V}(f) \in \mathcal{O}_X(V) = B$. We claim that $D_V(h) = D_U(f)$. Denote $\varphi : A \to B$ obtained by $V \subseteq U$. We then have that $\rho_{U,V} = \varphi$ and $h = \varphi(f)$. Thus $\mathfrak{q} \in D_V(h) \iff h \notin \mathfrak{q} \iff \varphi(f) \notin \mathfrak{q} \iff f \notin \varphi^{-1}(\mathfrak{q})$. As each $\mathfrak{p} \in D_U(f)$ is $\varphi^{-1}(\mathfrak{q})$ for some $\mathfrak{q} \in V$, therefore we are done. The last statement is immediate from above. $\square$

Closed subschemes are defined in not that obvious way in which we have defined open subschemes, but at any rate, they are natural. We motivate the need for ideal sheaves as follows. Let X be a scheme. Suppose a closed subset $C \hookrightarrow X$ intersects some collection of affine opens $\{\mathrm{Spec}\,(A_i)\}$ and moreover it happens that $C \cap \mathrm{Spec}\,(A_i) = C \cap \mathrm{Spec}\,(A_j)$ for some $i \neq j$. Now by Corollary 1.4.4.14 we may write $C \cap \mathrm{Spec}\,(A_i) = \mathrm{Spec}\,(A_i/\mathfrak{a}_i)$ and $C \cap \mathrm{Spec}\,(A_j) = \mathrm{Spec}\,(A_j/\mathfrak{a}_j)$ for some ideals $\mathfrak{a}_i \subseteq A_i$ and $\mathfrak{a}_j \subseteq A_j$. Hence, we get two different structure sheaves $\mathcal{O}_{\mathrm{Spec}(A_i/\mathfrak{a}_i)}$ and $\mathcal{O}_{\mathrm{Spec}(A_j/\mathfrak{a}_j)}$ on on open subset of C. Thus we have to systematically track such identifications in order to define a unique scheme structure on the closed set C. Indeed, we take the help of the rich amount of constructions that we can make on the category of sheaves over a space (for more information, see Section 3.5).

We first define closed immersions.

Definition 1.4.4.4. (**Closed immersions**) A map $f : X \to Y$ of schemes is a closed immersion if $f : X \to f(X)$ is a homeomorphism, $f(X) \subseteq Y$ is closed and $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$ is a surjective map.

Remark 1.4.4.5. Let $f : X \to Y$ be a closed immersion, so that $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$ is surjective. This is equivalent to saying that for each point $x \in X$, the map on stalks (see Theorem 20.3.0.6 and Lemma 20.5.0.5)

$$f^\flat_{f(x)} : \mathcal{O}_{Y,f(x)} \longrightarrow \mathcal{O}_{X,x}$$

is surjective. Observe that the above map is NOT the usual map on stalks $f^\sharp_x : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$. Further observe that since $f^\flat$ is surjective, therefore we have an ideal (see Section 3.5, Global algebra for more details) $\mathcal{I} = \mathrm{Ker}\left(f^\flat\right) \leq \mathcal{O}_Y$. We will later see that a closed subscheme is completely determined by this ideal sheaf and in-fact these ideal sheaves gives us a family of good examples of what will later be called quasicoherent modules over a scheme.

Remark 1.4.4.6. Let $f : X \to Y$ be a closed immersion. Then, the map $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$ is surjective. Pick any $x \in X$. Since we have the following commutative square for any open set $V \ni f(x)$ in Y

$$\begin{tikzcd} \mathcal{O}_Y(V) \arrow[r, "f^\flat_V"] \arrow[d] & \mathcal{O}_X(f^{-1}(V)) \arrow[d] \\ \mathcal{O}_{Y,f(x)} \arrow[r, "f^\sharp_x"'] & \mathcal{O}_{X,x} \end{tikzcd}.$$

It then follows from surjectivity of $f^\flat$ and $f : X \to f(X)$ being a homeomorphism that the local homomorphiosm $f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is surjective. It is also a simple exercise to see that surjectivity of $f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ for all $x \in X$ implies surjectivity of $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$.

Consequently, $f : X \to Y$ is a closed immersion if and only if f is a topological closed immersion and for all $x \in X$, the local homomorphism $f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is surjective.

A closed subscheme is then defined to be an isomorphism class of closed immersions.

Definition 1.4.4.7. (**Closed subscheme & ideal sheaf**) Let Y be a scheme. A closed subscheme of Y is an isomorphism class of closed immersions over Y. That is, a closed subscheme is the class $[f : X \to Y]$ of closed immersions where two closed immersions $f : X \to Y$ and $f' : X' \to Y$ are identified if there is an isomorphism $X \overset{\cong}{\to} X'$ such that the following commutes

$$\begin{tikzcd} X \arrow[rr, "\cong"] \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\ & Y & \end{tikzcd} .$$

For a closed subscheme $f : X \to Y$, we define kernel of $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$ to be the ideal sheaf corresponding to the closed subscheme f.

Remark 1.4.4.8. Note that this definition is not "unnatural" as every closed immersion $f : X \to Y$ defines a closed set $f(X) \subseteq Y$ and a scheme structure over it. We then just define a closed subscheme to be the data of this closed set together with its scheme structure that is given by f. Clearly to make such a definition via immersions, we would need to identify those immersions which give same scheme structure on $f(X) \subseteq Y$.

We define an immersion as follows.

Definition 1.4.4.9 (**Immersion**)**.** A map $f : X \to Z$ is said to be an immersion if f is an open immersion into a closed subscheme of Z.

We first understand closed subscheme structures in affine schemes.

Lemma 1.4.4.10. *Let* $X = \operatorname{Spec}(R)$ *be an affine scheme. Then every ideal* $\mathfrak{a} \leq R$ *defines a closed subscheme of* X.

Proof. Consider the closed set $Y = V(\mathfrak{a}) \subseteq X$. We endow Y with a scheme structure given by the isomorphism $Y \cong \operatorname{Spec}(R/\mathfrak{a})$. Now the inclusion map $i : (Y, \mathcal{O}_{\operatorname{Spec}(R/\mathfrak{a})}) \to X$ is clearly a topological closed immersion. Further, $i^\flat : \mathcal{O}_{\operatorname{Spec}(R)} \to i_*\mathcal{O}_{\operatorname{Spec}(R/\mathfrak{a})}$ is given on stalks (see Lemma 20.5.0.5) at point $x \in Y$ as $\mathcal{O}_{\operatorname{Spec}(R),x} \to \mathcal{O}_{\operatorname{Spec}(R/\mathfrak{a}),x}$ which is just $R_x \to (R/\mathfrak{a})_x$ which is surjective. Thus, $\mathfrak{a}$ defines a closed subscheme structure on Y. □

It is important to note that any other ideal $\mathfrak{b} \leq R$ such that $V(\mathfrak{a}) = V(\mathfrak{b})$ will define a possibly different closed subscheme structure on the underlying topological space. This is another example of the phenomenon that algebra has much more finer control over the geometric situation at hand. For example, for $X = \operatorname{Spec}(k[x])$, we have $\mathfrak{a}_n = \langle x^n \rangle$ and note that $V(\mathfrak{a}_n) = \{\langle x \rangle\} \subseteq X$. But each ideal $\mathfrak{a}_n$ defines a new closed subscheme structure on the same point $\langle x \rangle \in X$.

Properties of closed immersions

We discuss some general properties of closed immersions. We begin by observing that closed immersions are local on target.

Proposition 1.4.4.11. *Let $f : X \to Y$ be a morphism of schemes. Then the following are equivalent:*
1. *f is a closed immersion.*
2. *There is an affine open cover $\{V_i\}$ of Y such that $f : f^{-1}(V_i) \to V_i$ is a closed immersion for each i.*

Proof. (1. $\Rightarrow$ 2.) As f is a closed immersion, then $f(X) \subseteq Y$ is a closed subset and $f : X \to f(X)$ is a homeomorphism. Pick any open affine $V = \mathrm{Spec}\,(B) \subseteq Y$. Then, we wish to show that $f : f^{-1}(V) \to V$ is a closed immersion. Indeed, as f is a closed immersion, therefore $f : f^{-1}(V) \to V \cap f(X)$ is a homeomorphism. As $f(X)$ is closed in Y, therefore $V \cap f(X)$ is closed in V. This shows that $g := f|_{f^{-1}(V)}$ is a topological closed immersion.

Next, we wish to show that the map $g^\flat : \mathcal{O}_V \to g_*\mathcal{O}_{f^{-1}(V)}$ is a surjection. By Remark 1.4.4.6, it suffices to show that for any $x \in f^{-1}(V)$, the local morphism $g_x^\sharp : \mathcal{O}_{V,f(x)} \to \mathcal{O}_{f^{-1}(V),x}$ is a surjection. Since $g = f|_{f^{-1}(V)}$, therefore $g_x^\sharp = f_x^\sharp$ because stalks commute with restrictions. Consequently, we wish to show that $f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is a surjection, but this is true by Remark 1.4.4.6 and the fact that f is a closed immersion.

(2. $\Rightarrow$ 1.) We first wish to show that f is a topological closed immersion. We first establish that f is a homeomorphism onto its image. Indeed, we have $f_i = f|_{f^{-1}(V_i)} : f^{-1}(V_i) \to V_i \cap f(X)$ a homeomorphism for each i. Consequently, we have a map $g_i : V_i \cap f(X) \to f^{-1}(V_i)$ which is a continuous inverse of f_i. Clearly g_i forms a matching family for $f(X) = \bigcup_i V_i \cap f(X)$ and thus can be glued to form a global inverse $g : f(X) \to X$ of f. Consequently, $f : X \to f(X)$ is a homeomorphism.

We wish to show that $f(X)$ is closed in Y. As being a closed set is a local property, therefore we need only check that $V_i \cap f(X)$ is a closed set in V_i, but this is exactly what our hypothesis that $f_i : f^{-1}(V_i) \to V_i$ a closed immersion guarantees.

Finally, we wish to show, by Remark 1.4.4.6, that $f_x^\sharp : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is a surjection for each $x \in X$. Indeed, as taking germs commute with restrictions, therefore $f_x^\sharp$ is the same local homomorphism as $(f_i)_x^\sharp : \mathcal{O}_{V_i,f(x)} \to \mathcal{O}_{f^{-1}(V_i),x}$ where $f(x) \in V_i$, which is surjective as f_i is a closed immersion. $\square$

The following shows that closed immersions are stable under base change.

Proposition 1.4.4.12. [14] *Let $f : X \to Y$ be a closed immersion and $g : Y' \to Y$ be any other map. Then, the map $p : X \times_Y Y' \to Y'$ is a closed immersion.*

Proof. As $f : X \to Y$ is a closed immersion, therefore by Proposition 1.4.4.11, there is an affine open cover $\{V_i = \mathrm{Spec}\,(B_i)\}$ of Y such that $f : f^{-1}(V_i) \to V_i$ is a closed immersion. Consequently, $f^{-1}(V_i) \cong f(f^{-1}(V_i)) \subseteq V_i$ is a closed subscheme, thus $f^{-1}(V_i) \cong \mathrm{Spec}\,(B_i/\mathfrak{b}_i)$ (see Corollary 1.4.4.14). Consider $g^{-1}(V_i) \subseteq Y'$ and cover it by open affines U_{ij}. Hence, we obtain an affine open cover of Y' given by $\{U_{ij} = \mathrm{Spec}\left(B'_{ij}\right)\}_{i,j}$. We claim that $p^{-1}(U_{ij}) \to U_{ij}$ is a closed immersion. Indeed, by Lemma 1.6.4.8, we have $p^{-1}(U_{ij}) \cong U_{ij} \times_{V_i} f^{-1}(V_i) \cong \mathrm{Spec}\left(B'_{ij} \otimes_{B_i} B_i/\mathfrak{b}_i\right) \cong$

[14] Exercise II.3.11, a of Hartshorne.

Spec $\left(B'_{ij}/\mathfrak{b}_i B'_{ij}\right)$, which thus makes $p : p^{-1}(U_{ij}) \to U_{ij}$ equivalent to the scheme morphism Spec $\left(B'_{ij}/\mathfrak{b}_i B'_{ij}\right) \to$ Spec $\left(B'_{ij}\right)$ obtained by the natural quotient homomorphism (this follows from the tensor product square obtained by the fiber product $U_{ij} \times_{V_i} f^{-1}(V_i)$). Consequently, it is a closed immersion by Proposition 1.2.2.8, 3, as required. □

Closed subschemes and ideal sheaves

We now study closed subschemes of arbitrary schemes. To read the following results, see Section 1.9 on quasicoherent modules.

Proposition 1.4.4.13. *Let X be a scheme.*

1. *If $\mathcal{I} \leq \mathcal{O}_X$ is the ideal sheaf of a closed subscheme $Y \hookrightarrow X$, then $\mathcal{I}$ is a quasicoherent $\mathcal{O}_X$-module. If further X is Noetherian, then $\mathcal{I}$ is coherent.*
2. *If $\mathcal{I} \leq \mathcal{O}_X$ is an ideal of $\mathcal{O}_X$ such that it is quasicoherent, then $\mathcal{I}$ determines a unique closed subscheme $Y \hookrightarrow X$ where Y is given by* Supp $(\mathcal{O}_X/\mathcal{I})$.
3. *Consequently, we have a correspondence*

$$\left\{\begin{matrix}\text{Quasicoherent ideal} \\ \text{sheaves } \mathcal{I} \leq \mathcal{O}_X \text{ upto} \\ \text{isomorphism}\end{matrix}\right\} \cong \left\{\begin{matrix}\text{Closed subschemes } Y \hookrightarrow \\ X\end{matrix}\right\}.$$

Proof. 1. This follows from the following facts; closed subschemes are quasicompact separated maps, that direct image of quasicoherent is quasicoherent for such maps and that kernels of maps of quasicoherent modules is quasicoherent. The second statement follows from reducing to affine and using the fact that we know all quasicoherent modules over affine.
2. Pick an ideal sheaf $\mathcal{I} \leq \mathcal{O}_X$ which is quasicoherent and let $Y = \text{Supp}\,(\mathcal{O}_X/\mathcal{I}) := \{x \in X \mid \mathcal{O}_{X,x}/\mathcal{I}_x \neq 0\}$. Then consider $i : (Y, \mathcal{O}_X/\mathcal{I}) \hookrightarrow (X, \mathcal{O}_X)$. It is straightforward to see that the kernel of $i^\flat$ is exactly $\mathcal{I}$. We wish to show that this is a topological closed immersion and that the map $i^\flat$ is surjective. Clearly i is homeomorphic to its image, thus we need only show that its image is a closed set. This is a local property, so let $X = \text{Spec}\,(R)$, so that $\mathcal{I} = \tilde{\mathfrak{a}}$ for an ideal $\mathfrak{a} \leq R$. Now $Y = \{\mathfrak{p} \in \text{Spec}\,(R) \mid (R/\mathfrak{a})_\mathfrak{p} \neq 0\} = \{\mathfrak{p} \in \text{Spec}\,(R) \mid \mathfrak{p} \supseteq \mathfrak{a}\} = V(\mathfrak{a})$. Thus i is a topological closed immersion. Now the surjectivity of the map $i^\flat : \mathcal{O}_X \to i_*\mathcal{O}_X/\mathcal{I}$ follows from going to stalks via Lemma 20.5.0.5. The uniqueness of $(Y, \mathcal{O}_X/\mathcal{I})$ w.r.t. $\mathcal{I}$ is clear. □

Note that the main use of quasicoherence of $\mathcal{I}$ in statement 2 was to make sure that the support of $\mathcal{O}_X/\mathcal{I}$ is indeed closed. We have a straightforward, but important corollary.

Corollary 1.4.4.14. *Let $X =$ Spec (A) be an affine scheme. We have the following bijection*

$$\{\text{Closed subschemes } Y \hookrightarrow X\} \underset{(\text{Spec}(A/\mathfrak{a}), \widetilde{A/\mathfrak{a}}) \leftarrow\!\shortmid \mathfrak{a}}{\overset{\mathcal{I} \mapsto \Gamma(\mathcal{I}, X)}{\underset{\longleftarrow}{\overset{\longrightarrow}{\cong}}}} \{\text{Ideals } \mathfrak{a} \leq A\}/\cong .$$

Note that $\widetilde{A/\mathfrak{a}} \cong \mathcal{O}_{\text{Spec}(A/\mathfrak{a})}$.

Proof. Follows immediately from Proposition 1.4.4.13 and Corollary 1.9.1.12. □

1.5 Varieties

Most examples of schemes that we will encounter in the wild are quasi-projective/affine varieties. Therefore, we first cover them in a semi-classical setting not involving schemes. We will then show how to interpret them as finite type separated integral schemes over the base field. This will enable us to use the machinery we will be developing for schemes in the study of varieties. Indeed, by the end of this section, we will comfortably replace the definition of a variety to mean a separated, integral finite type scheme over an algebraically closed field.

1.5.1 Varieties over an algebraically closed field-I

We define varieties as zero sets of certain polynomials over an algebraically closed field k. We assume that the reader is aware of the Zariski topology that is present over $\mathbb{A}^n_k$. Let us first give the classical version of affine varieties.

Definition 1.5.1.1. (**Affine algebraic variety**) Let k be an algebraically closed field and let $\mathbb{A}^n_k$ be the affine n-space. An affine algebraic variety is an irreducible closed subset of $\mathbb{A}^n_k$.

We recall that the Hilbert Nullstellensatz further tells us that for any ideal $\mathfrak{a} \leq k[x_1, \dots, x_n]$, the zero set of the ideal $Z(\mathfrak{a}) \subseteq \mathbb{A}^n_k$ is such that the ideal it generates is equal to the radical of the ideal, $I(Z(\mathfrak{a})) = \sqrt{\mathfrak{a}}$.

Let $A \subseteq \mathbb{A}^n_k$ be an affine algebraic set. Then, the *affine coordinate ring* of A is defined to be the following finitely generated k-algebra

$$k[A] := \frac{k[x_1, \dots, x_n]}{I(A)}$$

where $I(A) \leq k[x_1, \dots, x_n]$ is the ideal generated by A. An important simple lemma to keep in mind for future is the following.

Lemma 1.5.1.2. *Let k be an algebraically closed field. Then B is a finitely generated k-algebra without nilpotent elements if and only if B is an affine coordinate ring of an algebraic set.*

Proof. One side is trivial and the other uses Nullstellensatz. □

Example 1.5.1.3. (*Hyperboloid of one sheet*) A recurring example that we choose to study in this notebook, amongst the others, is the hyperboloid of one sheet. This is given by the following equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

In the affine space over $\mathbb{R}$, $\mathbb{A}^3_{\mathbb{R}}$, we can draw it as shown in Figure 1.1.
We may simply call it a hyperboloid. This hyperboloid determines an affine variety given by the zero set of the polynomial

$$p(x, y, z) = x^2/a^2 + y^2/b^2 - z^2/c^2 - 1 \in k[x, y, z]$$

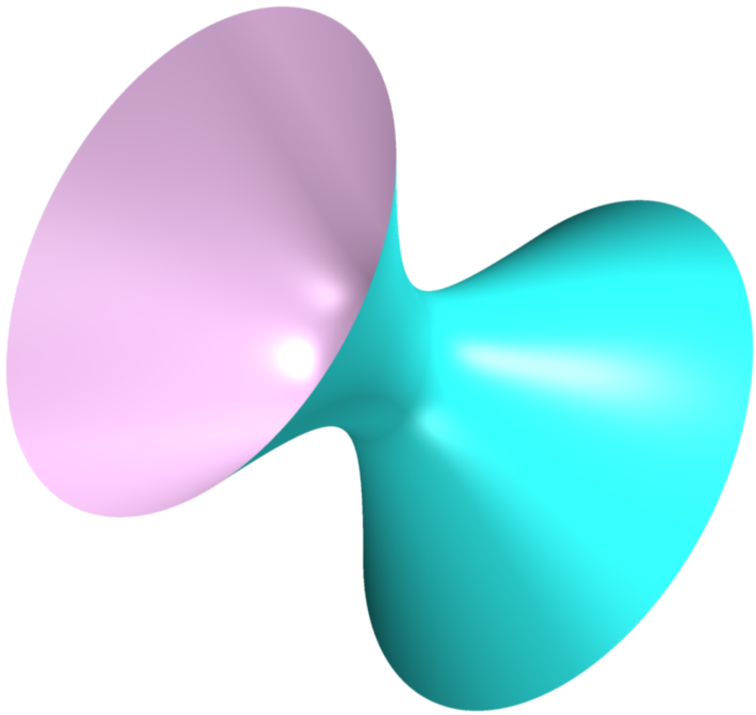


Figure 1.1: A hyperboloid of one sheet as a subvariety of $\mathbb{A}^3_{\mathbb{R}}$. The parameters are $a = 1.05, b = 1.05, c = 1$.

for any field k. Let $X = V(p) \subseteq \mathbb{A}^3_k$. The coordinate ring is given by

$$k[X] = \frac{k[x, y, z]}{I(V(p))}.$$

As we shall see, we will associate to the above variety $(X, \mathcal{O}_X)$ a scheme by considering the spectrum of the coordinate ring, Spec $(k[X])$.

We will understand this fantastic example in much more detail as we develop more tools to handle it.

We now define projective varieties. Consider an algebraically closed field. Then the *projective n-space* is defined to be the quotient $\mathbb{P}^n_k := \mathbb{A}^{n+1}_k / \sim$ where $(a_0, \dots, a_n) \sim (b_0, \dots, b_n)$ if and only if there exists $\lambda \in k^\times$ such that $a_i = \lambda b_i$ for all $i = 0, 1 \dots, n$. A point of $\mathbb{P}^n_k$ is denoted by $[a_0 : \dots : a_n]$ and this presentation of the point is called the homogeneous coordinates of the point. Assuming that the reader is aware about graded rings and the natural grading of $k[x_0, \dots, x_n]$, we observe that we can talk about the *zeroes of a homogeneous polynomial* $p(X) \in k[x_0, \dot; x_n]$ as follows:

$$Z(p) := \{P \in \mathbb{P}^n_k \mid p(P) = 0\}.$$

Indeed, one observes that a homogeneous polynomial is zero at a point $P \in P^n_k$ in a manner which is independent of the choice of representation of P in terms of the homogeneous coordinates of P.

With this in our hand, we further define the zero set of a homogeneous ideal $\mathfrak{a} \leq k[x_0, \ldots, x_n]$ as

$$Z(\mathfrak{a}) := \{P \in \mathbb{P}^n_k \mid f(P) = 0 \forall f \in T_\mathfrak{a}\}$$

where $T_\mathfrak{a}$ is the set of all homogeneous elements of $\mathfrak{a}$. Remember that an ideal in a graded ring is homogeneous if and only if it is generated by the set of all of its homogeneous elements.

Lemma 1.5.1.4. *Let k be a field. Then*

1. *For any two homogeneous ideals $\mathfrak{a}, \mathfrak{b} \leq k[x_0, \ldots, x_n]$, we have $Z(\mathfrak{a}\mathfrak{b}) = Z(\mathfrak{a}) \cup Z(\mathfrak{b})$.*
2. *For any family of homogeneous ideals $\{\mathfrak{a}_i\}_{\{}i \in I\}$, we have $\cap_{i \in I} Z(\mathfrak{a}_i) = Z(\sum_{i \in I} \mathfrak{a}_i)$.*

Proof. Straightforward unravelling of definitions. □

Therefore we obtain a topology on $\mathbb{P}^n_k$ where a set $Y \subseteq \mathbb{P}^n_k$ is closed if and only if $Y = Z(\mathfrak{a}_i)$ for a homogeneous ideal $\mathfrak{a}_i$ of $k[x_0, \ldots, x_n]$. This is called the Zariski topology of $\mathbb{P}^n_k$.

Definition 1.5.1.5. (**Projective algebraic variety**) Let k be an algebraically closed field. An irreducible algebraic set of $\mathbb{P}^n_k$ is said to be a projective algebraic variety in $\mathbb{P}^n_k$.

Let $V \subseteq \mathbb{P}^n_k$ be a projective algebraic variety. Then the *ideal generated by V* in $k[x_0, \ldots, x_n]$ is $I(V)$ which is the ideal generated by the following set of homogeneous polynomials: $\{f \in k[x_0, \ldots, x_n] \mid f \text{ is homogeneous and } f(P) = 0\}$.

For a projective algebraic set $Y \subseteq \mathbb{P}^n_k$, we define its *homogeneous coordinate ring* to be the following k-algebra

$$k[Y] := \frac{k[x_0, \ldots, x_n]}{I(Y)}$$

where $I(Y)$ is the homogeneous ideal of Y.

Definition 1.5.1.6 (Zero set and ideal of an algebraic set). Define for any set $T \subseteq k[x_0, \dot{;} x_n]$ of homogeneous elements the zero set of T as $Z(T) = \{p \in \mathbb{P}^n_k \mid f(p) = 0 \ \forall \ f \in T\}$. For any $Y \subseteq \mathbb{P}^n_k$, define $I(Y)$ as the ideal in $k[x_0, \ldots, x_n]$ generated by $\{f \in k[x_0, \ldots, x_n] \mid f \text{ is homogeneous } \& \ f(p) = 0 \ \forall p \in Y\}$.

To distinguish between affine and projective cases, we will reserve $Z(\mathfrak{a})$ for zero set of a homogeneous ideal in projective space and $V(\mathfrak{a})$ as the zero set of an ideal in the affine space.

We now show that how the projective space $\mathbb{P}^n_k$ is covered by $n+1$ copies of affine space $\mathbb{A}^n_k$. Before that we discuss few maps which allows us to treat affine case projectively.

Homogenization and dehomogenization

One way to move back and from affine to projective setting is to use to fundamental functions between $k[y_1, \ldots, y_i, \ldots, y_n]$ and $k[x_0, \ldots, x_n]_h$.

Definition 1.5.1.7. (**(De)homogenization**) Let k be an algebraically closed field and let $A := k[y_1, \ldots, y_n]$ and $B := k[x_0, \ldots, x_n]_h$, the set of all homogeneous polynomials in $k[x_0, \ldots, x_n]$. Consider the following two functions

$$\begin{aligned} d_i : B &\longrightarrow A \\ f(x_0, \ldots, x_n) &\longmapsto f(x_0, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) \\ h_i : A &\longrightarrow B \\ g(y_1, \ldots, y_n) &\longmapsto x_i^e g\left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i}\right) \end{aligned}$$

where e is the degree of g and $i = 0, \ldots, n$. The map h_i is called the i^{th}-homogenization map and d_i is called the i^{th}-dehomogenization map.

Using this, we can establish the result in question.

Proposition 1.5.1.8. *Let k be an algebraically closed field and consider the projective n-space over k, $\mathbb{P}_k^n$. Then, there exists $n+1$ open subspaces say $U_i \subseteq \mathbb{P}_k^n$, such that $\mathbb{P}_k^n = \bigcup_{i=0}^n U_i$ and for each i, U_i is homeomorphic to $\mathbb{A}_k^n$.*

Proof. Consider the $n+1$ open subspaces of $\mathbb{P}_k^n$ as follows:

$$U_i := \mathbb{P}_k^n \setminus H_i$$

where $H_i = Z(\langle x_i \rangle)$ is the algebraic set obtained by all those points whose i^{th} homogeneous coordinate is zero. Now consider the map

$$\begin{aligned} \varphi_i : U_i &\longrightarrow \mathbb{A}_k^n \\ [a_0 : \cdots : a_n] &\longmapsto \left(\frac{a_0}{a_i}, \ldots, \frac{a_{i-1}}{a_i}, \frac{a_{i+1}}{a_i}, \ldots, \frac{a_n}{a_i}\right). \end{aligned}$$

One can check that this pulls closed sets to closed sets by using the i^{th}-homogenization map. Conversely, one can define the map

$$\begin{aligned} \theta_i : \mathbb{A}_k^n &\longrightarrow U_i \\ (a_1, \ldots, a_n) &\longmapsto (a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n) \end{aligned}$$

and this can again be checked to be continuous by an application of i^{th} dehomogenization map. ☐

Corollary 1.5.1.9. *Let k be an algebraically closed field and $Y \subseteq \mathbb{P}_k^n$ be a projective algebraic variety. Then, in the notation of Proposition 1.5.1.8, for each $i = 0, \ldots, n$, $Y \cap U_i$ is an affine algebraic variety.*

Proof. This follows from the observation that $Y \cap U_i$ is a closed set of $U_i \cong \mathbb{A}_k^n$. The irreducibility follows from the fact that open subsets of irreducible spaces are irreducible. ☐

Properties of algebraic sets in $\mathbb{P}^n_k$

We now present some basic properties of algebraic sets in $\mathbb{P}^n_k$.

Lemma 1.5.1.10. [15] *(Homogeneous Nullstellensatz) Let k be an algebraically closed field and let $\mathfrak{a} \leq k[x_0, \dots, x_n]$ be a homogeneous ideal. Then,*

$$I(Z(\mathfrak{a})) = \sqrt{\mathfrak{a}}.$$

Proof. Denote by $V(\mathfrak{a}) \subseteq \mathbb{A}^{n+1}$ to be the vanishing set of $\mathfrak{a}$ in the affine $n+1$-space. This is called the *affine cone* of the ideal $\mathfrak{a}$ in $\mathbb{A}^{n+1}$. We claim that $I(Z(\mathfrak{a})) \hookrightarrow I(V(\mathfrak{a}))$ since if $f \in I(Z(\mathfrak{a}))$ is homogeneous, then $f(P) = 0$ for all $P \in Z(\mathfrak{a}) = \{P \in \mathbb{P}^n_k \mid g(P) = 0\ \forall g \in \mathfrak{a}\}$. Pick any point $Q \in V(\mathfrak{a}) \subseteq \mathbb{A}^{n+1}_k$. We see that $g(Q) = 0$ for all $g \in \mathfrak{a}$. We wish to show that $f(Q) = 0$. As any point $Q \in V(\mathfrak{a})$ determines a point $P \in Z(\mathfrak{a})$ by scaling, that is $P = \lambda Q$, we get by homogeneity of f that $f(Q) = f(\lambda P) = \lambda^d f(P) = 0$, that is, $f \in I(V(\mathfrak{a}))$, as required. By affine Nullstellensatz, it follows that $I(Z(\mathfrak{a})) \subseteq \sqrt{a}$. The converse is straightforward. □

The following tells us when is a projective algebraic set is empty.

Lemma 1.5.1.11. [16] *Let $\mathfrak{a} \leq k[x_0, \dots, x_n] = S$ be a homogeneous ideal. Then, the following are equivalent:*
1. *$Z(\mathfrak{a}) = \emptyset$ in $\mathbb{P}^n_k$,*
2. *$\sqrt{\mathfrak{a}}$ is either S or S_+,*
3. *$\mathfrak{a} \supseteq S_d$ for some $d > 0$.*

Proof. (1. $\Rightarrow$ 2.) The main idea here is again to reduce to affine case by considering the affine cone. Observe that if $Z(\mathfrak{a}) = \emptyset$, then $V(\mathfrak{a}) \subseteq \{0\}$ (where $V(\mathfrak{a})$ is the vanishing in $\mathbb{A}^{n+1}_k$ as in the proof of Lemma 1.5.1.13). Indeed, if not then there exists $p = (p_0, \dots, p_n) \in V(\mathfrak{a})$ such that $p \neq 0$. It follows that $[p_0 : \dots : p_n] \in Z(\mathfrak{a})$ since any homogeneous element f of $\mathfrak{a}$ vanishes at p in $\mathbb{A}^{n+1}_k$. Now if $V(\mathfrak{a}) = \emptyset$, then by the affine nullstellensatz, we get $\sqrt{\mathfrak{a}} = S$. If $V(\mathfrak{a}) = 0$, then $\sqrt{a} = I(0) = \langle x_0, \dots, x_n \rangle = S_+$.

(2. $\Rightarrow$ 1.) As $\sqrt{a} = I(V(\mathfrak{a})) = S$ or S_+, therefore $V(\sqrt{\mathfrak{a}}) = V(\mathfrak{a}) = \emptyset$ or 0. It follows again that $Z(\mathfrak{a}) = \emptyset$.

(2. $\Rightarrow$ 3.) **TODO.** □

Akin to affine varieties, we also have some basic results in projective algebraic sets.

Lemma 1.5.1.12. [17] *Let $\mathbb{P}^n_k$ be the projective n-space over k and let $S = k[x_0, \dots, x_n]$*
1. *If $Y_1 \subseteq Y_2$ in $\mathbb{P}^n_k$, then $I(Y_1) \supseteq I(Y_2)$.*
2. *If $T_1 \subseteq T_2$ in S be subsets of homegeneous elements, then $Z(T_1) \supseteq Z(Y_2)$.*
3. *If $Y_1, Y_2 \subseteq \mathbb{P}^n_k$, then $I(Y_1 \cup Y_2) = I(Y_1) \cap I(Y_2)$.*
4. *If $Y \subseteq \mathbb{P}^n_k$, then $Z(I(Y)) = \overline{Y}$.*

Proof. content... □

[15] Exercise I.2.1 of Hartshorne.
[16] Exercise I.2.2 of Hartshorne.
[17] Exercise I.2.3 of Hartshorne.

Some consequences of the homogeneous nullstellensatz yields us the familiar results as in the affine case.

Lemma 1.5.1.13. [18] *Let k be an algebraically closed field and consider the projective n-space $\mathbb{P}^n_k$. Then,*

1. *There is a bijection*

$$\{\text{All algebraic sets } Y \subseteq \mathbb{P}^n_k\} \underset{Z}{\overset{I}{\rightleftarrows}} \{\text{All homogeneous radical ideals of } k[x_0, \dots, x_n]\} \ .$$

2. *An algebraic set $Y \subseteq \mathbb{P}^n_k$ is irreducible if and only if $I(Y)$ is a prime ideal in $k[x_0, \dots, x_n]$.*
3. *$\mathbb{P}^n_k$ is a projective algebraic variety.*

Remark 1.5.1.14. A corollary of the above lemma is that one can look at projective algebraic varieties in $\mathbb{P}^n_k$ akin to homogeneous prime ideals in $k[x_0, \dots, x_n]$, thus telling us another hint at how the idea of schemes might have looked back in the days.

Proof of Lemma 1.5.1.13. 1. This is a direct consequence of homogeneous nullstellensatz (Lemma 1.5.1.10) and the fact that $Z(I(Y)) = \bar{Y}$ for any $Y \subseteq \mathbb{P}^n_k$.

2. (L $\Rightarrow$ R) Suppose $Y = Z(\mathfrak{a})$ is irreducible and $I(Z(\mathfrak{a})) = \sqrt{\mathfrak{a}}$ is not prime. Then there exists $f, g \notin \mathfrak{a}$ such that $fg \in \sqrt{\mathfrak{a}}$. Consider the ideals $\mathfrak{b} := \langle f, \sqrt{\mathfrak{a}} \rangle$ and $\mathfrak{c} := \langle g, \sqrt{\mathfrak{a}} \rangle$. We then observe that $Z(\mathfrak{b}), Z(\mathfrak{c}) \subseteq Z(\mathfrak{a})$ and $Z(\mathfrak{b}) \cup Z(\mathfrak{c}) = Z(\mathfrak{b}\mathfrak{c}) = Z(\sqrt{\mathfrak{a}}) = Z(\mathfrak{a})$, where we have used Lemma 1.5.1.10 in the second last equation and the fact that $fg \in \sqrt{a}$ in third last. This yields a contradiction to the irreducibility of Y.
(R $\Rightarrow$ L) Suppose $I(Y)$ is prime but Y is not irreducible. Consequently, there are proper closed sets $Y_1, Y_2 \subseteq Y$ such that $Y_1 \cup Y_2 = Y$. Further, we obtain that $I(Y_i) \gneq I(Y)$ for each $i = 1, 2$. It then follows that there exists $f_i \in I(Y_i) \setminus I(Y)$ such that $f_i \notin I(Y_j)$, $j \neq i$. Consequently, we have $f_1 f_2 \in k[x_0, \dots, x_n]$ such that $f_1 f_2(P) = f_1(P) f_2(P) = 0$ for all $P \in Y$, as $Y = Y_1 \cup Y_2$. We thus have a contradiction to primality of $I(Y)$.

3. Since $I(\mathbb{P}^n_k) = I(Z(\mathfrak{o})) = \sqrt{\mathfrak{o}} = \mathfrak{o}$, then by 2., $\mathbb{P}^n_k$ is irreducible. Note we have used the fact that $k[x_0, \dots, x_n]$ is an integral domain. $\square$

One of the reasons that one might be interested in projective varieties is that they "compactify" the question at hand, that is, there are no "missing points" in the ambient space. We will see more into this when we will see projective morphisms and invertible modules, but for now, it is good to keep in mind that reframing your question in the projective spaces/varieties may give you more handle (and of-course, machines) to solve the question at hand. In the same vein, we now see that every affine variety can be embedded compactly into a projective space, and this embedding is called the projective closure of the affine variety.

Definition 1.5.1.15. (**Projective closure of affine varieties**) Let k be an algebraically closed field and consider an affine variety $X \subseteq \mathbb{A}^n_k$. For any $i = 0, \dots, n$, consider the homeomorphism

$$\begin{aligned} \theta_i : \mathbb{A}^n_k &\longmapsto U_i \\ (a_1, \dots, a_n) &\longmapsto [1 : a_1 : \cdots : a_n], \end{aligned}$$

[18] Exercise I.2.4 of Hartshorne.

as we considered in Proposition 1.5.1.8. Then, the i^{th}-projective closure of X into $\mathbb{P}^n_k$ is given by the closure $\overline{\theta_i(Y)} \subseteq \mathbb{P}^n_k$ as a subspace in $\mathbb{P}^n_k$. We will usually say the 0th projective closure of X to be simply the *projective closure of X*.

Consider an affine variety $X \subseteq \mathbb{A}^n_k$ and consider $\overline{X} \subseteq \mathbb{P}^n_k$ to be the projective closure of X. Let $I(X) \leq k[y_1, \dots, y_n]$ be the affine ideal of X and let $I(\overline{X}) \leq k[x_0, \dots, x_n]$ be the homogeneous ideal of projective closure. A natural question is that how the homogeneous ideal $I(\overline{X})$ is connected to the affine ideal $I(X)$. The following proposition answers that.

Proposition 1.5.1.16. *Let k be an algebraically closed field and $X \subseteq \mathbb{A}^n_k$ be an affine variety. Let $I(X) \leq k[y_1, \dots, y_n]$ be the affine ideal of X and let $I(\overline{X}) \leq k[x_0, \dots, x_n]$ be the homogeneous ideal of projective closure. Then,*

$$I(\overline{X}) = \langle h_0(I(X)) \rangle$$

where $h_0 : k[y_1, \dots, y_n] \to k[x_0, \dots, x_n]$ is the 0th homogenization function (Definition 1.5.1.7).

Proof. Since $X \subseteq \mathbb{A}^n_k$ is irreducible and closure of irreducible is irreducible, therefore $\overline{X} \subseteq \mathbb{P}^n_k$ is irreducible. It would thus suffice to show that

$$\overline{X} = Z(h_0(I(X))).$$

Indeed, this would imply that $h_0(I(X))$ is a homogeneous prime ideal by Lemma 1.5.1.13, 1, thus applying $I(-)$ would yield the result. We therefore show the above equality. Consider any closed set $Y \supseteq X$ in $\mathbb{P}^n_k$. We then wish to show that $Y \supseteq Z(h_0(I(X)))$. Since $Y \subseteq \mathbb{P}^n_k$ is closed, therefore $Y = Z(\mathfrak{a})$ for some homogeneous ideal $\mathfrak{a}$ in $k[x_0, \dots, x_n]$. It would thus suffice to show that

$$\mathfrak{a} \hookrightarrow h_0(I(X)).$$

It would further suffice to show the above inclusion only for homogeneous elements, as $\mathfrak{a}$ is generated by homogeneous elements. Consequently, pick any homogeneous polynomial $f \in \mathfrak{a}$. We can write

$$f(x_0, \dots, x_n) = x_0^e g\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right)$$

for some $g \in k[y_1, \dots, y_n]$ and $e = \deg f$. In other words, $f = h_0(g)$. Now since $Y \supseteq X$, therefore $f(P) = 0 \forall P \in X \subseteq \mathbb{P}^n_k$, that is, if $P = [1 : a_1, \dots, a_n] \in X$, then $f(1, a_1, \dots, a_n) = 0$ and thus $g(a_1, \dots, a_n) = 0$. Hence $g \in I(X) \leq k[y_1, \dots, y_n]$. Thus $f = h_0(g)$ where $g \in I(X)$, that is, $f \in h_0(I(X))$, as required. $\square$

Dimension, hypersurfaces and complete intersections

Let us first understand how the notion of dimension plays out with the Krull dimension of homogeneous coordinate ring of a projective variety.

Proposition 1.5.1.17. *Let k be an algebraically closed field and $X \subseteq \mathbb{P}^n_k$ be a projective k-variety. Then,*

1. $\dim k[X] = \dim X + 1$,
2. $\dim X = \dim U_i \cap X$ *where $U_i \subseteq X$ is an affine open subset as in Proposition 1.5.1.8, for all $i = 0, \dots, n$.*

Proof. We will prove the two statements together. The main technique here is, as usual, to reduce the computations to one of the affine patches. Let $U_i \subseteq \mathbb{P}^n_k$ be the hyperplane where $x_i \neq 0$. We know that U_is covers $\mathbb{P}^n_k$ and each U_i is isomorphic to $\mathbb{A}^n_k$. Denote $X_i = U_i \cap X$ so that X_i is an open subvariety of X. Further denote $k[X]^h$ to be the homogeneous coordinate ring of X and $k[X_i]^a$ the affine coordinate ring of X_i. Note that $k[X_i]^a = k[x_0, \dots, \hat{x}_i, \dots, x_n]/d_i I(X)$ where d_i is the i^{th} dehomogenisation map. We would now like to note two things to move forward:

1. $\dim X = \dim X_j$ for some $j = 0, \dots, n$,
2. $k[X]^h_{x_i} \cong k[X_i]^a[x_i, 1/x_i]$[19].

The first statement is immediate from the fact that $\dim Y = \sup_i \dim U_i$ for any space Y with U_i an open covering. The second statement is the heart of the proof. Indeed, consider the map $k[X]^h_{x_i} \to k[X_i]^a[x_i, 1/x_i]$ which takes an element f/x_i^n and treats it as a polynomial in $x_i, 1/x_i$ with coefficients in $k[X_i]^a$. One immediately checks all the necessary conditions to ensure that this is an isomorphism.

Observe that if K/k is algebraic, then $K(x)/k(x)$ is algebraic. It follows that trdeg $k[X_i]^a[x_i, 1/x_i] = 1 + \text{trdeg } k[X_i]^a$. We now complete the proof. We may assume $\dim X = \dim X_0$. Consequently, via Proposition 1.5.3.10, 6 and Theorem 16.8.2.1, we obtain the following equalities:

$$\begin{aligned}\dim k[X]^h &= \text{trdeg } k[X]^h = \text{trdeg } k[X]^h_{x_0} = \text{trdeg } k[X_0]^a[x_0, 1/x_0] = 1 + \text{trdeg } k[X_0]^a \\ &= 1 + \dim k[X_0]^a = 1 + \dim X_0 = 1 + \dim X.\end{aligned}$$

The statement 2. follows from the following equalities:

$$\begin{aligned}\dim X_i &= \dim k[X_i]^a = \text{trdeg } k[X_i]^a_{x_i} = \text{trdeg } k[X_i]^a[x_i, 1/x_i] - 1 = \text{trdeg } k[X]^h_{x_0} - 1 \\ &= \text{trdeg } k[X]^h - 1 = \dim X + 1 - 1 = \dim X.\end{aligned}$$

□

We would now like to establish the following result, which will later motivate the definition of Weil divisors and of complete intersections.

Lemma 1.5.1.18. *Let k be an algebraically closed field and $X \subseteq \mathbb{P}^n_k$ be a projective k-variety. Then, the following are equivalent*

1. $\dim X = n - 1$.
2. *The homogeneous ideal $I(X) \leq k[x_1, \dots, x_n]$ is generated by a single irreducible homogeneous polynomial.*

Proof. (1. ⇒ 2.) By Proposition 1.5.1.17, 1, we have $\dim k[X] = n$, where $k[X] = k[x_0, \dots, x_n]/I(X)$. By Theorem 16.8.2.2, we have ht $I(X) = 1$. Since any height 1 prime ideal of a UFD is principal, therefore $I(X)$ is principal. Since $I(X)$ is homogeneous, therefore the statement 2. follows.
(2. ⇒ 1.) By Proposition 1.5.1.17, 2 and Theorem 16.8.2.2, we have

$$\dim X = \dim X_0 = \dim k[X_0]^a = n - \text{ht } d_0(I(X)).$$

We need only show that ht $d_0(I(X)) = 1$. Since $I(X) = \langle p(x_0, \dots, x_n) \rangle$, therefore $d_o(I(X)) = \langle p(1, x_1, \dots, x_n) \rangle$. Since $k[x_0, \dots, x_n]$ is a UFD and an easy observation about UFDs yields that height 1 prime ideals are exactly principal prime ideals, therefore the result follows. □

[19] This statement can be seen as a generalization of Lemma 1.5.3.11.

Cones

d-uple embedding

Veronese surface

Segre embedding

To complete d-uple, Vero Segre and q Chapter 1.

1.5.2 Morphism of varieties

We have defined affine and projective varieties so far. One would often, however, would like to know whether a subset of $\mathbb{A}^n$ or $\mathbb{P}^n$ is an open subspace of some affine or projective variety. Due to to this need, we define the following.

Definition 1.5.2.1. (**Quasi-affine/projective variety**) A subset X of $\mathbb{A}^n$ or $\mathbb{P}^n$ is said to be quasi-affine or quasi-projective if X is an open subset of an affine or projective variety, respectively.

Let X be a quasi-affine or projective variety. From our knowledge of geometry, we know that in a real C^α-manifold M, the right type of functions are those which are defined on open subsets of M as C^α-maps to $\mathbb{R}$, where the latter is treated as a C^α-manifold. Consequently, we are interested in the same type of maps to the affine line $\mathbb{A}^1_k$.

Definition 1.5.2.2. (**Regular maps**) This notion is defined differently for quasi-affine and quasi-projective varieties.

1. Let X be a quasi-affine variety. A function
$$\varphi : X \to \mathbb{A}^1_k$$
is said to be a *regular function* if for all $P \in X$, there exists an open subset $U \subseteq X$ such that $\varphi|_U = g/h$ where $g, h \in k[x_1, \dots, x_n]$ and $h(P) \neq 0\ \forall P \in U$.
2. Let X be a quasi-projective variety. A function
$$\varphi : X \to \mathbb{A}^1_k$$
is said to be a *regular function* if for all $P \in X$, there exists an open subset $U \subseteq X$ such that $\varphi|_U = g/h$ where $g, h \in k[x_0, \dots, x_n]$ are homogeneous polynomials of same degree and $h(P) \neq 0\ \forall P \in U$. Note that this defines a valid function to the affine line.

Indeed, regular maps are continuous.

Lemma 1.5.2.3. *Let X be a quasi-affine or quasi-projective variety and $\varphi : X \to \mathbb{A}^1_k$ be a regular function. Then φ is continuous.*

Proof. The Zariski topology on $\mathbb{A}^1_k$ is the cofinite topology, hence any closed set in $\mathbb{A}^1_k$ is a finite union of points of k. It thus suffices to show that for any $a \in k$, $Y := \varphi^{-1}(a) \subseteq X$ is closed. Since checking a set is closed is local in X, that is, $Y \subseteq X$ is closed if and only if there exists an open covering of X, say $\{U_\alpha\}$ such that $U_\alpha \cap Y$ is closed in U_α. We may thus replace X by an open subset of X where φ is represented as g/h for $g, h \in k[y_1, \dots, y_n]$ (in $k[x_0, \dots, x_n]$, homogeneous and of same degree in the projective case). Consequently, $\varphi^{-1}(a) \subseteq X$ is given by $\{P \in X \mid (g - ah)(P) = 0\}$ which in other words is $Z(g - ah)$ ($g - ah$ is homogeneous in the projective case). Thus $\varphi^{-1}(a)$ is closed, as required. $\square$

A simple corollary of above is the first striking result one learns in complex analysis for holomorphic maps (see Proposition 10.2.3.10).

Lemma 1.5.2.4. *(Identity principle) Let $\varphi, \xi : X \to \mathbb{A}^1_k$ be two regular maps over a quasi-affine or quasi-projective variety X. Then, $\varphi = \xi$ if and only if there exists an open set $U \subseteq X$ such that $\varphi = \xi$ over U.*

Proof. L $\Rightarrow$ R is easy. For R $\Rightarrow$ L, observe that for $\phi := \varphi - \xi$ is continuous by Lemma 1.5.2.3. Further, the set $\phi^{-1}(0) \subseteq X$ is closed and contains U. Since $\phi^{-1}(0) \supseteq U$ and U is an open set of an irreducible space, therefore U is dense in X. Consequently, $\phi^{-1}(0)$ is a closed and dense in X, hence is equal to X. □

We now define varieties in general.

Definition 1.5.2.5. (**Varieties**) Let k be an algebraically closed field. A variety over k is defined to be a quasi-affine or a quasi-projective variety in $\mathbb{A}^n_k$ or $\mathbb{P}^n_k$, respectively.

The notion of morphism of varieties is then given by functions which pulls regular functions back by pre-composition.

Definition 1.5.2.6. (**Map of varieties**) Let k be an algebraically closed field and let X, Y be two varieties over k. A map of varieties is a continuous function $f : X \to Y$ such that for any open set $V \subseteq Y$ and any regular function $\varphi : Y \to \mathbb{A}^1_k$, the function

$$\varphi \circ f : \varphi^{-1}(V) \to \mathbb{A}^1_k$$

is a regular function on the open set $\varphi^{-1}(V)$ of X. We may also call a map of varieties a *morphism of varieties*.

We therefore obtain the category of varieties over k, whose objects are varieties over k and arrows are maps of varieties. We will denote this category by

$$\mathbf{Var}_k.$$

Just like in topological spaces, it is not true in general that a bijective continuous map is a homeomorphism, similarly it is not true in general that a bijective map of varieties is an isomorphism of varieties, as the following example shows.

Example 1.5.2.7. Consider the affine line $\mathbb{A}^1_k$ and consider the affine variety $X := Z(y^2 - x^3) \subseteq \mathbb{A}^2_k$. The function

$$\begin{aligned} f : \mathbb{A}^1_k &\longrightarrow X \\ t &\longmapsto (t^2, t^3) \end{aligned}$$

is a map of varieties as for any open set $U \subseteq X$ and regular map $\varphi : X \to \mathbb{A}^1_k$, the composite $\varphi \circ f : \varphi^{-1}(U) \to \mathbb{A}^1_k$ is given by $t \mapsto \varphi(t^2, t^3)$ and then the regularity of this composite can be seen to be a result of regularity of φ. Further note that f induces an inverse continuous function

$$\begin{aligned} f^{-1} : X &\longrightarrow \mathbb{A}^1_k \\ (a, b) &\longmapsto ba^{-1}. \end{aligned}$$

Thus, $\mathbb{A}^1_k$ and X are homeomorphic as topological spaces. However, as varieties, they can not be isomorphic. Indeed, we shall soon see that coordinate rings are invariant of affine varieties and in our case $\mathbb{A}^1_k$ has $k[x]$ as its coordinate ring whereas X has $k[x,y]/\langle y^2 - x^3\rangle$ as its coordinate ring. These are not isomorphic as one is PID and the other is not.

We now construct some more algebraic gadgets on top of varieties and will prove how they will turn out to be invariants of the varieties under question. We have already seen one, the coordinate ring. We will now see the construction of others and we shall do it in a manner so that it is amenable to generalization to schemes, as is studied elsewhere in this chapter.

1.5.3 Varieties as locally ringed spaces

See Chapter 3, Foundational Geometry, for background on locally ringed spaces and basic global algebra. In this section, we would like to interpret varieties as locally ringed spaces, so that we can understand later that how a variety can be interpreted as a scheme. Clearly, for a variety X, we already have an underlying topological space X itself. To give X the structure of a locally ringed space, we need to consider a sheaf over X. We shall use regular functions over open sets of X for that.

Definition 1.5.3.1. (**Structure sheaf of a variety**) Let k be an algebraically closed field and X be a variety over k. For each open set $U \subseteq X$, consider the following set

$$\mathcal{O}_X(U) := \{f : U \to \mathbb{A}^1_k \mid f \text{ is regular}\}.$$

Further, for open $V \subseteq U$ in X, consider the function

$$\begin{aligned} \rho_{U,V} : \mathcal{O}_X(U) &\longrightarrow \mathcal{O}_X(V) \\ f &\longmapsto f|_V\,. \end{aligned}$$

This defines a sheaf of sets, as the following lemma shows.

Lemma 1.5.3.2. *The assignment $\mathcal{O}_X$ on open sets of a k-variety X as defined in Definition 1.5.3.1 defines a sheaf of sets over X.*

Proof. The locality axiom is straightforward as $\mathcal{O}_X(U)$ is a collection of functions, which thus can be checked locally for equality. It thus suffices to show that $\mathcal{O}_X$ satisfies the gluing axiom. Pick any open set U, an open covering $\{U_i\}_{i\in I}$ of U and a matching family $f_i \in \mathcal{O}_X(U_i)$ for each $i \in I$, that is $\rho_{U_i,U_i\cap U_j}(f_i) = \rho_{U_j,U_i\cap U_j}(f_j)$ for each $i,j \in I$. Consequently, we define $f : U \to \mathbb{A}^1_k$ given by $x \mapsto f_i(x)$ if $x \in U_i$. This is a well-defined function by the matching condition and further f is a regular function as for each point $x \in U$, f can be written as a rational function in some open neighborhood around x (essentially by regularity of f_is). Consequently, $\mathcal{O}_X$ is a sheaf. $\square$

Further, $\mathcal{O}_X$ is a sheaf of k-algebras if X is a k-variety.

Lemma 1.5.3.3. *Let k be an algebraically closed field and consider a k-variety X. The structure sheaf $\mathcal{O}_X$ of X is a sheaf of k-algebras.*

Proof. Indeed, $\mathcal{O}_X$ is a ring by point-wise addition and multiplication. Further, its a k-algebra via the injective ring homomorphism

$$\begin{aligned} k &\hookrightarrow \mathcal{O}_X(U) \\ c &\mapsto c : U \to \mathbb{A}^1_k \end{aligned}$$

where c is treated as the constant rational map. □

Hence, $(X, \mathcal{O}_X)$ is a k-ringed space. We now show that it is locally k-ringed.

Lemma 1.5.3.4. *Let k be an algebraically closed field and let X be a k-variety. Then, for all points $x \in X$, the stalk $\mathcal{O}_{X,x}$ is a local ring.*

Proof. We wish to show that $\mathcal{O}_{X,x}$ has a unique maximal ideal $\mathfrak{m}_x \leq \mathcal{O}_{X,x}$. Consider the set

$$\mathfrak{m}_x := \{(U, f) \in \mathcal{O}_{X,x} \mid f(x) = 0\}.$$

It then easily follows that $\mathfrak{m}_x$ an ideal and consequently is a maximal ideal because $\mathcal{O}_{X,x} \setminus \mathfrak{m}_x$ is jut the set of all units of $\mathcal{O}_{X,x}$. □

Remark 1.5.3.5. We have thus established that for any k-variety X we obtain a locally k-ringed space $(X, \mathcal{O}_X)$. We now observe how the data of a morphism of varieties can be represented as data of a morphism of underlying locally ringed spaces.

The notion of morphism of locally ringed spaces is elucidated in Definition 3.1.0.2.

Lemma 1.5.3.6. *Let k be an algebraically closed field and X, Y be two k-varieties. Then, there is an injective inclusion*

$$\mathrm{Hom}_{\mathbf{Var}_k}(X, Y) \hookrightarrow \mathrm{Hom}_{\mathbf{LRSpace}}(X, Y).$$

Proof. Indeed, consider the map

$$\begin{aligned} \theta : \mathrm{Hom}_{\mathbf{Var}_k}(X, Y) &\hookrightarrow \mathrm{Hom}_{\mathbf{LRSpace}}(X, Y) \\ f : X \to Y &\longmapsto (f, f^\flat) : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y) \end{aligned}$$

where $\theta(f)$ has the underlying continuous map same as f but the map on sheaves, $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$, is given on sections as follows: let $V \subseteq Y$ be an open set, then the map on sections over V is

$$\begin{aligned} f^\flat_V : \mathcal{O}_Y(V) &\longrightarrow \mathcal{O}_X(f^{-1}(V)) \\ (V, \varphi) &\longmapsto (f^{-1}(V), \varphi \circ f). \end{aligned}$$

The fact that $f^\flat$ as defined above is indeed a sheaf morphism is straightforward. We thus need only show that the adjoint map $f^\sharp$ of the above defines a map on stalks which is local. For this, we need only observe how the comorphism, $f^\sharp_x : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$, as defined in Definition 3.1.0.2, in this case turns out to be the following mapping

$$(V, \varphi)_x \longmapsto (f^{-1}(V), \varphi \circ f)_x.$$

Now if $(V, \varphi)_x \in \mathfrak{m}_{Y,f(x)}$, then $\varphi(f(x)) = 0$ by definition. Thus $(f^{-1}(V), \varphi \circ f) \in \mathfrak{m}_{X,x}$. With this, the fact that θ is injective is straightforward. □

Remark 1.5.3.7. We therefore have an inclusion

$$\mathbf{Var}_k \hookrightarrow \mathbf{LRSpace}.$$

Indeed, we now show that the notion of isomorphisms coincide here.

We will now define various algebraic gadgets out of the structure sheaf $\mathcal{O}_X$ of a variety X. Indeed, to some extent, that's the goal of algebraic geometry in general.

We now define an important field corresponding to each variety X, called its function field.

Definition 1.5.3.8. (**Function field of a variety**) Let k be an algebraically closed field and X be a k-variety. The function field of X, denoted $K(X)$, is obtained as the quotient of the set $\cup_{U \supseteq X,\text{ open}} \cup_{(U,\varphi) \in \mathcal{O}_X(U)} (U, f)$ by the following relation

$$(U, \varphi) \sim (V, \phi) \iff \exists \text{ open } W \subseteq U \cap V \text{ s.t. } \rho_{U,W}(\varphi) = \rho_{V,W}(\phi).$$

Indeed, this has an addition and a multiplication given by restriction to the open sets where they agree. This is further a field as any non-zero element $[(U, f)]$ can be inverted in a small enough open set $W \subseteq U$ (which will be non-empty as X is irreducible) where f is non-zero (otherwise the class $[(U, f)]$ is identically zero).

Remark 1.5.3.9. Note that we have the following ring homomorphisms for any k-variety X and $x \in X$

$$\begin{aligned} \Gamma(\mathcal{O}_X, X) &\longrightarrow \mathcal{O}_{X,x} \longrightarrow K(X) \\ (X, \varphi) &\longmapsto (X, \varphi)_x \longmapsto [(X, \varphi)]. \end{aligned}$$

In-fact, both these are injective by a simple use of the identity principle (Lemma 1.5.2.4). In this way, algebraic gadgets start taking a hold onto the geometry of varieties, which we will see further in this chapter.

We now give two results; one for affine and one for projective; which shows how the three algebraic gadgets introduced in Remark 1.5.3.9 can be realized more algebraically.

Proposition 1.5.3.10. *Let k be an algebraically closed field and let X be an affine k-variety. Let $\mathfrak{m}_p = \{f \in k[X] \mid f(p) = 0$ as a regular function$\}$. Then,*

1. *$\mathfrak{m}_p$ is a maximal ideal of $k[X]$ for every point $p \in X$,*
2. *$\mathrm{mSpec}\,(k[X]) \cong X$ as sets,*
3. *$k[X]_{\mathfrak{m}_p} \cong \mathcal{O}_{X,p}$,*
4. *$k[X]_{\langle 0 \rangle} \cong K(X)$,*
5. *$\Gamma(\mathcal{O}_X, X) \cong k[X]$,*
6. *$\dim X = \mathrm{trdeg}\, K(X)/k$[20],*
7. *$\dim X = \dim \mathcal{O}_{X,p}$ for all $p \in X$[21].*

[20]Thus the function field $K(X)/k$ holds important global information about the algebra and geometry of X.

[21]Thus the notion of dimension of varieties is detectable at the level of stalks. This is because, as the proof and the statement 3 shows, the local ring $\mathcal{O}_{X,p}$ holds almost all relevant information about the coordinate ring.

Proof. We give the main ideas of each. The main idea in the latter parts is to embed all the relevant rings inside the function field and do the relevant algebra there.

1. Since there is a correspondence between radical ideals of $k[X]$ and algebraic sets of X and since the correspondence is antitone, therefore minimal algebraic sets (point $p \in X$) of X correspond to maximal ideals of $k[X]$ vanishing at p. The result then follows.
2. This follows from 1. Explicitly, one considers the mapping $p \in X \mapsto \mathfrak{m}_p$.
3. Consider the canonical mapping
$$\begin{aligned} k[X]_{\mathfrak{m}_p} &\longrightarrow \mathcal{O}_{X,p} \\ \frac{f}{g} &\longmapsto (X \setminus Z(g), f/g)_p \end{aligned}$$
where $g(p) \neq 0$ (so $g \notin \mathfrak{m}_p$). This is a homomorphism by the fact that $Z(f) \cup Z(g) = Z(fg)$. This is injective because if $f/g = 0$, then $f = 0$ on some open subset $W \subseteq X \setminus Z(g)$. By an application of identity principle (Lemma 1.5.2.4), the injectivity follows. For surjectivity, observe that for any $(U, f)_p \in \mathcal{O}_{X,p}$, we can represent it by the rational function that f looks like around p, so $(U, f)_p = (W, g/h)_p$ where g/h is a rational function. Consequently, $g/h \mapsto (X \setminus Z(h), g/h)_p = (W, g/h)_p$. The result follows.
4. Observe first that if R is a domain and $\mathfrak{p} \leq R$ is an prime ideal of R, then $(R/\mathfrak{p})_{\langle 0 \rangle}$ is isomorphic to $R_{\langle 0 \rangle}$. Now, by 3, we obtain that $k[X]_{\langle 0 \rangle} \cong (k[X]_{\mathfrak{m}_p})\langle 0 \rangle \cong (\mathcal{O}_{X,p})_{\langle 0 \rangle}$. The map
$$\begin{aligned} (\mathcal{O}_{X,p})_{\langle 0 \rangle} &\longrightarrow K(X) \\ \frac{(U, f/g)_p}{(V, h/l)_p} &\longmapsto [(U \cap V, fl/gh)] \end{aligned}$$
can be seen to be a well-defined (use Lemma 1.5.2.4) isomorphism.
5. By Lemma 16.1.2.12, we have that $\bigcap_{\mathfrak{m} < k[X]} k[X]_{\mathfrak{m}} \cong k[X]$. By Remark 1.5.3.9, we have $\Gamma(\mathcal{O}_X, X) \hookrightarrow \mathcal{O}_{X,p}$ (in $K(X)$). We further have $k[X] \hookrightarrow \Gamma(\mathcal{O}_X, X)$. Consequently, we obtain via 3. the following
$$k[X] \hookrightarrow \Gamma(\mathcal{O}_X, X) \hookrightarrow \bigcap_{p \in X} \mathcal{O}_{X,p} \cong \bigcap_{p \in X} k[X]_{\mathfrak{m}_p} \hookrightarrow \bigcap_{\mathfrak{m} < k[X]} k[X]_{\mathfrak{m}} \cong k[X].$$
The result then follows.
6. We have $\dim X = \dim k[X]$ as any irreducible closed subset of X corresponds in a contravariant manner to a prime ideal of $k[X]$. By Theorem 16.8.2.1, we have $\dim k[X] = \operatorname{trdeg} K(X)/k$.
7. By 3, $\dim \mathcal{O}_{X,p} = \operatorname{ht} \mathfrak{m}_p$. By Theorem 16.8.2.2, we have $\operatorname{ht} \mathfrak{m}_p + \dim k[X]/\mathfrak{m}_p = \dim k[X]$. But since $k[X]/\mathfrak{m}_p \cong k$ by Nullstellensatz, therefore the above equation reduces to $\operatorname{ht} \mathfrak{m}_p = \dim k[X]$ and the right side is just $\dim X$.

□

We next do the projective case. See Chapter 16, Section 16.1.2 for homogeneous localization of graded rings.

Lemma 1.5.3.11. *Let k be an algebraically closed field and X be a projective k-variety in $\mathbb{P}^n_k$. Let $U_i = \mathbb{P}^n_k \setminus Z(x_i)$ and $X_i := X \cap U_i$. Then*
$$\varphi_i : k[X_i]^a \cong k[X]^h_{(x_i)}$$

where $k[X_i]^a$ denotes the affine coordinate ring of $X_i \subseteq \mathbb{A}^n_k$ and $k[X]^h$ denotes the homogeneous coordinate ring of $X \subseteq \mathbb{P}^n_k$. Further the localization above is homogeneous.

Proof. Consider the map $k[y_1, \dots, y_n] \to k[x_0, \dots, x_n]$ mapping as $f(y_1, \dots, y_n) \mapsto f\left(\frac{x_0}{x_i}, \dots, \frac{\hat{x_i}}{x_i}, \dots, \frac{x_n}{x_i}\right)$. This can easily be seen to be a well-defined ring isomorphism mapping the ideal $I(X_i) \mapsto I(X_i)^h = I(X)^h_{(x_i)}$. The result follows by quotienting. □

Proposition 1.5.3.12. *Let k be an algebraically closed field and X be a projective k-variety. Let $\mathfrak{m}_p = \langle \{f \in k[X] \mid f \text{ is homogeneous } \& \ f(p) = 0\}\rangle$ for any $p \in X$ and $k[X]$ be the homogeneous coordinate ring of X. Then,*

1. *$\mathfrak{m}_p$ is a maximal ideal of $k[X]$ for every element $p \in X$,*
2. *$k[X]_{(\mathfrak{m}_p)} \cong \mathcal{O}_{X,p}$,*
3. *$k[X]_{(\langle 0 \rangle)} \cong K(X)$,*
4. *$\Gamma(\mathcal{O}_X, X) \cong k$.*

Proof. Denote by $k[X]^h$ the homogeneous coordinate ring and $X_i := X \cap U_i$ where $U_i = \mathbb{P}^n_k \setminus Z(x_i)$. By Lemma **??**, $U_i \cong \mathbb{A}^n_k$ as varieties, therefore denote X_i^a to be the affine variety corresponding to $X_i \subseteq U_i$. We thus denote $k[X_i]^h$ for the homogeneous coordinate ring when $X_i \subseteq U_i$ and $k[X_i]^a$ to be the affine coordinate ring when $X_i \subseteq \mathbb{A}^n_k$. Let $R := k[X]^h$. The main idea of the last part is to use the theory of integral dependence together with algebraic closure of k.

1. Let $P \in X$, so $P \in X_i$ for some $i = 0, \dots, n$. Thus, let $P^a \in X_i^a$ and by Lemma 1.5.3.11 and Proposition 1.5.3.10, we obtain that $\mathfrak{m}_{P^a}$ is a maximal ideal of $k[X_i]^a$. Thus, $\varphi_i(\mathfrak{m}_{P^a}) = \mathfrak{m}_P k[X]^h_{x_i}$ is a maximal ideal of $k[X]^h_{x_i}$.
2. We simply have the following for any $p \in X$ by irreducibility of X, by Lemma 1.5.3.11 and by Proposition 1.5.3.10:
$$\mathcal{O}_{X,p} \cong \mathcal{O}_{X_i,p} \cong \mathcal{O}_{X_i^a,p^a} \cong k[X_i]^a_{\mathfrak{m}_{p^a}} \cong \left(k[X]^h_{x_i}\right)_{\mathfrak{m}_{p^a}} \cong k[X]^h_{\mathfrak{m}_{p^a}}.$$
3. By irreducibility of X, by Lemma 1.5.3.11 and by Proposition 1.5.3.10, we have the following identifications
$$K(X) \cong K(X_i) \cong K(X_i^a) \cong k[X_i]^a_{\langle 0 \rangle} \cong \left(k[X]^h_{x_i}\right)_{\langle 0 \rangle} \cong k[X]^h_{\langle 0 \rangle}.$$
4. First note that $k \hookrightarrow \Gamma(\mathcal{O}_X, X)$. It would thus suffice to show that $\Gamma(\mathcal{O}_X, X) \hookrightarrow k$. Pick any $f \in \Gamma(\mathcal{O}_X, X)$. We wish to show that $f \in k$. Let $R = k[X]^h$. Note that we can embed $\Gamma(X, \mathcal{O}_X)$ inside the (non-homogeneous) fraction field $L = k[X]^h_{\langle 0 \rangle}$. Consequently, by algebraic closure of k, it would suffice to show that $f \in L$ satisfies a polynomial with coefficients in k. Since f is a regular function on each of the X_i, therefore $f \in k[X_i]^a \cong k[X]^h_{x_i}$. Consequently, $f = g_i / x_i^{n_i}$ in L where $\deg g_i = n_i$ and thus $x_i^{n_i} f \in R_{n_i}$ for each $i = 0, \dots, n$. It thus follows that $\deg f = 0$ in L. Consequently, it would suffice to show that $f \in L$ is integral over R (as we can then obtain a polynomial in $k[x]$ whose zero is f by restricting to 0 degree coefficients). By Corollary **??**, it would thus suffice to show that $R[f]$ is a finitely generated R-module.

   It would thus suffice if we show that $\exists M \in \mathbb{N}$ such that $\forall N \geq M$, $R_N f^m \subseteq R_N$ for all $m \geq 0$. Indeed, for $M = \sum_i n_i$, we see that $R_N f \subseteq R_N$ as for any $g \in R_N$, we have that each term of g will have to have one x_i whose power is $\geq n_i$. Repeatedly applying $R_N f \subseteq R_N$ yields $R_N f^m \subseteq R_N$ for all $m \geq 0$, as needed.

□

Remark 1.5.3.13. Note that in Proposition 1.5.3.12, 1, the maximal ideal $\mathfrak{m}_P$ does not contain all of non-constant polynomials in $k[X]$ because $\mathfrak{m}_p$ is generated by homogeneous polynomials vanishing at $p \in X$ and a polynomial with non-zero constant terms cannot be in such an ideal, thus such an $\mathfrak{m}_p$ will exactly be the ideal of all non-constant polynomials in $k[X]$, but then $p \in \bigcap_{f \in k[X], f(0)=0} Z(f) = \emptyset$.

We now show that affine varieties are completely determined by their coordinate rings in the following sense

Theorem 1.5.3.14. *Let k be a algebraically closed field. Then the following*

$$\begin{aligned} k[-] : \mathbf{AfVar^{op}}_k &\longrightarrow \mathbf{FGIAlg}_k \\ X &\longmapsto k[X] \\ X \stackrel{\varphi}{\to} Y &\longmapsto k[Y] \stackrel{k[\varphi]}{\to} k[X] \end{aligned}$$

is a functor[22] *which induces an equivalence between the opposite category of affine varieties over k and finitely generated integral domains over k.*

Proof. **TODO.** □

We now show some examples of the machinery developed so far. We first show that any affine plane conic is isomorphic as a variety to either the parabola $y - x^2$ or the hyperbola $xy - 1$. Indeed, we use here the familiar high-school topic that one classifies conics on the basis of discriminant(!) This will further show that the usual substitutions that we so used to do in school days to reduce an algebraic equation into a simpler form can equivalently be stated in algebraic language as finding a correct automorphism of the corresponding ring in question.

the solu-
n notebook,
.

Subvarieties

[22]Note that by Proposition 1.5.3.10, this is just the global sections functor.

1.5.4 Varieties as schemes

In this section we show how to realize a k-variety (see Definition 1.5.2.5) as a scheme. This will be essential as it fulfill all the reasons to work with schemes as they generalize the concept of varieties to just the right level where all algebro-geometric questions can be asked and be attempted to be solved.

We first show a fully-faithful functor which embeds the category of k-varieties into the category of k-schemes (that is, schemes over k). This will hence show how to obtain a scheme from a variety because, as the following construction of the relevant functor will show, it is not straightforward how should one begin defining it[23].

Definition 1.5.4.1. (**Spectral space of** X) For every topological space X, we can associate a topological space

$$t(X) := \{\text{All non-empty closed irreducible subsets of } X\}$$

where any closed set is given by $t(Y) \subseteq t(X)$ for a closed set $Y \subseteq X$. The following lemma shows that this indeed defines a topology on $t(X)$. We will call $t(X)$ the spectral space of X.

Lemma 1.5.4.2. *Let X be a space and $Y, Z, Y_i \subseteq X$ be closed subsets of X. Then,*

1. $t(Y) \subseteq t(X)$,
2. $t(Y \cup Z) = t(Y) \cup t(Z)$,
3. $t\left(\bigcap_i Y_i\right) = \bigcap_i t(Y_i)$.

Proof. 1. Any closed irreducible subset of Y, where Y is closed in X, will again be closed and irreducible in X.
2. Any irreducible subset of $Y \cup Z$ cannot have non-empty intersection with both of them.
3. Follows from 1. □

Indeed, our main idea is to show that for a variety V, the space $t(V)$ will eventually become a scheme. We have few observations about spectral spaces, before we realize that idea.

Lemma 1.5.4.3. *Let X, X_1, X_2 be spaces and $f : X_1 \to X_2$ be a continuous map. Then,*

1. *there is a one-to-one correspondence between closed subsets of X and closed subsets of $t(X)$,*
2. *the following is a continuous map*

$$\begin{aligned} t(f) : t(X_1) &\longrightarrow t(X_2) \\ Y_1 &\longmapsto \overline{f(Y_1)}, \end{aligned}$$

3. *the following is a functor*

$$\begin{aligned} t : \mathbf{Top} &\longrightarrow \mathbf{Top} \\ X &\longmapsto t(X), \end{aligned}$$

4. *the following is a continuous map*

$$\begin{aligned} \alpha : X &\longrightarrow t(X) \\ x &\longmapsto \overline{\{x\}}. \end{aligned}$$

[23] However, one may take a hint (albeit quite vague) from Lemma 1.3.0.2 in the following construction.

Proof. 1. Follows from the definition of topology on the spectral space.
2. Let $Y_2 \subseteq X_2$ be closed so that $t(Y_2) \subseteq t(X_2)$ is closed. We wish to show that $(t(f))^{-1}(t(Y_2)) \subseteq t(X_1)$ is closed. This follows from the observation that for $Y_1 \in t(X_1)$, we have $\overline{f(Y_1)} \in t(Y_2) \iff Y_1 \in t(f^{-1}(Y_2))$.
3. Follows from 2.
4. Pick any closed $Y \subseteq X$ to thus obtain a closed $t(Y) \subseteq t(X)$. Then $\alpha^{-1}(t(Y)) = \{x \in X \mid \overline{\{x\}} \in t(Y)\} = \{x \in X \mid x \in Y\} = Y$. □

We now give scheme structure to the space $t(X)$. But first, we need a small lemma.

Lemma 1.5.4.4. *Let $A = k[V]$ be the coordinate ring of an affine k-variety V over an algebraically closed field k. Then, for any open set $U \subseteq \operatorname{Spec}(A)$, the set of all closed points of U are dense in U.*

Proof. Since all closed points of $\operatorname{Spec}(A)$ are its maximal ideals by Nullstellensatz, thus, any closed point of U is a maximal ideal of A as well. Consequently, we may assume $U = D(f)$ is a basic open set for $f \in A$. But since $D(f) \cong \operatorname{Spec}(A_f)$ and closed points of any affine scheme are always dense, the result follows. □

Theorem 1.5.4.5. *Let k be an algebraically closed field and $(V, \mathcal{O}_V)$ be a k-variety. Let $\alpha : V \to t(V)$ be the continuous map as defined in Lemma 1.5.4.3, 4. Then, $(t(V), \alpha_*\mathcal{O}_V)$ is a scheme over k which admits an affine open cover by $\operatorname{Spec}(A)$ for $A = k[W]$ where W is an affine open subvariety of V.*

Proof. For better clarity of this important proof, we break it in multiple acts.

Act 1 : *We may assume V is an affine k-variety.*

Since we wish to show that $t(V)$ is a scheme, hence we need to produce an open cover of $t(V)$ by affine schemes. Since V is covered by open affine k-varieties, thus if we can show that for an affine k-variety W, the space $t(W)$ is a scheme, then we would be done. Hence we may assume V is affine with coordinate ring $k[V] =: A$.

Act 2 : $t(V) \cong \operatorname{Spec}(A)$ *as topological spaces.*

Consider the usual maps that we know from our study of varieties:

$$\begin{tikzcd} t(V) \arrow[r, shift left, "I(-)"] & \operatorname{Spec}(A) \arrow[l, shift left, "Z(-)"] \end{tikzcd}$$

These are easily seen to be continuous inverses of each other by the correspondence between closed irreducible subsets of an affine variety and prime ideals of its coordinate ring (Lemma 1.2.1.1).

Act 3 : *The closed points of* $\operatorname{Spec}(A)$ *are points of V.*

We first construct the following map

$$\varphi : V \longrightarrow \operatorname{Spec}(A)$$
$$p \longmapsto \mathfrak{m}_p$$

where $\mathfrak{m}_p$ is defined together with some properties in Proposition 1.5.3.10. This is continuous by a small check on closed sets. Moreover, this is injective. Now, we claim that $\varphi(V) \subseteq \operatorname{Spec}(A)$ are all closed points of $\operatorname{Spec}(A)$. Indeed, this follows from the correspondence between closed points of $\operatorname{Spec}(A)$ and maximal prime ideals of A (Lemma 1.2.1.3). We will thus denote $\varphi(V)$ as the set of closed points of $\operatorname{Spec}(A)$.

Act 4 : *It is enough to show that* $\varphi_* \mathcal{O}_V \cong \mathcal{O}_{\operatorname{Spec}(A)}$.

Since we have the following commutative triangle

$$\begin{tikzcd} & V \arrow[dl, "\varphi"'] \arrow[dr, "\alpha"] & \\ \operatorname{Spec}(A) \arrow[rr, shift left, "Z(-)"] \arrow[rr, phantom, "\cong"] & & t(V) \arrow[ll, shift left, "I(-)"] \end{tikzcd},$$

thus $\alpha_* \mathcal{O}_V \cong (Z \circ \varphi)_* \mathcal{O}_V = Z_* \varphi_* \mathcal{O}_V$. Since Z is an isomorphism, thus the reduction is justified.

Act 5 : $\varphi_* \mathcal{O}_V \cong \mathcal{O}_{\operatorname{Spec}(A)}$.

Let $U \subseteq \operatorname{Spec}(A)$ be an open set. We will construct an isomorphism between $\mathcal{O}_{\operatorname{Spec}(A)}(U)$ and $\mathcal{O}_V(\varphi^{-1}(U))$. Consider the map

$$\begin{aligned} \eta_U : \mathcal{O}_{\operatorname{Spec}(A)}(U) &\longrightarrow \mathcal{O}_V(\varphi^{-1}(U)) \\ s : U \to \amalg_{\mathfrak{p} \in U} A_{\mathfrak{p}} &\longmapsto \eta_U(s) : \varphi^{-1}(U) \to k \end{aligned}$$

where for any $q \in \varphi^{-1}(U)$, we define $\eta_U(s)(q) = s(\mathfrak{m}_q)(q)$. It clearly is a ring homomorphism which commutes with appropriate restriction maps. Thus, we need to show the following three statements in order to conclude.

1. $\eta_U(s)$ is regular,
2. η_U has zero kernel,
3. η_U is surjective.

In-fact, the above three statements are at the technical heart of the proof. The main driving force behind this is the density of closed points of open sets in $\operatorname{Spec}(A)$ (Lemma 1.5.4.4) and the identity principle of regular maps on a variety (Lemma 1.5.2.4).

Statement 1. is immediate as s is regular. For statement 2., suppose that $\eta_U(s) = 0$ over $\varphi^{-1}(U)$. Thus $s(\mathfrak{m}_q)(q) = f_q(q)/g_q(q) = 0$ for all $q \in \varphi^{-1}(U)$. Thus, $\eta_U(s)$ around q is represented by rational function f_q/g_q. By Lemma 1.5.2.4 on $\eta_U(s)$, we obtain that $f_q = 0$ for all $q \in \varphi^{-1}(U)$. Thus s is zero at all closed points of U, which are exactly $\varphi(\varphi^{-1}(U))$. But since closed points of U are dense by Lemma 1.5.4.4 and s is a locally constant function, hence $s = 0$.
Finally, to see statement 3., pick any $f \in \mathcal{O}_V(\varphi^{-1}(U))$ and notice that $W := \varphi(\varphi^{-1}(U))$ is a dense subset of U (set of all closed points, Lemma 1.5.4.4). Thus, it is enough to define a locally constant function s over W whose extension $\tilde{s}$ over U is such that $\eta_U(\tilde{s}) = f$. Indeed, consider

$$\begin{aligned} s : W &\longrightarrow \coprod_{\mathfrak{p} \in U} A_{\mathfrak{p}} \\ \mathfrak{m}_q &\longmapsto g_q/h_q \end{aligned}$$

where g_q/h_q is the rational function representing f at the point $q \in V$. Clearly, the extension $\tilde{s}$ is in $\mathcal{O}_{\mathrm{Spec}(A)}(U)$ and it is mapped by η_U to f.

Act 6 : $(t(V), \alpha_* \mathcal{O}_V)$ *is a scheme over* k*.*

Now let V be a k-variety. We wish to show that $t(V)$ is a scheme over Spec (k). Thus we need to produce a map $t(V) \to \mathrm{Spec}\,(k)$, which is equivalent to a map $k \to \Gamma(\alpha_* \mathcal{O}_V, t(V))$ via the Theorem 1.3.0.5. Since $\Gamma(\alpha_* \mathcal{O}_V, t(V)) = \Gamma(\mathcal{O}_V, V) = A$ via Proposition 1.5.3.10, 5, the result follows.

This completes the proof. □

Remark 1.5.4.6. Theorem 1.5.4.5 yields that the functor t restricts to the following

$$\begin{aligned} t : \mathbf{Var}_k &\longrightarrow \mathbf{Sch}_k \\ (V, \mathcal{O}_V) &\longmapsto (t(V), \alpha_* \mathcal{O}_V). \end{aligned}$$

We will now show that this is a fully-faithful embedding. In other words, any map of $t(V_1) \to t(V_2)$ as of schemes over k is equivalent to a map $V_1 \to V_2$ of k-varieties.

Let us begin with some elementary properties of the residue fields of the k-scheme $t(V)$ attached to a k-variety V.

Lemma 1.5.4.7. *Let k be an algebraically closed field and let V be a k-variety. A point $p \in t(V)$ is closed if and only if $\kappa(p) = k$.*

Proof. (L $\Rightarrow$ R) Since $p \in t(V)$ is closed and closed points of $t(V)$ are exactly points of V, therefore $p \in V \subseteq t(V)$. Consequently, for an affine k-variety $X \subseteq V$ containing p, we obtain the following by Proposition 1.5.3.10, 3 and Nullstellensatz:

$$\kappa(p) = \mathcal{O}_{t(V),p}/\mathfrak{m}_{t(V),p} \cong \mathcal{O}_{V,p}/\mathfrak{m}_{V,p} \cong \mathcal{O}_{X,p}/\mathfrak{m}_{X,p} \cong k[X]_{\mathfrak{m}_p}/\mathfrak{m}_p k[X]_{\mathfrak{m}_p} \cong (k[X]/\mathfrak{m}_p k[X])_0 \cong k.$$

(R $\Rightarrow$ L) By Theorem 1.5.4.5, we have that for some open affine k-variety $X \subseteq V$, $p \in \mathrm{Spec}\,(k)[X]$. Consequently, $\kappa(p) = (k[X]/pk[X])_0 = k$ where p is treated as a prime ideal of $k[X]$. Consequently, we have that the domain $k[X]/pk[X] = k$ as we have inclusions $k \hookrightarrow k[X]/pk[X] \hookrightarrow (k[X]/pk[X])_0$. Thus $p \lneq k[X]$ is maximal. □

Proposition 1.5.4.8. *Let k be an algebraically closed field. Then there is a natural bijection*

$$\mathrm{Hom}_{\mathbf{Var}_k}(V_1, V_2) \cong \mathrm{Hom}_{\mathbf{Sch}_k}(t(V_1), t(V_2)).$$

That is, the functor t is a fully-faithful embedding of k-varieties into schemes over k.

e the proof
ding vari-
schemes,
.

Proof. Exercise 2.15 of Hartshorne Chapter 2. □

Let us now spell out all the properties that the scheme $t(V)$ satisfies for a k-variety V.

Proposition 1.5.4.9. *Let k be an algebraically closed field and V be a k-variety. Then, the scheme $t(V)$ over k is (for * properties, see Section 1.12)*

1. *integral,*
2. *noetherian,*
3. *finite type over k,*

4. *quasi-projective**,
5. *separated**.

Proof. 1. to 3. are immediate from the open covering by Spec $(k[W])$ of $t(V)$ where $W \subseteq V$ is an open affine subvariety (Theorem 1.5.4.5). Consequently $t(V)$ is covered by spectrum of finite type k-algebras.
4. is also immediate as any k-variety is an open subset of an affine or a projective k-variety by definition. Since any affine k-variety can be seen as a projective k-variety, consequently, we have an open immersion of V into a closed subvariety of some projective space over k. This extends to an open immersion of $t(V)$ into a closed subscheme of $\mathbb{P}^n_k$.
5. Follows from 4. and Theorem 1.12.8.2. □

We now state an important rectification result which precisely shows what type of schemes are those which are in the image of functor t as in Remark 1.5.4.6.

Corollary 1.5.4.10. *Let k be an algebraically closed field. Then, the functor of Remark 1.5.4.6*

$$t : \mathbf{Var}_k \longrightarrow \mathbf{QPISch}_k$$

establishes an equivalence between varieties over k and quasi-projective integral schemes over k. Further, the image of projective varieties under this functor is exactly the projective integral schemes over k.

Proof. By Proposition 1.5.4.8, we reduce to showing that t lands into quasi-projective schemes and is essentially surjective. Indeed, for a k-variety V, the scheme $t(V)$ is quasi-projective by Proposition 1.5.4.9, 4. Now, to show essential surjection, we first observe that open subschemes of $t(V)$ is in one-to-one bijection with open subsets of V. Consequently, it would suffice to show that any projective integral k-scheme X is in the essential image of t. Indeed, let V denote the closed points of X as a closed subscheme of some $\mathbb{P}^n_k$. Consequently, as closed points of a finite type k-scheme is dense (Lemma 1.12.2.6), therefore V is irreducible (note we are using irreducibility of X here), thus a projective variety in $\mathbb{P}^n_k$. Now, $t(V)$ and X have same underlying space. As a subspace of $\mathbb{P}^n_k$, $t(V)$ and X have both have the structure of a reduced scheme over the common underlying space. By uniqueness of reduced induced closed subscheme structure on a closed subset, we have that $t(V) \cong X$ (see Section 1.6.3). □

We now redefine varieties as schemes and use them as such for the remainder of the sections.

Definition 1.5.4.11. (**Abstract and classical varieties**) Let k be an algebraically closed field. An *abstract variety* or simply a variety, is a separated, integral finite type k-scheme. Those varieties which are furthermore quasi-projective are exactly the varieties we defined earlier by Corollary 1.5.4.10. We will further call the notion of varieties we defined earlier in Definition 1.5.2.5 by referring to them as *classical varieties*.

1.6 Fundamental constructions on schemes

In this section, we would like to understand some of the basic constructions which one can perform with a collection of schemes.

1.6.1 Points of a scheme

Let X be a scheme. Pick any point $x \in X$. We then have the residue field $\kappa(x) = \mathcal{O}_{X,x}/\mathfrak{m}_x$. Hence we have a projection map

$$\mathcal{O}_{X,x} \to \kappa(x).$$

Consider now an open affine $x \in \mathrm{Spec}\,(A) \subseteq X$. Consequently, we have $\mathcal{O}_{X,x} \cong \mathcal{O}_{\mathrm{Spec}(A),x} \cong A_x$. Thus, denoting the inclusion $j_x : \mathrm{Spec}\,(A_x) \hookrightarrow \mathrm{Spec}\,(A)$, we obtain the following composition:

$$i_x : \mathrm{Spec}\,(\kappa(x)) \to \mathrm{Spec}\,(\mathcal{O}_{X,x}) = \mathrm{Spec}\,(A_x) \stackrel{j_x}{\hookrightarrow} \mathrm{Spec}\,(A) \hookrightarrow X.$$

Remember that $\mathrm{Spec}\,(A_x)$ can be interpreted as the affine subset in $\mathrm{Spec}\,(A)$ which is "very close" to $x \in \mathrm{Spec}\,(A)$. The map j_x takes the singleton point in $\mathrm{Spec}\,(\kappa(x))$ to $x \in X$. This map is usually called the *canonical map of point* $x \in X$. The map on stalks that i_x yields is the natural projection $\mathcal{O}_{X,x} \to \kappa(x)$. This map is quite unique as it is universal amongst all those maps $\mathrm{Spec}\,(K) \to X$ which maps to x. Indeed, we have the following.

Lemma 1.6.1.1. *Let X be a scheme and let $x \in X$ be a point. If K is a field and $f : \mathrm{Spec}\,(K) \to X$ is a map, then*

1. *If $f(\star) = x$, then $\kappa(x) \hookrightarrow K$.*
2. *If $f(\star) = x$, then f factors via the canonical map i_x at point $x \in X$*

$$\begin{tikzcd} \mathrm{Spec}\,(\kappa(x)) \arrow[r, "i_x"] & X \\ \mathrm{Spec}\,(K) \arrow[u] \arrow[ur, "f"'] & \end{tikzcd}.$$

3. $\mathrm{Hom}_{\mathbf{Sch}}\,(\mathrm{Spec}\,(K), X) \cong \{x \in X \mid \kappa(x) \hookrightarrow K\}$.

Proof. 1. At the stalk, we have a local ring homomorphism $\varphi : \mathcal{O}_{X,x} \to K$. Consequently, $\mathrm{Ker}\,(\varphi) = \mathfrak{m}_{X,x}$. It then follows that $\kappa(x) = \mathcal{O}_{X,x}/\mathfrak{m}_{X,x} \hookrightarrow K$.
2. Clearly f factors as above as a continuous map. To check the commutativity of sheaf maps, we need only check at stalks (Theorem 20.3.0.7). This is straightforward, as we get on stalks the following commutative diagram:

$$\begin{tikzcd} \kappa(x) \arrow[d, hook] & \mathcal{O}_{X,x} \arrow[l] \arrow[dl, "f^\sharp_\star"] \\ K & \end{tikzcd}.$$

3. It suffices to show that a morphism $f : \mathrm{Spec}\,(K) \to X$ is equivalent to the data of a point $x \in X$ such that $\kappa(x) \hookrightarrow K$. By 1, one side is immediate. Now consider a point $x \in X$ and a

field extension $\kappa(x) \hookrightarrow K$. We wish to construct a map $f : \operatorname{Spec}(K) \to X$ such that the above data is obtained via the construction in 1 applied on f. Indeed, the map f on topological spaces is straightforward, $f(\star) = x$. On sheaves, it reduces to define a natural local ring homomorphism $\mathcal{O}_{X,x} \to K$. This is immediate, as we need only define this as $\mathcal{O}_{X,x} \to \kappa(x) \hookrightarrow K$. □

The above lemma shows that defining a map $\operatorname{Spec}(K) \to X$ is equivalent to taking a point $x \in X$ such that K extends $\kappa(x)$. There is another similar important characterization of maps from $\operatorname{Spec}\left(\frac{k[x]}{x^2}\right)$ into X, which characterizes all rational points of X together with "direction" (that is, together with an element of the tangent space). We first define a rational point of a k-scheme. Recall that by Corollary 1.3.0.7, $\kappa(x)$ is a field extension of k. Further observe the definition of Zariski tangent space T_xX of a scheme as defined in Definition 1.11.1.10.

Definition 1.6.1.2. (**Rational points**) Let X be a k-scheme. Then a point $x \in X$ is said to be rational if $\kappa(x) = k$.

Let us denote $k[\epsilon] = k[x]/x^2$. The ring $k[\epsilon]$ is usually called the ring of *dual numbers*.

Proposition 1.6.1.3. [24] *Let X be a scheme over a field k. Then, we have a bijection*

$$\operatorname{Hom}_{\mathbf{Sch}/k}(\operatorname{Spec}(k[\epsilon]), X) \cong \{(x, \xi) \mid x \in X \text{ is a rational point } \& \ \xi \in T_xX\}$$

Proof. ($\Rightarrow$) Take a scheme homomorphism $f : \operatorname{Spec}(k[\epsilon]) \to X$. Note that we have a map

$$k[\epsilon] \to k[\epsilon]/\epsilon \cong k.$$

Consequently, we get a map $g : \operatorname{Spec}(k) \to \operatorname{Spec}(k[\epsilon])$ which by composing by f, we get

$$\operatorname{Spec}(k) \xrightarrow{g} \operatorname{Spec}(k[\epsilon]) \xrightarrow{f} X.$$

Observe that $\operatorname{Spec}(k[\epsilon])$ is a one point scheme, therefore $f(\text{pt.}) = f \circ g(\text{pt.}) =: x$. We wish to show that x is a rational point. By Lemma 1.6.1.1, 3, we have $\kappa(x) \hookrightarrow k$. But since X is a scheme over k, therefore $k \hookrightarrow \kappa(x)$. We further deduce from the fact that X is a k-scheme that we have a triangle

$$\begin{tikzcd} \kappa(x) \arrow[r, hook] & k \\ k \arrow[u, hook] \arrow[ur, "\cong"'] & \end{tikzcd}.$$

This shows that horizontal arrow above is an isomorphism. Thus, $\kappa(x) = k$. We now wish to obtain an element of T_xX.

At the point $x \in X$, we have a map $f : \operatorname{Spec}(k[\epsilon]) \to X$. This yields a map on stalks given by

$$\varphi : A \to k[\epsilon]$$

where $A = \mathcal{O}_{X,x}$ is the local ring at point $x \in X$ and φ is furthermore a local k-algebra homomorphism. Let $\mathfrak{m}$ be the maximal ideal of the local ring A. Then, $A/\mathfrak{m} = \kappa(x)$, which is equal to k as x is a rational point. Thus, A is a rational local k-algebra (Definition 16.1.2.17). It follows from Proposition 16.1.2.18 that φ is equivalent to an element of the tangent space $\xi \in TA$ and by definition, $TA = T_xX$. This completes the proof. □

[24] Exercise II.2.8 of Hartshorne.

We now see that closed points of a finite-type k-scheme are those whose residue extension of k is algebraic.

Proposition 1.6.1.4. *Let X be a finite-type k-scheme. Then the following are equivalent:*

1. *$x \in X$ is a closed point.*
2. *$x \in X$ is such that $\kappa(x)/k$ is an algebraic (equivalently, finite).*

Proof. (1. $\Rightarrow$ 2.) Clearly, $\kappa(x)$ is a finitely type field extension of k. By essential Nullstellensatz, $\kappa(x)/k$ is algebraic.

(2. $\Rightarrow$ 1.) Pick an affine open Spec (A) containing x so that $\mathfrak{p} \in$ Spec (A) corresponds to x. We wish to show that $\mathfrak{p}$ is maximal. As $\kappa(x) = Q(A/\mathfrak{p})$ and

$$k \hookrightarrow A \twoheadrightarrow A/\mathfrak{p} \hookrightarrow Q(A/\mathfrak{p}) = \kappa(x),$$

thus, as $\kappa(x)/k$ is algebra, we deduce that $\kappa(x)$ is integral over $A/\mathfrak{p}$. Let B be a finite type k-domain such that $Q(B)$ is integral over B. One can check by writing down the relevant polynomials that this implies for any element $b \in B$, the inverse $b^{-1} \in Q(B)$ is in B by integrality. Using this for $B = A/\mathfrak{p}$, we deduce that $A/\mathfrak{p}$ is a filed, so $\mathfrak{p}$ is maximal, as required. □

1.6.2 Gluing schemes & strongly local constructions

We now show how to obtain new schemes from old by the gluing construction. Indeed, the idea is simple, glue the underlying topological spaces of a certain collection of schemes and identifications and define a new structure sheaf over the resultant space which canonically makes it into a scheme. We will further see that there is a universal property that is satisfied by such a glue. We suggest that the reader make a diagram of *blobs* and draw the corresponding maps in order to see the naturality of the following.

Definition 1.6.2.1. (**Gluing datum**) A tuple of data $(I, \{X_i\}_{i\in I}, \{U_{ij}\}_{i,j\in I}, \{\varphi_{ij}\}_{i,j\in I})$ of an index set I, schemes X_i for each $i \in I$, open subschemes $U_{ij} \subseteq X_i$ for each $i, j \in I$ and scheme isomorphisms $\varphi_{ij} : U_{ij} \to U_{ji}$ for each $i, j \in I$ is a gluing datum if it satisfies the following:

1. $U_{ii} = X_i$ for all $i \in I$,
2. $\varphi_{ji} = \varphi_{ij}^{-1}$,
3. $\varphi_{ii} = \mathrm{id}_{U_{ii}} = \mathrm{id}_{X_i}$,
4. the cocycle condition,

$$\varphi_{jk} \circ \varphi_{ij} = \varphi_{ik} \text{ on } U_{ij} \cap U_{ik}\ \forall i, j, k \in I.$$

We then have that there is a unique glue of the above.

Proposition 1.6.2.2. *For a gluing datum $(I, \{X_i\}_{i\in I}, \{U_{ij}\}_{i,j\in I}, \{\varphi_{ij}\}_{i,j\in I})$ of schemes, there exists a unique scheme X with the following properties:*

1. *there exists an open embedding of schemes*

$$\phi_i : X_i \to X \text{ for each } i \in I,$$

2. *$\phi_j \circ \varphi_{ij} = \phi_i$ on U_{ij} for all $i, j \in I$,*

3. $X = \bigcup_{i \in I} \phi_i(X_i)$,
4. $\phi_i(X_i) \cap \phi_j(X_j) = \phi_i(U_{ij}) = \phi_j(U_{ji})$ *for all* $i, j \in I$.

Proof. The underlying space of X is obtained by gluing the underlying spaces of X_i in the usual manner;

$$X := \coprod_{i \in I} X_i / \sim$$

where $x_i \sim \varphi_{ij}(x_i)$ for all $x_i \in U_{ij}$ and $i, j \in I$. Let $\phi_i : X_i \to X$ be the canonical inclusion map. The topology is given on X via the quotient topology; $U \subseteq X$ is open if and only if $\phi_i^{-1}(U) \subseteq X_i$ is open for each $i \in I$. Then to define the sheaf $\mathcal{O}_X$, pick any open $U \subseteq X$ and define the sections over it as follows (let us write $\varphi_{ij} : \mathcal{O}_{U_{ij}} \xrightarrow{\cong} \mathcal{O}_{U_{ji}}$ as well):

$$\mathcal{O}_X(U) = \left\{ [(\phi_i^{-1}(U), s_i)] \mid \forall i,\ s_i \in \mathcal{O}_{X_i}(\phi_i^{-1}(U)) \text{ s.t. } \varphi_{ij}(\rho_{\phi_i^{-1}(U), \phi_i^{-1}(U) \cap U_{ij}}(s_i)) = \rho_{\phi_j^{-1}(U), \phi_j^{-1}(U) \cap U_{ji}}(s_j) \right\}.$$

By local nature, this is again a sheaf (also called the glued sheaf). Now, ϕ_i is an open embedding as for any open $U \subseteq X$, it follows that $\mathcal{O}_X(\varphi_i(X_i) \cap U) \cong \mathcal{O}_{X_i}(\phi_i^{-1}(U))$. Thus, X is a scheme as for each $x \in X$, $x \in \phi_i(X_i)$ which is a scheme. $\square$

A lot of times we have the situation that a certain construction on a ring A leads to a map $\varphi : A \to \tilde{A}$. Consequently, we obtain maps $f : \operatorname{Spec}(\tilde{A}) \to \operatorname{Spec}(A)$. If X is a scheme, then for each open affine $V_i = \operatorname{Spec}(A_i)$, we get a map $X_i \to V_i$ given by $\operatorname{Spec}(\tilde{A}_i) \to \operatorname{Spec}(A_i)$. Consequently, we are interested in the conditions that the construction $A \to \tilde{A}$ must satisfy so that X_i glue together to give a scheme $\tilde{X}$ which represents the construction globally.

Definition 1.6.2.3 (**Construction on rings**). A construction on rings is a collection of maps $\{\varphi_A : A \to \tilde{A}\}$ one for each ring A such that for any isomorphism $\eta_{AB} : A \xrightarrow{\cong} B$, we have an isomorphism $\tilde{\eta}_{AB} : \tilde{A} \to \tilde{B}$ which is id if η is id, the diagram

$$\begin{tikzcd} A \arrow[r, "\varphi_A"] \arrow[d, "\eta_{AB}"', "\cong"] & \tilde{A} \arrow[d, "\tilde{\eta}_{AB}"', "\cong"] \\ B \arrow[r, "\varphi_B"'] & \tilde{B} \end{tikzcd}$$

commutes and if $\eta_{BC} \circ \eta_{AB} = \eta_{AC}$, then $\tilde{\eta}_{BC} \circ \tilde{\eta}_{AB} = \tilde{\eta}_{AC}$. That is, constructions are functorial on isomorphisms.

Definition 1.6.2.4 (**Strongly local constructions**). A construction on rings $\{\varphi_A : A \to \tilde{A}\}$ is said to be strongly local if it naturally commutes with localization. That is, for each $g \in A$ not in nilradical, there exists an isomorphism $\widetilde{A_g} \cong \tilde{A}_g$ such that

$$\begin{tikzcd} A \arrow[r] \arrow[d, "\varphi_A"'] & A_g \arrow[d, "(\varphi_A)_g"'] \arrow[dr, "\varphi_{A_g}"] & \\ \tilde{A} \arrow[r] & \tilde{A}_g \arrow[r, "\cong"'] & \widetilde{A_g} \end{tikzcd}$$

commutes where $(\varphi_A)_g : A_g \to \tilde{A}_g$ is the localization of map $\varphi_A : A \to \tilde{A}$ at the element $g \in A$ and the horizontal arrows of the square are localization maps.

Remark 1.6.2.5. Let $\eta : A_f \cong B_g$ be an isomorhism where $f \in A$ and $g \in B$. Then we get an isomorphism $\hat{\eta} : \tilde{A}_f \cong \tilde{B}_g$ as in the following commutative diagram:

$$\begin{tikzcd}
\tilde{A}_f \arrow[r, dashed, "\hat{\eta}"] & \tilde{B}_g \\
\widetilde{A_f} \arrow[u, "\cong"] \arrow[r, "\cong", "\tilde{\eta}"'] & \widetilde{B_g} \arrow[u, "\cong"'] \\
A_f \arrow[u, "\varphi_{A_f}"] \arrow[r, "\cong", "\eta"'] \arrow[uu, bend left, "(\varphi_A)_f"] & B_g \arrow[u, "\varphi_{B_g}"'] \arrow[uu, bend right, "(\varphi_B)_g"']
\end{tikzcd}.$$

Let X be a scheme. Our main goal is to show that strongly local constructions done on each affine open subset of X can be glued to give a scheme $\tilde{X}$ admitting a map $\tilde{X} \to X$.

We will achieve this in steps. We first translate strongly local property more geometrically.

Lemma 1.6.2.6. *Let $\{\varphi_A : A \to \tilde{A}\}$ be a strongly local construction on rings. For any ring A denote $\phi_A : \operatorname{Spec}(\tilde{A}) \to \operatorname{Spec}(A)$ to be the map corresponding to φ_A. Then, for any $f \in A$, the following diagram commutes:*

$$\begin{tikzcd}
\operatorname{Spec}(A_f) & & \operatorname{Spec}(\tilde{A}_f) \arrow[ll, "\phi_A|_{\operatorname{Spec}(\tilde{A}_f)}"'] \arrow[dl, "\cong"] \\
& \operatorname{Spec}\left(\widetilde{A_f}\right) \arrow[ul, "\phi_{A_f}"] &
\end{tikzcd}.$$

Proof. This is the translation of Definition 1.6.2.3 in $\operatorname{Spec}(-)$ where localization amounts to restricting to the corresponding open subscheme. $\square$

The following is an important observation which will help in checking the cocycle condition.

Lemma 1.6.2.7. *Let $\{\varphi_A : A \to \tilde{A}\}$ be a strongly local construction on rings and the following be a commutative triangle of isomorphisms*

$$\begin{tikzcd}
R_f \arrow[r] \arrow[dr] & S_g \arrow[d] \\
& T_h
\end{tikzcd}$$

for $f \in R$, $g \in S$ and $h \in T$. Then, the following triangle of isomorphisms as constructed in Remark 1.6.2.5 also commutes

$$\begin{tikzcd}
\tilde{R}_f \arrow[r] \arrow[dr] & \tilde{S}_g \arrow[d] \\
& \tilde{T}_h
\end{tikzcd}.$$

Proof. By definition of a construction, we get that the following triangle commutes

$$\begin{tikzcd} \widetilde{R_f} \arrow[r] \arrow[dr] & \widetilde{S_g} \arrow[d] \\ & \widetilde{T_h} \end{tikzcd}.$$

By the construction of isomorphism $\tilde{R}_f \to \tilde{S}_g$ and others as in Remark 1.6.2.5, we immediately get that the required triangle commutes. □

Lemma 1.6.2.8. *Let $X = \operatorname{Spec}(A)$ and $Y = \operatorname{Spec}(B)$ be two affine schemes. Let R be a ring with isomorphisms $A_f \cong R \cong B_g$ for some $f \in A$ and $g \in B$. Let $\{\varphi_S : S \to \tilde{S}\}$ be a strongly local construction on rings. Then there are open immersions $\operatorname{Spec}(\tilde{R}) \hookrightarrow \operatorname{Spec}(\tilde{A})$ and $\operatorname{Spec}(\tilde{R}) \hookrightarrow \operatorname{Spec}(\tilde{B})$ so that the following commutes*

$$\begin{tikzcd} \operatorname{Spec}(\tilde{A}) \arrow[d, "\phi_A"'] & \operatorname{Spec}(\tilde{R}) \arrow[l, hook'] \arrow[r, hook] \arrow[d, "\phi_R"'] & \operatorname{Spec}(\tilde{B}) \arrow[d, "\phi_B"] \\ \operatorname{Spec}(A) & \operatorname{Spec}(R) \arrow[l, hook'] \arrow[r, hook] & \operatorname{Spec}(B) \end{tikzcd}.$$

Proof. This follows from the following diagram

$$\begin{tikzcd} \operatorname{Spec}(\tilde{A}) \arrow[d, "\phi_A"'] & \operatorname{Spec}(\tilde{A}_f) \cong \operatorname{Spec}\left(\widetilde{A_f}\right) \arrow[l, hook'] \arrow[d, "\phi_A|_{\operatorname{Spec}(\tilde{A}_f)}"] & \operatorname{Spec}(\tilde{R}) \arrow[l, "\cong"'] \arrow[r, "\cong"] \arrow[d, "\phi_R"] & \operatorname{Spec}\left(\widetilde{B_g}\right) \cong \operatorname{Spec}(\tilde{B}_g) \arrow[r, hook] \arrow[d, "\phi_B|_{\operatorname{Spec}(\tilde{B}_g)}"] & \operatorname{Spec}(\tilde{B}) \arrow[d, "\phi_B"] \\ \operatorname{Spec}(A) & \operatorname{Spec}(A_f) \arrow[l, hook'] & \operatorname{Spec}(R) \arrow[l, "\cong"] \arrow[r, "\cong"'] & \operatorname{Spec}(B_g) \arrow[r, hook] & \operatorname{Spec}(B) \end{tikzcd}$$

the commutativity of which follows from Lemma 1.6.2.6 and the definition of a construction. □

Let X be a scheme and $U = \operatorname{Spec}(A)$ and $V = \operatorname{Spec}(B)$ be two open affines. We can now glue $\operatorname{Spec}(\tilde{A})$ and $\operatorname{Spec}(\tilde{B})$ along the intersection $U \cap V$ as follows.

Proposition 1.6.2.9. *Let X be a scheme and $U = \operatorname{Spec}(A)$ and $V = \operatorname{Spec}(B)$ be two open affines. Let $\{\varphi_S : S \to \tilde{S}\}$ be a strongly local construction on rings. Let $\phi_A : \tilde{U} = \operatorname{Spec}(\tilde{A}) \to \operatorname{Spec}(A)$ and $\phi_B : \tilde{V} = \operatorname{Spec}(\tilde{B}) \to \operatorname{Spec}(B)$ be the maps corresponding to φ_A and φ_B. Then, there exists an isomorphism of schemes*

$$\Theta : \phi_A^{-1}(U \cap V) \xrightarrow{\cong} \phi_B^{-1}(U \cap V)$$

such that the following commutes for any affine open $\operatorname{Spec}(R) \subseteq U \cap V$ which is basic in both U and V by the isomorphisms $A_f \cong R \cong B_g$ (see Lemma 1.4.4.3)

$$\begin{tikzcd} \phi_A^{-1}(U \cap V) \arrow[r, "\Theta", "\cong"'] & \phi_B^{-1}(U \cap V) \\ \operatorname{Spec}(\tilde{A}_f) \arrow[u, hook] \arrow[r, "\cong", "\Theta_f"'] & \operatorname{Spec}(\tilde{B}_g) \arrow[u, hook] \end{tikzcd}$$

where Θ_f is obtained from $\theta : A_f \cong B_g$ via˜construction (Remark 1.6.2.5).

Proof. Cover $U \cap V$ by open affines which are basic in both U and V (Lemma 1.4.4.3) and write $U \cap V = \bigcup_{i\in I} \mathrm{Spec}\,(A_{f_i}) = \bigcup_{i\in I} \mathrm{Spec}\,(B_{g_i})$ where $f_i \in A$ and $g_i \in B$. Consequently we may write

$$\phi_A^{-1}(U \cap V) = \bigcup_{i\in I} \phi_A^{-1}(\mathrm{Spec}\,(A_{f_i})) = \bigcup_{i\in I} \mathrm{Spec}\,(\tilde{A}_{f_i})$$

and thus similarly,

$$\phi_B^{-1}(U \cap V) = \bigcup_{i\in I} \mathrm{Spec}\,(\tilde{B}_{g_i}).$$

For each $i \in I$, Lemma 1.6.2.8 provides us with an isomorphism

$$\Theta_i : \mathrm{Spec}\,(\tilde{A}_{f_i}) \xrightarrow{\cong} \mathrm{Spec}\,(\tilde{B}_{g_i}) \hookrightarrow \tilde{V}.$$

We claim that Θ_i can be glued. Indeed, for $i \neq j$, we have $\mathrm{Spec}\,(\tilde{A}_{f_i}) \cap \mathrm{Spec}\,(\tilde{A}_{f_j}) = \mathrm{Spec}\,(\tilde{A}_{f_i f_j})$, therefore we reduce to showing that Θ_i and Θ_j are equal when restricted to $\mathrm{Spec}\,(\tilde{A}_{f_i f_j})$. Observe from Lemma 1.4.4.3 that for each $i \in I$, the isomorphism $A_{f_i} \cong B_{g_i}$ takes $f_i \mapsto g_i$. The above is now equivalent to showing that the isomorphisms $\theta_i : \tilde{A}_{f_i} \cong \tilde{B}_{g_i}$ and $\theta_j : \tilde{A}_{f_j} \cong \tilde{B}_{g_j}$ obtained from $A_{f_i} \cong B_{g_i}$ and $A_{f_j} \cong B_{g_j}$ fit in the following commutative diagram

$$\begin{tikzcd} \tilde{A}_{f_i f_j} \arrow[r, "(\theta_i)_{f_j}"] \arrow[d, equal, "\mathrm{id}"'] & \tilde{B}_{g_i f_j} \arrow[d, equal, "\mathrm{id}"] \\ \tilde{A}_{f_j f_i} \arrow[r, "(\theta_j)_{f_i}"'] & \tilde{B}_{g_j f_i} \end{tikzcd} .$$

But $\theta_i(f_j) = g_j$ and $\theta_j(f_i) = g_i$, as mentioned above. Therefore $\tilde{B}_{g_i f_j} = \tilde{B}_{g_i g_j} = \tilde{B}_{g_j g_i} = \tilde{B}_{g_j f_i}$ and the above square commutes, showing that Θ_i glues to give a map $\Theta : \phi_A^{-1}(U \cap V) \to \phi_B^{-1}(U \cap V)$, which is an isomorphism as locally it is an isomorphism (Proposition 1.3.1.5). $\square$

Using Proposition 1.6.2.9, we can now globalize a strongly local construction.

Theorem 1.6.2.10. *Let X be a scheme and $\{\varphi_S : S \to \tilde{S}\}$ be a strongly local construction on rings. Then there exists a scheme $\alpha : \tilde{X} \to X$ such that for any affine open $\mathrm{Spec}\,(A) \hookrightarrow X$, the following square commutes*

$$\begin{tikzcd} \mathrm{Spec}\,(\tilde{A}) \arrow[r, hook] \arrow[d, "\phi_A"'] & \tilde{X} \arrow[d, "\alpha"] \\ \mathrm{Spec}\,(A) \arrow[r, hook] & X \end{tikzcd} .$$

Proof. We first construct $\tilde{X}$ by gluing each $\mathrm{Spec}\,(\tilde{A})$. Indeed, let $\{V_i = \mathrm{Spec}\,(A_i)\}_{i\in I}$ be the collection of affine opens in X and let $\{\tilde{X}_i = \mathrm{Spec}\,(\tilde{A}_i)\}$ be the collection of corresponding $\tilde{-}$-constructions. Let $\phi_i : \tilde{X}_i \to V_i$ be the maps corresponding to φ_{A_i}.

For each $i \neq j \in I$ we wish to construct open subschemes $U_{ij} \subseteq \tilde{X}_i$ and isomorphisms $\varphi_{ij} : U_{ij} \to U_{ji}$ satisfying the gluing conditions of Definition 1.6.2.1. We let

$$U_{ij} = \phi_i^{-1}(V_i \cap V_j).$$

Then Proposition 1.6.2.9 provides us with an isomorphism

$$\varphi_{ij} : U_{ij} \xrightarrow{\cong} U_{ji}.$$

It is immediate that $U_{ii} = \tilde{X}_i$ and $\varphi_{ii} = \mathrm{id}_{U_{ii}}$. Moreover, $\varphi_{ji} = \varphi_{ij}^{-1}$ by construction. We now check the cocycle condition. Indeed, pick $i, j, k \in I$ and pick an open affine $\mathrm{Spec}\,(R) \subseteq V_i \cap V_j \cap V_k$ in X which is basic open in V_i, V_j and V_k (Lemma 1.4.4.3 such that we have isomorphisms $A_{i,f_i} \cong A_{j,f_j} \cong A_{k,f_k} \cong R$ so that the following triangle commutes

$$\begin{tikzcd} A_{i,f_i} \arrow[r, "\cong"] \arrow[dr, "\cong"'] & A_{j,f_j} \arrow[d, "\cong"] \\ & A_{k,f_k} \end{tikzcd} . \qquad (*)$$

By taking inverse images under ϕ_i, it follows that $\mathrm{Spec}\,(\tilde{A}_{i,f_i}) \subseteq U_{ij} \cap U_{ik}$ is basic open in both $\tilde{X}_i$ and $\tilde{X}_j$. We wish to show that φ_{ik} restricted to $\mathrm{Spec}\,(\tilde{A}_{i,f_i})$ is the composition $\varphi_{jk} \circ \varphi_{ij}$. By Proposition 1.6.2.9, we get that φ_{ik} on this open affine is an isomorphism to $\mathrm{Spec}\,(\tilde{A}_{k,f_k})$ and φ_{ij} is an isomorphism to $\mathrm{Spec}\,(\tilde{A}_{j,f_j})$. Consequently, we wish to show that the following triangle of isomorphisms commute

$$\begin{tikzcd} \mathrm{Spec}\,(\tilde{A}_{i,f_i}) \arrow[r, "\varphi_{ij}"] \arrow[dr, "\varphi_{ik}"'] & \mathrm{Spec}\,(\tilde{A}_{j,f_j}) \arrow[d, "\varphi_{jk}"] \\ & \mathrm{Spec}\,(\tilde{A}_{k,f_k}) \end{tikzcd} .$$

But these isomorphisms are obtained by the following isomorphisms on the localizations (Proposition 1.6.2.9):

$$\begin{tikzcd} \tilde{A}_{i,f_i} \arrow[r, "\cong"] \arrow[dr, "\cong"'] & \tilde{A}_{j,f_j} \arrow[d, "\cong"] \\ & \tilde{A}_{k,f_k} \end{tikzcd} .$$

Hence it suffices to show that the above triangle commutes. The Lemma 1.6.2.7 applied on $(*)$ yields the required commutativity. □

Definition 1.6.2.11 ($\tilde{}$**-fication).** Let $\{\varphi_S : S \to \tilde{S}\}$ be a strongly local construction of rings and let X be a scheme. The scheme $\tilde{X} \to X$ obtained in Theorem 1.6.2.10 is called the $\tilde{}$-fication of X.

1.6.3 Reduced scheme of a scheme

For any scheme X, we can obtain a scheme with the same underlying space but with reduced structure sheaf. This procedure is called *reducing a scheme* to a reduced scheme.

Construction 1.6.3.1. Let X be a scheme. Consider the sheaf associated to the presheaf $U \mapsto \mathcal{O}_X(U)/\mathfrak{n}_U$ where $\mathfrak{n}_U$ is the nilradical of $\mathcal{O}_X(U)$ and denote this sheaf by $\mathcal{O}_X^{\mathrm{red}}$. The pair $(X, \mathcal{O}_X^{\mathrm{red}})$ will be called the *associated reduced scheme* of the scheme $(X, \mathcal{O}_X)$, usually denoted by X_{red}. Indeed, $(X, \mathcal{O}_X^{\mathrm{red}})$ is a scheme as the following result shows.

Remark 1.6.3.2 (*Reducing a ring is a strongly local construction*)**.** It is easy to see that $A \to A/\mathfrak{n}$ for each ring A defines a strongly local construction on rings as in Definition 1.6.2.4. Consequently, by Theorem 1.6.2.10, we immediately get a scheme $\tilde{X}$ obtained by reducing each open affine by dividing by nilradical. Indeed, one checks that we get the same scheme as $(X, \mathcal{O}_X^{\mathrm{red}})$. However, we still give a proof of $(X, \mathcal{O}_X^{\mathrm{red}})$ being a scheme without appealing to Theorem 1.6.2.10.

Lemma 1.6.3.3. [25] *Let X be a scheme. Then,*

1. *the pair $(X, \mathcal{O}_X^{\mathrm{red}})$ is a scheme,*
2. *there exists a map of schemes $\varphi : (X, \mathcal{O}_X^{\mathrm{red}}) \to (X, \mathcal{O}_X)$ which is a homeomorphism on the spaces.*

Proof. 1. Let $(\mathrm{Spec}\,(A), \mathcal{O}_{\mathrm{Spec}(A)})$ be an open affine of X. We shall show that $(\mathrm{Spec}\,(A), \mathcal{O}^{\mathrm{red}}_{\mathrm{Spec}(A)})$ is isomorphic to $(\mathrm{Spec}\,(A_{\mathrm{red}}), \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})})$. First, the isomorphism on spaces is straightforward as every prime ideal contains nilradical (nilradical is the intersection of all prime ideals, Lemma 16.1.2.9). We thus need to produce a sheaf morphism $\mathcal{O}^{\mathrm{red}}_{\mathrm{Spec}(A)} \to \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})}$ which is an isomorphism. Let us denote the presheaf $U \mapsto \mathcal{O}_X(U)/\mathfrak{n}_U$ by F. We first immediately reduce to showing the existence of a map $F \to \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})}$ which is an isomorphism on basic open sets, as we then obtain a map of sheaves $\mathcal{O}^{\mathrm{red}}_{\mathrm{Spec}(A)} \to \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})}$ by the universal property of sheafification (Theorem 20.2.0.1) which is an isomorphism on stalks (Theorem 20.3.0.6, 4).

Since sheaves and sheaf morphisms are uniquely determined by defining them on a basis, thus we further reduce to defining a presheaf map $F \to \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})}$ with above properties on a basis. Since $\mathrm{Spec}\,(A)$ has a canonical basis, namely, $\mathcal{B} = \{\mathrm{Spec}\,(A)_f\}_{f \in A}$, consequently one sees that isomorphism $A_f/\mathfrak{n}_f \cong (A/\mathfrak{n})_f$ can be naturally extended to a presheaf map $F \to \mathcal{O}_{\mathrm{Spec}(A_{\mathrm{red}})}$, which is an isomorphism on the basis $\mathcal{B}$.

2. Consider the map $f : (X, \mathcal{O}_X^{\mathrm{red}}) \to (X, \mathcal{O}_X)$ which is given by id_X on spaces but by the following quotient map $\mathcal{O}_X(U) \to \mathcal{O}_X(U)/\mathfrak{n}_U \to \mathcal{O}_X^{\mathrm{red}}(U)$. ∎

There is a universal property of reduced schemes which says that a map out of a reduced scheme necessarily has to factor through the reduction of the codomain.

Proposition 1.6.3.4. [26] *Let $f : X \to Y$ be a map of schemes with X being a reduced scheme. Then there exists a unique map of schemes $g : X \to Y_{\mathrm{red}}$ such that the triangle commutes:*

$$\begin{tikzcd} Y & Y_{\mathrm{red}} \arrow[l, "\varphi"'] \\ X \arrow[u, "f"] \arrow[ur, "g"'] & \end{tikzcd}.$$

Proof. The map g on spaces is immediate; it shoould be identical to f as φ is identity on spaces. The map $g^\flat$ on the other hand can be constructed as follows. First observe that if A and B are rings with B being reduced, then any ring map $\eta : A \to B$ extends to a unique map $\tilde{\eta} : A_{\mathrm{red}} \to B$ given

[25] Exercise II.2.3.b of Hartshorne.
[26] Exercise II.2.3.c of Hartshorne.

by $a + \mathfrak{n} \mapsto \eta(a)$ which makes the triangle commute:

$$\begin{tikzcd} A \arrow[r, "\eta"] \arrow[d, two heads] & B \\ A_{\mathrm{red}} \arrow[ur, "\tilde{\eta}"'] & \end{tikzcd}.$$

In our case, we therefore get a unique map $\tilde{f_U^\flat}$ as below for any $U \subseteq Y$, which further gives us the required unique map $g_U^\flat : \mathcal{O}_Y^{\mathrm{red}}(U) \to \mathcal{O}_X(f^{-1}(U))$ which we need (by universality of sheafification, Theorem 20.2.0.1):

$$\begin{tikzcd} \mathcal{O}_Y(U) \arrow[r, "\varphi_U"] \arrow[rr, bend left, "f_U^\flat"] \arrow[dr, two heads] & \mathcal{O}_Y^{\mathrm{red}}(U) \arrow[r, "=:g_U^\flat"] & \mathcal{O}_X(f^{-1}(U)) \\ & \mathcal{O}_Y(U)/\mathfrak{n}_U \arrow[u] \arrow[ur, "\tilde{f_U^\flat}"'] & \end{tikzcd}.$$

One can then easily check that g as given above makes the triangle commute. □

For each closed set $Z \subseteq X$ of a scheme, we construct a unique closed reduced subscheme structure over it.

Construction 1.6.3.5 (*Reduced induced subscheme*)**.** Let X be a scheme and $Z \subseteq X$ be a closed set. We wish to define a natural scheme structure on the subspace Z. Indeed, if $X = \mathrm{Spec}\,(A)$ is affine and $Z \subseteq X$ is closed, then let $\mathfrak{a} = \bigcap_{\mathfrak{p}\in Z} \mathfrak{p}$ so that $Z = V(\mathfrak{a})$. Then we define the reduced induced subscheme structure on Z as that of $\mathrm{Spec}\,(A/\mathfrak{a})$. Observe that $(Z, \mathcal{O}_{\mathrm{Spec}(A/\mathfrak{a})})$ is a reduced scheme as $\mathfrak{a} \supseteq \mathfrak{n}$ where $\mathfrak{n} \leq A$ is the nilradical.

For an arbitrary scheme X and a closed subset $Z \subseteq X$, we proceed as follows. Let $\{U_i\}_{i\in I}$ be the collection of all open affines in X. Consider the intersections $Z_i = U_i \cap Z$ for each $i \in I$. As $Z_i \subseteq U_i$ are closed subsets in an affine scheme U_i, so by definition they carry the reduced induced subscheme structure on Z_i. We claim that the sheaves on each Z_i can be glued. Indeed, by usual argument involving Lemma 1.4.4.3, we reduce to checking that if $U = \mathrm{Spec}\,(A)$ is an open affine, $V = D(f) \subseteq U$ a basic open subset, $\mathcal{R}_U$ and $\mathcal{R}_V$ denote the sheaves obtained by reduced induced subscheme structures on $Z \cap U$ and $Z \cap V$ respectively, then

$$(\mathcal{R}_U)_{|Z\cap V} \cong \mathcal{R}_V.$$

Let $\mathfrak{a} = \bigcap_{\mathfrak{p}\in Z\cap U} \mathfrak{p}$ which gives the required structure on $Z \cap U$. Similarly, we have $\mathfrak{b} = \bigcap_{\mathfrak{p}\in Z\cap V} \mathfrak{p}$. We claim that $\mathfrak{b} = \mathfrak{a}A_f$. This would establish the required isomorphism between $A/\mathfrak{a}$ and $A_f/\mathfrak{b}$. Indeed, by definition, it is clear that $\mathfrak{b} \supseteq \mathfrak{a}A_f$. Conversely, pick $x/f^n \in \mathfrak{a}A_f$ where $x \in \mathfrak{a}$. We wish to show that $x/f^n \in \mathfrak{b}$. Pick any prime ideal $\mathfrak{q} \in Z \cap V$. We wish to show that $x/f^n \in \mathfrak{q}$. As $x \in \mathfrak{a}$, therefore $x \in \mathfrak{p}$ for each $\mathfrak{p} \in Z \cap U$. Thus, for $\mathfrak{p} \in D(f)$, $x \in \mathfrak{p}$. As each $\mathfrak{q} \in Z \cap V$ comes from $\mathfrak{p} \in Z \cap D(f)$, therefore $x/1 \in \mathfrak{b}$ and thus $x/f^n \in \mathfrak{b}$.

This completes the gluing procedure, to yield a subscheme structure on Z which we call the reduced induced subscheme structure on Z.

We now show the universal property of the above construction.

Proposition 1.6.3.6 (Universal property of reduced induced subscheme)**.** ***TODO.***

1.6.4 Fiber product of schemes

One of the most important tool in scheme theory is that of fiber product of schemes. This is essential as this is exactly the right notion using which one can define intersection of subschemes, which is one of the fundamental goals of this book.

Existence of fiber products is equivalent to saying that the category of schemes **Sch** have all pullbacks. In particular, it is equivalent to saying that for any two S-schemes X and Y, their product in $\mathbf{Sch}/S$ exists, called the *fiber product* denoted $X \times_S Y$.

However, we need to be more explicit than this abstract definition; we have to show that $X \times_S Y$ actually exists. Since we know how pushouts are constructed in the category of rings, their tensor products, therefore we can define it for affine schemes without much effort using the functor $\operatorname{Spec}(-) : \mathbf{Ring}^{\mathrm{op}} \to \mathbf{Sch}$ of Theorem 1.3.0.5.

Definition 1.6.4.1. (**Fiber product of affine schemes**) Let the following be a coCartesian[27] diagram of rings (or of R-algebras)

$$\begin{tikzcd} A \otimes_R B & B \arrow[l] \\ A \arrow[u] & R \arrow[l, "f"] \arrow[u, "g"'] \end{tikzcd}.$$

Since the $\operatorname{Spec}(-) : \mathbf{Ring}^{\mathrm{op}} \to \mathbf{Sch}$ of Theorem 1.3.0.5 is right adjoint to global sections, therefore it preserves all limits of $\mathbf{Ring}^{\mathrm{op}}$, and thus, takes the above pushout diagram of R-algebras to a pullback diagram of affine schemes over $\operatorname{Spec}(R)$:

$$\begin{tikzcd} \operatorname{Spec}(A \otimes_R B) \arrow[r] \arrow[d] & \operatorname{Spec}(B) \arrow[d, "\operatorname{Spec}(g)"] \\ \operatorname{Spec}(A) \arrow[r, "\operatorname{Spec}(f)"'] & \operatorname{Spec}(R) \end{tikzcd}.$$

We hence define $\operatorname{Spec}(A \otimes_R B)$ to be the fiber product of affine schemes $\operatorname{Spec}(A)$ and $\operatorname{Spec}(B)$ over $\operatorname{Spec}(R)$.

Definition 1.6.4.2 (Fiber product of schemes). Fiber product of S-schemes X and$Y$ is an S-scheme $X \times_S Y$ such that for any other S-scheme Z with map $f : Z \to X$ and $g : Z \to Y$ over S, there exists a unique map $u : Z \to X \times_S Y$ such that the diagram commutes

$$\begin{tikzcd} Z \arrow[dr, dashed, "u"] \arrow[drr, bend left, "g"] \arrow[ddr, bend right, "f"'] & & \\ & X \times_S Y \arrow[r] \arrow[d] & Y \arrow[d] \\ & X \arrow[r] & S \end{tikzcd}.$$

[27] another name for pushout diagrams.

The most important part in this construction is the description of the structure sheaf of $X \times_S Y$. We now show how to construct fiber products of arbitrary S-schemes. In the process, we give a rather explicit description of fiber products and its structure sheaf, which we may think of as an explicit definition of fiber product. We begin with the affine case. Recall the notion of compositum of fields in Definition **??**.

Proposition 1.6.4.3. *Let A, B be two R-algebras and let $X = \operatorname{Spec}(A), Y = \operatorname{Spec}(B)$ and $S = \operatorname{Spec}(R)$. Then, as a set, we have the following bijection*

$$X \times_S Y \cong \left\{ \begin{matrix} \text{Tuples } (\mathfrak{p}_A, \mathfrak{p}_B, L, \alpha, \beta) \text{ where } \mathfrak{p}_A \in X,\ \mathfrak{p}_B \in Y \\ \text{such that both have same inverse image } \mathfrak{p}_R \text{ in } S \\ \text{and } (L, \alpha, \beta) \text{ is the compositum of fields } \kappa(\mathfrak{p}_A) \\ \text{and } \kappa(\mathfrak{p}_B) \text{ over } \kappa(\mathfrak{p}_R). \end{matrix} \right\}$$

Proof. Pick any prime ideal $\mathfrak{p} \in X \times_S Y = \operatorname{Spec}(A \otimes_R B)$. We wish to construct the datum $(\mathfrak{p}_A, \mathfrak{p}_B, L, \alpha, \beta)$. **TODO**. □

We now construct the fiber product of two schemes. This is more of an exercise in gluing techniques rather than anything else, so is ommited.

Theorem 1.6.4.4. *Let X, Y be two S-schemes. The fiber product $X \times_S Y$ exists.* □

Remark 1.6.4.5. While working with fiber products, one of the most important tool is its universal property. Most of the results about fiber products rarely uses the point-set construction as laid out above, just like the construction of tensor product is rarely used. Consequently, one should/must prove results about fiber products only using universal properties.

We now portray some easy applications of the universal property of fiber products.

Lemma 1.6.4.6. *Let $f : X \to Y$ and $g : Z \to Y$ be scheme morphisms and $U \subseteq X$ be an open subscheme. If $p : X \times_Y Z \to X$ is the scheme over X obtained by base change under f, then $p^{-1}(U) \cong U \times_Y Z$.*

Proof. We claim that the open subscheme $p^{-1}(U)$ of $X \times_Y Z$ is isomorphic to $U \times_Y Z$ by showing that it satisfies the same universal property. Indeed, suppose we have the following diagram

$$\begin{tikzcd} T \arrow[drr, bend left, "k"] \arrow[ddr, bend right, "h"'] & & \\ & p^{-1}(U) \arrow[r, "q"] \arrow[d, "p"'] & Z \arrow[d, "g"] \\ & U \arrow[r, "f"'] & Y \end{tikzcd}$$

where $f \circ h = g \circ k$. By the universal property of fiber product $X \times_Y Z$, we get a unique map $\varphi : T \to X \times_Y Z$ such that $p \circ \varphi = h$ and $q \circ \varphi = k$. As $\operatorname{Im}(h) \subseteq U$, therefore $\operatorname{Im}(p \circ \varphi) \subseteq U$. Consequently, we have $\operatorname{Im}(\varphi) \subseteq p^{-1}(U)$, hence we may write $\varphi : T \to p^{-1}(U)$, where $p^{-1}(U)$ is an open subscheme of $X \times_Y Z$. Thus, we get a unique map $\varphi : T \to p^{-1}(U)$ which makes the above diagram a fiber product diagram, thus completing the proof. □

The following is an important technical result.

Lemma 1.6.4.7. *Let $X = \bigcup_\alpha U_\alpha$ be an open cover of the scheme X. Let $f : X \to Y$ and $g : Z \to Y$ be scheme morphisms. Then,*

$$X \times_Y Z \cong \bigcup_\alpha U_\alpha \times_Y Z.$$

Proof. Let $p : X \times_Y Z \to X$ be the fiber product scheme over X obtained by base change along f. Then,

$$p^{-1}\left(\bigcup_\alpha U_\alpha\right) = \bigcup_\alpha p^{-1}(U_\alpha).$$

By Lemma 1.6.4.6, we see that $p^{-1}(U_\alpha) \cong U_\alpha \times_Y Z$. It follows that

$$X \times_Y Z = p^{-1}(X) = \bigcup_\alpha p^{-1}(U_\alpha) \cong \bigcup_\alpha U_\alpha \times_Y Z,$$

as needed. □

Lemma 1.6.4.8. *Let $f : X \to Y$ and $g : Z \to Y$ be scheme morphisms and $U \subseteq X$ be an open subscheme such that $f(U) \subseteq V$ for some open subscheme $V \subseteq Y$ and let $W = g^{-1}(V)$ be an open subscheme in Z. If $p : X \times_Y Z \to X$ is the fiber product over X obtained by base change along f, then $p^{-1}(U) \cong U \times_Y Z \cong U \times_V W$.*

Proof. The first isomorphism is the content of Lemma 1.6.4.6. The second isomorphism follows from the simple observation that $U \times_Y Z$ satisfies the same universal property as that of $U \times_V W$. □

We portray some pathologies of fiber product in the following examples.

Example 1.6.4.9. We show that fiber product of one point schemes may have more than one point(!) Indeed, consider the schemes $X = Y = \text{Spec}\,(\mathbb{C})$ over $\text{Spec}\,(\mathbb{R})$. Observe that $X \times_{\text{Spec}(\mathbb{R})} Y \cong \text{Spec}\,(\mathbb{C} \otimes_\mathbb{R} \mathbb{C})$. But since we have

$$\begin{aligned} \mathbb{C} \otimes_\mathbb{R} \mathbb{C} &\cong \frac{\mathbb{R}[x]}{x^2+1} \otimes_\mathbb{R} \mathbb{C} \\ &\cong \frac{\mathbb{C}[x]}{x^2+1} \\ &\cong \mathbb{C} \times \mathbb{C} \end{aligned}$$

by Chinese remainder theorem. Consequently, $\text{Spec}\,(\mathbb{C} \times \mathbb{C}) \cong \text{Spec}\,(\mathbb{C}) \amalg \text{Spec}\,(\mathbb{C})$ which has 2 points.

1.6.5 Applications of fiber product

We would now like to portray some of the applications of fiber products, especially in endowing the fibers of a morphism with a scheme structure.

Inverse image of a closed subscheme

TODO.

Fibers of a map

Keep in mind the Lemma 20.5.0.3 and the surrounding remarks about stalks of sheaves for the remainder of this discussion. Let $f : X \to Y$ be a map and $y \in Y$ be a point. We endow $f^{-1}(y) \hookrightarrow X$ with a scheme structure. Define the *fiber* of f at y to be the following fiber product:

$$X_y := X \times_Y \operatorname{Spec}(\kappa(y)).$$

We at times denote it by $X \times_Y y$. Note that by natural map onto second factor, X_y is a scheme over $\kappa(y)$.

We now show that fiber of a scheme morphism as defined above matches with the usual notion of fiber in the sense that both spaces are homeomorphic. We first do this for affine schemes.

Proposition 1.6.5.1. *Let $X = \operatorname{Spec}(S)$, $Y = \operatorname{Spec}(R)$ and $f : X \to Y$ be the map associated to a ring homomorphism $\varphi : R \to S$. Let $y = \mathfrak{p} \in Y$ be a prime ideal of R. Then, X_y is homeomorphic to the subspace $f^{-1}(y)$ of Y.*

Proof. We have that $X_y = \operatorname{Spec}(S \otimes_R \kappa(\mathfrak{p}))$, that is, the fiber of φ at prime ideal $\mathfrak{p}$ (Definition 16.5.1.5). We now calculate $S \otimes_R \kappa(\mathfrak{p})$. Indeed, we have

$$\begin{aligned} S \otimes_R \kappa(\mathfrak{p}) = S \otimes_R F(R/\mathfrak{p}) &\cong S \otimes_R (R/\mathfrak{p} \otimes_R R_{\mathfrak{p}}) \\ &\cong S/\mathfrak{p}S \otimes_R R_{\mathfrak{p}} \\ &\cong (S/\mathfrak{p}S)_{\varphi(R \setminus \mathfrak{p})} . \end{aligned}$$

It follows from Lemma 16.1.2.3 that $\operatorname{Spec}(S \otimes_R \kappa(\mathfrak{p}))$ is exactly the subspace of X consisting of those primes $\mathfrak{q}$ such that $\mathfrak{q} \supseteq \varphi(\mathfrak{p})$ and does not intersects $\varphi(R \setminus \mathfrak{p})$. This is equivalent to saying that $\varphi^{-1}(\mathfrak{q}) \supseteq \mathfrak{p}$ and $\varphi^{-1}(\mathfrak{q}) \subseteq \mathfrak{p}$, that is, $\varphi^{-1}(\mathfrak{q}) = \mathfrak{p}$, as needed. □

We now do the general case. The main idea is just to reduce to the affine case as above.

Lemma 1.6.5.2. *Let $f : X \to Y$ be a scheme morphism and $y \in Y$. Then, $f^{-1}(y)$ as a subspace of X is homeomorphic to X_y.*

Proof. Let $V = \operatorname{Spec}(B)$ be an open affine of Y containing y. Then, by definition of fiber products, we immediately see that $f^{-1}(V) \cong X \times_Y V$. Clearly, $f^{-1}(y) \subseteq f^{-1}(V)$. Cover $f^{-1}(V)$ by open affines $\{U_\alpha = \operatorname{Spec}(R_\alpha)\}$. By Proposition 1.6.5.1, we see that $f^{-1}(y) \cap U_\alpha \cong \operatorname{Spec}(R_\alpha \otimes_B \kappa(y)) = U_\alpha \times_V \operatorname{Spec}(\kappa(y))$. Since

$$\begin{aligned} X_y = X \times_Y \operatorname{Spec}(\kappa(y)) &\cong f^{-1}(V) \times_V \operatorname{Spec}(\kappa(y)) \\ &= \left(\bigcup_\alpha U_\alpha\right) \times_V \operatorname{Spec}(\kappa(y)) \\ &\cong \bigcup_\alpha (U_\alpha \times_V \operatorname{Spec}(\kappa(y))) \\ &\cong \bigcup_\alpha f^{-1}(y) \cap U_\alpha \\ &= f^{-1}(y), \end{aligned}$$

as needed. $\square$

Example 1.6.5.3. We calculate explicit fibers of a map at every point of a familiar map. *Write solution of Exercise 3.10 of Hartshorne Chapter 2, written in notebook.*

The fibers of $\operatorname{Spec}(\mathbb{Z}[x]) \to \operatorname{Spec}(\mathbb{Z})$

We know that $\operatorname{Spec}(\mathbb{Z})$ is the final object in the category of schemes **Sch**. We also know that there is the canonical inclusion $\mathbb{Z} \hookrightarrow \mathbb{Z}[x]$. This induces a map

$$\varphi : \operatorname{Spec}(\mathbb{Z}[x]) \longrightarrow \operatorname{Spec}(\mathbb{Z}).$$

Understanding the fibers of this map will allow us to understand the affine arithmetic surface $\operatorname{Spec}(\mathbb{Z})$ (as $\mathbb{Z}[x]$ is a 2-dimensional ring). Note that we can already understand $\operatorname{Spec}(\mathbb{Z}[x])$ by the results surrounding Gauss' lemma as done in Theorem 16.1.5.3, but the following is a more geometric way of understadning this.

Proposition 1.6.5.4. *The prime ideals of $\mathbb{Z}[x]$ can be categorized into following three types.*

1. *$\langle p\rangle$ where $p \in \mathbb{Z}$ is a prime,*
2. *$\langle f(x)\rangle$ where $f(x) \in \mathbb{Z}[x]$ is an irreducible polynomial,*
3. *$\langle p, f(x)\rangle$ where $p \in \mathbb{Z}$ is a prime and $f(x) \in \mathbb{Z}[x]$ irreducible in $\mathbb{Z}[x]$ which remains irreducible in $\mathbb{Z}/p\mathbb{Z}$,*

Proof. We will prove this by analyzing the fibers of $f : \operatorname{Spec}(\mathbb{Z}[x]) \to \operatorname{Spec}(\mathbb{Z})$. Pick a prime $p \in \mathbb{Z}$ and denote $X = \operatorname{Spec}(\mathbb{Z}[x])$. The fiber $X_p = \operatorname{Spec}(\mathbb{Z}[x]) \times_{\operatorname{Spec}(\mathbb{Z})} \operatorname{Spec}(\kappa(p))$. As $\kappa(p) = \mathbb{F}_p$, finite field with p elements, therefore we have that $X_p = \operatorname{Spec}(\mathbb{Z}[x] \otimes_{\mathbb{Z}} \mathbb{F}_p) = \operatorname{Spec}(\mathbb{F}_p[x])$. Note that for same reasons we have $X_o = \operatorname{Spec}(\mathbb{Q}[x])$.

As fibers of f covers the whole scheme, it follows that any point in $\mathbb{Z}[x]$ looks like one of the following:

1. a prime ideal in $\mathbb{Q}[x]$,
2. a prime ideal in $\mathbb{F}_p[x]$.

Moreover, we have the following diagrams

$$\begin{tikzcd}
\operatorname{Spec}(\mathbb{F}_p[x]) \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & \operatorname{Spec}(\mathbb{Z}[x]) \arrow[d] \arrow[dr, phantom, "\leftrightsquigarrow"] & \mathbb{F}_p[x] \arrow[dr, phantom, "\ulcorner"] & \mathbb{Z}[x] \arrow[l] \\
\operatorname{Spec}(\mathbb{F}_p) \arrow[r] & \operatorname{Spec}(\mathbb{Z}) & \mathbb{F}_p \arrow[u] & \mathbb{Z} \arrow[l] \arrow[u]
\end{tikzcd}.$$

$$\begin{tikzcd}
\operatorname{Spec}(\mathbb{Q}[x]) \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & \operatorname{Spec}(\mathbb{Z}[x]) \arrow[d] \arrow[dr, phantom, "\leftrightsquigarrow"] & \mathbb{Q}[x] \arrow[dr, phantom, "\ulcorner"] & \mathbb{Z}[x] \arrow[l] \\
\operatorname{Spec}(\mathbb{Q}) \arrow[r] & \operatorname{Spec}(\mathbb{Z}) & \mathbb{Q} \arrow[u] & \mathbb{Z} \arrow[l] \arrow[u]
\end{tikzcd}$$

Observe that $\mathbb{Z}[x] \to \mathbb{F}_p[x]$ is the mod-p map. Since every prime ideal of $\mathbb{Z}[x]$ now is a inverse image of a prime ideal by $\mathbb{Z}[x] \to \mathbb{F}_p[x]$ and $\mathbb{Z}[x] \to \mathbb{Q}[x]$, we get the desired result. $\square$

Geometric properties

Cover geometric reducibility and etc etc from Hartshorne exercises.

1.6.6 Normal schemes and normalization

Do mainly Exercise 3.7, 3.8 of Chapter 2 of Hartshorne. Also do Exercise 3.17, 3.18 of Chapter 1 of Hartshorne.

We now study a class of schemes which globalizes the notion of integral closure from algebra (Definition 16.7.1.11). These will find its main use in arithmetic where normal domains fundamental.

Definition 1.6.6.1 (**Normal schemes**)**.** A scheme X is said to be normal if for all $x \in X$, the local ring $\mathcal{O}_{X,x}$ is a normal domain.

The following is immediate from local nature of normal domains (Proposition 16.7.2.10).

Lemma 1.6.6.2. *Let X be an integral scheme. Then the following are equivalent:*

1. *X is a normal scheme.*
2. *For all open affine* $\operatorname{Spec}(A) \subseteq X$*, the ring A is a normal domain.*

Proof. As X is integral, therefore for every open affine $\operatorname{Spec}(A)$ of X, A is a domain by Lemma 1.4.2.2. As X is normal iff $\mathcal{O}_{X,x}$ is a normal domain for all $x \in X$, the result follows from Proposition 16.7.2.10. □

The main result in normal schemes is that any integral scheme induces a unique normal scheme obtained by normalizing each open affine.

Theorem 1.6.6.3. [28] *Let X be an integral scheme. Then there exists a scheme $\tilde{X} \to X$ over X where $\tilde{X}$ is a normal integral scheme such that for any normal integral scheme Z and a dominant map $f : Z \to X$, there exists a unique map $\tilde{f} : Z \to \tilde{X}$ such that the following commutes*

$$\begin{tikzcd} \tilde{X} \arrow[d] & Z \arrow[l, dashed, "\tilde{f}"'] \arrow[dl, "f"] \\ X & \end{tikzcd}.$$

The scheme $\tilde{X} \to X$ is called the normalization of X and is unique upto isomorphism.

We first see this for affine domains.

Lemma 1.6.6.4. *Let $X = \operatorname{Spec}(A)$ be an integral affine scheme and $Z = \operatorname{Spec}(B)$ be a normal integral affine scheme. Let $\tilde{X} = \operatorname{Spec}(\tilde{A})$ be the normalization of X and denote the natural map $\pi : \tilde{X} \to X$. If $f : Z \to X$ is any dominant map, then there exists a map $\tilde{f} : Z \to \tilde{X}$ such that $\pi \circ \tilde{f} = f$.*

$$\begin{tikzcd} \operatorname{Spec}(\tilde{A}) \arrow[d, "\pi"'] & \operatorname{Spec}(B) \arrow[l, dashed, "\tilde{f}"'] \arrow[dl, "f"] \\ \operatorname{Spec}(A) & \end{tikzcd}.$$

[28]Exercise II.3.8 of Hartshorne.

Proof. Indeed, by Proposition 16.7.2.12, this follows immediately. □

Remark 1.6.6.5. By Remark 16.7.2.11, it follows that normalization is a strongly local property. Thus Theorem 1.6.6.3 holds.

Proof of Theorem 1.6.6.3. By Remark 16.7.2.11, it follows that normalization is a strongly local construction for domains. Let $A \hookrightarrow \tilde{A}$ be the normalization map for any domain A. Therefore by Theorem 1.6.2.10, we have a scheme $\alpha : \tilde{X} \to X$ such that for any open affine $\mathrm{Spec}\,(A) \hookrightarrow X$, the following diagram commutes

$$\begin{tikzcd}
\mathrm{Spec}\,(\tilde{A}) \arrow[r, hook] \arrow[d] & \tilde{X} \arrow[d, "\alpha"] \\
\mathrm{Spec}\,(A) \arrow[r, hook] & X
\end{tikzcd}$$

where the left vertical map is the map corresponding to normalization $A \hookrightarrow \tilde{A}$. This shows the construction of $\alpha : \tilde{X} \to X$.

Now let Z be an arbitrary normal integral scheme and $f : Z \to X$ be a dominant map. Pick any open affine $\mathrm{Spec}\,(A) \subseteq X$ and consider the non-empty (f is dominant) open subset $f^{-1}(\mathrm{Spec}\,(A))$. Write

$$f^{-1}(\mathrm{Spec}\,(A)) = \bigcup_{i \in I} \mathrm{Spec}\,(B_i)$$

where $\mathrm{Spec}\,(B_i) \subseteq Z$ are open affine. As Z is normal integral, therefore B_i are normal domains from Lemma 1.6.6.2. By restriction we thus have the map

$$f|_{\mathrm{Spec}(B_i)} : \mathrm{Spec}\,(B_i) \to \mathrm{Spec}\,(A)$$

for each $i \in I$. Observe that $\alpha^{-1}(\mathrm{Spec}\,(A)) \supseteq \mathrm{Spec}\,(\tilde{A})$. By Lemma 1.6.6.4, it follows that we have a unique map $\tilde{f}_i : \mathrm{Spec}\,(B_i) \to \mathrm{Spec}\,(\tilde{A})$ such that the following commutes

$$\begin{tikzcd}
\mathrm{Spec}\,(\tilde{A}) \arrow[d, "\alpha|_{\mathrm{Spec}(\tilde{A})}"'] & \mathrm{Spec}\,(B_i) \arrow[l, dashed, "\tilde{f}_i"'] \arrow[dl, "f|_{\mathrm{Spec}(B_i)}"] \\
\mathrm{Spec}\,(A) &
\end{tikzcd}.$$

It thus follows that for every open affine $\mathrm{Spec}\,(B_{ij}) \subseteq \mathrm{Spec}\,(B_i)$, we have a map $\tilde{f}_i : \mathrm{Spec}\,(B_i) \to \mathrm{Spec}\,(\tilde{A})$ by restriction. Hence by Lemma 1.6.6.4, we have that this is unique. As $\mathrm{Spec}\,(A) \subseteq X$ is arbitrary open affine, therefore we have an open affine covering $\{\mathrm{Spec}\,(A_i)\}_{i \in I}$ of X which by inverse image gives an open affine covering $\{\mathrm{Spec}\,(B_{ij})\}$ of Z and a collection of open affines $\{\mathrm{Spec}\,(\tilde{A}_i)\}$ of $\tilde{X}$ such that for each i, we have a unique map $\tilde{f}_{ij} : \mathrm{Spec}\,(B_{ij}) \to \tilde{X}$ such that

$$\begin{tikzcd}
\tilde{X} \arrow[d, "\alpha"'] & \mathrm{Spec}\,(\tilde{A}_i) \arrow[l, hook'] \arrow[d, "\alpha"'] & \mathrm{Spec}\,(B_{ij}) \arrow[l, dashed, "\tilde{f}_{ij}"'] \arrow[dl, "f"] \\
X & \mathrm{Spec}\,(A_i) \arrow[l, hook'] &
\end{tikzcd}$$

commutes. We claim that $\tilde{f}_{ij}$ can be glued to a unique map $\tilde{f} : Z \to \tilde{X}$, which would complete the proof. First, for a fixed i, we glue $\tilde{f}_{ij}$ and $\tilde{f}_{il}$. Indeed, covering the intersection $\operatorname{Spec}(B_{ij}) \cap \operatorname{Spec}(B)_{il}$ by open affines $\operatorname{Spec}(C_p)$, we immediately by restriction get maps $\tilde{f}_{ij} : \operatorname{Spec}(C_p) \to \operatorname{Spec}(\tilde{A}_i)$ and $\tilde{f}_{il} : \operatorname{Spec}(C_p) \to \operatorname{Spec}(\tilde{A}_i)$ which are thus equal by uniqueness. Hence, for each i, we may glue the maps $\{\tilde{f}_{ij}\}_j$ to obtain a unique map $\tilde{f}_i : Z_i = f^{-1}(\operatorname{Spec}(A_i)) \to \operatorname{Spec}(\tilde{A}_i)$ as in

$$\begin{tikzcd} \operatorname{Spec}(\tilde{A}_i) \arrow[d, "\alpha"'] & Z_i \arrow[l, dashed, "\tilde{f}_i"'] \arrow[dl, "f"] \\ \operatorname{Spec}(A_i) & \end{tikzcd}.$$

We now wish to glue these $\tilde{f}_i$. To this end, pick an affine open $\operatorname{Spec}(C) \subseteq Z_i \cap Z_k = f^{-1}(\operatorname{Spec}(A_i) \cap \operatorname{Spec}(A_k))$ and observe $\alpha^{-1}(\operatorname{Spec}(A_i) \cap \operatorname{Spec}(A_k)) \supseteq \operatorname{Spec}(\tilde{A}_i) \cap \operatorname{Spec}(\tilde{A}_k)$. We thus have the following diagram

$$\begin{tikzcd} \operatorname{Spec}(\tilde{A}_i) \arrow[d, "\alpha"'] & \operatorname{Spec}(C) \arrow[l, "\tilde{f}_i"'] \arrow[r, "\tilde{f}_k"] \arrow[d, "f"] & \operatorname{Spec}(\tilde{A}_k) \arrow[d, "\alpha"] \\ \operatorname{Spec}(A_i) & \operatorname{Spec}(A_i) \cap \operatorname{Spec}(A_k) \arrow[l, hook'] \arrow[r, hook] & \operatorname{Spec}(A_k) \end{tikzcd}.$$

By Lemma 1.6.6.4, it then suffices to show that $\tilde{f}_i(\operatorname{Spec}(C)), \tilde{f}_k(\operatorname{Spec}(C)) \subseteq \operatorname{Spec}(\tilde{A}_i) \cap \operatorname{Spec}(\tilde{A}_k)$, as then uniqueness would imply $\tilde{f}_i$ and $\tilde{f}_k$ are equal over $\operatorname{Spec}(C)$. By symmetry, it suffices to show this for $\tilde{f}_i$. Since $\alpha \circ \tilde{f}_i(\operatorname{Spec}(C)) \subseteq \operatorname{Spec}(A_i) \cap \operatorname{Spec}(A_k)$, therefore $\tilde{f}_i(\operatorname{Spec}(C)) \subseteq \alpha^{-1}(\operatorname{Spec}(A_i) \cap \operatorname{Spec}(A_k)) \cap \operatorname{Spec}(\tilde{A}_i) \subseteq \operatorname{Spec}(\tilde{A}_i) \cap \operatorname{Spec}(\tilde{A}_k)$, as required. Hence $\tilde{f}_i$ can be glued to a unique map $\tilde{f} : Z \to \tilde{X}$, thus completing the proof. □

The following is the globalization of the fact that normalization of a finite type algebra is again a finite type algebra, over a field (Noether's Theorem **??**).

Corollary 1.6.6.6. *If X is a finite type integral scheme, then the the normalization $\tilde{X} \to X$ is a finite map.*

Proof. **TODO**. □

1.7 Dimension & components of a scheme

Do from Vakil, Hartshorne Exercise 3.20, 3.21, 3.22 of Chapter 2.
The notion of dimension of a geometric object serves as an essential tool for any attempt at its understanding. Schemes are no different and we have a notion of dimension for them. However, we also have a notion of dimension of rings. This section explores how these two interrelates and thus facilitate understanding of geometry of schemes.

1.7.1 General properties

Before moving to schemes that we will encounter the most, let us first give a general review of the notion of dimension of topological spaces and some general properties of dimension of schemes. Recall that the dimension of a topological space is the supremum of the length of the strictly decreasing chains of finite length of closed irreducible subsets of the space. Further for a space X and a closed irreducible subset $Z \subseteq X$, the codimension of Z in X is defined to be the supremum of the length of strictly increasing chains of closed irreducible subsets starting from Z. For an arbitrary closed subset $Y \subseteq X$, we define codim $(Y, X) = \inf_{Z \subseteq Y}$ codim (Z, X) where Z varies over all closed irreducible subsets of Y. For any closed set $Y \subseteq X$, if $\dim X < \infty$, we always have codim $(Y, X) \leq \dim X$.

Proposition 1.7.1.1. *Let X be a topological space. Then,*

1. *If $Y \subseteq X$ is a subspace, then $\dim Y \leq \dim X$.*
2. *If $\{U_i\}_{i \in I}$ is an open covering of X, then $\dim X = \sup_i \dim U_i$.*
3. *Let $Y \subseteq X$ be a closed subspace and X be of finite dimension. If X is irreducible and $\dim Y = \dim X$, then $Y = X$.*

Proof. The main tool in all of them is just a clear understanding of the definition of dimension and of closed irreducible sets. We establish some terminologies to work with in this proof. For any space X a strictly decreasing chain of finite length of closed irreducible subsets will be called a *finite chain* of X and set of all finite chains will be denoted by $FC(X)$. We denote a chain by $Z_\bullet \in FC(X)$ and its length by $l(Z_\bullet)$. Consequently, $\dim X = \sup_{Z_\bullet \in FC(X)} l(Z_\bullet)$.

1. First observe that if Y is closed then the result is immediate as any finite chain of Y will be a finite chain of X. Consequently, we reduce to showing that $\dim Y \leq \dim \bar{Y}$. In particular, we reduce to showing that if Y is dense in X, then $\dim Y \leq \dim X$. It further suffices to show existence of a length preserving map $FC(Y) \to FC(X)$. Indeed, for any $Z_\bullet \in FC(Y)$, one observes that $\mathrm{Cl}_X(Z_i)$ is a closed subset of X which is further irreducible in X. Consequently, $\mathrm{Cl}(Z_\bullet)$ is a finite chain of X of same length as of $Z_\bullet$[29].
2. By 1. we already have $\sup_i \dim U_i \leq \dim X$ so we need only show that $\dim X \leq \sup_i \dim U_i$. It suffices to show that for each $Z_\bullet \in FC(X)$, there exists $i \in I$ and $W_\bullet \in FC(U_i)$ such that $l(Z_\bullet) \leq l(W_\bullet)$. Let $r = l(Z_\bullet)$ and $i \in I$ be such that $U_i \cap Z_r \neq \emptyset$. Then, $W_\bullet = U_i \cap Z_\bullet$ forms a finite chain of U_i of same length as $Z_\bullet$. To see this, observe that if $U_i \cap Z_a = U_i \cap Z_b$ where we may assume $Z_a \supsetneq Z_b$, then the open set $U_i \cap Z_a$ of Z_a is contained in the closed set Z_b of Z_a, hence the closure of $U_i \cap Z_a$ in Z_a is inside Z_b. But since Z_a is irreducible so $U_i \cap Z_a$ must be dense in Z_a, a contradiction.

[29] Actually we didn't needed the reduction to Y being dense in X.

3. Let $r = \dim X = \dim Y$. Suppose $Y \subsetneq X$. Let $Z_0 \supsetneq Z_1 \supsetneq \cdots \supsetneq Z_r$ be a maximal finite chain of Y. Then the chain $X \supsetneq Z_0 \supsetneq Z_1 \supsetneq \cdots \supsetneq Z_r$ is a finite chain in X as Y is closed. Thus $\dim X \geq 1 + r$, therefore $r = \dim Y \geq r + 1$, a contradiction.

$\square$

The following technical lemma was employed in proving the statement 2 of above, but is good to keep in handy.

Lemma 1.7.1.2. *Let X be a topological space and $Z_\bullet \in FC(X)$ a finite chain (in the terminology of Proposition 1.7.1.1) of length $l(Z_\bullet) = r$. If $U \subseteq X$ is an open set such that $U \cap Z_r \neq \emptyset$, then $U \cap Z_\bullet$ is a finite chain of length r in U.* $\square$

The following result gives a connection between all closed irreducible containing a given point and prime ideals of the local ring at that point.

Proposition 1.7.1.3. *Let X be a scheme and $x \in X$ be a point. We obtain an order reversing bijection*

$$\{\textit{Closed irreducibles } Y \textit{ of } X \textit{ containing } x\} \cong \operatorname{Spec}(\mathcal{O}_{X,x}).$$

Proof. Denote the collection of all closed irreducibles of X containing x as I. Let $U = \operatorname{Spec}(A)$ be an affine open containing $x \in X$ so that $\mathcal{O}_{X,x} \cong A_x$. Consequently, we wish to show a bijection $I \cong \operatorname{Spec}(A_x)$, which is further equivalent to showing that I is bijective to all prime ideals of A contained in x. As all prime ideals of A contained in x is further bijective to all closed irreducible of U containing x by Lemma 1.2.1.1, we thus reduce to showing existence of a bijection between I and closed irreducibles of U containing x, denoted J.

Consider the following function

$$\begin{aligned} \varphi : I &\longrightarrow J \\ Y &\longmapsto Y \cap U. \end{aligned}$$

Indeed, this map is well-defined as for any $Y \in I$, $\varphi(Y) = Y \cap U$ is first irreducible as any open subset of an irreducible set is irreducible. Further, it is closed in U as Y is closed. In order to show injectivity, we need only recall that any open subset of an irreducible set is dense. Finally, for surjectivity, take any $Z \in J$ so that Z is a closed irreducible in U containing x. Now let Y to be the closure of Z in X. We thus need only show that Y is irreducible in X. That follows immediately from the fact that closure of irreducible is again an irreducible, which in turn follows immediately from a simple observation on open subsets of the closure. $\square$

One then observes the following general result which will be used heavily in the future.

Lemma 1.7.1.4. *Let X be a scheme and Y be an irreducible closed subscheme of X with $\eta \in Y$ being its generic point. Then,*

$$\operatorname{codim}(Y, X) = \dim \mathcal{O}_{X,\eta}.$$

Proof. This is immediate from Proposition 1.7.1.3 as Y is the smallest closed irreducible containing η. $\square$

1.7.2 Dimension of finite type k-schemes

In this section, we prove various results surrounding the relationship between dimension of a given integral finite type k-scheme as a topological space and Krull dimension of various local rings.

Theorem 1.7.2.1. [30] *Let k be a field and X be a finite type integral k-scheme.*

1. *If $U, V \subseteq X$ are two open affines which are spectra of finite type k-domains, then* $\dim U = \dim V$.
2. *If $\{U_i\}_{i=1}^n$ is any finite open affine covering by spectra of finite type k-domains, then* $\dim U_i = \dim X$ *for all $i = 1, \dots, n$.*
3. *If $p \in X$ is a closed point, then* $\dim X = \dim \mathcal{O}_{X,p}$.
4. *Let $K(X)$ be the function field of X. Then* $\dim X = \operatorname{trdeg} K(X)/k$.
5. *If Y is a closed subset of X, then* $\operatorname{codim}(Y, X) = \inf_{p\in Y} \dim \mathcal{O}_{X,p}$.
6. *If Y is a closed subset of X, then* $\dim Y + \operatorname{codim}(Y, X) = \dim X$.

Proof. The main tools are the Theorems 16.8.2.1 and 16.8.2.2.

1. Observe that since X is irreducible, therefore U and V are dense open subsets of X, so $U \cap V \neq \emptyset$. Consequently, it will suffice to show that any dense affine open subset $W \subseteq U$ has same dimension as U. Indeed, U is spectra of finite type k-domain, so it is a separated finite type integral affine scheme, that is, an abstract affine variety. Consequently, by Proposition I.1.10 of cite[Hartshorne], $\dim W = \dim \overline{W} = \dim U$.
2. Follows from Proposition 1.7.1.1, 2 and statement 1.
3. As X is finite type, it admits a finite open affine covering by spectra of finite type k-domains. Let $U = \operatorname{Spec}(A)$ be one such open affine such that $p \in U$. Consequently, $p = \mathfrak{m} \in \operatorname{Spec}(A)$ represents a maximal ideal of R (Lemma 1.2.1.3). Thus, $\mathcal{O}_{X,p} \cong A_\mathfrak{m}$ and so $\dim \mathcal{O}_{X,p} = \dim A_\mathfrak{m}$. Note that A is a finite type k-algebra which is an integral domain. It thus follows by Theorem 16.8.2.2 that we have $\operatorname{ht}\mathfrak{m} + \dim A/\mathfrak{m} = \dim A$ and since $\dim A/\mathfrak{m} = 0$, therefore $\operatorname{ht}\mathfrak{m} = \dim A$. Further, since $\dim A_\mathfrak{m} = \operatorname{ht}\mathfrak{m}$, therefore we have $\dim \mathcal{O}_{X,p} = \dim A_\mathfrak{m} = \dim A = \dim U$. By statement 2, $\dim U = \dim X$ and the result follows.
4. Function field is defined to be the local ring at the generic point of X, say $\eta \in X$ (Remark 1.4.2.5). Let $\eta \in \operatorname{Spec}(A)$ where $\operatorname{Spec}(A)$ is a member of an open affine cover of X by spectra of finite type k-domains. Observe that $\operatorname{Spec}(A)$ has η as its generic point as well. Consequently, $\dim \operatorname{Spec}(A) = \dim A = \operatorname{trdeg} K(A)/k$ and since $K(A) = \mathcal{O}_{\operatorname{Spec}(A),\eta} \cong \mathcal{O}_{X,\eta} = K(X)$, therefore $\dim \operatorname{Spec}(A) = \operatorname{trdeg} K(X)/k$. By statement 2, $\dim \operatorname{Spec}(A) = \dim X$ and the result follows.
5. First observe that for any closed irreducible $Z \subseteq X$, we have $\operatorname{codim}(Z, X) \leq \dim X$. By statement 3, therefore, we have $\inf_{p\in Y} \dim \mathcal{O}_{X,p} = \inf_{p\in Y \text{ non-closed}} \mathcal{O}_{X,p}$. We will now show that for any closed irreducible subset $Z \subseteq X$ with $\eta \in Z$ its generic paint (schemes are *sober*[31]), we have $\dim \mathcal{O}_{X,\eta} = \operatorname{codim}(Z, X)$. By taking infimum, the result would then follow, so it would suffice to show the above claim.
   Let $\{\operatorname{Spec}(A_\alpha)\}$ be a finite open affine cover of X where A_α is a finite type k-domain. Observe that if $Z \cap \operatorname{Spec}(A_\alpha) \neq \emptyset$, then $\eta \in \operatorname{Spec}(A_\alpha)$. Now, $\eta \in \operatorname{Spec}(A_\alpha)$ is a point whose closure in $\operatorname{Spec}(A_\alpha)$ is $Z \cap \operatorname{Spec}(A_\alpha)$ so $Z \cap \operatorname{Spec}(A_\alpha)$ is a closed irreducible subspace of $\operatorname{Spec}(A_\alpha)$ whose generic point is η and thus $Z \cap \operatorname{Spec}(A_\alpha) \cong \operatorname{Spec}(A_\alpha/\eta)$, where we treat

[30] Exercise II.3.20 of Hartshorne.

[31] a space where all closed irreducibles have a unique generic point.

$\eta \lneq A_\alpha$ as a prime ideal of A_α. Consequently, $\dim \mathcal{O}_{X,\eta} = \dim \mathcal{O}_{\mathrm{Spec}(A_\alpha),\eta} = \dim(A_\alpha)_\eta = \mathrm{ht}\,\eta$. Since A_α is a finite type k-domain, therefore by Theorem 16.8.2.2, we obtain that $\mathrm{ht}\,\eta + \dim A_\alpha/\eta = \dim A_\alpha$, which thus yields $\mathrm{ht}\,\eta = \dim X - \dim A_\alpha/\eta$ by statement 2. It thus suffices to show that for some index α we get $\dim A_\alpha/\eta = \dim Z$ as then we would obtain $\dim \mathcal{O}_{X,\eta} = \dim X - \dim Z = \mathrm{codim}\,(Z, X)$.
Indeed, since $\{\mathrm{Spec}\,(A_\alpha/\eta)\}$ forms a finite open affine cover of Z, therefore by Proposition 1.7.1.1, 2 we get such an index α.

6. Observe that since $\mathrm{codim}\,(Y, X) < \infty$, therefore there exists a maximal closed irreducible $Z \subseteq Y$ such that $\mathrm{codim}\,(Y, X) = \mathrm{codim}\,(Z, X)$. Consequently, we have a finite chain of X, say $Z_\bullet$, ending at Z such that $l(Z_\bullet) = \mathrm{codim}\,(Y, X)$.
Let $U = \mathrm{Spec}\,(A)$ be an open affine where A is a finite type k-domain such that $U \cap Z \neq \emptyset$. Further, $\dim U \cap Y = \dim Y$. Consequently, by Lemma 1.7.1.2, we have $\mathrm{codim}\,(Y, X) = \mathrm{codim}\,(Z \cap U, U)$. Since $U \cap Y$ is a closed subscheme of U, therefore we may write $U \cap Y = \mathrm{Spec}\,(A/I)$ for an ideal $I \leq A$. Consequently, $\mathrm{codim}\,(Y, X) = \mathrm{codim}\,(\mathrm{Spec}\,(A/I), \mathrm{Spec}\,(A))$.
It is immediate from first definitions that

$$\begin{aligned}\mathrm{codim}\,(\mathrm{Spec}\,(A/I), \mathrm{Spec}\,(A)) &= \inf_{\mathfrak{p} \supseteq I} \mathrm{codim}\,(\mathrm{Spec}\,(A/\mathfrak{p}), \mathrm{Spec}\,(A)) \\ &= \inf_{\mathfrak{p} \supseteq I} \mathrm{ht}\,\mathfrak{p}.\end{aligned}$$

Now by Theorem 16.8.2.2 and above, we further obtain that

$$\begin{aligned}\mathrm{codim}\,(\mathrm{Spec}\,(A/I), \mathrm{Spec}\,(A)) &= \inf_{\mathfrak{p} \supseteq I} (\dim A - \dim A/\mathfrak{p}) \\ &= \dim A - \sup_{\mathfrak{p} \supseteq I} \dim A/\mathfrak{p} \\ &= \dim X - \dim U \cap Y \\ &= \dim X - \dim Y\end{aligned}$$

where $\dim A = \dim X$ because of statement 2.

□

Corollary 1.7.2.2. *Let X be a variety over a field k. Then* $\dim X < \infty$.

Proof. As X is a finite type integral k-scheme, therefore by Theorem 1.7.2.1, 3, $\dim X = \dim \mathcal{O}_{X,p}$ for any closed point $p \in X$. Fixing a closed point $p \in X$ in an open affine $\mathrm{Spec}\,(A)$ of X, we first deduce that A is a finite type k-domain. Let p be the maximal ideal $\mathfrak{m} \leq A$. Hence, $\mathcal{O}_{X,p} \cong A_\mathfrak{m}$. Hence $\dim \mathcal{O}_{X,p} = \mathrm{ht}\,\mathfrak{m}$ in ring A. By Theorem 16.8.2.2, $\mathrm{ht}\,\mathfrak{m} = \dim A - \dim A/\mathfrak{m} = \dim A$ as $A/\mathfrak{m}$ is a field. As A is a finite type k-domain, therefore its dimension is finite, as required. □

Corollary 1.7.2.3. *Let k be a field $\mathbb{A}^n_k$ be the affine n-space over k. Let H be a hyperplane in $\mathbb{A}^n_k$, that is $H = V(f)$ where $f \in k[x_1, \dots, x_n]$ is a linear polynomial. Then* $\dim H = n - 1$.

Proof. As $H = \mathrm{Spec}\,(A/\langle f \rangle)$, and $\langle f \rangle$ is a prime ideal as any linear polynomial is irreducible in $k[x_1, \dots, x_n]$ and since the latter is a UFD, therefore f prime as well. By Theorem 16.8.2.2, we have $\dim H = \dim k[x_1, \dots, x_n] - \mathrm{ht}\,\langle f \rangle = n - 1$, as required. □

An important observation about varieties is as follows.

Remark 1.7.2.4 (Dimension of scheme theoretic image)**.** Let $f : X \to Y$ be a dominant morphism of varieties. Consequently, there is an induced map on function fields as generic point maps to generic point by dominance. Let $f^\flat : K(Y) \to K(X)$ be this map. As $f^\flat$ is an injection, we thus have the inequality

$$\text{trdeg } K(Y)/k \leq \text{trdeg } K(X)/k.$$

Consequently, we deduce that

$$\dim Y \leq \dim X.$$

1.7.3 Dimension of fibers

In this section, we discuss the question of how the dimension of fibers of a morphism varies. We'll see that certain nice geometric situations are encoded in the maps for which the dimension of fibers is not too erratic.

1.7.4 Irreducible components

In this short section, we describe the decomposition of a closed subscheme of a locally noetherian scheme into finitely many irreducible components. Let us begin by the following basic observation.

Remark 1.7.4.1 (Integral closed subschemes by points)**.** Let X be a scheme and $x \in X$ be a point and $Z = \overline{\{x\}}$ to be the closed irreducible subspace of X. Giving Z the reduced induced subscheme structure on Z, thus making $Z \hookrightarrow X$ an integral closed subscheme of X.

The following proposition shows that a minimal prime in an affine open subset gives an irreducible component of the whole scheme!

Proposition 1.7.4.2. *Let X be a scheme, $U = \text{Spec}\,(A) \subseteq X$ an open affine and $\mathfrak{p} \in U$. Denote $Y = \overline{\{\mathfrak{p}\}}$ to be the closed irreducible subspace of X. Then the following are equivalent:*

1. *$\mathfrak{p}$ is a minimal prime of A.*
2. *Y is an irreducible component of X.*

Proof. (L $\Rightarrow$ R) Suppose $Y \subsetneq Z$ is a closed irreducible set of X properly containing Y. Denote $\eta \in Z$ to be its unique generic point. As $U \cap Z$ is a non-empty open subset of Z, therefore it is dense in Z. Consequently, $\eta = \mathfrak{q} \in U$. If $Y \cap U \subsetneq Z \cap U$, then $Z \cap U = V(\mathfrak{q}) \subseteq U$ properly contains $Y \cap U = V(\mathfrak{p})$. By Lemma 1.2.0.1, we get $\sqrt{\mathfrak{q}} \subsetneq \sqrt{\mathfrak{p}}$, so that $\mathfrak{q} \subsetneq \mathfrak{p}$, contradicting the minimality of $\mathfrak{p}$.

(R $\Rightarrow$ L) Let Y be a maximal closed irreducible set. If $\mathfrak{q} \subsetneq \mathfrak{p}$, then $V(\mathfrak{q}) \supseteq V(\mathfrak{p})$ in U. Denote $Z = \overline{\{\mathfrak{q}\}}$. If $V(\mathfrak{p}) = V(\mathfrak{q})$, then $Z \cap U = Y \cap U$ and hence $Y = Z$. Hence we may assume $Y \cap U \subsetneq Z \cap U$. As $\mathfrak{p} \in Y \cap U \subseteq Z \cap U$, therefore $\overline{\{\mathfrak{p}\}} = Y \subseteq Z$. But since $Y \cap U \neq Z \cap U$, therefore $Y \subsetneq Z$, a contradiction to maximality of Y. □

Construction 1.7.4.3 (Irreducible & embedded components of a subscheme)**.** Let X be a locally noetherian scheme and $Y \hookrightarrow X$ be a closed subscheme. Cover X by open affines $\{U_\alpha\}_{\alpha \in I}$ where

$U_\alpha = \operatorname{Spec}(A_\alpha)$, A_α is a noetherian ring. Fixing $\alpha \in I$, we see that $Y \cap U_\alpha \subseteq U_\alpha$ is a closed subscheme. It follows that $Y \cap U_\alpha = \operatorname{Spec}(A_\alpha/\mathfrak{a}_\alpha) = V(\mathfrak{a}_\alpha)$ for some ideal $\mathfrak{a}_\alpha \leq A_\alpha$.

By primary decomposition theorem (Theorem 16.1.7.7 or Corollary 16.4.0.10), it follows that there are distinct primes $\mathfrak{p}_{\alpha,i}$, $i = 1, \dots, N_\alpha$ of A_α such that

$$\mathfrak{a}_\alpha = \bigcap_{i=1,\dots,N_\alpha} \mathfrak{q}_{\alpha,i}.$$

where $\mathfrak{q}_{\alpha,i}$ is a $\mathfrak{p}_{\alpha,i}$-primary ideal. Moreover, by Theorem 16.1.7.7, 2, we also get

$$V(\mathfrak{a}_\alpha) = \bigcup_{i=1,\dots,N_\alpha} V(\mathfrak{q}_{\alpha,i}) = \bigcup_{i=1,\dots,N_\alpha} V(\mathfrak{p}_{\alpha,i}).$$

We thus get points $S = \{\mathfrak{p}_{\alpha,i}\}_{\alpha,i}$ of X which we relabel by $S = \{\mathfrak{p}_\beta\}_{\beta \in J}$ and remove repeated points. By above remark, we thus get closed integral subschemes $\{Y_\beta\}_{\beta \in J}$ of X, which we call *components*. Moreover, as each $\mathfrak{p}_\beta \in Y$ and Y is closed, therefore each Y_β is a closed integral subscheme of Y itself. The minimal/isolated primes amongst S correspond to *irreducible components* of Y via Proposition 1.7.4.2. The others correspond to *embedded components* of Y.

If X is moreover noetherian, then indexing set I is finite and thus Y has finitely many irreducible and embedded components.

1.8 Projective schemes

The most important type of examples that we will encounter in our study of algebraic geometry are subvarieties of projective space $\mathbb{P}^n_k$. Indeed, this is a construction which is fundamental because of the many nice properties enjoyed by realizing familiar constructions in it. One of them being this classical observation that any two straight lines are bound to intersect at atleast one point in the projective space. We shall see more equally nice results, not to mention the quadrics with which we wish to spend some considerable time as the main motivating example for us (Example 1.5.1.3) is itself realized as a quadric in projective space.

We recall that the notion of projective varieties, whose generalization we shall embark now on, has been covered in Section 1.5.

We first begin by defining the space $\mathrm{Proj}(S)$ of a graded ring $S = \bigoplus_{d\geq 0} S_d$.

Definition 1.8.0.1. (**Projective spectrum of a graded ring**) Let $S = \bigoplus_{d\geq 0} S_d$ be a graded ring and let $S_+ = \bigoplus_{d>0} S_d$ be the ideal generated by non-zero degree elements. Denote

$$\mathrm{Proj}(S) := \{\mathfrak{p} \lneq S \mid \mathfrak{p} \text{ is homogeneous prime ideal \& } \mathfrak{p} \not\supseteq S_+\}.$$

The set $\mathrm{Proj}(S)$ is called the projective spectrum of the graded ring S.

Note that the latter condition is motivated by Remark 1.5.3.13. This is also used in a technical manner to show existence of a nice basis over $\mathrm{Proj}(S)$ in Lemma 1.8.1.3 and in other proofs as well. We now show that there is a natural topology over $\mathrm{Proj}(S)$, akin to the affine case.

Lemma 1.8.0.2. *Let S be a graded ring and denote for a homogeneous ideal $\mathfrak{a} \leq S$, the following subset of* $\mathrm{Proj}(S)$*:*

$$V(\mathfrak{a}) = \{\mathfrak{p} \in \mathrm{Proj}(S) \mid \mathfrak{p} \supseteq \mathfrak{a}\}.$$

Then, for any homogeneous ideals $\mathfrak{a}, \mathfrak{b}, \mathfrak{a}_i$ of S, we obtain

1. $V(\mathfrak{a}) \cup V(\mathfrak{b}) = V(\mathfrak{a}\mathfrak{b})$,
2. $\bigcap_i V(\mathfrak{a}_i) = V\left(\sum_i \mathfrak{a}_i\right)$.

Proof. Same as Lemma 1.2.0.1. □

We thus obtain a topological space $\mathrm{Proj}(S)$ where a set is closed if and only if it is of the form $V(\mathfrak{a})$ for a homogeneous ideal $\mathfrak{a} \leq S$. This is called the Zariski topology over the $\mathrm{Proj}(S)$.

We now give some more topological properties of $\mathrm{Proj}(S)$.

1.8.1 Topological properties of Proj(S)

The first obvious question is how does the inclusion $\mathrm{Proj}(S) \hookrightarrow \mathrm{Spec}\,(S)$ looks topologically?

Lemma 1.8.1.1. *Let S be a graded ring. The topology of* $\mathrm{Proj}(S)$ *is obtained by subspace topology of* $\mathrm{Spec}\,(S)$*. Thus, there is a continuous inclusion*

$$\mathrm{Proj}(S) \hookrightarrow \mathrm{Spec}\,(S).$$

Proof. Immediate from definitions. □

We further note that for a graded ring S, the degree zero elements S_0 form a subring of S by the virtue of the fact that $S_d \cdot S_e \subseteq S_{d+e}$. Thus, we obtain a continuous map as the following shows.

Lemma 1.8.1.2. *Let S be a graded ring. Then the following is a continuous map*

$$\varphi : \mathrm{Proj}(S) \longrightarrow \mathrm{Spec}\,(S_0)$$
$$\mathfrak{p} \longmapsto \mathfrak{p} \cap S_0.$$

Proof. Pick any ideal $\mathfrak{a} \leq S_0$ and notice that it is already homogeneous in S. Consequently, $\varphi^{-1}(V(\mathfrak{a})_a) = V(\mathfrak{a})_h$ where $V(\mathfrak{a})_a \subseteq \mathrm{Spec}\,(S_0)$ and $V(\mathfrak{a})_h \subseteq \mathrm{Proj}(S)$. □

We now find a collection of open sets which forms a basis for $\mathrm{Proj}(S)$. This is akin to Lemma 1.2.1.4.

Lemma 1.8.1.3. *Let S be a graded ring and $f, g \in S_d$ for some $d > 0$ be homogeneous elements. Denote*

$$D_+(f) := \{\mathfrak{p} \in \mathrm{Proj}(S) \mid f \notin \mathfrak{p}\}.$$

Then,

1. *$D_+(f)$ is an open subset of* $\mathrm{Proj}(S)$*,*
2. *$D_+(f) \cap D_+(g) = D_+(fg)$,*
3. *$\{D_+(f)\}_{f \in S_d, d>0}$ forms a basis of* $\mathrm{Proj}(S)$*.*

Proof. 1. Since $D_+(f) = \mathrm{Proj}(S) \setminus V(f)$, thus $D_+(f)$ is open.
2. Straightforward.
3. Since for any $\mathfrak{p} \in \mathrm{Proj}(S)$, there exists $f \in S_d$ for some $d > 0$ such that $f \notin \mathfrak{p}$ as $\mathfrak{p}$ does not contain all of S_+, thus $\bigcup_{f \in S_d, d>0} D_+(f) = \mathrm{Proj}(S)$. The rest follows by 2. □

Remark 1.8.1.4. As tempting as it might be to think, but not all projective schemes are quasi-compact. An example is given by the graded ring $S = \mathbb{Z}[x_1, x_2, \dots]$, the polynomial ring over $\mathbb{Z}$ with countably infinitely many indeterminates. Then one observes that $\mathrm{Proj}(S) = \bigcup_{n=1}^{\infty} D_+(x_n)$. Moreover, as for any $\mathfrak{p} \in \mathrm{Proj}(S)$ can not contain S_+, therefore $\mathfrak{p}$ necessarily has to not contain some x_i, otherwise it contains S_+. Consequently, we cannot form a finite subcover of the above cover, showing that $\mathrm{Proj}(S)$ is not quasi-compact.

However, the following lemma might be helpful in checking when a projective scheme has a finite cover by basic open sets.

Lemma 1.8.1.5. *Let S be a graded ring and consider $X = \mathrm{Proj}(S)$. Let $f = f_0 + \cdots + f_n$ be a decomposition of $f \in S$ into homogeneous elements $f_d \in S_d$. Then,*

$$D(f) \cap X = (D(f_0) \cap X) \cup \bigcup_{d=1}^{n} D_+(f_d)$$

where we view $X \subseteq \mathrm{Spec}\,(S)$ and $D(f), D(f_0) \subseteq \mathrm{Spec}\,(S)$.

Proof. This is a rather straightforward proof. To show ($\subseteq$), consider a point $\mathfrak{p} \in D(f) \cap X$ so that $f \notin \mathfrak{p}$. It follows from $f = f_0 + \cdots + f_n$ that for some $d = 0, \dots, n$, $f_d \notin \mathfrak{p}$, which is in turn equivalent to stating that $\mathfrak{p} \in D_+(f_i)$ if $i \geq 1$ or $\mathfrak{p} \in D(f_0) \cap X$ if $d = 0$.

Conversely, pick $\mathfrak{p} \in (D(f_0) \cap X) \cup \bigcup_{d=1}^{n} D_+(f_d)$. We obtain that for some $d = 0, \dots, n$, $f_d \notin \mathfrak{p}$. It follows from $\mathfrak{p} = \bigoplus_{i \geq 0} \mathfrak{p} \cap S_i$ that if $f \in \mathfrak{p}$, then we get by uniqueness of representatives of the direct sum that $f_d \in \mathfrak{p}$, a contradiction. □

1.8.2 The structure sheaf $\mathcal{O}_{\text{Proj}(S)}$ and projective schemes

We have studied some basic properties of the topological space $\text{Proj}(S)$ so far, we now construct a structure sheaf over it and make it, first, into a locally ringed space and, second, into a scheme. We first define the structure sheaf of projective spectrum, in which there is nothing new in comparison to projective varieties (see Definition 1.5.3.1).

Definition 1.8.2.1. **(The structure sheaf $\mathcal{O}_{\text{Proj}(S)}$)** Let S be a graded ring. Let $U \subseteq \text{Proj}(S)$ be an open set of the projective spectrum of S. Define the following set

$$\mathcal{O}_{\text{Proj}(S)}(U) := \left\{ s : U \to \coprod_{\mathfrak{p} \in U} S_{(\mathfrak{p})} \mid \forall \mathfrak{p} \in U, s(\mathfrak{p}) \in S_{(\mathfrak{p})} \ \&\ \exists \text{ open } \mathfrak{p} \in V \subseteq U \ \&\ f, g \in S_d,\ d \geq 0 \text{ s.t. } \forall \mathfrak{q} \in V,\ g \notin \mathfrak{q} \ \&\ s(\mathfrak{q}) = f/g \right\}.$$

From the fact that its elements are functions locally defined, one immediately obtains that $\mathcal{O}_{\text{Proj}(S)}$ is a sheaf with obvious restriction maps. By appropriate restrictions on the domain, one further sees that under pointwise addition and multiplication, $\mathcal{O}_X(U)$ forms a commutative ring with 1.

Let us now show that $\text{Proj}(S)$ is a scheme over $\text{Spec}(S_0)$ in a natural manner.

Lemma 1.8.2.2. *Let S be a graded ring. Then $\text{Proj}(S)$ is a scheme over $\text{Spec}(S_0)$.*

Proof. We need only define a map $\text{Proj}(S) \to \text{Spec}(S_0)$. By Theorem 1.3.0.5, we need only construct a homomorphism $S_0 \to \Gamma(\text{Proj}(S), \mathcal{O}_{\text{Proj}(S)})$. This is straightforward, as we can interpret each $a \in S_0$ as a homogeneous regular function $s : \text{Proj}(S) \to \amalg_{\mathfrak{p} \in \text{Proj}(S)} S_{(\mathfrak{p})}$ mapping as $\mathfrak{p} \mapsto a/1$. □

Thus, $(\text{Proj}(S), \mathcal{O}_{\text{Proj}(S)})$ is a ringed space. We now see that the stalk of this sheaf is isomorphic to the homogeneous localization. This will thus show that $(\text{Proj}(S), \mathcal{O}_{\text{Proj}(S)})$ is a locally ringed space (Lemma 16.2.1.3).

Lemma 1.8.2.3. *Let S be a graded ring and consider the ringed space $(\text{Proj}(S), \mathcal{O}_{\text{Proj}(S)})$. For each $\mathfrak{p} \in \text{Proj}(S)$, we have*

$$\mathcal{O}_{\text{Proj}(S), \mathfrak{p}} \cong S_{(\mathfrak{p})}.$$

Proof. Consider the following map

$$\begin{aligned} \varphi : \mathcal{O}_{\text{Proj}(S), \mathfrak{p}} &\longrightarrow S_{(\mathfrak{p})} \\ (U, s)_{\mathfrak{p}} &\longmapsto s(\mathfrak{p}). \end{aligned}$$

It is straightforward to see that φ is a well-defined ring homomorphism. To see injectivity, suppose $(U, s)_{\mathfrak{p}} \mapsto 0$. Thus $s(\mathfrak{p}) = 0$. Consequently, for some open $V \subseteq U$ containing $\mathfrak{p}$ where s is given by f/g for $f, g \in S_d$, $d \geq 0$, we obtain $s(\mathfrak{q}) = f/g = 0$ for all $\mathfrak{q} \in V$. Thus $s = 0$ on V and hence $(U, s)_{\mathfrak{p}} = (V, \rho_{U,V}(s))_{\mathfrak{p}} = 0$. To see surjectivity, pick any $f/g \in S_{(\mathfrak{p})}$. Observe that $g \notin \mathfrak{p}$. Thus consider $(D_+(g), f/g)_{\mathfrak{p}} \in \mathcal{O}_{\text{Proj}(S), \mathfrak{p}}$. □

We now show that the locally ringed space $(\text{Proj}(S), \mathcal{O}_{\text{Proj}(S)})$ is a scheme. For this purpose we would need to show that $\text{Proj}(S)$ is covered by affine opens. Indeed, we have the following lemma.

Lemma 1.8.2.4. *Let S be a graded ring and consider the locally ringed space $(\mathrm{Proj}(S), \mathcal{O}_{\mathrm{Proj}(S)})$. For each $f \in S_d$, $d > 0$, we have the following isomorphism of locally ringed spaces*

$$\left(D_+(f), \mathcal{O}_{\mathrm{Proj}(S)|D_+(f)}\right) \cong \left(\mathrm{Spec}\left(S_{(f)}\right), \mathcal{O}_{\mathrm{Spec}\left(S_{(f)}\right)}\right).$$

Proof. Consider the map

$$\begin{aligned} \varphi : D_+(f) &\longrightarrow \mathrm{Spec}\left(S_{(f)}\right) \\ \mathfrak{p} &\longmapsto (\mathfrak{p} \cdot S_f)_0. \end{aligned}$$

By Lemma 16.2.1.8, it follows that φ is a bijection. To show that φ is an isomorphism it is sufficient to show that φ is a closed map. This is immediate as $\mathfrak{p} \supseteq \mathfrak{a}$ in $D_+(f)$ if and only if $(\mathfrak{p}\cdot S_f)_0 \supseteq (\mathfrak{a}\cdot S_f)_0$ in $\mathrm{Spec}\left(S_{(f)}\right)$.

We now wish to show isomorphism of corresponding sheaves. For this, we construct a map

$$\varphi^\flat : \mathcal{O}_{\mathrm{Spec}\left(S_{(f)}\right)} \longrightarrow \varphi_* \mathcal{O}_{D_+(f)}$$

and show that this is an isomorphism. Indeed, we first observe a canonical isomorphism on stalks

$$\mathcal{O}_{D_+(f),\mathfrak{p}} \cong S_{(\mathfrak{p})} \xrightarrow{\eta_\mathfrak{p}} \left(S_{(f)}\right)_{(\mathfrak{p}\cdot S_f)_0} \cong \mathcal{O}_{\mathrm{Spec}\left(S_{(f)}\right),\varphi(\mathfrak{p})}.$$

Then one can construct the above isomorphism $\varphi^\flat$ by observing the following square for sections of the relevant sheaves over open $U \subseteq \mathrm{Spec}\left(S_{(f)}\right)$ and the corresponding $\varphi^{-1}(U) \subseteq D_+(f)$:

$$\begin{tikzcd} U \arrow[d, "s"'] & \varphi^{-1}(U) \arrow[l, "\varphi"', "\cong"] \arrow[d, "t"] \\ \coprod_{\mathfrak{p}\in\varphi^{-1}(U)}\left(\left(S_{(f)}\right)_{\varphi(\mathfrak{p})}\right) & \coprod_{\mathfrak{p}\in\varphi^{-1}(U)} S_\mathfrak{p} \arrow[l, "\amalg_\mathfrak{p}\eta_\mathfrak{p}"', "\cong"] \end{tikzcd},$$

where $s \in \mathcal{O}_{\mathrm{Spec}\left(S_{(f)}\right)}(U)$ and $t \in \mathcal{O}_{D_+(f)}(\varphi^{-1}(U))$ its image under $\varphi^\flat$ (which is defined by the above square). One can indeed check that $\varphi^\flat$ as defined is natural w.r.t restrictions. □

Remark 1.8.2.5. Thus, for a graded ring S, we obtain a scheme $(\mathrm{Proj}(S), \mathcal{O}_{\mathrm{Proj}(S)})$, which is called the projective scheme associated to a graded ring S.

We now give some more properties of $\mathrm{Proj}(S)$.

Proposition 1.8.2.6. [32] *Let S be a graded ring. Then,*

1. $\mathrm{Proj}(S) = \emptyset$ *if and only if* $\forall s \in S_+$, *s is a nilpotent element of S.*
2. *Let $\varphi : S \to T$ be a graded map of graded rings. Then $U = \{\mathfrak{q} \in \mathrm{Proj}(T) \mid \mathfrak{q} \not\supseteq \varphi(S_+)\}$ is an open set and the natural map*

$$\begin{aligned} f : U &\longrightarrow \mathrm{Proj}(S) \\ \mathfrak{q} &\longmapsto \varphi^{-1}(\mathfrak{q}) \end{aligned}$$

defines a map of schemes.

[32] Exercise II.2.14 of Hartshorne.

3. *Let $\varphi : S \to T$ be a graded map of graded rings for which there exists $d_0 \in \mathbb{N}$ such that $\varphi_d : S_d \to T_d$ is an isomorphism for all $d \geq d_0$. Then, $U = \mathrm{Proj}(T)$ and $f : \mathrm{Proj}(T) \to \mathrm{Proj}(S)$ is an isomorphism.*

Proof. 1. The R $\implies$ L is immediate. Otherwise take an element $s \in S_d$. By Lemmas 1.8.1.3, 3 and 1.8.2.4, we obtain that $\mathrm{Spec}\left(S_{(s)}\right) = \emptyset$. Consequently, any prime ideal of $\mathrm{Spec}\left(S_s\right)$ has no zero degree terms, which can be seen to be not true. Consequently, $D(s) = \mathrm{Spec}\left(S_s\right) = \emptyset$. It follows from Lemma 1.2.2.7 that s is nilpotent.

2. The fact that U is open depends on φ being graded, i.e. $\varphi(S_d) \subseteq T_d$ for all $d \geq 0$. The continuity of f follows from the same observation. The map on sheaves is given by extending the natural map on stalks $\varphi_{(\mathfrak{q})} : S_{(\varphi^{-1}(\mathfrak{q}))} \to T_{(\mathfrak{q})}$, whose well-definedness, again, uses the fact that φ is graded.

3. The main trick here is to observe that if $s \in S_d$ for $d < d_0$, then raising some high enough power of s will make $s^n \in S_e$ where $\deg s^n \geq d_0$. For showing isomorphism on stalks $\varphi_{(\mathfrak{q})} : S_{(\varphi^{-1}(\mathfrak{q}))} \to T_{(\mathfrak{q})}$, it comes down to observing the following: let $s/t \in S_{(\varphi^{-1}(\mathfrak{q}))}$, then $s/t = st^n/t^{n+1}$ for any $n \in \mathbb{N}$. Then use the trick above. $\square$

Remark 1.8.2.7. The above Proposition 1.8.2.6 shows that the mapping $S \mapsto \mathrm{Proj}(S)$ is NOT functorial! However the statement 3. might give some hint how to fix this.

Next, we understand all closed subschemes of $\mathrm{Proj}(S)$ in the following two results (Corollary 1.8.2.10)

Proposition 1.8.2.8. *Let S, T be a graded rings.*

1. *If $\varphi : S \to T$ is a surjective graded map, then the open set $U = \mathrm{Proj}(T)$ and $f : \mathrm{Proj}(T) \to \mathrm{Proj}(S)$ is a closed immersion (see Proposition 1.8.2.6, 2).*
2. *Let $I \leq S$ be a homogeneous ideal and consider the ideal $I' = \bigoplus_{d \geq d_0} I_d$. Then, I, I' defines the same closed subscheme of $\mathrm{Proj}(S)$.*
3. *Let $I \leq S$ be a homogeneous ideal and let $\pi : S \to S/I$ be the natural projection. Then the closed subscheme $f : \mathrm{Proj}(S/I) \to \mathrm{Proj}(S)$ (as in 1.) has the ideal sheaf given by $\widetilde{I} \leq \mathcal{O}_{\mathrm{Proj}(S)}$.*

Proof. 1. $U = \mathrm{Proj}(T)$ because $\varphi(S_+) = T_+$. The fact that f is a topological immersion follows from the observations that $f(\mathrm{Proj}(T)) = V(\mathrm{Ker}\,(\varphi))$ where $\mathrm{Ker}\,(\varphi)$ is homogeneous and that for any ideal $\mathfrak{q} \leq T$, it follows from surjectivity that $\varphi(\varphi^{-1}(\mathfrak{q})) = \mathfrak{q}$. To show surjectivity of sheaves, it reduces to showing surjectivity of localization maps $S_{\varphi^{-1}(\mathfrak{q})} \xrightarrow{\varphi_{(\mathfrak{q})}} T_{(\mathfrak{q})}$, which is immediate from surjectivity of φ.
2. We wish to show an isomorphism as in the following commutative diagram:

$$\begin{tikzcd} \mathrm{Proj}(S/I) \arrow[rr, "\cong"] \arrow[dr] & & \mathrm{Proj}(S/I') \arrow[dl] \\ & \mathrm{Proj}(S) & \end{tikzcd}.$$

Now since $(S/I)_d = S_d/I_d$ for all $d \geq 0$, therefore we have an isomorphism $\varphi : (S/I)_d \to (S/I')_d$ given by $s_d + I_d \mapsto s_d + I'_d$. The result follows from Proposition 1.8.2.6, 3.
3. We wish to show that $\mathrm{Ker}\left(f^\flat : \mathcal{O}_{\mathrm{Proj}(S)} \to f_*\mathcal{O}_{\mathrm{Proj}(S/I)}\right)$ is given by $\widetilde{I}$. It suffices to check this

on basic open sets $D_+(g)$, $g \in S_d, d > 0$, by uniqueness of the sheaf defined on a basis. Indeed it follows that $f^\flat$ on $D_+(g)$ is given by the localisation map $S_{(g)} \to S_{(g)}/I_{(g)}$, whose kernel is $I_{(g)} = \widetilde{I}(D_+(g))$. ☐

Proposition 1.8.2.9. *Let $S = A[x_0, \dots, x_r]$ for a ring A and let $X = \mathrm{Proj}(S)$.*

1. *Let $I \leq S$ be a homogeneous ideal and denote $\bar{I} = \{s \in S \mid \forall i = 0, \dots, r,\ \exists n_i \text{ s.t. } x_i^{n_i} s \in I\}$ to be the saturation of I. Then, $\bar{I}$ is homogeneous.*
2. *Let $I, J \leq S$ be two homogeneous ideals. Then $\mathrm{Proj}(S/I) \cong \mathrm{Proj}(S/J)$ if and only if $\bar{I} = \bar{J}$.*
3. *Let $Y \hookrightarrow \mathrm{Proj}(S)$ be a closed subscheme. Then, $\Gamma_*(\mathcal{I}_Y)$ is a saturated ideal of S.*

Proof. 1. This follows from a simple consideration of the uniqueness of homogeneous decomposition of each element in a graded ring.
2. We may reduce to showing that I and $\bar{I}$ defines the same closed subscheme. We already have $I \hookrightarrow \bar{I}$ which translates to $V(\bar{I}) \hookrightarrow V(I)$. Conversely, pick $\mathfrak{p} \in V(I) \subseteq \mathrm{Proj}(S)$. We wish to show $\mathfrak{p} \supseteq \bar{I}$. Pick any $s \in \bar{I}$. Assume that $s \notin \mathfrak{p}$. It then follows that $\mathfrak{p} = \langle x_0, \dots, x_r \rangle$ which is a prime ideal which contains S_+, thus $\mathfrak{p} \notin \mathrm{Proj}(S)$, a contradiction.
We then wish to show isomorphism of sheaves. Going to basic opens, this reduces to showing surjection is an injection:

$$\begin{aligned} (S/I)_{(f)} &\longrightarrow (S/\bar{I})_{(f)} \\ \frac{s+I}{f^n} &\longmapsto \frac{s+\bar{I}}{f^n}. \end{aligned}$$

This follows from the fact that $\bar{I}$ is saturated[33].
3. Pick a homogeneous element $s \in S_d$ such that for each $i = 0, \dots, r$, there exists $n_i \in \mathbb{N}$ such that $x_i^{n_i} s \in \Gamma(\mathcal{I}_Y(d+n_i), X)$. We wish to show that $s \in \Gamma_*(\mathcal{I}_Y)$. Note that $s \in \Gamma(\mathcal{O}_X(d), X)$. Cover X by $D_+(x_i)$ and consider the restrictions $x_i^{n_i} s \in \mathcal{I}_Y(d+n_i)(D_+(x_i))$. Multiplying (tensoring) $x_i^{n_i} s$ with $x_i^{-n_i} \in \mathcal{O}_X(-n_i)(D_+(x_i))$ yields $s \in \mathcal{I}_Y(d+n_i) \otimes_{\mathcal{O}_X} \mathcal{O}_X(-n_i) \cong \mathcal{I}_Y(d)$ over $D_+(x_i)$. Thus, gluing these sections up from each $D_+(x_i)$, we get $s \in \Gamma(\mathcal{I}_Y(d), X) \subseteq \Gamma_*(\mathcal{I}_Y)$, as required. ☐

Using the above result, it is possible to find a characterization of closed subschemes of $\mathrm{Proj}(S)$ in terms of algebraic data.

Corollary 1.8.2.10. *Let $S = A[x_0, \dots, x_r]$ be a graded ring for a ring A. Then there is a correspondence:*

$$\{\textit{All closed subschemes of } \mathrm{Proj}(S)\} \quad \cong \quad \{\textit{All saturated ideals of } S\}\ .$$

Proof. Follows from Proposition 1.8.2.9. ☐

Next, let us show how projective n-spaces over a ring changes with extension of scalars.

Definition 1.8.2.11. (**Projective n-space over a ring**) Let A be a ring. The projective n-space over A is defined to be $\mathbb{P}^n_A := \mathrm{Proj}(A[x_0, \dots, n])$. By Lemma 1.8.2.2, $\mathbb{P}^n_A$ is a scheme over $\mathrm{Spec}\,(A)$.

We now see how $\mathbb{P}^n_A$ behaves under extension of scalars.

[33] In-fact, this step shows exactly why the definition of saturation would've been made!

Lemma 1.8.2.12. *Let $A \to B$ be a map of rings and* $\mathrm{Spec}\,(B) \to \mathrm{Spec}\,(A)$ *be the corresponding map of affine schemes. Then,*

$$\mathbb{P}^n_B \cong \mathbb{P}^n_A \times_{\mathrm{Spec}(A)} \mathrm{Spec}\,(B).$$

Proof. Observe that $D_+(x_i) \subseteq \mathbb{P}^n_A$ for $i = 0, \dots, n$ covers $\mathbb{P}^n_A$ as $A[x_0, \dots, x_n]$ is finitely generated by x_i as an A-algebra. By Lemma 1.6.4.7 together with Lemma 1.8.2.4 we obtain the following:

$$\begin{aligned}
\mathbb{P}^n_A \times_{\mathrm{Spec}(A)} \mathrm{Spec}\,(B) &= \left(\bigcup_{i=0}^n D_+(x_i)\right) \times_{\mathrm{Spec}(A)} \mathrm{Spec}\,(B) \\
&\cong \bigcup_{i=0}^n D_+(x_i) \times_{\mathrm{Spec}(A)} \mathrm{Spec}\,(B) \\
&\cong \bigcup_{i=0}^n \mathrm{Spec}\,\left(A[x_0, \dots, x_n]_{(x_i)} \otimes_A B\right) \\
&\cong \bigcup_{i=0}^n \mathrm{Spec}\,\left(A[x_0/_i, \dots, \widehat{x_i/x_i}, \dots, x_n/x_i] \otimes_A B\right) \\
&\cong \bigcup_{i=0}^n \mathrm{Spec}\,\left(B[x_0, \dots, x_n]_{(x_i)}\right) \\
&\cong \mathbb{P}^n_B.
\end{aligned}$$

□

Remark 1.8.2.13. Since any ring A is a $\mathbb{Z}$-algebra and $\mathbb{P}^n_A$ is naturally a $\mathbb{Z}$-scheme, therefore $\mathbb{P}^n_A \cong \mathbb{P}^n_{\mathbb{Z}} \times_{\mathrm{Spec}(\mathbb{Z})} \mathrm{Spec}\,(A)$, where the projection map $\mathbb{P}^n_A \to \mathrm{Spec}\,(A)$ is the usual structure map. This further motivates the construction of a projective space over any scheme.

Definition 1.8.2.14. (**Projective n-space over a scheme**) Let X be a scheme. The projective n-space over X is defined to be

$$\mathbb{P}^n_X := \mathbb{P}^n_{\mathbb{Z}} \times_{\mathrm{Spec}(\mathbb{Z})} X.$$

The natural projection map thus makes $\mathbb{P}^n_X$ a scheme over X.

1.8.3 Blowups

Do from Chapter 3 of Mumford and Hartshorne.

1.9 $\mathcal{O}_X$-modules

We will now cover certain types of important $\mathcal{O}_X$-modules that we will need in our study. Note that we defined $\mathcal{O}_X$-modules and various other algebraic constructions on them in Chapter 3, thus we assume the basic notion of $\mathcal{O}_X$-modules and its global algebra being known and we will thus specialize to the case of X being a scheme. The main goal is to define and study coherent and quasi-coherent modules over a scheme X. Its importance will manifest later in our study of projective schemes and their cohomology, the latter of which is an extremely powerful and versatile tool for doing geometry over schemes.

1.9.1 Coherent and quasi-coherent modules on schemes

Quasi-coherent sheaves form an integral part of the backbone of an attempt at doing geometry on schemes. Even though the definitions here makes sense in the setting of locally ringed spaces, but this theory is much more better behaved in the setting of schemes; for schemes, such sheaves have nice description on affine opens. This is the reason it is not included in Foundational Geometry, Chapter 3.

We first define the notion of quasicoherent modules on schemes.

Definition 1.9.1.1. (**Quasicoherent and coherent $\mathcal{O}_X$-modules**) Let X be a scheme. Then an $\mathcal{O}_X$-module $\mathcal{F}$ is called quasicoherent if there exists an affine open cover $\{U_i := \operatorname{Spec}(R_i)\}_{i\in I}$ of X and $\{M_i\}_{i\in I}$ where M_i is an R_i-module such that $\mathcal{F}_{|U_i} \cong \widetilde{M_i}$ for all $i \in I$. Further, $\mathcal{F}$ is said to be a coherent module if each M_i is a finitely generated R_i-module for each $i \in I$.

Remark 1.9.1.2. There are five basic properties of quasi-coherent sheaves on a scheme, which we point out now.

1. Quasicoherence of a module can be checked locally.
2. The global sections functor of a quasicoherent module over an affine scheme is exact[34].
3. The image of the functor $\widetilde{(-)} : \mathbf{Mod}(R) \to \mathbf{Mod}(\mathcal{O}_{\operatorname{Spec}(R)})$ (see Definition 1.2.3.1 and remarks surrounding it) is precisely all quasicoherent modules over $\operatorname{Spec}(R)$.
4. Quasicoherence is preserved under inverse image. It is further preserved under direct image if domain is a Noetherian scheme or if the map is quasi-compact and separated.
5. The category of all quasicoherent modules
$$\mathbf{QCoh}(\mathcal{O}_X)$$
is a Grothendieck-abelian category.

We will come to these results one by one. We first discuss some basic properties and examples.

Examples of quasicoherent modules

Lemma 1.9.1.3. *Let $X = \operatorname{Spec}(A)$ be an affine scheme and $\mathfrak{a} \leq A$ be an ideal. Consider the corresponding closed immersion*
$$i : \operatorname{Spec}(A/\mathfrak{a}) = Y \hookrightarrow \operatorname{Spec}(A) = X.$$

[34]This in cohomological language means that the first cohomology group $H^1(X, \mathcal{F}) = 0$, as we shall see after few sections.

Then,

1. *$i_*\mathcal{O}_Y$ is a coherent $\mathcal{O}_X$-module,*
2. *$i_*\mathcal{O}_Y \cong \widetilde{A/\mathfrak{a}}$.*

Proof. 1. Consider the following

$$\varphi : \mathcal{O}_X \times i_*\mathcal{O}_Y \longrightarrow i_*\mathcal{O}_Y$$

which on a basic open $D(f) \subseteq X$ for $f \in A$ is given by

$$\varphi_{D(f)} : A_f \times (A/\mathfrak{a})_{\bar{f}} \to (A/\mathfrak{a})_{\bar{f}}$$

as the usual A_f-module structure over $(A/\mathfrak{a})_{\bar{f}}$. Indeed, as the above maps are natural w.r.t. restrictions, this suffices by Theorem 20.6.1.3. Thus, $i_*\mathcal{O}_Y$ is an $\mathcal{O}_X$-module. Further, $i_*\mathcal{O}_Y$ is a coherent $\mathcal{O}_X$-module as the open cover $\{D(f)\}_{f\in A}$ of X is such that $i_*\mathcal{O}_Y(D(f)) \cong (A/\mathfrak{a})_{\bar{f}}$ is an $\mathcal{O}_X(D(f)) \cong A_f$-module generated by $\bar{1}$ (in the case when $f \in A$, we have $i_*\mathcal{O}_Y(D(f)) = 0$ and so trivially a finitely generated A_f-module).

2. Again, by the use of above mentioned lemma, we may reduce to working over a basis of X. Choosing $\{D(f)\}_{f\in A}$ to be such a basis, we see that $i_*\mathcal{O}_Y(D(f)) \cong (A/\mathfrak{a})_f$ and $\widetilde{A/\mathfrak{a}}(D(f)) \cong (A/\mathfrak{a})_f$. Hence, we may define a map $i_*\mathcal{O}_Y \to \widetilde{A/\mathfrak{a}}$ which on basic opens is identity. Consequently, this map on stalks is identity. This by above used lemma again yields a unique sheaf morphism $\varphi : i_*\mathcal{O}_Y \to \widetilde{A/\mathfrak{a}}$ which is an isomorphism as at stalks it is an isomorphism. □

Example 1.9.1.4. Let X be an integral noetherian scheme and let K be its function field. Let $\mathcal{K}$ be the constant sheaf of field K over X. Then $\mathcal{K}$ is a quasi-coherent $\mathcal{O}_X$-module.

As X is noetherian, therefore let $X = \bigcup_{i=1}^n \operatorname{Spec}(A_i)$ where A_i are noetherian rings. As X is integral, therefore by Lemma 1.4.2.2, each A_i is a noetherian domain. Thus we deduce that $K \cong Q(A_i)$ for each i, where $Q(A_i)$ is the fraction field of A_i. This is because $\operatorname{Spec}(A_i)$ are open and X irreducible. We now show that $\mathcal{K}$ is an $\mathcal{O}_X$-module.

Pick any open $U \subseteq X$. Recall from Chapter 20 that a section of $\mathcal{K}(U)$ is a continuous map $U \to K$ with K in discrete topology. For any point $p \in \operatorname{Spec}(A_i)$, as A_i is a domain, we see that $(A_i)_p \hookrightarrow Q(A_i) \cong K$ for each i. We thus deduce that K is an $\mathcal{O}_{X,x}$-algebra for each $x \in X$ in a natural way and $\mathcal{O}_{X,x} \subseteq K$. So we may now define

$$\begin{aligned} \mathcal{O}_X(U) \times \mathcal{K}(U) &\to \mathcal{K}(U) \\ (c, s) &\mapsto c \cdot s \end{aligned}$$

where $c \cdot s : U \to K$ is defined by $c(x)s(x) \in K$, $c(x) \in \mathcal{O}_{X,x} \subseteq K$. This is continuous as each $c \in \mathcal{O}_X(U)$ is seen to be a continuous map $c : U \to \coprod_{x\in U} \mathcal{O}_{X,x} \subseteq K$ as it is locally constant (Remark 20.2.0.4 and that locally around each point we have an affine open inside every open). This is automatically compatible with restrictions. Consequently, $\mathcal{K}$ is an $\mathcal{O}_X$-module.

Next, to see this is quasi-coherent, we claim that the affine open cover $\{U_i = \operatorname{Spec}(A_i)\}_{i=1,\dots,n}$ is such that $\mathcal{K}_{|U_i}$ is isomorphic to $\widetilde{K}$. Consequently, we reduce to proving the following claim : Let $X = \operatorname{Spec}(A)$ be an affine scheme where A is a noetherian domain and let $K = Q(A)$ be its fraction field. Then, the constant sheaf $\mathcal{K}$ associated to K is isomorphic to the $\mathcal{O}_X$-module $\widetilde{K}$.

It suffices to construct a map $\varphi : \mathcal{K} \to \widetilde{K}$ defined only on a basis such that on basics it is an isomorphism. For this, we notice that since localization of a domain at an element is again a domain, therefore for each $g \in R$, the open $D(g) \subseteq X$ is connected. Hence, $\mathcal{K}(D(g)) = K$ and $\widetilde{K}(D(g)) \cong K_g = K$. Thus, we may define $\varphi_{D(g)} : K \to K$ to be identity which is then easily seen to be a sheaf morphism. Hence, these sheaves are isomorphic as $\mathcal{O}_X$-modules.

Example 1.9.1.5. We now discuss a specific example of quasi-coherent modules over Spec $(\mathbb{Z})$, which brings to light the constraints put on by quasi-coherence on an $\mathcal{O}_X$-module. We ask the following question : *What are all quasicoherent skyscraper $\mathcal{O}_{\mathbb{Z}}$-modules over* Spec $(\mathbb{Z})$ *supported at non-zero prime $p \in \mathbb{Z}$?* We claim that these are in bijection with all p^∞-torsion $\mathbb{Z}$-modules, that is, every element of the module is annihilated by some power of p:

$$\left\{\begin{matrix}\text{Quasicoherent } \mathcal{O}_{\mathbb{Z}}\text{−modules } \mathcal{F} \\ \text{skyscraper at } p \in \mathbb{Z}\end{matrix}\right\} \cong \left\{\begin{matrix}\text{Abelian groups } M \text{ which are } p^\infty\text{-} \\ \text{torsion.}\end{matrix}\right\}$$

Indeed, let $\mathcal{F}$ be a quasicoherent skyscraper module at prime $p \in \mathbb{Z}$. Let us invoke the Corollary 1.9.1.12, to conclude that $\mathcal{F} = \widetilde{M}$ for some $\mathbb{Z}$-module M. As it is skyscraper, therefore for any open $U \ni p$ in Spec $(\mathbb{Z})$, we have $\mathcal{F}(U) = G$ where G is a fixed abelian group and $\mathcal{F}(U) = 0$ if $p \notin U$. Consequently, we have that $\mathcal{F}_x = 0$ if $x \neq p$ and $\mathcal{F}_p = G$. As $\mathcal{F} = \widetilde{M}$, therefore we have

$$\Gamma(\mathcal{F}, X) = G \cong M.$$

Further, for any basic open $D(f) \subseteq$ Spec $(\mathbb{Z})$ containing prime p, we deduce that $\mathcal{F}(D(f)) \cong M_f \cong G \cong M$. This, when unravelled, yields that for any integer $f \in \mathbb{Z}$ such that $f \notin \langle p \rangle \iff p \not| f$, we have $M_f \cong M$. Further, if $\langle p \rangle \notin D(f) \iff p|f$, then $M_f = 0$. Now fix any $m \in M$. We claim that some power of p annihilates m. Indeed, consider $D(p)$ which does not contain $\langle p \rangle$ as $p \in \langle p \rangle$. Thus, by above, we have that $\frac{m}{1} = 0$ in M_p. Consequently, for some $k \in \mathbb{N}$, we have $p^k m = 0$, as required. Hence, $T_{p^\infty}(M) = M$.

Conversely, consider a p^∞-torsion abelian group M. We wish to show that the quasicoherent module associated to M, $\widetilde{M}$, is skyscraper at $p \in \mathbb{Z}$. Let $D(f) \subseteq$ Spec $(\mathbb{Z})$ be a basic open not containing $\langle p \rangle$, equivalently, $p|f$. Then, we see that $\widetilde{M}(D(f)) \cong M_f$. Now pick any $\frac{m}{f^k} \in M_f$. Let $p^n m = 0$. Thus, $f^n m = 0$ as $p|f$. Consequently, we may write $\frac{m}{f^k} = \frac{1}{f^k}\frac{m}{1} = \frac{1}{f^k}\frac{f^n m}{f^n} = \frac{1}{f^k}\frac{0}{f^n} = 0$. Thus, $M_f = 0$.

Let $D(f)$ now be a basic open set which contains $\langle p \rangle$, equivalently, $p \not| f$. Then, $\widetilde{M}(D(f)) \cong M_f$ and we wish to show that $M_f \cong M$. Indeed, observe that since $p \not| f$, therefore $\gcd(p^k, f^l) = 1$ for all $k, l \geq 1$. It follows that there exists $a_k, b_l \in \mathbb{Z}$ such that

$$a_k p^k + b_l f^l = 1.$$

Thus, for any $\frac{m}{f^n} \in M_f$, where $p^k m = 0$, we obtain $a_k, b_n \in \mathbb{Z}$ such that $a_k p^k + b_n f^n = 1$. Using this on module M, we yield $a_k p^k m + b_n f^n m = m$, that is, $b_n f^n m = m$. Consequently, we may write

$$\frac{m}{f^n} = \frac{b_n f^n m}{f^n} = \frac{b_n m}{1}$$

in M_f. It follows that the localization map $\varphi : M \to M_f$ is surjective. We thus need only establish the injectivity of φ. Indeed, if $\varphi(m) = \frac{m}{1} = 0$ in M_f, then $f^n m = 0$ for some $n \in \mathbb{N}$. By above, we have $b_n \in \mathbb{Z}$ such that $b_n f^n m = m$. Consequently, $m = b_n f^n m = 0$, that is, Ker $(\varphi) = 0$, as required. Thus, $\varphi : M \to M_f$ is the required isomorphism. This completes the proof.

Locality of quasicoherence

We now discuss some more results which would culminate in the proofs of statement 1 in Remark 1.9.1.2.

Lemma 1.9.1.6. *Let $\mathcal{F}$ be a quasicoherent module over an affine scheme $X = \mathrm{Spec}\,(R)$. Then X admits a finite open affine cover $\{D(g_i)\}_{i=1}^n$ such that $\mathcal{F}_{|D(g_i)} \cong \widetilde{M_i}$ where M_i is an R_{g_i}-module.*

Proof. Since $\mathcal{F}$ is quasicoherent, therefore there exists an open affine cover $\{U_i = \mathrm{Spec}\,(S_i)\}_i$ of X such that $\mathcal{F}_{|U_i} \cong \widetilde{M_i}$ where M_i is an S_i-module. Since subsets of the form $D(g)$ forms a basis of X therefore for $D(g) \subseteq U_i$ we obtain via Lemma 1.2.3.4, 2, that $\mathcal{F}_{|D(g)} \cong \widetilde{R_g \otimes_{S_i} M_i}$ as $D(g) \cong \mathrm{Spec}\,(R_g)$. Since $N_i := R_g \otimes_{S_i} M_i$ is an R_g-module, so we have a cover of X by finitely many $D(g_i)$ by Lemma 1.2.1.6 such that $\mathcal{F}_{|D(g_i)} \cong \widetilde{N_i}$ where N_i is an R_{g_i}-module. □

Using the above, we first show a technical lemma, which will be generalized later on, which will be used to show locality of quasi-coherent modules[35].

Lemma 1.9.1.7. *Let $X = \mathrm{Spec}\,(A)$ be an affine scheme and $\mathcal{F} \in \mathbf{QCoh}(X)$ be a quasi-coherent module. Let $D(f) \subseteq X$ be a basic open set for some $f \in A$.*

1. *If $s \in \Gamma(\mathcal{F}, X)$ is a global section of the module $\mathcal{F}$ such that s restricted on $D(f)$ is 0, then there exists $n > 0$ such that $f^n s = 0$ over X.*
2. *If $t \in \mathcal{F}(D(f))$, then there exists $n > 0$ such that $f^n t \in \mathcal{F}(D(f))$ extends to a global section of the module $\mathcal{F}$.*

Proof. 1. By Lemma 1.9.1.6, there exists a finite open cover $D(g_i)$ of X such that $\mathcal{F}_{|D(g_i)} \cong \widetilde{M_i}$. Denoting the restriction of s to $D(g_i)$ as $s_i \in M_i$, we see that the image of s_i is zero in $(M_i)_f$ when restricted to $D(fg_i) = D(f) \cap D(g_i)$. Consequently, for some $n_i > 0$, we have $f^{n_i} s_i = 0$ over $D(g_i)$. As g_i are finitely many, taking large enough n, we obtain $f^n s_i = 0$ over each $D(g_i)$. It follows that the global section $f^n s$ of the module $\mathcal{F}$ is such that it's restriction to each open set of an open cover of X is 0. By sheaf axioms, it follows that $f^n s = 0$ over X.

2. Fix the finite open affine cover $\{D(g_i)\}_{i=1}^n$ of X coming from Lemma 1.9.1.6. Consider all the finitely many intersections $D(g) \cap D(g_i) = D(fg_i)$. Restricting t from $D(f)$ to $D(fg_i)$, we obtain $t_i \in (M_i)_f$ for each i. Hence, for each i, there is some $n_i > 0$ such that $f^{n_i} t_i \in M_i = \mathcal{F}(D(g_i))$. By multiplying by large f^k to each $f^{n_i} t_i$ which are finitely many, we may arrange that $f^n t_i \in \mathcal{F}(D(g_i))$.

We now form a matching family for the module $\mathcal{F}$ over the open cover $\{D(g_i)\}$ which would glue up to give the required global section. Indeed, fix two $D(g_i)$ and $D(g_j)$. Restrict $f^n t_i$ and $f^n t_j$ to $D(g_i) \cap D(g_j) = D(g_i g_j)$. Observe that over the even smaller open $D(fg_i g_j)$, the section $f^n t_i - f^n t_j$ is zero as $t_i = t_j = t$ over $D(fg_i g_j) \subseteq D(f)$. Hence by item 1 applied over $D(g_i g_j)$, there exists $m_{ij} > 0$ such that $f^{m_{ij}}(f^n t_i - f^n t_j) = 0$, hence $f^{n+m_{ij}}(t_i - t_j) = 0$ over $D(fg_i g_j)$. As i and j are finitely many, so taking m large enough, we obtain $f^{n+m} t_i = f^{n+m} t_j$ over $\mathcal{F}(D(g_i g_j))$ for each i and j. Thus, the family $\{f^{n+m} t_i\}$ is a matching family which glues up to give $s \in \Gamma(\mathcal{F}, X)$ such that its restriction over $D(f)$ is $f^{n+m} t$[36]. □

[35] The result is similar in flavour to Proposition 1.3.1.4.

[36] Note that we have implicitly used the fact the restriction maps of $\mathcal{F}$ preserves the respective module structures (see remarks surrounding Definition 3.5.0.1)

Remark 1.9.1.8. Let $X = \operatorname{Spec}(R)$ be an affine scheme and $\mathcal{F}$ be a quasicoherent $\mathcal{O}_X$-module over X. Then, we obtain a map

$$\alpha : \widetilde{\Gamma(\mathcal{F}, X)} \longrightarrow \mathcal{F}$$

which on a basic open set $D(f) \subseteq X,\ f \in R$ is given by $\Gamma(\mathcal{F}, X)_f \to \mathcal{F}(D(f))$ mapping as $m/f^n \mapsto \rho_{X,D(f)}(m)/f^n$. Indeed, this is a $\mathcal{O}_X$-linear homomorphism which on stalks yields the $R_\mathfrak{p}$-linear map

$$\Gamma(\mathcal{F}, X)_\mathfrak{p} \longrightarrow \mathcal{F}_\mathfrak{p}$$

which is given by

$$\Gamma(\mathcal{F}, X) \otimes_R R_\mathfrak{p} \cong \Gamma(\mathcal{F}, X) \otimes_R \varinjlim_{f \notin \mathfrak{p}} R_f \cong \varinjlim_{f \notin \mathfrak{p}} \Gamma(\mathcal{F}, X) \otimes_R R_f = \varinjlim_{D(f) \ni \mathfrak{p}} \widetilde{\Gamma(\mathcal{F}, X)}(D(f)) \to \varinjlim_{D(f) \ni \mathfrak{p}} \mathcal{F}(D(f)).$$

We will see that this map α would become an isomorphism, especially due to the fact that quasicoherent modules behave very nicely on open affines of the form $D(f)$, as the Lemma 1.9.1.6 shows.

Corollary 1.9.1.9. *Let $\mathcal{F}$ be a quasicoherent sheaf over an affine scheme $X = \operatorname{Spec}(A)$. Then, there is a natural isomorphism of A_f-modules for each $f \in A$*

$$\Gamma(\mathcal{F}, X)_f \xrightarrow{\cong} \mathcal{F}(D(f))$$

given by $m/f^n \mapsto \rho(m)/f^n$ where ρ is the restriction map of $\mathcal{F}$ from X to $D(f)$

Proof. Follows from Lemma 1.9.1.7. $\square$

Using the above, one proves the local nature of quasicoherence.

Proposition 1.9.1.10. *Let $\mathcal{F}$ be an $\mathcal{O}_X$-module over a scheme X. Then the following are equivalent:*

1. *$\mathcal{F}$ is quasicoherent.*
2. *For all open affine $U = \operatorname{Spec}(A) \subseteq X$ we have $\mathcal{F}_{|U} \cong \widetilde{M}$ where M is an A-module.*

Proof. We need to only show $2 \Rightarrow 1$. Let $\mathcal{F}$ be quasicoherent and $U = \operatorname{Spec}(A)$ open affine. We may assume $X = \operatorname{Spec}(A)$. Thus we need to show $\mathcal{F} \cong \widetilde{M}$ for an A-module M. By Lemma 1.9.1.6, we obtain an open affine cover $D(g_i)$ of X where $\mathcal{F}_{|D(g_i)} \cong \widetilde{M_i}$ for an A_{g_i}-module M_i. Let $M = \Gamma(\mathcal{F}, X)$, which is an A-module. By Corollary 1.9.1.9, we obtain a natural isomorphism $M_i \cong M_{g_i}$. Thus we have the required result. $\square$

A similar result is true for coherent modules.

Proposition 1.9.1.11. *Let $\mathcal{F}$ be an $\mathcal{O}_X$-module over a Noetherian scheme X. Then the following are equivalent:*

1. *$\mathcal{F}$ is coherent.*
2. *For all open affine $U = \operatorname{Spec}(A) \subseteq X$ we have $\mathcal{F}_{|U} \cong \widetilde{M}$ where M is a finitely generated A-module.*

Proof. See Proposition 5.4, Chapter 2 [Hartshorne]. $\square$

Corollary 1.9.1.12. *The image of the functor $\widetilde{(-)}$ of Definition 1.2.3.1 is exactly all quasicoherent modules over* $\operatorname{Spec}(R)$. *In other words,* $\mathbf{Mod}(R) \equiv \widetilde{\mathbf{Mod}}(\mathcal{O}_{\operatorname{Spec}(R)}) = \mathbf{QCoh}(\mathcal{O}_{\operatorname{Spec}(R)})$. *Further, if R is noetherian, then this restricts to* $\mathbf{Mod}(R)^{f.g.} \equiv \mathbf{Coh}(\mathcal{O}_{\operatorname{Spec}(R)})$. $\square$

Quasicoherence and exactness of global sections

We next see the exactness of global sections functor.

Proposition 1.9.1.13. *Let X be an affine scheme and $0 \to \mathcal{F}' \to \mathcal{F} \to \mathcal{F}'' \to 0$ be an exact sequence of $\mathcal{O}_X$-modules. If $\mathcal{F}'$ is quasicoherent, then*

$$0 \to \Gamma(X, \mathcal{F}') \to \Gamma(X, \mathcal{F}) \to \Gamma(X, \mathcal{F}'') \to 0$$

is exact.

Proof. Proposition 5.6, Chapter 2 [Hartshorne]. □

The category $\mathbf{QCoh}(X)$

The category of quasicoherent modules is further a Grothendieck-abelian category.

Theorem 1.9.1.14. *Let X be a scheme. The category $\mathbf{QCoh}(\mathcal{O}_X)$ is a Grothendieck-abelian category. Consequently, it is an abelian category which has all coproducts.*

Proof. Tag 077P, [Stacksproject]. □

Apart from $\mathbf{QCoh}(\mathcal{O}_X)$ being abelian in its own right, it is also an exact category, where the underlying abelian category is $\mathbf{Mod}(\mathcal{O}_X)$.

Proposition 1.9.1.15. *Let X be a scheme. If $0 \to \mathcal{F} \to \mathcal{G} \to \mathcal{H} \to 0$ is an exact sequence of $\mathcal{O}_X$-modules where $\mathcal{F}$ and $\mathcal{H}$ are quasicoherent, then so is $\mathcal{G}$. Consequently, quasicoherence is preserved under extensions of modules.*

Quasicoherence, direct and inverse images

We now see behavior of quasicoherence and coherence under inverse and direct images.

Lemma 1.9.1.16. *Let $f : X \to Y$ be a morphism of schemes and let $\mathcal{G}$ be a quasicoherent $\mathcal{O}_Y$-module. Then $f^*\mathcal{G}$ is a quasicoherent $\mathcal{O}_X$-module. If X, Y are noetherian schemes and $\mathcal{G}$ is coherent, then $f^*\mathcal{G}$ is coherent.*

Proof. The first question is local in both X and Y by Proposition 1.9.1.10. Indeed, pick $x \in X$ and an open affine $V \ni f(x)$ in Y. Then by continuity there is an open affine $U \ni x$ in X such that $f(U) \subset V$. This shows that we may assume X and Y to be affine. The result is now immediate by Corollary 1.9.1.12 and Lemma 1.2.3.4, 2. The same technique works for coherent case. □

For stability under direct image, we need some conditions on the map if noetherian conditions need to be dropped.

Lemma 1.9.1.17. *Let $f : X \to Y$ be a morphism of schemes and $\mathcal{F}$ be a quasicoherent $\mathcal{O}_X$-module. Then $f_*\mathcal{F}$ is a quasicoherent $\mathcal{O}_Y$-module if any of the following holds:*

1. *X is noetherian,*
2. *f is quasi-compact and separated.*

Proof. By Proposition 1.9.1.10, we may assume that Y is affine. First, if X is noetherian, then $X = \bigcup_{i=1}^n U_i$, $U_i = \operatorname{Spec}(A_i)$ and A_i noetherian. By Proposition 1.4.1.8, $U_i \cap U_j$ is quasi-compact and thus can be covered by finitely many open affines, say U_{ijk}. On the other hand if f is q.c.s. then again $X = \bigcup_{i=1}^n U_i$ and by separatedness $U_i \cap U_j$ is affine, in which case we let $U_{ijk} = U_i \cap U_j$. By sheaf axioms of $f_*\mathcal{F}$, we have an exact sequence as in

$$0 \longrightarrow f_*\mathcal{F} \longrightarrow \bigoplus_i f_*\left(\mathcal{F}_{|U_i}\right) \longrightarrow \bigoplus_{i,j,k} f_*\left(\mathcal{F}_{|U_{ijk}}\right)$$

where all maps are induced by restrictions. As $\mathcal{F} \in \mathbf{QCoh}(\mathcal{O}_X)$, thus by local criterion of Proposition 1.9.1.10, we get that $\mathcal{F}_{|U_i} = \widetilde{M_i}$ for some A_i-module M_i. By Lemma 1.2.3.4, 1, $f_*\mathcal{F}_{|U_i}$ and $f_*\mathcal{F}_{|U_{ijk}}$ are quasicoherent. By Theorem 1.9.1.14, the middle and the right term in the above exact sequence are quasicoherent. By the same theorem again, the left term, $f_*\mathcal{F}$ is quasicoherent, as required. $\square$

More properties

As promised earlier, we state a general result about invertible modules and quasicoherent modules over schemes. This is a fundamental result and will be used to portray the simplicity of the techniques developed so far. Moreover, its proof showcases the simplicity of the sheaf language and is thus a good exercise.

Lemma 1.9.1.18. *Let X be a scheme, $\mathcal{L} \in \operatorname{Pic}(X)$, $\mathcal{F} \in \mathbf{QCoh}(X)$, $f \in \Gamma(\mathcal{L}, X)$ and $s \in \Gamma(X, \mathcal{F})$. Denote by $X_f \subseteq X$ the open subset $X_f := \{x \in X \mid f_x \notin \mathfrak{m}_x\mathcal{L}_x\}$.*

1. *If X is quasicompact and s is such that $s|_{X_f} = 0$, then there exists $n \in \mathbb{N}$ such that $f^n s = 0$ in $\Gamma(\mathcal{L}^{\otimes n} \otimes \mathcal{F}, X)$.*
2. *If X admits a finite affine open cover $\{U_i\}$ where $\mathcal{L}_{|U_i}$ is free (of rank 1) and $U_i \cap U_j$ is quasicompact, then for any $t \in \mathcal{F}(X_f)$, there exists $n \in \mathbb{N}$ such that $f^n t \in (\mathcal{L}^{\otimes n} \otimes \mathcal{F})(X_f)$ extends to a global section $s \in \Gamma(\mathcal{L}^{\otimes n} \otimes \mathcal{F}, X)$.*

Proof. 1. Cover X by finitely many affine open sets $U = \operatorname{Spec}(A)$ which satisfies $\varphi : \mathcal{L}_{|U} \cong \mathcal{O}_{X|U} \cong \mathcal{O}_{\operatorname{Spec}(A)}$. Further, denote $\mathcal{F}_{|U} = \widetilde{M}$ where M is an A-module (Corollary 1.9.1.12). By restricting f to U, we may write $g = \varphi_U(f) \in A$ and by restricting s to U, we may write $s \in M$. Since $s|_{X_f} = 0$ and $X_f \cap U = D(g)$, therefore $s/1 = 0$ in M_g. Consequently, there exists $n \in \mathbb{N}$ such that $g^n s = 0$ in $M = \Gamma(U, \mathcal{F}_{|U})$. We then observe the following isomorphisms (see Lemma 20.2.0.5):

$$(\mathcal{L}^{\otimes n} \otimes \mathcal{F})_{|U} \cong \mathcal{O}_{X|U}^{\otimes n} \otimes \mathcal{F}_{|U} \cong \mathcal{O}_{\operatorname{Spec}(A)}^{\otimes n} \otimes \widetilde{M} \cong \widetilde{M}.$$

Consequently, we get isomorphisms in sections over U which yields that $f^{\otimes n} \otimes s \mapsto g^n s = 0$. hence $f^{\otimes n} \otimes s = 0$ in $\mathcal{L}^{\otimes n} \otimes \mathcal{F}$ over U. Since this happens for all finitely many Us, therefore taking large enough n, we observe that $f^{\otimes n} \otimes s = 0$ in $\mathcal{L}^{\otimes n} \otimes \mathcal{F}$ over X.

2. Pick $t \in \mathcal{F}(X_f)$. For each of the finitely many i, let $U_i = \operatorname{Spec}(A_i)$. As $\mathcal{L}_{|U_i} \cong \mathcal{O}_{X|U_i}$, therefore $X_f \cap U_i = \{\mathfrak{p} \in \operatorname{Spec}(A_i) \mid f_\mathfrak{p} \notin \mathfrak{p}A_\mathfrak{p}\} = D(f)$ where we interpret $f \in \mathcal{L}(U_i)$ by restricting the global section f. By locality of quasicoherence (Proposition 1.9.1.10), we have an A_i-module M_i such that $\mathcal{F}_{|U_i} \cong \widetilde{M_i}$. As $t \in \mathcal{F}(X_f)$, therefore by restriction, we have $t_i \in \mathcal{F}(U_i \cap X_f) = \mathcal{F}(D(f)) \cong M_f$

(Proposition 1.2.3.3). It follows that for some n_i, we have $f^{n_i}t \in M_i = \mathcal{F}(U_i)$. Since U_i are atmost finite, so we may take a large enough n so that $f^n t \in M_i = \mathcal{F}(U_i)$.

Observe that

$$(\mathcal{L}^{\otimes n} \otimes \mathcal{F})_{|U_i} \cong \mathcal{O}^{\otimes n}_{\mathrm{Spec}(A_i)} \otimes \mathcal{F}_{|U_i} \cong \mathcal{F}_{|U_i} \cong \widetilde{M_i}$$

where $f^n t \in \mathcal{F}(U_i)$ corresponds to $f^{\otimes n} \otimes t \in \mathcal{L}^{\otimes n} \otimes \mathcal{F}(U_i)$. As $t_i = t = t_j \in \mathcal{F}(U_i \cap U_j \cap X_f)$, therefore $f^n(t_i - t_j) = 0$ in $\mathcal{F}(U_i \cap U_j \cap X_f)$. Our hypothesis that $U_i \cap U_j$ is quasicompact ensures by item 1 that there exists $k > 0$ such that $f^{n+k}(t_i - t_j) = 0$ in $\mathcal{F}(U_i \cap U_j)$, for all i, j. It follows that $f^{n+k} t_i \in \mathcal{F}(U_i) = \mathcal{L}^{\otimes n+k} \otimes \mathcal{F}(U_i)$ is a matching family. It follows that there exists $s \in \Gamma(\mathcal{L}^{\otimes n+k} \otimes \mathcal{F}, X)$ which on X_f is $f^{n+k}t$, as required. $\square$

These were some of the basic results on quasicoherent modules. We now do perhaps the most important application of $\mathcal{O}_X$-modules, that when X is a projective scheme.

1.9.2 Modules over projective schemes

Let S be a graded ring and M a graded S-module. We attach a sheaf $\widetilde{M}$ to M over Proj$S$.

Definition 1.9.2.1. ($\widetilde{M}$) Let S be a graded ring and M be a graded S-module. Then we define a sheaf $\widetilde{M}$ over $\mathrm{Proj}(S)$ given on an open set $U \subseteq \mathrm{Proj}(S)$ by

$$\widetilde{M}(U){:=}\left\{s{:}U{\to}\coprod_{\mathfrak{p}\in U} M_{(\mathfrak{p})} \mid \forall \mathfrak{p}\in U,\ s(\mathfrak{p})\in M_{(\mathfrak{p})} \&\exists \text{ open } \mathfrak{p}\in V\subseteq U \ \&\ m\in M_d, f\in S_d \text{ s.t. } f\notin \mathfrak{q} \ \&\ s(\mathfrak{q})=m/f \forall \mathfrak{q}\in V\right\}.$$

The restrictions are the obvious ones. It is clear that if we treat S as a graded S-module, then $\widetilde{S} \cong \mathcal{O}_{\mathrm{Proj}(S)}$ where we treat $\mathcal{O}_{\mathrm{Proj}(S)}$ as an $\mathcal{O}_{\mathrm{Proj}(S)}$-module.

Remark 1.9.2.2. Over a projective scheme $X = \mathrm{Proj}(S)$, the theory of quasi-coherent modules is the most useful. In particular, we will have the following observations to make about them:

1. Any graded S-module gives an $\mathcal{O}_X$-module $\widetilde{M}$ which is furthermore quasicoherent.
2. Any $\mathcal{O}_X$-module $\mathcal{F}$ gives a graded S-module $\Gamma_*(\mathcal{F})$.
3. For X being the projective n-space over a ring A, we have $\Gamma_*(\mathcal{O}_X) \cong A[x_0, \dots, x_n]$.
4. Assume S is furthermore finitely generated by degree 1 elements. If $\mathcal{F}$ is a quasicoherent $\mathcal{O}_X$-module, then $\widetilde{\Gamma_*(\mathcal{F})} \cong \mathcal{F}$.
5. All projective schemes over Spec(A) is of the form $\mathrm{Proj}(S)$ where $S_0 = A$ and S is finitely generated as by S_1 as an S_0-algebra.

These are the main takeaways from the general theory of quasicoherent $\mathcal{O}_{\mathrm{Proj}(S)}$-modules.

We now attend to these results one-by-one. We first have analogous results to the affine case on the behaviour of $\widetilde{M}$ on basis, on stalks and its quasicoherence.

Proposition 1.9.2.3. *Let S be a graded ring, M be a graded S-module, $X = \mathrm{Proj}(S)$ be the projective scheme over S and $\widetilde{M}$ to be the associated sheaf of M over X. Then,*

1. *for any $\mathfrak{p} \in X$,*
$$(\widetilde{M})_{\mathfrak{p}} \cong M_{(\mathfrak{p})},$$
2. *for any $f \in S_d$, $d > 0$ and basic open $D_+(f)$,*
$$\widetilde{M}_{|D_+(f)} \cong \widetilde{M_{(f)}},$$
3. *the sheaf $\widetilde{M}$ is an $\mathcal{O}_X$-module which is furthermore quasicoherent,*
4. *if S is a noetherian ring and M is finitely generated, then $\widetilde{M}$ is coherent.*

Proof. 1. and 2. follows from repeating Lemma 1.8.2.4. Statement 3. follows from local property of quasicoherence (Proposition 1.9.1.10), the fact that sets of the form $D_+(f)$ for $f \in S_d$, $d > 0$ forms a basis of X (Lemma 1.8.1.3) and statement 2 above. Statement 4 follows from coherence being a local property for Noetherian schemes (Proposition 1.9.1.11) and statement 2 above. □

Remark 1.9.2.4. The theory of $\mathcal{O}_{\mathrm{Proj}(S)}$-modules is rich because of various constructions which interrelates the category $\mathbf{grMod}(S)$ of graded S-modules and graded maps and the category $\mathbf{Mod}(\mathcal{O}_{\mathrm{Proj}(S)})$. Indeed, these constructions is what we will study now, and these will be absolutely indispensable to do geometry in projective spaces $\mathrm{Proj}(k[x_0, \dots, x_n]/f)$ for a homogeneous polynomial f.

Remark 1.9.2.5. The construction of $\mathcal{O}_X$-modules is functorial ($X = \mathrm{Proj}(S)$):

$$\begin{aligned} \widetilde{(-)} : \mathbf{grMod}(S) &\longrightarrow \mathbf{QCoh}(\mathcal{O}_X) \\ M &\longmapsto \widetilde{M} \\ M \xrightarrow{\varphi} N &\longmapsto \widetilde{M} \xrightarrow{\eta} \widetilde{N} \end{aligned}$$

where η on a basic open $D_+(f)$ is given by the localization maps $\eta_{D_+(f)} : M_{(f)} \to N_{(f)}$.

We first begin by twisting each $\mathcal{O}_{\mathrm{Proj}(S)}$-module.

Twists and Serre twists

Definition 1.9.2.6. (**Twists**) Let S be a graded ring and $X = \mathrm{Proj}(S)$. For each $n \in \mathbb{Z}$, we define the n^{th}*-Serre twist* to be $\mathcal{O}_X(n)$ which is defined to be $\widetilde{S(n)}$, the sheaf associated to the n-th twisted graded S-module $S(n)$ (Definition 16.2.1.7). For each $\mathcal{O}_X$-module $\mathcal{F}$, we then define the n^{th}*-twist of* $\mathcal{F}$ to be $\mathcal{F}(n) = \mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X(n)$.

Some obvious questions are: what happens to n^{th}-twist of $\widetilde{M}$ for a graded S-module M, what is so special about $\mathcal{O}_X(n)$ in relation to $\mathcal{O}_X$? We answer these in the following result.

Proposition 1.9.2.7. *Let S be a graded ring generated by S_1 as an S_0-algebra and $X = \mathrm{Proj}(S)$. Then,*

1. *$\mathcal{O}_X(n)$ is an invertible module for all $n \in \mathbb{Z}$,*
2. *for any graded S-modules M, N and $n \in \mathbb{Z}$,*
   (a) *$\widetilde{M \otimes_S N} \cong \widetilde{M} \otimes_{\mathcal{O}_X} \widetilde{N}$,*
   (b) *$\widetilde{M}(n) \cong \widetilde{M(n)}$.*
3. *Let T be a graded ring generated by T_1 as a T_0-algebra. Let M be a graded S-module and N a graded T-module. Let $\varphi : S \to T$ be a graded map and $f : U \to \mathrm{Proj}(S)$ be the corresponding map (Proposition 1.8.2.6). Then,*
   (a) *$f_*(\widetilde{N}_{|U}) \cong \widetilde{{}_SN}$,*
   (b) *$f^*(\widetilde{M}) \cong \widetilde{(M \otimes_S T)}_{|U}$,*
   (c) *$f_*(\mathcal{O}_{\mathrm{Proj}(T)|U}) \cong \widetilde{T}$, where T is treated to be an S-module via φ.*
4. *Let φ and f as in 3 and let $Y = \mathrm{Proj}(T)$. Then,*
   (a) *$f_*(\mathcal{O}_Y(n)_{|U}) \cong f_*(\mathcal{O}_{Y|U})(n)$,*
   (b) *$f^*(\mathcal{O}_X(n)) \cong \mathcal{O}_Y(n)_{|U}$.*
5. *For all $n, m \in \mathbb{Z}$,*

$$\mathcal{O}_X(n) \otimes_{\mathcal{O}_X} \mathcal{O}_X(m) \cong \mathcal{O}_X(n+m).$$

Proof. 1. Cover X by basic open sets of the form $D_+(f)$ for $f \in S_1$. One then easily reduces to showing that $S(n)_{(f)}$ is a free $S_{(f)}$-module of rank 1. Indeed, one shows that the following map is an $S_{(f)}$-linear map which is an isomorphism as $S_{(f)}$-modules: $S_{(f)} \to S(n)_{(f)}$ given by $s/f^k \mapsto sf^n/f^k$. One really needs f to be of degree 1 to be able to show that this is an isomorphism.
2. This reduces to finding natural isomorphism $M_{(f)} \otimes_{S_{(f)}} N_{(f)} \cong (M \otimes_S N)_{(f)}$ which one can do constructing two sided inverses. One of these maps well-definedness will use the fact that degree

of f is 1.
3. See Lemma 1.2.3.4 for $a)$ and $b)$ and observe that $f^{-1}(D_+(g)) = D_+(\varphi(g))$ for $g \in S$ homogeneous by a simple unravelling of definition of $U \subseteq \mathrm{Proj}(T)$. The statement $c)$ is immediate by looking the respective sections on a basic open set $D_+(g)$.
4. Statement $a)$ follows from $3.a)$ and $3.c)$ is immediate from $3.b)$.
5. Follows from $2.a)$. □

Remark 1.9.2.8. The twisting functor given by

$$\mathbf{Mod}(\mathcal{O}_X) \longrightarrow \mathbf{Mod}(\mathcal{O}_X)$$
$$\mathcal{F} \xrightarrow{f} \mathcal{G} \longmapsto \mathcal{F}(n) \xrightarrow{f\otimes \mathrm{id}} \mathcal{G}(n)$$

is exact. This is immediate as $\mathcal{F}(n) = \mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X(n)$ and thus localizing at a point x, we get $\mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{O}_X(n)_x \cong \mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{O}_{X,x}$ where the latter isomorphism follows from Proposition 1.9.2.7, 1.
In general this tells us also that the stalks of all twisted sheaves $\mathcal{F}(n)$ is identical to that of $\mathcal{F}$.

Remark 1.9.2.9. Let S be a graded ring and $X = \mathrm{Proj}(S)$ be the corresponding projective scheme with $\mathcal{F} \in \mathbf{QCoh}(X)$. Our goal in the next few pages is to understand how we can recover $\mathcal{F}$ by the global sections of all the twisted sheaves $\mathcal{F}(n)$. This is recorded in Propositions 1.9.2.12 and 1.9.2.13.

Associated graded S-module

We now associate to each $\mathcal{O}_X$-module $\mathcal{F}$ a graded S-module, where $X = \mathrm{Proj}(S)$.

Definition 1.9.2.10. (**Associated graded S-module**) Let S be a graded ring, $X = \mathrm{Proj}(S)$ and $\mathcal{F}$ an $\mathcal{O}_X$-module. Define the associated graded S-module to be

$$\Gamma_*(\mathcal{F}) := \bigoplus_{n\in\mathbb{Z}} \Gamma(\mathcal{F}(n), X)$$

where the S-module structure is given as follows: we need only define the scalar multiplication for homogeneous elements, so let $s_d \in S_d$ and $t_n \in \Gamma(\mathcal{F}(n), X)$. Then define $s_d \cdot t_n$ to be the image of $s_d \otimes t_n \in \Gamma(\mathcal{O}_X(d) \otimes_{\mathcal{O}_X} \mathcal{F}(n), X)$ under the isomorphism $\mathcal{O}_X(d) \otimes_{\mathcal{O}_X} \mathcal{F}(n) \cong \mathcal{F}(n+d)$ via Proposition 1.9.2.7, 5, in order to obtain an element of $\Gamma(\mathcal{F}(n+d), X)$, as needed.

Remark 1.9.2.11. There are two main results about associated graded S-modules.

1. Let $S = A[x_0, \dots, x_r]$ for a ring A and $r \geq 1$. Then $\Gamma_*(\mathcal{O}_X) \cong S$ for $X = \mathrm{Proj}(S)$. The relevance of this result is as follows. We know that the global sections of the structure sheaf over a projective scheme doesn't recover the homogeneous coordinate ring back. This result tells us that looking only at global sections of structure sheaf won't suffice (hopefully obvious by now), we need to instead look at global sections of all twists of the structure sheaf in order to recover the coordinate ring. For example, consider the quadric $xy - wz$ in $\mathbb{P}^5_k$. The corresponding coordinate ring is $S = k[w,x,y,z,a,b]/xy - wz \cong \frac{k[w,x,y,z]}{xy-wz}[a,b]$ and the corresponding scheme is $X = \mathrm{Proj}\,(S)$. Consequently, we can write $S = A[a,b]$ for $A = \frac{k[w,x,y,z]}{xy-wz}$ and thus this result would yield that S is isomorphic to $\Gamma_*(\mathcal{O}_X)$. Note that to use this result, we have to force ourselves to go 2 dimensions up.

2. Let S be a graded ring which is *finitely* generated by S_1 as an S_0-algebra and let $X = \mathrm{Proj}(S)$. Then, for any quasicoherent $\mathcal{O}_X$-module $\mathcal{F}$, we obtain a natural isomorphism $\widetilde{\Gamma_*(\mathcal{F})} \cong \mathcal{F}$. This result therefore tells us that the functor $M \mapsto \widetilde{M}$ of Remark 1.9.2.5 from graded S-modules to quasicoherent $\mathcal{O}_X$-modules is essentially surjective.

We'll later see that these results will allow us to obtain an equivalent criterion of when is a scheme over an affine scheme projective.

We now state these results and sketch their proofs.

Proposition 1.9.2.12. *Let $S = A[x_0, \dots, x_r]$ for a ring A and $r \geq 1$ and denote $X = \mathrm{Proj}(S)$. Then $\Gamma_*(\mathcal{O}_X) \cong S$.*

Proof. The main idea is to keep reducing the problem to a problem about graded ring S. Since S is generated by x_is as an A-algebra, therefore $D_+(x_i)$ for $i = 0, \dots, r$ covers X. An element in $\Gamma_*(\mathcal{O}_X)$ is given by a sum of elements $t_n \in \Gamma(\mathcal{O}_X(n), X)$. Let $t_n \in \Gamma(\mathcal{O}_X(n), X)$. The data of t_n is equivalently represented by the $(t_{n,0}, \dots, t_{n,r})$ where $t_{n,i} \in \mathcal{O}_X(n)(D_+(x_i)) = S(n)_{(x_i)}$ are the corresponding restrictions. Thus, $t_{n,i} \in S_{x_i}$ is a homogeneous element of degree n. Thus, $t = \sum_n t_n$ is equivalently represented by the tuple $(t_0, \dots, t_r)$ where $t_i = \sum_n t_{n,i}$ such that the image of t_i under $S_{x_i} \to S_{x_i,x_j}$ is same as the image of t_j under $S_{x_j} \to S_{x_j,x_i}$ for all $i, j = 0, \dots, r$. Note each of these S_{x_i,x_j} for varying i, j are contained in $R = S_{x_0,\dots,x_r}$. Now, we have injective maps $S \to S_{x_i} \to S_{x_i,x_j} \to R$ and thus $t = (t_0, \dots, t_r)$ as above is contained in $\bigcap_{i=0}^r S_{x_i} \hookrightarrow R$. In fact, any element of this intersection also corresponds to an element of $\Gamma_*(\mathcal{O}_X)$. Consequently, $\Gamma_*(\mathcal{O}_X) = \bigcap_{i=0}^r S_{x_i}$. It is straightforward to see that this intersection is exactly S by writing a general homogeneous element of R and observing what it needs to satisfy to be in the intersection. $\square$

We now show the essential surjectivity of $\widetilde{(-)}$. The proof of this result nicely shows the elegance of the techniques developed so far.

Proposition 1.9.2.13. *Let S be a graded ring which is finitely generated by S_1 as an S_0-algebra and $X = \mathrm{Proj} S$.*

1. *For each $\mathcal{O}_X$-module $\mathcal{F}$, there is a natural map*

$$\beta : \widetilde{\Gamma_*(\mathcal{F})} \longrightarrow \mathcal{F}.$$

2. *For each quasicoherent $\mathcal{O}_X$-module $\mathcal{F}$, the above map β is an isomorphism, that is,*

$$\beta : \widetilde{\Gamma_*(\mathcal{F})} \xrightarrow{\cong} \mathcal{F}.$$

Proof. 1. Since $D_+(f)$ for $f \in S_1$ covers X, we may define β naturally only on $D_+(f)$. This is done as follows:

$$\begin{tikzcd} \Gamma_*(\mathcal{F})_{(f)} \arrow[r, "\beta_{D_+(f)}"] \arrow[dr] & \mathcal{F}(D_+(f)) \\ & \Gamma(\mathcal{F}(d)_{|D_+(f)} \otimes \mathcal{O}_X(-d), X) \arrow[u, "\varphi"', "\cong"] \end{tikzcd}$$

where the diagonal map is given by

$$\frac{m}{f^d} \mapsto m \otimes \frac{1}{f^d}$$

and the isomorphism φ is given by restrictions.
2. In the above, we need to show that the diagonal map is an isomorphism. Suppose for some $m/f^d \in \Gamma_*(\mathcal{F})_{(f)}$, we have that $m \otimes 1/f^d = 0$ in $\Gamma(\mathcal{F}(d)_{|D_+(f)} \otimes \mathcal{O}_X(-d), X)$. Denote $\mathcal{G} = \mathcal{F}(d) \otimes \mathcal{O}_X(-d)$, which is quasicoherent. Note that $m \otimes 1/f^d = 0$ as an element in $\mathcal{G}(D_+(f))$ and also note that $D_+(f) = X_f$. Consequently from Lemma 1.9.1.18, 1, there exists $n \in \mathbb{N}$ such that $f^n \otimes m \otimes 1/f^d$ is zero as a global section of $\mathcal{O}_X(1)^{\otimes n} \otimes \mathcal{G} \cong \mathcal{F}(n)$. Hence, $f^{n-d}m = 0$ in $\Gamma(\mathcal{F}(n), X)$ and thus $\frac{m}{f^d} = \frac{f^{n-d}m}{f^n}$ is zero in $\Gamma_*(\mathcal{F})_{(f)}$. This shows injectivity. We now show surjectivity. Pick $t \in \mathcal{F}(D_+(f))$. By Lemma 1.9.1.18, 2, (which applies here as $D_+(f)$s are affine and finitely many whose intersection is again affine), we obtain a section $f^n t$ of $\mathcal{O}_X(1)^{\otimes n} \otimes \mathcal{F} \cong \mathcal{O}_X(n) \otimes \mathcal{F} \cong \mathcal{F}(n)$ over $D_+(f)$ which extends to a global section of $\mathcal{F}(n)$, say s. Consider s/f^n in $\Gamma_*(\mathcal{F})_{(f)}$, which maps to $s \otimes 1/f^n = t$ in $\mathcal{F}(D_+(f))$, as needed. □

Closed subschemes of $\mathbb{P}^n_A$

We can use these results to obtain a nice characterization of closed subschemes of projective schemes and an equivalent characterization of projective schemes over affine schemes. Denote by $\mathbb{P}^r_A = \text{Proj}(A[x_0, \dots, x_r])$ for a ring A.

Proposition 1.9.2.14. *Let $Y \hookrightarrow \mathbb{P}^r_A$ be a closed subscheme with ideal sheaf $\mathcal{I}_Y$ of the projective r-space over a ring A. Then $I = \Gamma_*(\mathcal{I}_Y)$ is a homogeneous ideal of $A[x_0, \dots, x_r]$ and we have*

$$Y \cong \text{Proj}\left(A[x_0, \dots, x_r]/I\right).$$

Proof. Let $S = A[x_0, \dots, x_r]$. The fact that I is a homogeneous ideal of S follows from exactness twisting functor (Remark 1.9.2.8), left exactness of global sections and $\Gamma_*(\mathcal{O}_X) = S$ of Proposition 1.9.2.12. In order to show that $Y \cong \text{Proj}(S/I)$, it is enough to show that they both define isomorphic ideal sheaves (Proposition 1.4.4.13, 3). The ideal sheaf of $\text{Proj}(S/I)$ is $\widetilde{I}$ by Proposition 1.8.2.8, 3 and the ideal sheaf of Y is $\mathcal{I}_Y$. Since $I = \Gamma_*(\mathcal{I}_Y)$, therefore the result follows from Proposition 1.9.2.13, 2. □

Proposition 1.9.2.15. *Let A be a ring. A scheme $Y \to \text{Spec}\,(A)$ is projective if and only if $Y \cong \text{Proj}(S)$ for a graded ring S with $S_0 = A$ and which is finitely generated by S_1 as an S_0-algebra.*

Proof. (L ⇒ R) We have a closed immersion $Y \to \mathbb{P}^r_A$. From Proposition 1.9.2.14, it follows that $Y \cong \text{Proj}(S)$ where $S = A[x_0, \dots, x_r]/I$, but S_0 might not be A. By Proposition 1.8.2.8, 2, we may assume I to not have any degree 0 component. Thus, S as defined will satisfy the necessary criterion.
(R ⇒ L) We have $S \cong A[x_0, \dots, x_r]/I$, so by Proposition 1.8.2.8, 1, we have a closed immersion $\text{Proj}(S) \to \mathbb{P}^r_A$. □

Very ample invertible modules

We now study modules which determine when a scheme is projective.

Definition 1.9.2.16 (Twisting modules). Let X be a scheme and consider $\mathbb{P}^n_X \to X$ to be the projective n-scheme over X. Consider the projection $p : \mathbb{P}^n_X \to \mathbb{P}^n_{\mathbb{Z}}$. The k^{th}-Serre twist sheaf over $\mathbb{P}^n_X$ are defined to be $p^*(\mathcal{O}(k))$ where $\mathcal{O}(k)$ is the k^{th}-Serre twist sheaf over the projective scheme $\mathbb{P}^n_{\mathbb{Z}}$.

What we have defined above is indeed a generalization of usual twisted sheaves available on projective schemes, as the following lemma shows.

Lemma 1.9.2.17. *Let $X = \text{Spec}(A)$. Denote $p : \mathbb{P}^n_X \to \mathbb{P}^n_{\mathbb{Z}}$ the projection map. Then,*

1. *$\mathbb{P}^n_X \cong \mathbb{P}^n_A$,*
2. *The twisting module $p^*(\mathcal{O}(k)) \cong \mathcal{O}_{\mathbb{P}^n_A}(k)$ under the above isomorphism.*

Proof. Item 1 follows from Lemma 1.8.2.12. For item 2, observe that the map p is obtained by composing with the isomorphism $\mathbb{P}^n_X \cong \mathbb{P}^n_A$ the canonical map $q : \mathbb{P}^n_A \to \mathbb{P}^n_{\mathbb{Z}}$, which is induced from the canonical map $\varphi : \mathbb{Z}[x_0, \dots, x_n] \to A[x_0, \dots, x_n]$ (see Proposition 1.8.2.6). Thus we wish to show that $q^*(\mathcal{O}(k)) \cong \mathcal{O}_{\mathbb{P}^n_A}(k)$. Denote $S = \mathbb{Z}[x_0, \dots, x_n]$ so that $\mathcal{O}(k) = \widetilde{S(k)}$. Hence, by Proposition 1.9.2.7, 4, we have

$$q^*(\mathcal{O}(k)) \cong \mathcal{O}_{\mathbb{P}^n_A}(k),$$

as needed. □

Definition 1.9.2.18 (Very ample invertible module). Let $X \to Y$ be a scheme over Y. An invertible module $\mathcal{L}$ over X is said to be very ample over Y if there is an immersion (Definition 1.4.4.9) $i : X \to \mathbb{P}^n_Y$ such that $i^*(\mathcal{O}(1)) \cong \mathcal{L}$.

Proposition 1.9.2.19. *Let Y be a Noetherian scheme. Then the following are equivalent:*

1. *Scheme $f : X \to Y$ is projective.*
2. *Scheme $f : X \to Y$ is proper and there exists a very ample invertible sheaf over X relative to Y.*

1.10 Divisors

The notion of divisors is one of the central tools for understanding the geometrical properties of a given scheme. Indeed, in the special case of curves in projective plane, a (Weil) divisor is just a formal linear combination of points of the curve. From this data, one can in-fact recover the embedding of the curve in the projective plane. Hence the data that divisors of a scheme stores is rich in geometric information.

We will first define the notion of Weil divisors in those schemes in which the points lying on codimension 1 subset of the scheme are regular (see Lemma 1.7.1.4 for a motivation behind the definition).

Definition 1.10.0.1. (**Regular in codimension 1**) A scheme X is said to be regular in codimension 1 if the local rings $\mathcal{O}_{X,p}$ which are of dimension 1 are regular.

Remark 1.10.0.2. All non-singular abstract varieties are regular in codimension 1 as all local rings are regular. In this section, we will be working with schemes which are noetherian integral separated and regular in codimension 1. We call them *Weil schemes*. All non-singular abstract varieties are Weil schemes.

1.10.1 Weil divisors & Weil divisor class group

We will define the notion of Weil divisors and divisor class group on Weil schemes.

Definition 1.10.1.1. (**Weil divisors**) Let X be a Weil scheme. A *prime divisor* is an integral closed subscheme of codimension 1. A *Weil divisor* is an element of the free abelian group generated by the set of all prime divisors, denoted Div (X). A Weil divisor is denoted $\sum_{i=1}^{k} n_i Y_i \in \text{Div } X$. A Weil divisor $\sum_i n_i Y_i$ is *effective* if $n_i \geq 0$ for all i.

Example 1.10.1.2 (Prime divisors of Spec $(\mathbb{Z})$)**.** Consider $X = \text{Spec}\,(Z)$. One can check immediately that a prime divisor $Y \hookrightarrow X$ is equivalent to the data of a non-zero prime p, where $Y = V(\langle p \rangle)$. Indeed, as X is of dimension 1, therefore any codimension 1 integral closed subscheme is just the closure of a non-generic point, which are all of the closed points of X.

We now look at a foundational result which will guide the further development. Its proof is important as it combines a lot of our previous knowledge.

Proposition 1.10.1.3. *Let X be a Weil scheme and $Y \subseteq X$ be a prime divisor with $\eta \in Y$ be its generic point. Then there is an injective map*

$$\text{PDiv}(X) \to \text{DVal}(K(X))$$

where PDiv *is the set of all prime divisors of X, $K(X)$ is the function field of X and* $\text{DVal}(K(X))$ *is the set of all discrete valuations over $K(X)$.*

Proof. Note that if Y is a prime divisor, then there is a generic point $\eta \in Y$. Since codim $(Y, X) = \dim \mathcal{O}_{X,\eta}$ by Lemma 1.7.1.4, therefore we obtain that $\mathcal{O}_{X,\eta}$ is a regular local ring of dimension 1. Now there is a special result for such rings, which in particular establishes equivalences of such rings with a lot of other type of rings. Indeed, by Theorem 16.10.1.8 we obtain that in our case

$\mathcal{O}_{X,\eta}$ is a DVR. As the fraction field of $\mathcal{O}_{X,\eta}$ (it is a domain as X is integral) is $K(X)$, the function field of X (Lemma 1.3.1.2), therefore we have a valuation $v : K(X) \to \mathbb{Z}$ whose valuation ring is $\mathcal{O}_{X,\eta}$. By Lemma 1.12.4.8 (which holds for Y as Y is separated by Corollary 1.12.4.5, 2), the valuation v uniquely determines the point $\eta \in X$ as v has center η because the valuation ring $\mathcal{O}_{X,\eta}$ of v dominates the local ring $\mathcal{O}_{X,\eta}$. As the information of point $\eta \in X$ yields the closed set $Y \subseteq X$, therefore the valuation $v : K(X) \to \mathbb{Z}$ uniquely determines the prime divisor Y. □

Remark 1.10.1.4. As a consequence, we can study a prime divisor via the valuation that comes through the Proposition 1.10.1.3. Indeed, for a prime divisor $Y \subseteq X$ and the associated discrete valuation $v_Y : K(X) \to \mathbb{Z}$, we can think of the value $v_Y(f)$ for some $f/g \in K(X) \setminus \{0\}$ to be telling us the number of poles that f/g has along Y. We can justify this via the proof of Proposition 1.10.1.3 as follows. For a prime divisor $Y \subseteq X$ with generic point $\eta \in Y$, the corresponding valuation is obtained by the fact that $\mathcal{O}_{X,\eta}$ is a DVR in our case. For a DVR R with fraction field K, the corresponding discrete valuation $v : K \to \mathbb{Z}$ can be thought of as an abstraction of the idea that we want to know how many poles a fraction $f/g \in K$ has and v provides that data to us. In particular, we think that if $v(f/g)$ is positive, then that tells us f/g has that many zeros in Y and if $v(f/g)$ is negative then that many poles in Y. We mostly have only this idea in mind when dealing with valuations.

Our next goal is to assign a divisor to any regular function in the function field of a Weil scheme X. To this end, the following proposition is essential for its well-definedness. We begin with the following observation.

Lemma 1.10.1.5. *Let X be a noetherian integral scheme. If $Z \subsetneq X$ is a proper closed subset, then there are finitely many prime divisors of X in Z.*

Proof. Let $Y \subseteq Z$ be a prime divisor of X. As $Z \subsetneq X$, therefore $\operatorname{codim} Z \geq 1$. Thus $1 = \operatorname{codim} Y \geq \operatorname{codim} Z \geq 1$, from which it follows that $\operatorname{codim} Y = \operatorname{codim} Z = 1$. Thus by Proposition 1.7.1.1, $Y = Z_\alpha$ where Z_α is an irreducible component of Z. As X is noetherian, then so is Z. It follows that Z has finitely many irreducible components and hence prime divisors in Z are also finite, as required. □

The following shows that a regular function can only have zeroes at finitely many prime divisors on an affine Weil scheme.

Lemma 1.10.1.6. *Let $X = \operatorname{Spec}(A)$ be an affine Weil scheme. If $f \in A$ is a regular function on X, then $v_Y(f) \neq 0$ only for finitely many prime divisors $Y \subseteq X$.*

Proof. Let $K = Q(A)$ be the fraction field of A. Pick a prime divisor $Y = V(\mathfrak{p}) \subseteq X$ with generic point $\mathfrak{p} \in X$ and $v_Y : K \to \mathbb{Z}$ be the corresponding discrete valuation. As $f \in A$, therefore $v_Y(f) \geq 0$. Note from definition of valuation ring that $v_Y(f) > 0$ if and only if $f/1 \in \mathfrak{p}A_{\mathfrak{p}}$. The latter condition is equivalent to $f \in \mathfrak{p}$. Hence, $v_Y(f) > 0$ if and only if $V(f) \supseteq V(\mathfrak{p})$. By Lemma 1.10.1.5, it follows that $V(f)$ contains only finitely many prime divisors. Consequently, $v_Y(f) > 0$ only for finitely many primes Y, as required. □

Proposition 1.10.1.7. *Let X be a Weil scheme. Denote by $v_Y : K(X) \to \mathbb{Z}$ the associated discrete valuation corresponding to a prime divisor $Y \subseteq X$. Then for each $f \in K(X)^\times$, the integer $v_Y(f)$ is non-zero only for finitely many prime divisors $Y \subseteq X$.*

Proof. Pick $f \in K(X)^\times$ to be a rational function. Let $\eta \in X$ be the generic point and $U = \mathrm{Spec}\,(A)$ be an open affine of X, which will thus contain η. As $K(X) \cong K(U)$, therefore $f = \frac{g}{h}$ in $K(U)$. Consider $D(h) \subseteq U$, where h doesn't vanish. We deduce that $f \in \mathcal{O}_X(D(h)) = A_h$ is a regular function, that is, f is regular over an affine open $D(h) \subseteq X$. Replace U by $D(h)$ to assume that f is regular over U.

Next, consider a prime divisor $Y \hookrightarrow X$. It suffices to show there are finitely many prime divisors contained in $X - U$ and finitely prime divisors intersecting U. The former follows from Lemma 1.10.1.5. For the latter, first observe that prime divisors of X intersecting U corresponds to $\eta \in U$ which are non-closed points whose local rings are regular in codimension 1. Thus, they correspond to prime divisors of U, and since f is a regular function on U, it follows by Lemma 1.10.1.6 that there are finitely many prime divisors of U, as required. ☐

Having deduced the above results, we can now define a fundamental equivalence relation on Div (X).

Definition 1.10.1.8 (Principal divisors, linear equivalence & class group). Let X be a Weil scheme and $f \in K(X)^\times$ be a non-zero rational function. Define the following effective divisor on X which is well-defined by Proposition 1.10.1.7:

$$\langle f \rangle := \sum_{Y \in \mathrm{PDiv}(X)} v_Y(f) \cdot Y.$$

We call $\langle f \rangle$ the principal divisor generated by f. This defines a group homomorphism

$$\begin{aligned} r^1 : K(X)^\times &\longrightarrow \mathrm{Div}\,(X) \\ f &\longmapsto \langle f \rangle. \end{aligned}$$

Indeed, $\langle fg \rangle = \langle f \rangle + \langle g \rangle$ follows from definition of valuations. Any principal divisor is said to be linearly equivalent to 0. We then define class group of X to be the cokernel of $r^1 : K(X)^\times \to \mathrm{Div}\,(X)$:

$$\mathrm{Cl}(X) := \mathrm{Div}\,(X)/\mathrm{Im}\,(r^1).$$

The abelian group $\mathrm{Cl}(X)$ is also called Weil divisor class group.

1.10.2 Weil divisors on affine schemes

We will study Weil divisors over affine Weil schemes (see Remark 1.10.0.2), to portray the type of information that they contain. However, we will introduce something more general (Krull domains) but will show later that with noetherian hypothesis, we have done no extra work (see Corollary 1.10.2.5).

Definition 1.10.2.1 (Weil & Krull domains). Let R be a domain. We call R to be a Weil domain if $R_\mathfrak{p}$ is a DVR (equivalently regular, or Dedekind by Theorem 16.10.1.8) for all $\mathfrak{p} \in \mathrm{Spec}\,(R)$ such that $\mathrm{ht}\,(\mathfrak{p}) = 1$ (equivalently, $\dim R_\mathfrak{p} = 1$ by Lemma 16.8.1.4). A Weil domain is a Krull domain if moreover $R = \bigcap_{\mathrm{ht}\,(\mathfrak{p})=1} R_\mathfrak{p} = 1$ in $F = Q(R)$ and every non-zero $r \in R$ is in only finitely many prime ideals of height 1. Hence any Krull domain is in particular a Weil domain.

The first observation is that noetherian normal domains are Krull domains.

Proposition 1.10.2.2. *Let R be a ring.*

1. *If R is a noetherian normal domain, then R is a Weil domain.*
2. *If R is a noetherian Weil domain, then R is a Krull domain.*

Proof. 1. Let $\mathfrak{p} \in \operatorname{Spec}(R)$ be a height 1 prime ideal. As localization of normal domains is normal (Proposition 16.7.2.8), therefore $R_{\mathfrak{p}}$ is a noetherian local domain of dimension 1 which is normal. By Theorem 16.10.1.8, we deduce that $R_{\mathfrak{p}}$ is a DVR, as required.

2. We first wish to show that $R = \bigcap_{\text{ht}(\mathfrak{p})=1} R_{\mathfrak{p}}$. We need only show ($\supseteq$). Indeed, let $\frac{r}{s} \in R_{\mathfrak{p}}$ for all $\mathfrak{p}$ of height 1. It follows that $s \notin \mathfrak{p}$ for all height 1 prime ideals. We claim that any non-zero element of R is contained in a height 1 prime ideal. Indeed, this is the content of Krull's principal ideal theorem (Theorem 16.8.3.2). Thus $s \in R$ is a unit, hence $\frac{r}{s} = rs^{-1} \in R$.

Finally, we wish to show that any non-zero element $r \in R$ is in only finitely many height 1 primes. By going modulo rR and recalling that R is a domain, we need only show that there are finitely many height 0 primes in $S = R/rR$. As S is noetherian and minimal primes are equivalent to height 0 primes, so we reduce to showing that a noetherian ring has finitely many minimal primes. Indeed, this is the content of Lemma 1.2.1.8. $\square$

Example 1.10.2.3. A Dedekind domain is therefore a Krull domain of dimension 1.

Any Krull domain is normal.

Lemma 1.10.2.4. *If R is a Krull domain, then R is a normal domain.*

Proof. As $R = \bigcap_{\text{ht}(\mathfrak{p})=1} R_{\mathfrak{p}}$ where each $R_{\mathfrak{p}}$ is a normal domain in particular, therefore the integral closure of R in $F = Q(R)$ is contained in each $R_{\mathfrak{p}}$, thus is contained in $\bigcap_{\text{ht}(\mathfrak{p})=1} R_{\mathfrak{p}}$ which is R, as required. $\square$

In summary, we have now shown the following important equivalences.

Corollary 1.10.2.5. *Let R be a ring. Then the following are equivalent:*

1. *R is a noetherian normal domain.*
2. *R is a noetherian Weil domain.*
3. *R is a noetherian Krull domain.*

Proof. Follows from Proposition 1.10.2.2 and Lemma 1.10.2.4 $\square$

Remark 1.10.2.6. Consequently, a Krull domain can be thought of as an abstraction of all the nice "divisorial properties" that we expect from noetherian normal domains, so that we can talk about it in non-noetherian settings.

The following is what we expect and the following is indeed true.

Proposition 1.10.2.7. *Let $X = \operatorname{Spec}(R)$ be an affine scheme. Then the following are equivalent:*

1. *X is an affine Weil scheme.*
2. *R is a noetherian Weil domain.*

Proof. Suppose X is an affine Weil scheme. Then X is in particular noetherian and integral. By Lemmas 1.4.1.3 and 1.4.2.2, we deduce that R is a noetherian domain. As X is regular in codimension 1, $R_{\mathfrak{p}}$ is regular if ht $(\mathfrak{p}) = 1$. Consequently, $R_{\mathfrak{p}}$ is a noetherian local domain of dimension 1. Hence by Theorem 16.10.1.8 we deduce that $R_{\mathfrak{p}}$ is a DVR, as required.

Now if R is a noetherian Weil domain, then X is a noetherian integral scheme regular in codimension 1. As affine schemes are separated, so we conclude the proof. $\square$

As for most purposes we do enforce noetherian hypothesis, thus by above result it is noetherian normal domains, i.e. noetherian Weil domains, which are most important for us. We now define Weil divisors on Weil domains just as we did in Definition 1.10.1.1.

Definition 1.10.2.8 (**Weil divisors on Weil domains**)**.** Let R be a Weil domain. A *prime divisor* is a height 1 prime ideal. A *Weil divisor* is an element of the free abelian group generated by all prime divisors. A Weil divisor is denoted by $D = \sum_i n_i[\mathfrak{p}_i]$. The group of all Weil divisors over R is denoted by Div (R). An *effective Weil divisor* is $D = \sum_i n_i[\mathfrak{p}_i]$ where $n_i \geq 0$ for all i. We denote by PDiv(R) the set of prime divisors, that is, the generating set of Div (R).

Every Weil domain comes equipped with a discrete valuation at height 1 prime.

Definition 1.10.2.9 (**$\mathfrak{p}$-adic valuation**)**.** Let R be a Weil domain and $\mathfrak{p}$ be any height 1 prime ideal of R. Then $R_{\mathfrak{p}}$ is a DVR and thus has a function

$$\begin{aligned} \nu_{\mathfrak{p}} : \mathrm{Cart}(R) &\longrightarrow \mathbb{Z} \\ I &\longmapsto \nu_{\mathfrak{p}}(I) \end{aligned}$$

where since $R_{\mathfrak{p}}$ is a PID in particular (Proposition 16.10.1.9), so we can write the invertible ideal $I_{\mathfrak{p}}$ as $I_{\mathfrak{p}} = \mathfrak{p}^{\nu_{\mathfrak{p}}(I)}R_{\mathfrak{p}}$. We call $\nu_{\mathfrak{p}}$ as the $\mathfrak{p}$-adic valuation of R for height 1 prime ideal $\mathfrak{p}$.

Remark 1.10.2.10. Indeed, the name is justified by the simple observation that if we consider the usual valuation $\nu : Q(R_{\mathfrak{p}}) \to \mathbb{Z}$ given by $f \mapsto \nu(f)$ such that $ut^{\nu(f)} = f$ where $t \in R_{\mathfrak{p}}$ is the local parameter of the DVR, then $\nu_{\mathfrak{p}}$ is the function $\mathrm{Cart}(R) \to \mathrm{Cart}(R_{\mathfrak{p}}) = Q(R_{\mathfrak{p}})^{\times} \xrightarrow{\nu} \mathbb{Z}$ which is given by $I \mapsto I_{\mathfrak{p}} = fR_{\mathfrak{p}} \mapsto \nu(f)$.

Construction 1.10.2.11 (The Cart-Div homomorphism)**.** Let R be a Krull domain. We define a group homomorphism

$$\begin{aligned} \nu : \mathrm{Cart}(R) &\longrightarrow \mathrm{Div}\ (R) \\ I &\longmapsto \sum_{\mathrm{ht}\ (\mathfrak{p})=1} \nu_{\mathfrak{p}}(I)[\mathfrak{p}] \end{aligned}$$

which is well defined as $\nu_{\mathfrak{p}}(I) \neq 0$ only for finitely many $\mathfrak{p}$ of height 1 by the third axiom of Krull domains[37]. This is a group homomorphism as $\nu(IJ) = \sum \nu_{\mathfrak{p}}(IJ)[\mathfrak{p}] = \sum(\nu_{\mathfrak{p}}(I) + \nu_{\mathfrak{p}}(J))[\mathfrak{p}] = \nu(I) + \nu(J)$. We call this the Cart-Div homomorphism. We also call the divisor $\nu(I)$ corresponding to an invertible ideal to be the *associated divisor of I*.

[37] One may see this as follows. As $I \subseteq \frac{a}{b}R$, $a \neq b$, therefore for any height 1 prime $\mathfrak{p}$, we get $I_{\mathfrak{p}} = t_{\mathfrak{p}}^{\nu_{\mathfrak{p}}}R_{\mathfrak{p}} \subseteq \frac{a}{b}R_{\mathfrak{p}}$ where $t_{\mathfrak{p}} \in R_{\mathfrak{p}}$ is the local parameter of the DVR. Consequently, $\frac{a}{b}t_{\mathfrak{p}}^{\mu_{\mathfrak{p}}} = t_{\mathfrak{p}}^{\nu_{\mathfrak{p}}}$. Note that $\nu_{\mathfrak{p}} = 0$ if and only if $\frac{a}{b} \cdot R_{\mathfrak{p}} = R_{\mathfrak{p}}$ which in turn happens if and only if $a \notin \mathfrak{p}$ and $b \notin \mathfrak{p}$, i.e. if and only if $\frac{a}{b}$ is a unit of $R_{\mathfrak{p}}$. Hence, $\nu_{\mathfrak{p}} \neq 0$ if and only if $\frac{a}{b}$ or $\frac{b}{a}$ is a non-unit of $R_{\mathfrak{p}}$. As $\frac{a}{b}$ is a non-unit of $R_{\mathfrak{p}}$ if and only if $a \notin \mathfrak{p}$ and there are only finitely many height 1 primes containing a, thus there are only finitely many height 1 primes $\mathfrak{p}$ such that $\frac{a}{b} \in R_{\mathfrak{p}}$ is non-unit. Similarly for $\frac{b}{a}$. This shows that only for finitely many height 1 primes do we have that $\nu_{\mathfrak{p}}(I) \neq 0$.

Definition 1.10.2.12 (Principal divisors). Let R be a Krull domain. For any invertible ideal $I \in \mathrm{Cart}(R)$, the divisor $\nu(I) = \sum \nu_{\mathfrak{p}}(I)[\mathfrak{p}]$ is called the principal divisor of I. Any divisor in the image of ν will be called a principal divisor.

An immediate question will be to see what do effective principal divisors correspond to. Indeed, they exactly correspond to ideals of R.

Lemma 1.10.2.13. *Let R be a Krull domain. Then there is a bijection*

$$\{\textit{Effective principal divisors on } R\} \leftrightsquigarrow \{\textit{Ideals of } R\}.$$

Proof. Indeed, if $D = \nu(I)$ is an effective principal divisor, then $\nu_{\mathfrak{p}}(I) \geq 0$ for each $\mathfrak{p}$ of height 1. Consequently, $I_{\mathfrak{p}}$ is an ideal of $R_{\mathfrak{p}}$ for each $\mathfrak{p}$ of height 1. Thus, $I \subseteq \bigcap_{\mathrm{ht}\,(\mathfrak{p})=1} I_{\mathfrak{p}} \subseteq \bigcap_{\mathrm{ht}\,(\mathfrak{p})=1} R_{\mathfrak{p}} = R$, as required. For the converse, repeat the same in reverse. □

Lemma 1.10.2.14. *Let R be a Krull domain. Then the Cart-Div homomorphism*

$$\nu : \mathrm{Cart}(R) \to \mathrm{Div}\,(R)$$

is injective.

Proof. Indeed, if $\nu(I) = 0$, then $I_{\mathfrak{p}} = R_{\mathfrak{p}}$ for all height 1 primes $\mathfrak{p}$. As there is an invertible ideal J such that $IJ = R$, therefore $\nu(I) + \nu(J) = 0$, from which it follows that $\nu(J) = 0$ as well. As $I \subseteq \bigcap_{\mathrm{ht}\,(\mathfrak{p})=1} I_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\,(\mathfrak{p})=1} R_{\mathfrak{p}} = R$ and similarly for J, thus, I, J are ideals of R such that $IJ = R$. Consequently, as $R = IJ \subseteq I \cap J \subseteq I, J$, thus we get that $I = J = R$, as required. □

We now define the divisor class group of a Krull domain.

Definition 1.10.2.15 (Divisor class group & Pic-Cl map). Let R be a Krull domain. The Weil divisor class group $\mathrm{Cl}(R)$ of R is defined to be the cokernel of the composite

$$F^\times \xrightarrow{\mathrm{div}} \mathrm{Cart}(R) \xrightarrow{\nu} \mathrm{Div}\,(R).$$

That is,

$$\mathrm{Cl}(R) = \frac{\mathrm{Div}\,(R)}{\mathrm{Im}\,(\nu \circ \mathrm{div})}.$$

We write $\nu \circ \mathrm{div} : F^\times \to \mathrm{Div}\,(R)$ as div as well.

As $\mathrm{Pic}(R) = \mathrm{CoKer}\,(\mathrm{div} : F^\times \to \mathrm{Cart}(R))$, $\mathrm{Cl}(R) = \mathrm{CoKer}\,(\mathrm{div} : F^\times \to \mathrm{Div}\,(R))$ and we have a map $\nu : \mathrm{Cart}(R) \to \mathrm{Div}\,(R)$, thus by universal property of cokernels, we get a map

$$\tilde{\nu} : \mathrm{Pic}(R) \longrightarrow \mathrm{Cl}(R)$$

which we call the Pic-Cl map.

Remark 1.10.2.16 (Div-Cl sequence)**.** We summarize the whole discussion by observing the following exact sequence:

$$1 \to R^\times \to F^\times \stackrel{\mathrm{div}}{\to} \mathrm{Div}\,(R) \to \mathrm{Cl}(R) \to 0.$$

The exactness at $F^\times$ follows from the second axiom of Krull domains.

Its an easy observation by 4-lemma that the Pic-Cl map is injective as well.

Lemma 1.10.2.17 (Cart-Pic to Div-Cl)**.** *Let R be a Krull domain. Then the Pic-Cl map $\tilde{\nu}$: $\mathrm{Pic}(R) \to \mathrm{Cl}(R)$ is injective. Moreover, the following is commutative:*

$$\begin{tikzcd}
1 \arrow[r] & R^\times \arrow[r] \arrow[d, "\mathrm{id}"'] & F^\times \arrow[r, "\mathrm{div}"] \arrow[d, "\mathrm{id}"'] & \mathrm{Cart}(R) \arrow[r] \arrow[d, hook, "\nu"'] & \mathrm{Pic}(R) \arrow[r] \arrow[d, hook, "\tilde{\nu}"'] & 0 \\
1 \arrow[r] & R^\times \arrow[r] & F^\times \arrow[r, "\mathrm{div}"] & \mathrm{Div}\ (R) \arrow[r] & \mathrm{Cl}(R) \arrow[r] & 0
\end{tikzcd}.$$

Proof. Commutativity is clear. Injectivity of $\tilde{\nu}$ is from 4-lemma. □

A simple corollary yields when Picard group and Weil divisor class groups are same.

Corollary 1.10.2.18. *Let R be a Krull domain. Then the following are equivalent:*
1. $\nu : \mathrm{Cart}(R) \to \mathrm{Div}\ (R)$ *is an isomorphism.*
2. $\tilde{\nu} : \mathrm{Pic}(R) \to \mathrm{Cl}(R)$ *is an isomorphism.*

Proof. Both sides are immediate from 5-lemma. □

Relative Weil divisors

As in homology, it is necessary at times to study *relative invariants*. Indeed, same is true for Weil divisors, as some results mentioned later will show.

Definition 1.10.2.19 (**Relative Weil divisors**)**.** Let R be a Krull domain and $S \subseteq R$ be a multiplicative set. Define Div $(R, S^{-1}R)$ to be the free abelian group generated by height 1 primes $\mathfrak{p}$ such that $\mathfrak{p} \cap S \neq \emptyset$. We call Div $(R, S^{-1}R)$ Weil divisors on R relative to S.

Remark 1.10.2.20. It is immediate that

$$\mathrm{Div}\ (R) = \mathrm{Div}\ (S^{-1}R) \oplus \mathrm{Div}\ (R, S^{-1}R).$$

Furthermore, we may define the map div as usual ($F = Q(R)$):

$$\begin{aligned} \mathrm{div} : F^\times &\longrightarrow \mathrm{Div}\ (R, S^{-1}R) \\ f &\longmapsto \sum_{\mathrm{ht}\ \mathfrak{p}=1, \mathfrak{p}\cap S \neq \emptyset} \nu_{\mathfrak{p}}(fR)[\mathfrak{p}]. \end{aligned}$$

The main result here is the following relative version of Div-Cl exact sequence.

Proposition 1.10.2.21. *Let R be a Krull domain and $S \subseteq R$ be a multiplicative set. Then, the following sequence is exact:*

$$1 \to R^\times \to (S^{-1}R)^\times \stackrel{\mathrm{div}}{\to} \mathrm{Div}\ (R, S^{-1}R) \to \mathrm{Cl}(R) \to \mathrm{Cl}(S^{-1}R) \to 0.$$

Proof. To check exactness at $(S^{-1}R)^{\times}$, we see that if $\mathrm{div}(f) = 0$ for some $f \in S^{-1}R$, then $\nu_{\mathfrak{p}}(fR) = 0$ for all $\mathfrak{p}$ of height 1 and $\mathfrak{p} \cap S \neq \emptyset$. Thus, $fR_{\mathfrak{p}} = R_{\mathfrak{p}}$ for all such primes. As $f \in (S^{-1}R)^{\times}$, consequently if $f = \frac{a}{b}$, then $a, b \in S$. Hence if $\mathfrak{p} \cap S = \emptyset$, then $S \subseteq R \setminus \mathfrak{p}$ and thus $a, b \in R \setminus \mathfrak{p}$ so that $f \in R_{\mathfrak{p}}$ is a unit. It follows that

$$fR = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1} fR_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} fR_{\mathfrak{p}} \cap \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S\neq\emptyset} fR_{\mathfrak{p}}.$$

By above, $\bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} fR_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} R_{\mathfrak{p}}$ and since $f \in (S^{-1}R)^{\times}$, so $\bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S\neq\emptyset} fR_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S\neq\emptyset} R_{\mathfrak{p}}$. Finally, we get $fR = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1} R_{\mathfrak{p}} = R$, so that $f \in R^{\times}$, as required. Conversely, if $u \in R^{\times}$, then $\mathrm{div}(u) = 0$.

To see exactness at $\mathrm{Div}\ (R, S^{-1}R)$, observe that if $D_S \in \mathrm{Div}\ (R, S^{-1}R)$ is a relative divisor such that $[D_S] = 0$ in $\mathrm{Cl}(R)$, then $D_S = \nu(fR)$ for some $f \in F^{\times}$, i.e. D_S is a principal divisor. Then, since $D_S = \nu(fR) \in \mathrm{Div}\ (R, S^{-1}R)$ and is a principal divisor as a divisor on R, thus we deduce that $\nu_{\mathfrak{p}}(fR) = 0$ for all $\mathfrak{p}$ of height 1 such that $\mathfrak{p} \cap S = \emptyset$, that is, $f \in R_{\mathfrak{p}}$ is a unit for all such primes. Consequently, $fR_{\mathfrak{p}} = R_{\mathfrak{p}}$ for all $\mathfrak{p} \cap S = \emptyset$ of height 1. Now, as $S^{-1}R$ is a Krull domain, we get

$$f \cdot (S^{-1}R) = f \cdot \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} S^{-1}R_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} S^{-1}fR_{\mathfrak{p}} = \bigcap_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S=\emptyset} S^{-1}R_{\mathfrak{p}} = S^{-1}R,$$

as required. Conversely, if $f \in (S^{-1}R)^{\times}$, then $\mathrm{div}(f)$ is by construction a principal divisor on R, so that its image in $\mathrm{Cl}(R)$ will be 0.

Exactness at $\mathrm{Cl}(R)$ is clear as if $[D]$ is 0 in $\mathrm{Cl}(S^{-1}R)$, then D can be written as a sum of two divisors, one on $S^{-1}R$, say D_1, and other a relative on R, say D_2, where D_1 is principal. Hence, $[D] = [D_2]$ and since D_2 is a relative divisor on R, thus its in image of $\mathrm{Div}\ (R, S^{-1}R) \to \mathrm{Cl}(R)$.

Finally, exactness at $\mathrm{Cl}(S^{-1}R)$ is clear as any $[D_S] \in \mathrm{Cl}(S^{-1}R)$ such that $D_S = \sum_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\cap S\neq\emptyset} n_{\mathfrak{p}}[\mathfrak{p}]$ is also a divisor on R, so that $D_S \in \mathrm{Div}\ (R)$ and thus defines the class $[D_S] \in \mathrm{Cl}(R)$, whose image in $\mathrm{Cl}(S^{-1}R)$ is $[D_S]$. This completes the proof. $\square$

We can now prove an important result.

Proposition 1.10.2.22. *Let R be a Krull domain and $f \in R$ be a prime element (that is, fR is a prime ideal). Then*

$$\mathrm{Cl}(R) \cong \mathrm{Cl}(R_f).$$

Proof. By relative Weil divisors exact sequence of Propsosition 1.10.2.21, we get that the map

$$\mathrm{Div}\ (R, R_f) \to \mathrm{Cl}(R) \to \mathrm{Cl}(R_f) \to 0$$

is exact. We need only show that the image of $\mathrm{Div}\ (R, R_f)$ in $\mathrm{Cl}(R)$ is 0. Indeed, let $D \in \mathrm{Div}\ (R, R_f)$ be a relative divisor. Thus

$$D = \sum_{\mathrm{ht}\ (\mathfrak{p})=1, \mathfrak{p}\ni f} n_{\mathfrak{p}}[\mathfrak{p}].$$

Now $fR \subseteq \mathfrak{p}$ and fR is prime by hypothesis. As $\mathrm{ht}\ (\mathfrak{p}) = 1$ and R is a domain, thus $fR = \mathfrak{p}$. Hence, only height 1 prime containing f is fR. Hence $\mathrm{Div}\ (R, R_f) = \mathbb{Z}([fR]) \cong \mathbb{Z}$. As fR is a principal ideal, thus it can be shown that its image in $\mathrm{Cl}(R)$ is 0 as $[fR] = \nu(fR)$ (fR is prime), and thus the image of $\mathrm{Div}\ (R, R_f) \to \mathrm{Cl}(R)$ is 0, as required. $\square$

1.10.3 Cartier divisors & Cartier divisor class group

Definition 1.10.3.1 (Cartier divisor and $\mathrm{CaCl}(X)$**).** Let X be a scheme. By $\mathcal{K}$, denote the sheaf associated to the presheaf

$$K : U \mapsto Q(\mathcal{O}_X(U))$$

where $Q(A)$ for a ring A is the total quotient ring, obtained by localizing A at the multiplicative set S of all non zero-divisors of A. Note that the resulting map $A \to S^{-1}A$ is injective. If A is a domain then it is the usual quotient field. We call $\mathcal{K}$ the total quotient sheaf. Note that the map $\mathcal{O}_X^\times \to \mathcal{K}^\times$ induced by localization as above is injective. Thus we have the short exact sequence

$$0 \to \mathcal{O}_X^\times \to \mathcal{K}^\times \to \mathcal{K}^\times/\mathcal{O}_X^\times \to 0.$$

Applying global sections, we get a map

$$\pi : \Gamma(\mathcal{K}^\times) \to \Gamma(\mathcal{K}^\times/\mathcal{O}_X^\times).$$

A Cartier divisor on X is a global section of $\mathcal{K}^\times/\mathcal{O}_X^\times$. The group of Cartier divisors of X, denoted $\mathrm{Cart}(X)$, is defined to be $\Gamma(\mathcal{K}^\times/\mathcal{O}_X^\times)$. A Cartier divisor is principal if it is in the image of π. The cokernel of the map π is defined to be the Cartier class group

$$\mathrm{CaCl}(X) := \mathrm{CoKer}\,(\pi) = \frac{\Gamma(\mathcal{K}^\times/\mathcal{O}_X^\times)}{\mathrm{Im}\,(\pi)}.$$

Here are some first observations.

Lemma 1.10.3.2. *If X is integral, then the total quotient sheaf $\mathcal{K}$ is isomorphic to the constant sheaf $\underline{K(X)}$.*

Proof. As the total quotient presheaf is given by $K : U \mapsto Q(\mathcal{O}_X(U))$ since each $\mathcal{O}_X(U)$ is a domain (Lemma 1.4.2.2), therefore there is a map of presheaves $\varphi : K \to \underline{K(X)}$ as there is an isomorphism of $K(U) \cong K(X)$ for any open affine. Consequently, φ is an isomorphism on stalks. By universal property, there is a map $\tilde{\varphi} : \mathcal{K} \to \underline{K(X)}$ which is isomorphism on stalks. It follows that $\tilde{\varphi}$ is an isomorphism, as required. □

Lemma 1.10.3.3. *Let X be a scheme. Then the following are equivalent:*

1. *D is a Cartier divisor on X.*
2. *There is an open cover $\{U_i\}$ of X and non-zero rational functions $f_i \in \mathcal{K}^\times(U_i)$ such that on $U_i \cap U_j$, there exists $c_{ij} \in \mathcal{O}_X^\times(U_i \cap U_j)$ and*

$$f_i = c_{ij} f_j \text{ in } \mathcal{K}^\times(U_i \cap U_j).$$

Proof. We have a short exact sequence of sheaves

$$0 \to \mathcal{O}_X^\times \to \mathcal{K}^\times \to \mathcal{K}^\times/\mathcal{O}_X^\times \to 0.$$

Hence by definition of surjectivity of the map $\pi : \mathcal{K}^\times \to \mathcal{K}^\times/\mathcal{O}_X^\times$, we get that a Cartier divisor D is gives an open cover $\{U_i\}$ together with $f_i \in \mathcal{K}^\times(U_i)$ such that $\pi_{U_i}(f_i) = D|_{U_i}$. It follows that on $U_i \cap U_j$, we have $\pi_{U_i \cap U_j}(f_i) = \pi_{U_i \cap U_j}(f_j)$ and thus $f_i f_j^{-1} \in \mathrm{Ker}\,(\pi_{U_i}) = \mathrm{Ker}\,(\pi)(U_i) = \mathcal{O}_X^\times(U_i)$, as required. Conversely, such a data gives rise to a matching family for the sheaf $\mathcal{K}^\times/\mathcal{O}_X^\times$, as required. □

1.10.4 Cartier divisors on affine schemes

We discuss Cartier divisors first on affine schemes. Here, we will see that a Cartier divisor is nothing but an invertible ideal.

Definition 1.10.4.1 (**Fractional & invertible ideals**). Let R be a domain and $F = Q(R)$. A fractional ideal is a non-zero R-module $I \subseteq F$ such that there exists $f \in F$ for which $I \subseteq f \cdot R$ in F. That is, I consists of some R-multiples of a fraction $f \in F$. Note that if I and J are fractional, then so is IJ. A fractional ideal I is said to be invertible, if there exists a fractional ideal J such that $IJ = R$. The set of all fractional ideals form an abelian group with identity being R which we denote by $\mathrm{Cart}(R)$ defined to be the abelian group of Cartier divisors on Spec (R) (or just R).

Remark 1.10.4.2. An invertible ideal over domain R can equivalently be defined to be an R-module $I \subseteq F$ such that there exists $b \in R$ for which $bI \leq \langle a \rangle$ for some $a \in R$. That is, bI is an ideal of R which is contained in some principal ideal.

Example 1.10.4.3. Let R be a domain and $n \in \mathbb{Z} \setminus \{0\}$ with $\mathrm{char}(R) \neq n$. Denote $(\frac{1}{n}) = \frac{1}{n}R \subseteq F$ and $(n) = nR \subseteq F$ be two fractional ideals (where $n = 1 + \cdots + 1$, n-times). Clearly $(\frac{1}{n}) \cdot (n) = R$. Thus, $(\frac{1}{n})$ is a Cartier divisor on R.

Remark 1.10.4.4 (Divisor map). For any domain R with fraction field F, we have a group homomorphism

$$\begin{aligned} \mathrm{div} : F^{\times} &\longrightarrow \mathrm{Cart}(R) \\ f &\longmapsto fR. \end{aligned}$$

It is interesting to note when is this an isomorphism. An immediate calculation shows that it is so when R is a PID. Thus, $\mathrm{Cart}(R)$ has information about factorization in R.

By analyzing kernel and cokernel of the divisor map, we get a useful exact sequence.

Theorem 1.10.4.5 (Cart-Pic sequence). *Let R be a domain.*

1. *The map*

$$\begin{aligned} \mathrm{Cart}(R) &\longrightarrow \mathrm{Pic}(R) \\ I &\longmapsto [I] \end{aligned}$$

   is a group homomorphism. That is, $I \otimes_R J \cong IJ$, for any two $I, J \in \mathrm{Cart}(R)$. This is also true for any two line bundles $I, J \in \mathrm{Pic}(R)$ such that $I, J \subseteq F$.
2. *We have* $\mathrm{Ker}\,(\mathrm{div} : F^{\times} \to \mathrm{Cart}(R)) = R^{\times}$.
3. *The following is an exact sequence*

$$1 \to R^{\times} \to F^{\times} \stackrel{\mathrm{div}}{\to} \mathrm{Cart}(R) \to \mathrm{Pic}(R) \to 0.$$

Proof. 1. We first show well-definedness of $\mathrm{Cart}(R) \to \mathrm{Pic}(R)$. To this end, we need to show that any invertible ideal is a rank 1 projective module. Indeed, as there is an invertible ideal J such that $IJ = R$, thus there exists $\{x_i\} \subseteq I$ and $\{y_i\} \subseteq I$ finitely many such that $1 = x_1y_1 + \cdots + x_ny_n$. Using this, we immediately get maps $I \to R^n \to I$ whose composite is identity. Consequently, we get that I is a direct summand of R^n, as required.

As R is a domain, so Spec (R) is in particular connected. Consequently, we need only find $\dim_F(I \otimes_R F)$. As $I \otimes_R F = I \otimes_R R_\circ = I_\circ = F$, thus, $\mathrm{rank}_\circ(F) = 1$ and by connectedness, $\mathrm{rank}(I)$ is a constant map to 1. This shows that I is a line bundle, hence $[I] \in \mathrm{Pic}(R)$.

Next, we wish to show that for any $I, J \in \mathrm{Cart}(R)$, we get $I \otimes_R J \cong IJ$. This will show that the above map is a group homomorphism, as required. To this end, observe that we have a map

$$\begin{aligned} \varphi : I \otimes_R J &\longrightarrow IJ \\ x \otimes y &\longmapsto xy. \end{aligned}$$

We claim that φ is an isomorphism. Indeed, as I is a line bundle, therefore I is a projective R-module. Consequently it is flat and thus $I \otimes_R -$ is exact. As $J \hookrightarrow F$ is the inclusion map, therefore $I \otimes_R J \hookrightarrow I \otimes_R F \cong F$ is also an inclusion. Note that $I \otimes_R F \cong F$ is the map given by $x \otimes y \mapsto xy$. Hence, we have shown that φ is injective. Surjectivity of φ is immediate, so that φ is an isomorphism.

2. Let $f \in F^\times$ be such that $\mathrm{div}(f) = fR = R$, then f is a unit of R, as required.

3. We need only show exactness at $\mathrm{Cart}(R)$ and surjectivity of $\mathrm{Cart}(R) \to \mathrm{Pic}(R)$. We first show the former. An invertible ideal $I \in \mathrm{Cart}(R)$ is in the kernel iff $I \cong R$ as an R-module. If $\varphi : R \to I$ is the isomorphism, then $I \cong fR$ where $f = \varphi(1)$, as required.

Next we show surjectivity of $\mathrm{Cart}(R) \to \mathrm{Pic}(R)$. To this end, we have to show that any line bundle L over R is isomorphic to an invertible module I on R. Indeed, as L is rank 1, therefore $L \otimes_R F \cong F$. As $R \hookrightarrow F$ and L is projective hence flat, thus $L \cong L \otimes_R R \to L \otimes_R F \cong F$ is injective. Let the image of L in F be I. We claim that $I \subseteq F$ is an invertible module. Indeed, as I is finitely generated, therefore $I = f_1 R + \cdots + f_n R$ for $\frac{a_i}{b_i} = f_i \in F$, which we may write as $I \subseteq \frac{1}{b_1 \ldots b_n} R$, so I is fractional. To see that I is invertible, let $J \subseteq F$ be the fractional ideal corresponding to $\check{L}$. As $L \otimes_R \check{L} \cong R$ in $\mathrm{Pic}(R)$, it follows that $L \otimes_R \check{L} \cong I \otimes_R J \cong IJ$ where the last isomorphism is obtained from item 1. Hence $IJ \cong R$, where $I, J \subseteq F$ so that $IJ \subseteq F$. Consequently, IJ is a free R-module of rank 1 in F. It follows that $IJ = uR$ for some $u \in R^\times$, so that $I(u^{-1}J) = R$, as required. □

To end this section, we show that the two notions of Cartier divisors and Cartier class group of a scheme specializes to the notions introduced in this section.

Theorem 1.10.4.6. *Let A be a domain. Then the Cartier divisor group as defined in Definitions 1.10.4.1 and 1.10.3.1 are isomorphic.*

1.10.5 Divisors and invertible modules

1.10.6 Divisors on curves

1.11 Smoothness & differential forms

In this section, we would like to understand the notion of smoothness in algebraic geometry. We will first begin by defining a non-singular point of a variety over an algebraically closed field, which would be an extrinsic definition. However, by a fundamental observation of Zariski, we can have an intrinsic definition of non-singular points, which would be in terms of regular local rings. The main thrust behind this latter definition will be the expectation that over non-signular points, the dimension of the tangent space is equal to the dimension of the variety (which is true in the case of, say manifolds). We would further see that for a variety over an algebraically closed field, the set of singular points is closed and proper.

We would then introduce the important notion of sheaf of differentials over a scheme. This will again allow us to characterize non-singular points of a variety, and much more.

1.11.1 Non-singular varieties

To start investigating the notion of non-singularity, we first investigate it in the setting of classical affine varieties (Definition 1.5.4.11). We will then proceed to abstract varieties.

Definition 1.11.1.1. (**Non-singular points of a classical affine variety**) Let k be an algebraically closed field and X be a classical affine k-variety with $I(X) = \langle f_1, \dots, f_m \rangle \lneq k[x_1, \dots, x_n]$. A point $p \in X$ is said to be non-singular if the $n \times m$ Jacobian matrix

$$[J_p]_{n\times m} = \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{ij}$$

is of rank $n - \dim X$.

The first obvious question is whether the above definition is independent of the choice of the generators of prime ideal $I(X)$. The following lemma says yes.

Lemma 1.11.1.2. *Let k be an algebraically closed field and X be a classical affine k-variety. The definition of a non-singular point $p \in X$ is independent of the choice of the generating set of $I(X)$.*

Proof. Let $I(X) = \langle f_1, \dots, f_m \rangle = \langle g_1, \dots, g_l \rangle$. We wish to show that

$$\operatorname{rank}\left[\frac{\partial f_i}{\partial x_j}(p)\right]_{ij} = \operatorname{rank}\left[\frac{\partial g_i}{\partial x_j}(p)\right]_{ij}.$$

This follows immediately after writing $f_i = \sum_{a=1}^{l} c_{i_a} g_a$, $c_{i_a} \in k[x_1, \dots, x_n]$, differentiating it and observing that $g_a(p) = 0$ for all $a = 1, \dots, l$. □

Remark 1.11.1.3. In geometry, one notes that at a smooth point, the dimension of tangent space equals the dimension of the manifold itself. We would like to do a similar construction here. Indeed, if $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth map and 0 is regular for f, then we know by implicit function theorem that $M = Z(f) \subseteq \mathbb{R}^n$ is a smooth manifold with normal vector field $(\nabla f) : \mathbb{R}^n \to \mathbb{R}^n$. Consequently, one can define the tangent space $T_x M$ for $x \in M$ to be the set of all those vectors which are normal to $(\nabla f)_x$. We mimic this definition for classical affine varieties.

Definition 1.11.1.4. (**Tangent space of a classical affine variety**) Let k be an algebraically closed field and let X be a classical affine k-variety in $\mathbb{A}^n_k$ with $I(X) = \langle f_1, \dots, f_m \rangle$. Denote for each $f \in k[x_1, \dots, x_n]$ and $p \in \mathbb{A}^n_k$ the following linear functional

$$\begin{aligned} (df)_p : k^n &\longrightarrow k \\ v &\longmapsto \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) v_i. \end{aligned}$$

For a point $p \in X$, we define the tangent space T_pX as the following k-vector space

$$\begin{aligned} T_pX &:= \{v \in k^n \mid (df_i)_p(v) = 0 \; i = 1, \dots, m\} \\ &= \{v \in k^n \mid (df)_p(v) = 0 \; \forall f \in I(X)\}. \end{aligned}$$

We now show that this definition of tangent space is intrinsic. Indeed, we will show that the $T_pX = T\mathcal{O}_{X,p} := \operatorname{Hom}_k\left(\mathfrak{m}/\mathfrak{m}^2, k\right)$, where $(\mathcal{O}_{X,p}, \mathfrak{m})$ is the local ring at $p \in X$ and $\kappa(p) = k$ in this case (see Definition 16.1.2.14). Let us begin with a series of observations.

Lemma 1.11.1.5. *Let k be an algebraically closed field and $p \in \mathbb{A}^n_k$. Then the k-linear map*

$$\begin{aligned} \theta_p : k[x_1, \dots, x_n] &\longrightarrow k^n \\ f &\longmapsto \left(\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_n}(p) \right) \end{aligned}$$

induces a k-linear isomorphism

$$\mathfrak{m}_p/\mathfrak{m}_p^2 \cong k^n$$

where $\mathfrak{m}_p = \langle x_1 - p_1, \dots, x_n - p_n \rangle$ is the maximal ideal of $k[x_1, \dots, x_n]$ corresponding to the point p.

Proof. Let $p = (p_1, \dots, p_n)$. Observe that $\{\theta_p(x_i - p_i)\}_{i=1,\dots,n}$ forms a basis of k^n. Consequently, θ_p restricts to a surjective k-linear map $\hat{\theta}_p : \mathfrak{m}_p \to k^n$. Now one observes that $f \in \operatorname{Ker}(\theta)_p$ if and only if $f \in \mathfrak{m}_p^2$. Thus we have by first isomorphism theorem that $\mathfrak{m}_p/\mathfrak{m}_p^2 \cong k^n$. □

Lemma 1.11.1.6. *Let k be an algebraically closed field and X be a classical affine k-variety in $\mathbb{A}^n_k$ with $p \in X$. Let $(\mathcal{O}_{X,p}, \mathfrak{m})$ denote the local ring of X at p and $I \leq k[x_1, \dots, x_n]$ be the ideal of X. Then,*

$$\mathfrak{m}/\mathfrak{m}^2 \cong \mathfrak{m}_p/(\mathfrak{m}_p^2 + I).$$

Proof. Let $A = k[x_1, \dots, x_n]$. By Proposition 1.5.3.10, 3, we have $\mathfrak{m} = (\mathfrak{m}_p)_{\mathfrak{m}_p}/I_{\mathfrak{m}_p}$ and $\mathfrak{m}^2 = \left((\mathfrak{m}_p^2)_{\mathfrak{m}_p} + I_{\mathfrak{m}_p}\right)/I_{\mathfrak{m}_p}$. By quotienting, we obtain

$$\begin{aligned} \mathfrak{m}/\mathfrak{m}^2 &\cong \frac{(\mathfrak{m}_p)_{\mathfrak{m}_p}}{(\mathfrak{m}_p^2 + I)_{\mathfrak{m}_p}} \\ &\cong \left(\frac{\mathfrak{m}_p}{\mathfrak{m}_p^2 + I} \right)_{\mathfrak{m}_p/(\mathfrak{m}_p^2 + I)} \\ &\cong \frac{\mathfrak{m}_p}{\mathfrak{m}_p^2 + I}. \end{aligned}$$

□

Recall the notion of regular local ring from Definition 16.1.2.16. We now see that non-singular points are classified by the local ring being regular.

Theorem 1.11.1.7. *Let k be an algebraically closed field and X be a classical affine k-variety and let $p \in X$. The following are equivalent:*
1. *The point $p \in X$ is non-singular.*
2. *The local ring $\mathcal{O}_{X,p}$ is regular.*

Proof. Let $\mathfrak{m}$ be the maximal ideal of the local ring $\mathcal{O}_{X,p}$. By definition, we have $\mathcal{O}_{X,p}$ is regular if and only if $\dim_k \mathfrak{m}/\mathfrak{m}^2 = \dim \mathcal{O}_{X,p}$. By Proposition 1.5.3.10, 7, we further have that $\mathcal{O}_{X,p}$ is regular if and only if $\dim_k \mathfrak{m}/\mathfrak{m}^2 = \dim X$. Whereas by Lemmas 1.11.1.6 and 1.11.1.5, we observe

$$\begin{aligned}\dim_k \mathfrak{m}/\mathfrak{m}^2 &= \dim_k \frac{\mathfrak{m}_p}{\mathfrak{m}_p^2 + I}\\ &= \dim_k \left(\frac{\frac{\mathfrak{m}_p}{\mathfrak{m}_p^2}}{\frac{\mathfrak{m}_p^2+I}{\mathfrak{m}_p^2}}\right)\\ &= \dim_k \frac{\mathfrak{m}_p}{\mathfrak{m}_p^2} - \dim_k \frac{\mathfrak{m}_p^2 + I}{\mathfrak{m}_p^2}\\ &= n - \dim_k \frac{\mathfrak{m}_p^2 + I}{\mathfrak{m}_p^2}.\end{aligned}$$

With these two observations, we thus reduce to proving that

$$\dim_k \frac{\mathfrak{m}_p^2 + I}{\mathfrak{m}_p^2} = \operatorname{rank} J_p$$

where $J_p = \left[\frac{\partial f_i}{\partial x_j}\right]$ for $I = \langle f_1, \dots, f_m\rangle$. This now follows by the following two rather straightforward observations; in the notations of Lemma 1.11.1.5 and its proof, one observes
1. $\hat{\theta}_p^{-1}(\theta_p(I))$ is isomorphic as k-vector space to $I + \mathfrak{m}_p^2$,
2. $\dim_k \theta_p(I) = \operatorname{rank} J_p$.

The result now follows. □

With the above result, we formulate the following definition of non-singular abstract varieties.

Definition 1.11.1.8. (**Non-singular abstract variety**) Let k be an algebraically closed field. A variety X over k is said to be non-singular if for all $x \in X$, the local ring $\mathcal{O}_{X,x}$ is a regular local ring.

Remark 1.11.1.9. Note that in the definition of non-singular varieties, it is sufficient to demand that $\mathcal{O}_{X,x}$ is a regular local ring for all closed points $x \in X$ only. Indeed, by Lemma 1.3.1.1, local ring at a non-closed point is obtained by localizing the local ring at a closed point at a prime ideal. As the localization of a regular local ring at a prime ideal is again a regular local ring by Theorem **??**, the result follows.

We now define the Zariski (co)tangent space of a scheme at a point.

Definition 1.11.1.10. (**Zariski (co)tangent space**) Let X be a scheme and $x \in X$ be a point and let κ be the residue field at point x. Then

1. the Zariski cotangent space at x is defined to be the κ-vector space
$$T_x^* X := \mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2,$$
2. the Zariski tangent space at x is defined to be the κ-vector space
$$T_x X := \mathrm{Hom}_\kappa\left(\mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2, \kappa\right).$$

These are the analogues to the case in algebra (see Definition 16.1.2.14).

TODO : State how this is related to usual definition of tangent spaces

1.11.2 Regular schemes

".... *Of what use is it to know the definition of a scheme if one does not realize that a ring of integers in an algebraic number field, an algebraic curve, and a compact Riemann surface are all examples of a 'regular scheme of dimension 1'?*"- Hartshorne.

Definition 1.11.2.1 (**Regular schemes**)**.** A locally noetherian scheme X is said to be regular if the local rings $\mathcal{O}_{X,x}$ for all $x \in X$ is a regular local ring.

Observe that any smooth affine curve over a field is spectrum of a Dedekind domain, so is in particular a regular scheme (as local rings of Dedekind domains are regular, as noted in Theorem 16.10.1.8).

Proposition 1.11.2.2. *Let C be a smooth affine plane curve over a field k. Then the coordinate ring of C is a Dedekind domain.*

Proof. Let $C = \mathrm{Spec}\,(R)$. As C is a curve, therefore C is an integral finite type scheme of dimension 1 over k. We thus deduce that R is a finite type k-domain of dimension 1. By Hilbert basis theorem (Theorem 16.3.0.6), we deduce that R is noetherian. Smoothness of C yields that $R_\mathfrak{p}$ is a regular local ring for all $\mathfrak{p} \in C$. As $\dim R = 1$, we deduce that $R_\mathfrak{p}$ is a noetherian local domain of dimension 1 which is also regular. By Theorem 16.10.1.8, we deduce that $R_\mathfrak{p}$ is normal for all $\mathfrak{p} \in C$. By local criterion of normal domains (Proposition 16.7.2.10), we deduce that R is normal. Hence R is noetherian normal domain of dimension 1, as required. □

1.11.3 Cotangent bundle on affine schemes

We next study the analogues of tangent and cotangent bundles on affine schemes, before moving to general schemes. We begin by contemplating the following two constructions.

Construction 1.11.3.1 (Algebraic differential of a map)**.** Let $f : B \to C$ be a map of A-algebras so that we have a map $\psi : \mathrm{Spec}\,(C) \to \mathrm{Spec}\,(B)$. Let $\Omega_{B/A}$ and $\Omega_{C/A}$ be the module of differentials over B and C relative to A, respectively. We wish to construct a map $\varphi : \mathrm{Spec}\,(\mathrm{Sym}_C\,\Omega_{C/A}) \to \mathrm{Spec}\,(\mathrm{Sym}_B\,\Omega_{B/A})$, which we will later show is equivalent to the derivative of the map ψ;

$$\begin{tikzcd}
\mathrm{Spec}\,(\mathrm{Sym}_C\,\Omega_{C/A}) \arrow[r, dashed, "\varphi"] \arrow[d] & \mathrm{Spec}\,(\mathrm{Sym}_B\,\Omega_{B/A}) \arrow[d] \\
\mathrm{Spec}\,(C) \arrow[r, "\psi"'] & \mathrm{Spec}\,(B)
\end{tikzcd}$$

where the vertical maps are induced from inclusions into degree 0-term of the respective symmetric algebras.

Indeed, we need only define a map of A-algebras $\mathrm{Sym}_B \Omega_{B/A} \to \mathrm{Sym}_C \Omega_{C/A}$. To this end, by the cotangent sequence (Proposition 16.15.0.7), we have a map of C-modules induced by f given

$$\Omega_{B/A} \otimes_B C \to \Omega_{C/A}.$$

By composing with B-linear map $\Omega_{B/A} \to \Omega_{B/A} \otimes_B C$, we get the A-linear map

$$g : \Omega_{B/A} \to \Omega_{C/A}.$$

This gives the required map $\mathrm{Sym}\, g : \mathrm{Sym}_B \Omega_{B/A} \to \mathrm{Sym}_C \Omega_{C/A}$. Applying $\mathrm{Spec}\,(-)$ gives the required map φ in the diagram above. We call the map φ the *algebraic differential* of the map ψ.

On the other hand, we have the following construction.

Construction 1.11.3.2 (Geometric differential at a point)**.** Let $f : B \to C$ be a map of A-algebras so that we have a map $\psi : \mathrm{Spec}\,(C) \to \mathrm{Spec}\,(B)$. Assume that $A = k$ is a field and B, C are rational k-algebras[38]. Let $\mathfrak{p} \in \mathrm{Spec}\,(C)$ be a fixed point, $\mathfrak{q} = \psi(\mathfrak{p})$ and $\mathfrak{m}_\mathfrak{p} \leq C_\mathfrak{p}$, $\mathfrak{m}_\mathfrak{q} \leq B_\mathfrak{q}$ be the maximal ideals of the corresponding local rings. Note that by Proposition **??**, we have

$$\Omega_{C/A} \otimes_C C/\mathfrak{m}_\mathfrak{p} \cong \mathfrak{m}_\mathfrak{p}/\mathfrak{m}_\mathfrak{p}^2$$
$$\Omega_{B/A} \otimes_B B/\mathfrak{m}_\mathfrak{q} \cong \mathfrak{m}_\mathfrak{q}/\mathfrak{m}_\mathfrak{q}^2$$

That is, the fiber of $\mathrm{Spec}\,(\mathrm{Sym}_C \Omega_{C/A}) \to \mathrm{Spec}\,(C)$ at $\mathfrak{p}$ is the cotangent space $T_\mathfrak{p}^*\mathrm{Spec}\,(C)$, similarly for B at $\mathfrak{q}$. Using ψ we now wish to construct a map $d\psi_\mathfrak{p} : T_\mathfrak{q}^*\mathrm{Spec}\,(B) \to T_\mathfrak{p}^*\mathrm{Spec}\,(C)$ as follows. We have the canonical localization map

$$\psi_\mathfrak{p}^\sharp : \mathcal{O}_{\mathrm{Spec}(B),\mathfrak{q}} = B_\mathfrak{q} \to C_\mathfrak{p} = \mathcal{O}_{\mathrm{Spec}(C),\mathfrak{p}}$$

induced by ψ which is local homomorphism. As B and C are rational local, therefore $B_\mathfrak{q} = k \oplus \mathfrak{m}_\mathfrak{q}$ and $C_\mathfrak{p} = k \oplus \mathfrak{m}_\mathfrak{p}$ (Lemma 16.23.0.7). As the map is k-linear, so going modulo k, we get a map

$$\psi_\mathfrak{p}^\sharp : \mathfrak{m}_\mathfrak{q} \to \mathfrak{m}_\mathfrak{p}.$$

By standard algebra, this descends to a map

$$d\psi_\mathfrak{p} := \bar{\psi}_\mathfrak{p}^\sharp : T_\mathfrak{q}^*\mathrm{Spec}\,(B) \to T_\mathfrak{p}^*\mathrm{Spec}\,(C)$$

[38] that is, every local ring of B and C have residue field k again (see Definition 16.1.2.17).

1.12 Morphism of schemes

The main use of schemes to answer geometric questions begin with defining various types of situations that one usually finds himself/herself in algebraic geometry. We discuss them here one by one by giving examples.

We begin by elucidating some basic facts about the maps induced on local rings. First, the behaviour of maps with respect to schemes over k and residue fields.

Lemma 1.12.0.1. *Let $f : X \to Y$ be a map of schemes over k. If $p \in X$ is such that $\kappa(p) = k$, then $\kappa(f(p)) = k$.*

Proof. By the map induced on stalks, if we mod out the maximal ideal (using the fact that the map is local) we get the following diagram

$$\begin{tikzcd} \kappa(f(p)) \arrow[rr] & & k \\ & k \arrow[ul] \arrow[ur, "\cong"'] & \end{tikzcd}.$$

The result then follows. $\square$

The following in particular says that a map of varieties induces a finite extension of function fields.

Proposition 1.12.0.2. *Let k be a field and $f : X \to Y$ be a dominant map of integral finite type k-schemes where $\dim X = \dim Y$. Then the induced map on function fields $f^\flat : K(Y) \to K(X)$ is a finite extension.*

Proof. Note that as X and Y are finite type k-schemes, therefore $K(X)$ and $K(Y)$ are fraction fields of finite type k-algebras so they are finitely generated field extensions of k. As $\dim X = \text{trdeg } K(X)/k = \text{trdeg } K(Y)/k = \dim Y$, therefore by Lemma 16.6.9.8, we deduce that

$$\text{trdeg } K(X)/K(Y) = 0.$$

It follows that $K(X)/K(Y)$ is an algebraic extension. As $K(X)$ and $K(Y)$ are finitely generated extensions of k, therefore by tower law, $K(X)/K(Y)$ is a finitely generated extension. By algebraicity of $K(X)/K(Y)$, we deduce that $K(X)/K(Y)$ is finite, as required. $\square$

1.12.1 Elementary types of morphism

We first cover some basic type of maps between schemes.

Definition 1.12.1.1. (**Quasi-compact & maps**) A map $f : X \to Y$ of schemes is said to be quasi-compact if there exists an affine open cover $\{V_i\}$ of Y such that the space $f^{-1}(V_i) \subseteq X$ is quasi-compact for each i.

Remark 1.12.1.2. Observe that a scheme X over k has a quasi-compact structure map $X \to \text{Spec}(k)$ if and only if X is quasi-compact.

We now see that quasi-compact maps are local on target.

Proposition 1.12.1.3. [39] *A map $f : X \to Y$ is quasi-compact if and only if for each open affine $V \subseteq Y$, the space $f^{-1}(V) \subseteq X$ is quasi-compact.*

Proof. The ($\Leftarrow$) is immediate. For ($\Rightarrow$), pick any open affine $V \subseteq Y$. We wish to show that $f^{-1}(V)$ is quasi-compact. Let $V_i = \operatorname{Spec}(B_i)$ be the collection of open affines covering Y such that $f^{-1}(V_i)$ is quasi-compact. We now obtain a finite covering of V by affine opens which are affine open in V_i for some i as well. Indeed, by Lemma 1.4.4.3, we may cover $V \cap V_i$ by open affines which are basic open in both V and V_i. Doing this for each i, we obtain a cover of V by basic opens. As V is affine, so by Lemma 1.2.1.6 we obtain a finite collection of basic opens $\{D(g_i)\}_{i=1}^n$ where $g_i \in B_i$ such that $D(g_i)$ is a basic open in V as well.

We now have that $f^{-1}(V) = \bigcup_{i=1}^n f^{-1}(D(g_i))$. Hence it suffices to show that $f^{-1}(D(g_i))$ is a quasi-compact subspace. To this end, we first immediately reduce to assuming that X is quasi-compact (by replacing X by $f^{-1}(V_i)$) and $Y = \operatorname{Spec}(B)$ is affine (by replacing Y by V_i). We now wish to prove that for any $g \in B$, $f^{-1}(D(g))$ is quasi-compact.

As X is quasi-compact, therefore there exists a finite affine open cover of X by $\operatorname{Spec}(A_i)$. It suffices to show that $\operatorname{Spec}(A_i) \cap f^{-1}(D(g))$ is a quasicompact space. Observe that $f|_{\operatorname{Spec}(A_i)} : \operatorname{Spec}(A_i) \to \operatorname{Spec}(B)$ is a morphism of affine schemes. It follows from Corollary 1.3.0.6 that $f|_{\operatorname{Spec}(A_i)}$ is induced from a ring map $\varphi_i : B \to A_i$. As $\operatorname{Spec}(A_i) \cap f^{-1}(D(g)) = (f|_{\operatorname{Spec}(A_i)})^{-1}(D(g)) = D(\varphi_i(g))$, which is an affine open, therefore by Lemma 1.2.1.6, we deduce that $\operatorname{Spec}(A_i) \cap f^{-1}(D(g))$ is quasi-compact, as required. $\square$

Example 1.12.1.4 (A non quasi-compact scheme)**.** Let $A = k[x_1, x_2, \dots]$ and $X = \operatorname{Spec}(A)$ be the infinite affine space over k. Consider the open subscheme obtained by removing the origin; $U = X - 0$ where 0 is the maximal ideal $\mathfrak{m}_0 = \mathfrak{x}_\mathfrak{i} \mid \mathfrak{i} \geq 1$. We claim that this is not quasi-compact. Indeed, each of the opens $D(x_i)$ is in U. Moreover, they cover whole of U as any prime of A not equal to $\mathfrak{m}_0$ will necessarily not contain some x_i, hence is in $D(x_i)$. Moreovoer, no finite subcover of $\{D(x_i)\}_i$ covers U.

Definition 1.12.1.5. (**Quasi-finite maps**) A map $f : X \to Y$ of schemes is quasi-finite if for each $y \in Y$, the fiber X_y is a finite set.

Example 1.12.1.6. Let k be an algebraically closed field. Consider the map

$$f : X = \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3}\right) \longrightarrow \mathbb{A}^1_k$$

obtained by the map $k[x] \to \frac{k[x,y]}{y^2-x^3}$ given by $x \mapsto x + \langle y^2 - x^3 \rangle$. Take any point $\mathfrak{p} = \langle p(x) \rangle \in \mathbb{A}^1_k$. Hence the fiber is

$$\begin{aligned} X_\mathfrak{p} &= \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3} \otimes_k \kappa(\mathfrak{p})\right) \\ &= \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3} \otimes_k \frac{k[x]_\mathfrak{p}}{\mathfrak{p}k[x]_\mathfrak{p}}\right) \\ &\cong \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3} \otimes_k \left(\frac{k[x]}{\mathfrak{p}}\right)_\mathfrak{p}\right). \end{aligned}$$

[39] Exercise II.3.2 of Hartshorne.

Let $\mathfrak{p} \neq \mathfrak{o}$. As $k[x]$ is a PID, therefore $\mathfrak{p}$ is a maximal ideal. Consequently, we have $\kappa(\mathfrak{p}) = k[x]/\mathfrak{p}$. Hence,

$$\begin{aligned} X_{\mathfrak{p}} &\cong \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3} \otimes_k \frac{k[x]}{p(x)}\right) \\ &\cong \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3, p(x)}\right). \end{aligned}$$

As k is algebraically closed, therefore by weak Nullstellensatz, we obtain that $p(x) = x - a$ for some $a \in k$. Consequently, if we have $a \neq 0$ then

$$\begin{aligned} X_{\mathfrak{p}} &\cong \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3, x - a}\right) \\ &\cong \operatorname{Spec}\left(\frac{k[y]}{y^2 - a^3}\right) \\ &\cong \operatorname{Spec}\left(\frac{k[y]}{(y + a^{3/2})(y - a^{3/2})}\right) \\ &\cong \operatorname{Spec}(k \times k) \\ &\cong \operatorname{Spec}(k) \amalg \operatorname{Spec}(k). \end{aligned}$$

If $a = 0$, then

$$X_{\mathfrak{p}} \cong \operatorname{Spec}\left(\frac{k[y]}{y^2}\right)$$

and we know that $k[y]/y^2$ has only one prime ideal, the one generated by $y + \langle y^2 \rangle \in k[y]/y^2$. Hence $X_{\mathfrak{p}}$ consists of two points at all non-zero closed points and of a single point at the origin, showing that f has finite fibers at all closed points. However, at the generic point $\mathfrak{p} = \mathfrak{o}$, we have a more complicated story:

$$\begin{aligned} X_{\mathfrak{o}} &\cong \operatorname{Spec}\left(\frac{k[x,y]}{y^2 - x^3} \otimes_k k(x)\right) \\ &\cong \operatorname{Spec}\left(\frac{k(x)[x,y]}{y^2 - x^3}\right) \\ &\cong \operatorname{Spec}\left(\frac{k(x)[y]}{y^2 - x^3}\right). \end{aligned}$$

As $k(x)[y]$ is a PID, therefore points of $X_{\mathfrak{o}}$ are thus in bijective correspondence with prime ideals of $k(x)[y]$ containing $y^2 - x^3$, which in turn is in bijection with the set of irreducible factors of $y^2 - x^3$ in $k(x)[y]$. As $k(x)[y]$ is a UFD, therefore there can atmost be finitely many such irreducible factors. Hence, $X_{\mathfrak{o}}$ is finite.

Hence all fibers are finite, making f quasi-finite.

1.12.2 Finite type

We already considered one example of such maps in the case of schemes over a field in Section 1.4.3

Definition 1.12.2.1. (**Locally finite type**) Let $f : X \to Y$ be a map of schemes. Then f is said to be locally of finite type if there is an affine open cover $V_i = \operatorname{Spec}(B_i)$, $i \in I$ of Y such that for each $i \in I$, $f^{-1}(V_i)$ has an open affine cover $U_{ij} = \operatorname{Spec}(A_{ij})$, $j \in J$ such that for each j, the ring A_{ij} is finite type[40] B_i-algebra.

Definition 1.12.2.2. (**Finite type**) Let $f : X \to Y$ be a map of schemes. Then f is said to be of finite type if there is an open affine cover $V_i = \operatorname{Spec}(B_i)$, $i \in I$ of Y such that for each $i \in I$, $f^{-1}(V_i)$ has a finite open affine cover $U_{ij} = \operatorname{Spec}(A_{ij})$, $j = 1, \dots, n$ such that for each j, A_{ij} is a finite type B_i-algebra.

It is an important observation that both the above definitions are local on target.

Proposition 1.12.2.3. [41] *A map $f : X \to Y$ is locally of finite type if and only if for all open affine $V = \operatorname{Spec}(B)$ in Y, there is an open affine cover $U_i = \operatorname{Spec}(A_i)$ of $f^{-1}(V)$ in X such that each A_i is a finite type B-algebra.*

Proof. The R $\Rightarrow$ L follows immediately. Let $V_i = \operatorname{Spec}(B_i)$ be an open affine cover of Y such that $f^{-1}(V_i)$ is covered by open affines $U_{ij} = \operatorname{Spec}(A_{ij})$ where each A_{ij} is a finite type B_i-algebra. Pick any affine open $V = \operatorname{Spec}(B)$ in Y and a point $x \in f^{-1}(V)$. We wish to find an open affine $x \in U = \operatorname{Spec}(A)$ inside $f^{-1}(V)$ such that A is a finite type B-algebra.

Consider $f(x) \in V$ and let $f(x) \in V \cap V_i$. Consequently, $x \in f^{-1}(V)$ will be contained in some U_{ij}, so $x \in f^{-1}(V) \cap U_{ij}$. By continuity of f, there exists a basic open $D(g) \subseteq V \cap V_i$ for some $g \in B_i$ which contains $f(x)$ such that $f^{-1}(D(g)) \subseteq f^{-1}(V) \cap U_{ij}$ is open. Restricting f to U_{ij}, we have $f : U_{ij} \to V_i$ which induces a map $\varphi : B_i \to A_{ij}$ which is of finite type. Denote $U = f^{-1}(D(g)) = D(\varphi(g)) = \operatorname{Spec}\left((A_{ij})_{\varphi(g)}\right) \subseteq f^{-1}(V) \cap U_{ij}$. We therefore get that the restriction of f on U, which is given by $f : U \to D(g)$, induces the localization map on algebras $\varphi_g : (B_i)_g \to (A_{ij})_{\varphi(g)}$. As localization of algebras are finite type, therefore φ_g makes $(A_{ij})_{\varphi(g)}$ a finite type $(B_i)_g$-algebra.

By Lemma 1.4.4.3, we have an isomorphism $B_h \to (B_i)_g$. Thus, we have

$$B \to B_h \stackrel{\cong}{\to} (B_i)_g \to (A_{ij})_{\varphi(g)}$$

where each map is of finite type. Since composite of finite type maps is of finite type, therefore $(A_{ij})_{\varphi(g)}$ is a finite type B-algebra, as required.

$\square$

We next see that finite type maps are also local on target and a nice property that they satisfy which says that finite type property descends to every open affine inside the inverse image of an open affine.

Theorem 1.12.2.4. [42] *Let $f : X \to Y$ be a map of schemes. Then,*

1. *f is of finite type if and only if f is locally of finite type and quasi-compact,*
2. *f is of finite type if and only if for every open affine $V = \operatorname{Spec}(B)$, the space $f^{-1}(V)$ can be covered by finitely many open affines $U_i = \operatorname{Spec}(A_i)$ where each A_i is a finite type B-algebra,*

[40] finite type algebra := finitely generated as an algebra.

[41] Exercise II.3.1 of Hartshorne.

[42] Exercise II.3.3 of Hartshorne.

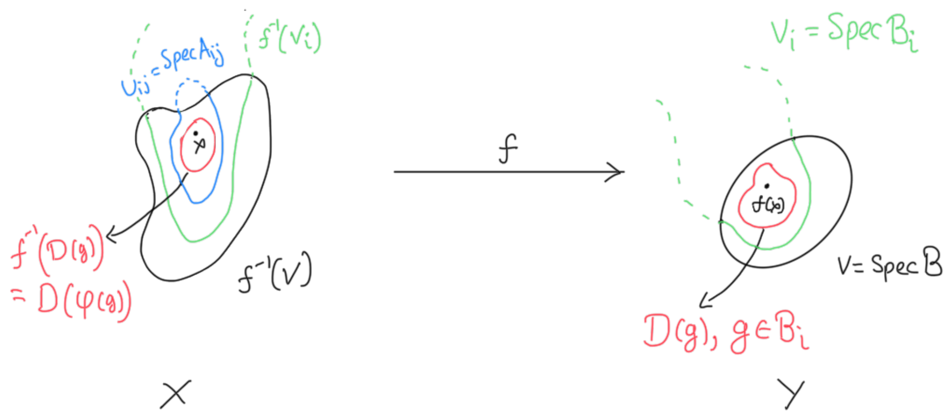


Figure 1.2: Sketch of the proof of Proposition 1.12.2.3

3. *if f is of finite type, then for any open affine $V = \text{Spec}\,(B) \subseteq Y$ and any open affine $U = \text{Spec}\,(A) \subseteq f^{-1}(V)$, A is a finite type B-algebra.*

Proof. 1. (L $\Rightarrow$ R) As f is of finite type, therefore there exists an open affine cover $\{V_i = \text{Spec}\,(B_i)\}$ of Y such that $f^{-1}(V)$ can be covered by finitely many $U_{ij} = \text{Spec}\,(A_{ij})$ where A_{ij} is a finite type B_i-algebra. Consequently, f is locally finite type. As each affine scheme is quasi-compact (Lemma 1.2.1.6) and finite union of quasi-compact spaces is quasi-compact, therefore we deduce that $f^{-1}(V)$ is quasi-compact.

(R $\Rightarrow$ L) As f is locally of finite type, therefore there exists an open affine cover $\{V_i = \text{Spec}\,(B_i)\}$ of Y such that $f^{-1}(V_i)$ is covered by open affines $U_{ij} = \text{Spec}\,(A_{ij})$ where each A_{ij} is a finite type B_i-algebra. As f is quasicompact, therefore we have a finite sub-cover $U_{i1}, \dots, U_{in}$ covering $f^{-1}(V)$, as required.

2. (R $\Rightarrow$ L) Immediate from definition.

(L $\Rightarrow$ R) Pick an open affine $V = \text{Spec}\,(B)$ in Y. We wish to show that $f^{-1}(V)$ is covered by finitely many open affines each of which is spectrum of a finite type B-algebra. Indeed, as f is quasi-compact by statement 1 above, therefore by Proposition 1.12.1.3, we see that $f^{-1}(V)$ is quasi-compact. Also by statement 1, f is of locally finite type. Hence by Proposition 1.12.2.3, $f^{-1}(V)$ is covered by spectra of finite type B-algebras. As $f^{-1}(V)$ is quasi-compact, we get a finite subcover, as required.

3. Pick any open affine $V = \text{Spec}\,(B)$ in Y and an open affine $U = \text{Spec}\,(A) \subseteq f^{-1}(V)$. As f is of finite type, therefore by statement 2 above, we obtain a finite collection $U_i = \text{Spec}\,(A_i)$ of open affines covering $f^{-1}(V)$. Observe that $U \cap U_i$ is an open set of U_i. By virtue of Lemma 1.4.4.3, we may cover $U \cap U_i$ by basic open sets of U_i which are basic open in U as well. Doing this for each i furnishes us with an open cover of U. As U is quasi-compact as it is affine (Lemma 1.2.1.6), consequently we get a finitely many elements $h_1, \dots, h_n \in A$ such that $D(h_i) \subseteq U$ covers U and furthermore for each $i = 1, \dots, n$, $D(h_i) \cong D(g_i)$ where $D(g_i) \subseteq U_i$ and $g_i \in A_i$. In particular, we have $A_{h_i} \cong (A_i)_{g_i}$. Now for each $i = 1, \dots, n$, we have

$$B \to A_i \to (A_i)_{g_i} \cong A_{h_i}$$

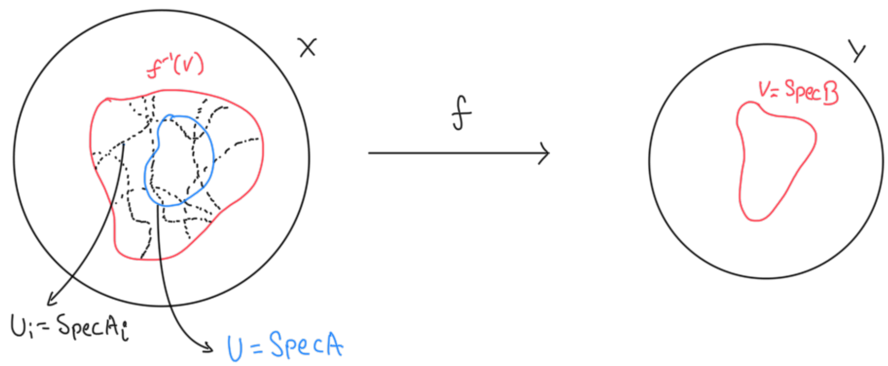


Figure 1.3: Sketch of the proof of Theorem 1.12.2.4, 3.

where each of the arrows makes the codomain a finite type algebra over the domain. Hence, A_{h_i} is a finite type B-algebra. Consequently, we have $h_1, \dots, h_n \in A$ such that $\cup_{i=1}^n D(h_i) = U$ (which is equivalent to saying that h_is generate the unit ideal of A by Lemma 1.2.1.5, 2) and A_{h_i} is a finite type B-algebra. It follows from Lemma 16.1.2.11 that A is a finite type B-algebra. This completes the proof. $\square$

We now list out some properties of finite type maps as we shall encounter them quite frequently.

Proposition 1.12.2.5. [43] *Properties of finite type maps.*

1. *A closed immersion $X \to Y$ is of finite type.*
2. *A quasicompact open immersion $X \to Y$ is of finite type.*
3. *Composition of finite type maps $X \to Y \to Z$ is of finite type.*
4. *Product of finite type schemes $X \to S$ and $Y \to S$ in $\mathbf{Sch}/S$ denoted $X \times_S Y \to S$ is of finite type.*
5. *Maps of finite type are stable under base extensions.*
6. *If $X \to Y$ is quasicompact and the composite $X \to Y \to Z$ is of finite type, then $X \to Y$ is of finite type.*
7. *If $X \to Y$ is of finite type and Y is noetherian, then X is noetherian.*

Proof. **TODO!** $\square$

The following is something we all expect, which indeed holds true for finite type schemes.

Lemma 1.12.2.6. *Let k be a field and X be a finite type k-scheme. The set of all closed points of X is dense in X.*

Proof. **TODO**. $\square$

Example 1.12.2.7. We give a number of examples of finite type maps.

1. Let k be a field. Consider the projection map $\pi : \mathbb{A}^2_k \to \mathbb{A}^1_k$ defined by the k-algebra map $k[x] \to k[x, y]$ mapping as $x \mapsto x$. Note that π is a finite type map of schemes as the open covering of $\mathbb{A}^1_k$ as itself yields that $\pi^{-1}(\mathbb{A}^1_k) = \mathbb{A}^2_k$ and $\mathbb{A}^2_k$ is spectra of $k[x, y]$ which is a finite type $k[x]$ algebra via the above map. Indeed, $k[x, y]$ is generated by $\{y\}$ as a $k[x]$-algebra. We deduce that projection maps $\mathbb{A}^n_k \to \mathbb{A}^1_k$ are finite type maps for any $n \in \mathbb{N}$.

[43] Exercise II.3.13 of Hartshorne.

2. We next consider a family of curves parameterized by a parameter t. Consider the map
$$\mathbb{C}[t] \longrightarrow \mathbb{C}[t][x,y]\langle y^2 - x^3 - t\rangle.$$
This yields the following map at the level of schemes
$$X := \operatorname{Spec}\left(\frac{\mathbb{C}[t][x,y]}{\langle y^2 - x^3 - t\rangle}\right) \to \operatorname{Spec}(\mathbb{C}[t]).$$
Pick the closed point corresponding to $a \in \mathbb{C}$ in $\operatorname{Spec}(\mathbb{C}[t])$. As $\mathbb{C}[t][x,y]/\langle y^2 - x^3 - t\rangle$ is a finite type $\mathbb{C}[t]$-algebra, therefore the above map of schemes is of finite type.
Observe that the fiber of X at $a \in \operatorname{Spec}(\mathbb{C}[t])$ (by abuse of notation) is given by
$$X_a = X \times_{\operatorname{Spec}(\mathbb{C}[t])} \operatorname{Spec}(\kappa(a)).$$
As $\kappa(a)$ is the fraction field of $\mathbb{C}[t]/\langle t-a\rangle$, which is $\mathbb{C}[a] = \mathbb{C}$, therefore we get the following
$$\begin{aligned} X_a &= \operatorname{Spec}\left(\frac{\mathbb{C}[t][x,y]}{\langle y^2 - x^3 - t\rangle} \otimes_{\mathbb{C}[t]} \mathbb{C}[a]\right) \\ &\cong \operatorname{Spec}\left(\frac{\mathbb{C}[x,y]}{\langle y^2 - x^3 - a\rangle}\right). \end{aligned}$$
Hence, we get the curve $y^2 - x^3 - a$ back as the fiber at the point $a \in \operatorname{Spec}(\mathbb{C}[t])$.
3. Consider the map
$$k[t] \longrightarrow \frac{k[t][w,x,y,z]}{\langle (w-y)^2 + (x-z)^2 - t^2\rangle}$$
which yields the map on geometric level as
$$X := \operatorname{Spec}\left(\frac{k[t][w,x,y,z]}{\langle (w-y)^2 + (x-z)^2 - t^2\rangle}\right) \longrightarrow \operatorname{Spec}(k[t]).$$
Again, this is a finite type map and for a closed point $a \in k$ corresponding to the ideal $\langle t-a\rangle \lneq k[t]$, the fiber is
$$\begin{aligned} X_a &\cong \operatorname{Spec}\left(\frac{k[t][w,x,y,z]}{\langle (w-y)^2 + (x-z)^2 - t^2\rangle} \otimes_{k[t]} k[a]\right) \\ &\cong \operatorname{Spec}\left(\frac{k[w,x,y,z]}{\langle (w-y)^2 + (x-z)^2 - a^2\rangle}\right) \end{aligned}$$
on $\mathbb{A}^2_{\mathbb{R}}$).
4. Any projective variety $X \to \mathbb{P}^n_k$ will by definition be finite type over k (Theorem 1.12.8.2, 1). For example, the projective parabola $X = \operatorname{Proj}\left(\frac{k[x,y,z]}{y^2 - xz}\right)$ is a finite type scheme over k. Indeed, by Proposition 1.8.2.8, 1, we get that the natural map $X \to \operatorname{Proj}(k[x,y,z]) = \mathbb{P}^3_k$ coming from the quotient $k[x,y,z] \twoheadrightarrow \frac{k[x,y,z]}{y^2 - xz}$ (which is a graded map) is a closed immersion. Hence, it defines a closed subscheme of projective 3-space $\mathbb{P}^3_k$ over k.

Over non-noetherian rings, we rather use finite "presentation" rather than finite type.

Definition 1.12.2.8 (Finite presentation). A map of schemes $f : X \to Y$ is of locally finite presentation if there exists an open affine cover $V_\alpha = \mathrm{Spec}\,(A)$ of Y such that $f^{-1}(V_\alpha)$ is covered by open affine $U_\beta = \mathrm{Spec}\,(B_\beta)$ where each B_β is finitely presented A-algebra via f. If f is quasi-compact, quasi-separated and locally finitely presented, then we say f is of finite presentation.

Clearly, over noetherian schemes, finite presentation and finite type are same.

1.12.3 Finite

This is a more stronger version of finite type maps discussed in previous section.

Definition 1.12.3.1. (**Finite**) Let $f : X \to Y$ be a map of schemes. Then f is said to be finite if there is an open affine covering $V_i = \mathrm{Spec}\,(B_i)$, $i \in I$ of Y such that $f^{-1}(V_i)$ is equal to an open affine $\mathrm{Spec}\,(A_i)$ where A_i is a finite B_i-algebra[44].

We see that finite maps are local on target.

Proposition 1.12.3.2. *A map $f : X \to Y$ of schemes is finite if and only if for each open affine $V =$ $\mathrm{Spec}\,(B)$, we have $f^{-1}(V)$ is an open affine $\mathrm{Spec}\,(A)$ in X such that $B \to A$ makes A a finite B-algebra.*

Proof. (R $\Rightarrow$ L) is immediate from definitions.

(L $\Rightarrow$ R) Pick any open affine $V = \mathrm{Spec}\,(B)$ in Y. We first wish to show that $U = f^{-1}(V)$ is an affine scheme. We may employ criterion for affineness, Proposition 1.3.1.6, for this purpose. Hence for showing that U is affine, we reduce to finding $g_1, \dots, g_n \in \Gamma(\mathcal{O}_{X|U}, U) = \mathcal{O}_X(U)$ such that U_{g_i} is affine and $\langle g_1, \dots, g_n \rangle = \mathcal{O}_X(U)$.

As f is finite, therefore there exists an open affine covering $V_i = \mathrm{Spec}\,(B_i)$ of Y such that $f^{-1}(V_i) = \mathrm{Spec}\,(A_i) = U_i$ is affine and A_i is a finite B_i-algebra. Observe that $V \cap V_i$ forms an open covering of V. As V is affine, so it is quasi-compact (Lemma 1.2.1.6). Consequeently, we obtain a finite cover of V by V_is. Now cover each $V \cap V_i$ by basic opens which are basic in both V and V_i (Lemma 1.4.4.3). Doing this for each of the finitely many i, we obtain a cover of V by basic open sets. As V is quasi-compact (Lemma 1.2.1.6), therefore we have obtained a cover of V by finitely many basics $D(k_i)$ for $k_i \in B$ such that $D(k_i) \cong D(l_i)$ where $D(l_i) \subseteq V_i$ and $l_i \in B_i$ for $i = 1, \dots, n$. Consequently by Lemma 1.2.1.5, the ideal generated by $k_1, \dots, k_n$ in B is the unit ideal.

As we have

$$U = f^{-1}(V) = f^{-1}\left(\bigcup_{i=1}^{n} D(k_i)\right) = \bigcup_{i=1}^{n} f^{-1}(D(k_i)),$$

therefore by Lemma 1.3.1.3, we may write

$$U = \bigcup_{i=1}^{n} U_{\varphi(k_i)}$$

where $\varphi : B \to \mathcal{O}_X(U)$ is the map induced by the restricted map $f : U \to V$ on the global sections. Furthermore, as $\sum_{i=1}^{n} k_i B = B$, therefore $\sum_{i=1}^{n} \varphi(k_i)\mathcal{O}_X(U) = \mathcal{O}_X(U)$. Hence, it now suffices to

[44] finite algebra := finitely generated as a module.

show that each $U_{\varphi(k_i)}$ is affine.

We have $U_{\varphi(k_i)} = f^{-1}(D(k_i)) \cong f^{-1}(D(l_i)) = D(\varphi_i(l_i))$ where $\varphi_i : B_i \to A_i$ is the map on global sections obtained by the restriction $f : U_i \to V_i$. As $D(\varphi_i(l_i))$ is affine, thus, so is $U_{\varphi(k_i)}$. This shows that indeed, $f^{-1}(V)$ is an open affine.

We may now write $U = \mathrm{Spec}\,(A)$. We reduce now to showing that A is a finite B-algebra. For this observe that in the above, we obtained a finite open cover of U given by $U_{\varphi(k_i)} \cong D(\varphi_i(l_i))$ where $D(\varphi_i(l_i)) \subseteq U_i$. As $U = \mathrm{Spec}\,(A)$, therefore $\mathcal{O}_X(U) = A$, so we may let $\varphi(k_i) = g_i$ for $i = 1, \dots, n$. Now, since $U_{\varphi(k_i)} = U_{g_i} = D(g_i) \cong D(\varphi_i(l_i))$, therefore we have $A_{g_i} \cong (A_i)_{\varphi_i(l_i)}$. As A_i is a finite B_i-algebra, therefore by Lemma 16.1.2.10, $(A_i)_{\varphi_i(l_i)}$ is a finite $(B_i)_{l_i}$-algebra. Further, as we saw in the beginning that $D(k_i) \cong D(l_i)$, hence we get $B_{k_i} \cong B_{l_i}$. We thus obtain a map $B_{k_i} \to A_{g_i}$ as in

$$\begin{tikzcd} (A_i)_{\varphi_i(l_i)} \arrow[d, phantom, "\cong" rotate=-90] & (B_i)_{l_i} \arrow[l, "(\varphi_i)_{l_i}"'] \arrow[d, phantom, "\cong" rotate=-90] \\ A_{g_i} & B_{k_i} \arrow[l, dashed] \end{tikzcd}$$

which thus makes A_{g_i} a finite B_{k_i}-algebra, in particular, a finitely generated B_{k_i}-module. This is for each of the $i = 1, \dots, n$, and since we have that $k_1, \dots, k_n$ generates the unit ideal in B, hence by another application of Lemma 16.1.2.10, we deduce that A is a finite B-algebra, as required. $\square$

Base change preserves finiteness.

Proposition 1.12.3.3. *Let $f : X \to S$ be a finite map of schemes. If $g : S' \to S$ is any map, then the map $f' : X' \to S'$ as in the base change*

$$\begin{tikzcd} X' \arrow[r] \arrow[d, "f'"'] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "f"] \\ S' \arrow[r, "g"'] & S \end{tikzcd}$$

is finite.

One important property of finite maps is that their fibers are finite.

Proposition 1.12.3.4. [45] *Let $f : X \to Y$ be a finite morphism. Then f is quasi-finite.*

Proof. Pick any point $y \in Y$ and an affine open $V = \mathrm{Spec}\,(B) \ni y$ in Y. As f is finite, therefore by restriction we have map $f : f^{-1}(V) \to V$ where $f^{-1}(V) = U = \mathrm{Spec}\,(A)$ and A is a finite B-algebra. Thus $f^{-1}(y) \subseteq U$ and hence we reduce to the affine case $X = \mathrm{Spec}\,(A)$ and $Y = \mathrm{Spec}\,(B)$.

Fix $\mathfrak{q} \in Y$. As the fiber $X_{\mathfrak{q}}$ is the base change of $f : X \to Y$ under the inclusion $\mathrm{Spec}\,(\kappa(\mathfrak{q})) \hookrightarrow Y$, thus by Proposition 1.12.3.3 we deduce that $X_{\mathfrak{q}} = \mathrm{Spec}\,(A \otimes_B \kappa(\mathfrak{q}))$ is finite over $\mathrm{Spec}\,(\kappa(\mathfrak{q}))$. In particular, $C = A \otimes_B \kappa(\mathfrak{q})$ is a finite $\kappa(q)$-algebra. By cite[AMD], Exercise 8.3, C is an Artinian ring, as required. $\square$

Another nice property enjoyed by finite maps is that they are closed.

Proposition 1.12.3.5. [46] *Let $f : X \to Y$ be a finite morphism. Then f is a closed map.*

[45] Exercise II.3.5, a) of Hartshorne.
[46] Exercise II.3.5, b) of Hartshorne.

Proof. Our goal is to reduce to the affine case as much as possible, where we have many algebraic results to use. Let $Z \subseteq X$ be a closed subset of X. We wish to show that $f(Z)$ is closed in Y. It first suffices to show that for every open affine $V \subseteq Y$, $V \cap f(Z)$ is closed in V. By definition, $U = f^{-1}(V)$ is an open affine. Consider then the restricted map

$$f : U \cap Z \longrightarrow V \cap f(Z).$$

As $U \cap Z$ is closed in U and $f(U \cap Z) = V \cap f(Z)$, we hence reduce to the assumption that $X = \operatorname{Spec}(A)$ and $Y = \operatorname{Spec}(B)$ are affine. Let $Z \subseteq X$ be a closed subscheme. Then $Z = V(\mathfrak{a})$ for some ideal $\mathfrak{a} \leq A$. Considering the restriction $f : V(I) \cong \operatorname{Spec}(A/I) \to Y$ and the fact that $A \twoheadrightarrow A/I$ is a finite map, we immediate reduce to the further assumption that $Z = X$.

Consider $X = \operatorname{Spec}(A)$ and $Y = \operatorname{Spec}(B)$ and $f : X \to Y$ a finite map corresponding to $\varphi : B \to A$. As required, we claim that f has a closed image. Indeed, consider $I = \operatorname{Ann}_B(A)$ to be the annihilator ideal of B-module A. We claim that $\operatorname{Im}(f) = V(I)$. For ($\subseteq$), pick any $\mathfrak{p} \in \operatorname{Spec}(A)$. We wish to show that $\varphi^{-1}\mathfrak{p} \supseteq I$. It suffices to show that $\mathfrak{p} \supseteq \varphi(I)$. Indeed, for any $b \in I$, we must show $\varphi(b) \in \mathfrak{p}$. As $\varphi(b) \cdot A = 0$, therefore $\varphi(b) \cdot \varphi(b) = 0 \in \mathfrak{p}$ and thus $\varphi(b) \in \mathfrak{p}$, as required. Conversely for ($\supseteq$), fix a prime $\mathfrak{q} \in V(I)$. We wish to find $\mathfrak{p} \in X$ such that $\mathfrak{q} = \varphi^{-1}\mathfrak{p}$. Indeed, consider the map

$$B \overset{\varphi}{\twoheadrightarrow} \operatorname{Im}(\varphi) =: B' \subseteq A.$$

Note that as φ is finite, therefore φ is integral (Proposition 16.7.1.9). Note that $B' \cong B/\operatorname{Ker}(\varphi)$, induced by $B \twoheadrightarrow B/\operatorname{Ker}(\varphi)$. Observe that as $\operatorname{Ker}(\varphi) \subseteq I$, therefore $\bar{\mathfrak{q}} \leq B'$ is a prime containing $\bar{I} \leq B'$. It follows by Cohen-Seidenberg theorems (Theorem **??**) that there exists $\mathfrak{p} \in \operatorname{Spec}(A)$ such that $\mathfrak{p} \cap B' = \bar{\mathfrak{q}}$. Hence it follows at once that $\varphi^{-1}(\mathfrak{p}) = \mathfrak{q}$, as required. $\square$

Remark 1.12.3.6. [47] As tempting it might be to say that, but it is not true that a surjective, finite type, quasi-finite map is finite.

Indeed, let k be an algebraically closed field. Consider the map

$$f : \operatorname{Spec}\left(\frac{k[x,y]}{xy-1}\right) \amalg \operatorname{Spec}(k) \longrightarrow \mathbb{A}^1_k$$

induced by $k[x] \to \frac{k[x,y]}{xy-1} \times k$ given by $x \mapsto (x + \langle xy-1 \rangle, 0)$. As a nice exercise, one checks that (we write $a \in \mathbb{A}^1_k$ to mean $\langle x - a \rangle \in \mathbb{A}^1_k$)

1. $f^{-1}(0)$ is a singleton ($\operatorname{Spec}(k)$),
2. $f^{-1}(a)$ is a singleton, given by point (a, a^{-1}) (in particular, the point $\langle x-a, xy-1 \rangle$),
3. the generic fiber $f^{-1}(\mathfrak{o})$ is isomorphic to $\operatorname{Spec}(k(x))$ in $\operatorname{Spec}\left(\frac{k[x,y]}{xy-1}\right)$, hence a singleton.

Consequently, f is surjective, quasi-finite and furthermore of finite type. But still, $\frac{k[x,y]}{xy-1} \times k$ is not a finite $k[x]$-algebra.

Remark 1.12.3.7. Let k be a field. Observe that $\mathbb{A}^1_k \times_k \mathbb{A}^1_k \cong \mathbb{A}^2_k$. However, the underlying set in $\mathbb{A}^2_k$ is not the product of underlying set of $\mathbb{A}^1_k$ with itself. Indeed, this is essentially due to the fact that every prime ideal of $k[x,y]$ is not of form $\mathfrak{p}_1 \times \mathfrak{p}_2$ where $\mathfrak{p}_1 \in k[x]$ and $\mathfrak{p}_2 \in k[y]$, as the prime ideal $xy-1$ in $k[x,y]$ shows.

[47] Exercise II.3.5, c) of Hartshorne.

Example 1.12.3.8. Consider the canonical map $k[t] \to \frac{k[t,x]}{x^n - t}$ and the corresponding map $X = \mathrm{Spec}\left(\frac{k[t,x]}{x^n-t}\right) \to \mathrm{Spec}\,(k[t]) = \mathbb{A}^1_k$. As $\frac{k[t,x]}{x^n-t}$ is a finite $k[t]$-algebra of rank n, therefore $X \to \mathbb{A}^1_k$ is a finite map. Note that for each closed point $a \in \mathbb{A}^1_k$, the fiber $X_a \cong \mathrm{Spec}\left(\frac{k[x]}{x^n-a}\right)$, which has n closed points if k is algebraically closed and $a \neq 0$.

Any closed immersion is a finite map.

Proposition 1.12.3.9. *Let $i : Z \hookrightarrow X$ be a closed immersion. Then i is a finite map.*

Proof. By Proposition 1.4.4.11, we have an open affine cover $\{V_k\}$ of X such that $i : i^{-1}(V_k) \to V_k$ is a closed immersion. Write $V_k = \mathrm{Spec}\,(A_k)$. Since $i^{-1}(V_k) = Z \cap V_k$ and $Z \cap V_k$ is a closed subscheme of V_k, therefore $Z \cap V_k = \mathrm{Spec}\,(A_k/I_k)$ and the map $\mathrm{Spec}\,(A_k/I_k) \to \mathrm{Spec}\,(A_k)$ is induced from the quotient map $\pi : A_k \twoheadrightarrow A_k/I_k$, which is finite. Hence i is a finite map, as required. $\square$

Generic finiteness

Definition 1.12.3.10 (Generically finite map). Let $f : X \to Y$ be a map of schemes such that Y is irreducible. The map f is said to be generically finite if $f^{-1}(\eta)$ for $\eta \in Y$ the generic point is a finite set.

The following is an important result in this regard, which says, like many statements about generic points, that a generically finite dominant map is *almost* like a finite map.

Theorem 1.12.3.11. [48] *Let X, Y be integral schemes and $f : X \to Y$ be a dominant, generically finite and finite type map. Then there exists a dense open $V \subseteq Y$ such that $f|_{f^{-1}(V)} : f^{-1}(V) \to V$ is a finite map.*

Proof. We first prove this for X and Y affine integral schemes. We will later reduce to this case. Let $X = \mathrm{Spec}\,(A)$ and $Y = \mathrm{Spec}\,(B)$ be affine schemes where A, B are domains. Let $f : \mathrm{Spec}\,(A) \to \mathrm{Spec}\,(B)$ be a finite type, dominant, generically finite map so that A is a finite type B-algebra. Let this be induced by a finite type ring homomotphism $\varphi : B \to A$. Our first goal is to show that the generic point of X is mapped to generic point of Y and that the induced map of function fields $K(Y) \hookrightarrow K(X)$ is a finite extension.

Indeed, let $\xi \in X$ and $\eta \in Y$ be the generic point of X and Y respectively. By continuity of f, we have $f(\bar{\xi}) \subseteq \overline{f(\xi)}$. As $\bar{\xi} = X$, we have $f(X) \subseteq \overline{f(\xi)}$. As $f(X)$ is dense in Y by dominance of f, we deduce that $Y \subseteq \overline{f(\xi)}$, that is, $f(\xi)$ is a generic point of Y. As schemes are sober and in our case Y is irreducible, therefore Y has a unique generic point which is η. It follows that $f(\xi) = \eta$. Dominance of f further shows that φ is injective since $\xi = \mathfrak{o} \in f^{-1}(\eta) = f^{-1}(\mathfrak{o})$. As $f^{-1}(\mathfrak{o}) = \{\mathfrak{p} \in \mathrm{Spec}\,(A) \mid \varphi^{-1}(\mathfrak{p}) = \mathfrak{o}\}$ therefore if $\mathfrak{o} \in f^{-1}(\mathfrak{o})$, then it follows that $\mathrm{Ker}\,(\varphi) = 0$, that is, φ is injective.

Thus, by considering the comorphism at stalks, we get a map

$$\varphi_{\mathfrak{o}} = f^{\sharp}_{\xi} : \mathcal{O}_{Y,f(\xi)} = K(Y) = Q(B) \longrightarrow \mathcal{O}_{X,\eta} = K(X) = Q(A).$$

Note that this map is the field homomorphism induced by $\varphi : B \hookrightarrow A$ on the fraction fields. As this is a map of fields, therefore $\varphi : K(Y) \to K(X)$ is injective. By replacing $K(Y)$ by the image of φ, we may assume φ is an inclusion. We wish to show that $K(X)/K(Y)$ is a finite extension.

[48] Exercise II.3.7 of Hartshorne.

To this end, we first observe the following by generic finiteness. Let $A = B[\alpha_1, \dots, \alpha_n]$. The fiber at η is

$$\begin{aligned} f^{-1}(\eta) &= \mathrm{Spec}\,(A \otimes_B \kappa(\eta)) \\ &= \mathrm{Spec}\,(A \otimes_B Q(B)) \\ &= \mathrm{Spec}\,(B[\alpha_1, \dots, \alpha_n] \otimes_B Q(B)) \\ &= \mathrm{Spec}\,(Q(B)[\alpha_1, \dots, \alpha_n]) \end{aligned}$$

Thus by generic finiteness, $\mathrm{Spec}\,(Q(B)[\alpha_1, \dots, \alpha_n])$ is a finite set. We wish to show that it is discrete so that $f^{-1}(\eta)$ is a finite discrete affine scheme, that is, $Q(B)[\alpha_1, \dots, \alpha_n]$ is an artinian ring. It would thus follow that $Q(B)[\alpha_1, \dots, \alpha_n]$ is an artinian finite type $Q(B)$-algebra and is thus a finite $Q(B)$-algebra. Now, finiteness is preserved under going to fraction fields (Lemma 16.6.1.4) thus $Q(Q(B)[\alpha_1, \dots, \alpha_n])$ is a finite extension of $Q(B)$. But $Q(Q(B)[\alpha_1, \dots, \alpha_n]) = Q(A)$. Hence, $Q(A)$ is a finite extension of $Q(B)$, as required. We thus reduce to proving that the finite spectrum $\mathrm{Spec}\,(Q(B)[\alpha_1, \dots, \alpha_n])$ is discrete. We wish to show that all finitely many points of it are open. To this end it suffices to show that all finitely many primes of $Q(B)[\alpha_1, \dots, \alpha_n]$ are incomparable. **TODO.**

Thus we have shown that $K(X)/K(Y)$ is a finite extension. Using this, we now find the required open subset $V \subseteq Y$. Indeed, we find a basic open $V = D(b) \subseteq Y$ where $b \in B \subseteq A$ and $f^{-1}(D(b)) = D(b) \subseteq X$ is such that $f : f^{-1}(D(b)) \to D(b)$ is a finite map. That is, we wish to show that there exists $b \in B$ such that $\varphi_b : B_b \hookrightarrow A_b$ is a finite map, using the fact that $Q(B) \hookrightarrow Q(A)$ is a finite extension. Indeed, let $\frac{a_1}{a'_1}, \dots, \frac{a_n}{a'_n}$ be a $Q(B)$-basis of $Q(A)$. Observe that we have

$$Q(A) = Q(B)\frac{a_1}{a'_1} + \dots + Q(B)\frac{a_n}{a'_n}.$$

Thus, multiplying both sides by a'_i, we get that there exists $a_1, \dots, a_N \in A$ such that $Q(A)$ is a $Q(B)$-span of $a_1, \dots, a_N$. Denote $A = B[\alpha_1, \dots, \alpha_n]$. Observe that for any $\alpha_i \in A$, the set $\{1, \alpha_i, \dots, \alpha_i^{N-1}, \alpha_i^N\}$ is linearly dependent as its size is greater than the degree $N = [Q(A) : Q(B)]$. Consequently, we see that every α_i^k for $k \geq N$ is a linear combination of $\{1, \alpha_i, \dots, \alpha_i^{N-1}\}$. Now consider any $0 \leq i_1, \dots, i_n$ and the term $\alpha_1^{i_1} \dots \alpha_n^{i_n}$. Then this can be written as linear combination of various $\alpha_1^{j_1} \dots \alpha_n^{j_n}$ where $0 \leq j_1, \dots, j_n \leq N$.

Thus we have a finite collection of terms $\{\alpha_1^{i_1} \dots \alpha_n^{i_n}\}_{0 \leq i_1, \dots, i_n \leq N-1}$ in A. In $Q(A)$, we thus get the following expression for each of them:

$$\alpha_1^{i_1} \dots \alpha_n^{i_n} = \sum_{k=1}^{N} \frac{b_{i_1 \dots i_N, k}}{b'_{i_1 \dots i_N, k}} a_k$$

where $b_{i_1 \dots i_N, k}, b'_{i_1 \dots i_N, k} \in B$. Collect all the finitely many denominators $\{b'_{i_1 \dots i_N, k}\}_{i_1, \dots, i_N, k}$ and consider their product $b \in B$. We claim that the induced map $\varphi_b : B_b \hookrightarrow A_b$ is a finite map.

Indeed, pick any $\frac{a}{b^p} \in A_b$. Then, $a = \sum c_{i_1 \dots i_n} \alpha_1^{i_1} \dots \alpha_n^{i_n}$ for $c_{i_1 \dots i_n} \in B$. Consequently, we have

$$\begin{aligned}
a &= \sum_{i_1, \dots, i_n} c_{i_1 \dots i_n} \alpha_1^{i_1} \dots \alpha_n^{i_n} \\
&= \sum_{i_1, \dots, i_n} c_{i_1 \dots i_n} \left(\sum_{k=1}^{N} \frac{b_{i_1 \dots i_N, k}}{b'_{i_1 \dots i_N, k}} a_k \right) \\
&= \sum_{k=1}^{N} \underbrace{\left(\sum_{i_1, \dots, i_n} c_{i_1 \dots i_n} \frac{b_{i_1 \dots i_N, k}}{b'_{i_1 \dots i_N, k}} \right)}_{d_k} a_k \\
&= \sum_{k=1}^{N} d_k a_k
\end{aligned}$$

where $d_k \in Q(B)$. Observe that denominator of d_k is some product of elements of $\{b'_{i_1 \dots i_N, k}\}_{i_1, \dots, i_N, k}$. Consequently, we get that in A_b, we will have

$$\frac{a}{b^p} = \sum_{k=1}^{N} \frac{d_k}{b^p} a_k$$

where d_k / b^p is an element of B_b since $d_k \in B_b$. Hence, we have shown that there exists elements $a_1, \dots, a_N \in A$ such that A_b is finite over B_b. This completes the proof for affine case.

TODO : General case. $\square$

1.12.4 Separated

This notion corresponds to the Hausdorff property for topological spaces. Recall that a space X is Hausdorff if and only if the diagonal $\Delta : X \to X \times X$ is closed. We shall mimic this in the category of schemes.

Definition 1.12.4.1. (**Separated**) A map $f : X \to Y$ of schemes is said to be separated if the diagonal $\Delta : X \to X \times_Y X$ is a closed immersion. A scheme X is said to be separated if $X \to \operatorname{Spec}(\mathbb{Z})$ is separated.

It follows that any map of affine schemes is separated.

Lemma 1.12.4.2. *Let $f : \operatorname{Spec}(A) \to \operatorname{Spec}(B)$ be a map of affine schemes. Then f is separated.*

Proof. By Corollary 1.3.0.6, f corresponds to a map of rings $\varphi : B \to A$. Similarly, the diagonal map $\Delta : \operatorname{Spec}(A) \to \operatorname{Spec}(A) \times_{\operatorname{Spec}(B)} \operatorname{Spec}(A)$ corresponds to the B-algebra structure map over A, given by $m : A \otimes_B A \to A$, which is surjective. Consequently, by Corollary 1.4.4.14, Δ is a closed immersion. $\square$

Since any scheme locally is affine, we get a nice consequence of the above lemma.

Lemma 1.12.4.3. *Let $f : X \to Y$ be a map of schemes. Then the following are equivalent.*

1. *f is separated.*
2. *The diagonal $\Delta : X \to X \times_Y X$ has closed image.*

Proof. (1. $\Rightarrow$ 2.) Immediate.
(2. $\Rightarrow$ 1.) By the definition of diagonal, it is immediate that $\Delta : X \to X \times_Y X$ is a homeomorphism onto its image, which is further closed by the given hypothesis. Thus, we need only show that $\Delta^\flat : \mathcal{O}_{X\times_Y X} \to \Delta_* \mathcal{O}_X$ is a surjective map. By Theorem 20.3.0.6, 3, this is a local property. Consequently, we further reduce to showing that for any point $x \in X$ there is an open set $f(x) \in V \subseteq X \times_Y X$ such that $\Delta^\flat_{|V} : \mathcal{O}_{V,f(x)} \to (\Delta_*\mathcal{O}_{\Delta^{-1}(V)})_{f(x)}$ is surjective. Now we may choose by continuity of f a small affine open $x \in U$ such that $f(U)$ is contained in an affine open V in Y. Consequently, $U \times_V U$ is an affine open subset of $X \times_Y X$ containing $f(x)$. We thus reduce to showing that $\mathcal{O}_{U\times_V U, f(x)} \to (\Delta_*\mathcal{O}_U)_x$ is surjective, which follows immediately from Lemma 1.12.4.2. $\square$

Next, we state an important characterization of separatedness which allows us to derive some very important and convenient results about it.

Theorem 1.12.4.4. *(Valuative criterion of separatedness) Let $f : X \to Y$ be a map of schemes where X is noetherian. Then the following are equivalent,*

1. *f is separated.*
2. *Pick any field K and any valuation ring R with fraction field K (see Section 16.10). Let $i : \operatorname{Spec}(K) \to \operatorname{Spec}(R)$ be the map corresponding to $R \hookrightarrow K$. For all $g : \operatorname{Spec}(R) \to Y$ and $h : \operatorname{Spec}(K) \to X$ such that the square commutes, there exists atmost one lift of g along f as to make the following diagram commute:*

$$\begin{tikzcd} \operatorname{Spec}(K) \arrow[r, "h"] \arrow[d, "i"'] & X \arrow[d, "f"] \\ \operatorname{Spec}(R) \arrow[r, "g"'] \arrow[ur, dashed] & Y \end{tikzcd}.$$

Proof. See Theorem 4.3, Chapter 2 of cite[Hartshorne]. □

The following important corollaries can now easily be derived from this characterization.

Corollary 1.12.4.5. *Let us work in the category of noetherian schemes. Then,*

1. *separated maps are stable under base extension,*
2. *open and closed immersions are separated*[49]*,*
3. *composition of separated maps is separated,*
4. *for a base scheme S, product of any two separated maps is separated in* $\mathbf{Sch}/S$*,*
5. *if the composite $X \to Y \to Z$ is separated, then $X \to Y$ is separated,*
6. *a map $f : X \to Y$ is separated if and only if there is an open cover V_i of Y such that the restricted maps $f|_{f^{-1}(V_i)} : f^{-1}(V_i) \to V_i$ is separated*[50]*.*

Proof. **TODO: From notebook.** □

An important result about separate schemes is that the intersection of any two open affines is again an open affine.

Lemma 1.12.4.6. [51] *Let X be a separated scheme. If $U, V \subseteq X$ are two open affines then $U \cap V$ is again an open affine.*

Proof. Let $U = \operatorname{Spec}(A)$ and $V = \operatorname{Spec}(B)$. We may replace X by $U \cup V$ and X would still be separated by Corollary 1.12.4.5, 6. Now let $W = U \cap V$. Then again by Corollary 1.12.4.5, 6, we have that W is separated. Consequently, we get that $\Delta : W \to W \times_{\mathbb{Z}} W$ is a closed immersion. We now claim that $W \times_{\mathbb{Z}} W \cong U \times_{\mathbb{Z}} V$. Indeed, this follows immediately from the universal property of fiber product. It follows that $\Delta : W \to \operatorname{Spec}(A \otimes_{\mathbb{Z}} B)$ is a closed immersion. By Corollary 1.4.4.14, W is spectrum of a quotient of $A \otimes_{\mathbb{Z}} B$. Consequently, W is affine, as needed. □

Separatedness of projective schemes

We next see that any projective scheme is separated.

Lemma 1.12.4.7. *Let S be a graded ring. Then,* $\operatorname{Proj}(S) \to \operatorname{Spec}(\mathbb{Z})$ *is separated.*

Proof. We need only check that the diagonal $\Delta : \operatorname{Proj}(S) \to \operatorname{Proj}(S) \times_{\operatorname{Spec}(\mathbb{Z})} \operatorname{Proj}(S)$ has closed image (Lemma 1.12.4.3). Since one can check a closed set locally and sets of the form $D_+(f) \times D_+(g)$ forms an open cover of $\operatorname{Proj}(S) \times_{\operatorname{Spec}(\mathbb{Z})} \operatorname{Proj}(S)$ for $f, g \in S_+$ homogeneous, therefore we reduce to checking that for $C = \Delta^{-1}(D_+(f) \times D_+(g))$, the restriction $\Delta|_C : C \to D_+(f) \times D_+(g)$ has closed image.

Since $C = D_+(fg) \cong \operatorname{Spec}\left(S_{(fg)}\right)$ and $D_+(f) \times D_+(g) \cong \operatorname{Spec}\left(S_{(f)} \otimes_{\mathbb{Z}} S_{(g)}\right)$, therefore we reduce to showing that the induced map $S_{(f)} \otimes_{\mathbb{Z}} S_{(g)} \to S_{(fg)}$ is surjective. This is clear, as for any $u/f^n g^n \in S_{(fg)}$ where let us denote $k = \deg f, l = \deg g$, for any m large enough such that all exponents in the below are positive, we obtain that

$$\frac{ug^{mk-n}}{f^{ml+n}} \otimes \frac{f^{ml}}{g^{mk}} \mapsto \frac{u}{f^n g^n}.$$

[49] in-fact, any topological immersion is separated, as is clear from the proof.
[50] This doesn't require the noetherian hypothesis.
[51] Exercise II.4.3 of Hartshorne.

Thus, the image of Δ is closed[52]. $\square$

Uniqueness of centers of valuations for varieties

We show a curious property for abstract varieties that any valuation defined over its function field has a unique *center*, if it exists[53]. See Definition 1.4.2.9 for definition of center points of a valuation over function field of an integral scheme.

Lemma 1.12.4.8. [54] *Let X be an integral scheme of finite type over k with function field K. If X is separated, then any valuation over K has a unique center if it exists.*

Proof. We will use the valuative criterion for this. Let $v : K \to G$ be a valuation over K with valuation ring $R \subseteq K$. Let $x, y \in X$ be two centers of v. As $K \subseteq K$, therefore by Lemma 1.6.1.1, 3, there exists a unique map $\operatorname{Spec}(K) \to X$ mapping $\star \mapsto \eta$, where η is the generic point of X. It follows that we have the following commutative square

$$\begin{tikzcd} \operatorname{Spec}(K) \arrow[r] \arrow[d] & X \arrow[d] \\ \operatorname{Spec}(R) \arrow[r] & \operatorname{Spec}(\mathbb{Z}) \end{tikzcd} . \tag{*}$$

As R is a local ring, therefore by Lemma 01J6 of StacksProject, we have a bijection between maps $\operatorname{Spec}(R) \to X$ and tuples (z, φ) where $z \in X$ and $\varphi : \mathcal{O}_{X,z} \to R$ is a local ring homomorphism. Consequently, as $\mathcal{O}_{X,x}$ and $\mathcal{O}_{X,y}$ are dominated by R, we obtain two tuples (x, ι_x) and (y, ι_y) where $\iota_x : \mathcal{O}_{X,x} \hookrightarrow R$ and $\iota_y : \mathcal{O}_{X,y} \hookrightarrow R$ are the two domination maps. Note that by the definition of domination, these two maps are local ring homomorphisms. Consequently, we get two maps $\operatorname{Spec}(R) \to X$ which makes the (*) commute. By the valuative crieterion of Theorem 1.12.4.4, the two maps $\operatorname{Spec}(R) \to X$ are same, and thus so are the tuples (x, ι_x) and (y, ι_y), proving that $x = y$. $\square$

[52]in-fact we have also shown in the process that Δ is a closed immersion, thus we may not use Lemma 1.12.4.3

[53]It will always exist (and thus be unique) if the variety is proper, as is shown in the next section.

[54]Exericse II.4.5 a) of Hartshorne.

1.12.5 Affine morphisms and global Spec

In this section, we cover important global generalization of Spec $(-)$. In particular, let X be a scheme and $\mathcal{F}$ be a quasicoherent $\mathcal{O}_X$-algebra, that is, an $\mathcal{O}_X$-module which is a sheaf of rings as well. Then we will construct a scheme $\mathbf{Spec}(\mathcal{F})$ over X which will behave as if it is constructed out of open affine subschemes U of X and the corresponding algebras $\mathcal{F}(U)$.

This construction will be used to show how locally free sheaf of constant rank actually corresponds to vector bundles. They are used elsewhere as well.

Definition 1.12.5.1 (**Affine morphism**). A map $f : X \to Y$ of schemes is called an affine morphism if there is an affine open cover $\{V_\alpha\}$ of Y such that $f^{-1}(V_\alpha)$ is an open affine scheme.

Remark 1.12.5.2. It follows from definition that any finite morphism is affine.

The first major property of affine maps is that they are local on target.

Proposition 1.12.5.3. *Let $f : X \to Y$ be a map. Then the following are equivalent.*

1. *f is affine.*
2. *For any open affine $V \subseteq Y$, $f^{-1}(V)$ is an open affine in X.*

Proof. We need only do 1 $\Rightarrow$ 2. This has been done in the proof of Proposition 1.12.3.2. □

Lemma 1.12.5.4. *Let $f : X \to Y$ be an affine morphism. Then f is quasicompact and separated.*

Proof. The fact that f is quasicompact is immediate by definition. Separatedness follows from Corollary 1.12.4.5, 6 and Lemma 1.12.4.2. □

The main theorem for affine maps is that they all come from quasicoherent algebras over the structure sheaf. Indeed, we have the following construction to obtain a scheme over Y by a quasi-coherent $\mathcal{O}_Y$-algebra.

Theorem 1.12.5.5. [55] *Let Y be a scheme and $\mathcal{A}$ be a quasicoherent $\mathcal{O}_Y$-algebra over Y. Then there exists a scheme*

$$f : \mathbf{Spec}(\mathcal{A}) \to Y$$

unique with respect to the property that for any open affine $V \subseteq Y$, we have $f^{-1}(V) \cong \mathrm{Spec}\,(\mathcal{A}(V))$ and for any inclusion $U \hookrightarrow V$ of open affines, the map $f^{-1}(U) \to f^{-1}(V)$ is induced by the restriction map $\rho : \mathcal{A}(V) \to \mathcal{A}(U)$.

Proof. Let $V = \mathrm{Spec}\,(B) \subseteq Y$ be an open affine in Y. Then we have a ring homomorphism $B \to \mathcal{A}(V)$ as $\mathcal{A}(V)$ is a B-algebra. Consequently, we get the map $\pi_V : \mathrm{Spec}\,(\mathcal{A}(V)) \to Y$ factoring through V. Observe that for any open affine $U \hookrightarrow V$, we have the following commutative triangle

$$\begin{tikzcd} \mathcal{A}(V) \arrow[dr, "\rho"] & \\ B \arrow[u] \arrow[r] & \mathcal{A}(U) \end{tikzcd} .$$

[55] Exercise II.5.17, c) of Harthsorne.

We now wish to glue the affine schemes $\pi_i : \mathrm{Spec}\,(\mathcal{A}(V_i)) \to Y$ where V_i varies over open affines of Y. Indeed, let $X_i = \mathrm{Spec}\,(\mathcal{A}(V_i))$ and $U_{ij} = \pi_i^{-1}(V_i \cap V_j)$ an open subscheme of X_i. We claim that there is a natural isomorphism $\varphi_{ij} : U_{ij} \to U_{ji}$ which satisfies the cocycle condition, so that we can glue these schemes together by Proposition 1.6.2.2 to get the desired scheme unique with the given properties. Indeed, to find φ_{ij}, we first observe that for $\pi_V : \mathrm{Spec}\,(\mathcal{A}(V)) \to Y$, we have that $\pi_{V*}\mathcal{O}_{\mathrm{Spec}(\mathcal{A}(V))} \cong \mathcal{A}_{|V}$. This is where quasicoherence is used and follows from checking on basis and using globalized restriction of scalars (Lemma 1.2.3.4). Using this isomorphism, we see that $\mathcal{O}_{X_i}(\pi_i^{-1}(V_i \cap V_j)) \cong \mathcal{A}(V_i \cap V_j) \cong \mathcal{O}_{X_j}(\pi_j^{-1}(V_j \cap V_i))$. Consequently, we have a commutative triangle where $V_i = \mathrm{Spec}\,(B_i)$

$$\begin{tikzcd}
\mathcal{A}(V_i) \arrow[dr, "\rho"] & \\
B_i \arrow[u] \arrow[r] & \mathcal{O}_{X_j}(\pi_j^{-1}(V_i \cap V_j))
\end{tikzcd}.$$

By Theorem 1.3.0.5, we get the following commutative triangle

$$\begin{tikzcd}
X_i \arrow[d, "\pi_i"'] & \\
V_i & \pi_j^{-1}(V_i \cap V_j) \arrow[l] \arrow[ul, "\varphi_{ji}"']
\end{tikzcd}$$

By commutativity of this triangle, it follows that the unique morphism φ_{ji} factors through $\pi_i^{-1}(V_i \cap V_j)$. Interchanging i and j we get that φ_{ji} is an isomorphism. By uniqueness of φ_{ij}, we further get the cocycle condition, as required. □

We see from the proof the following.

Corollary 1.12.5.6. *Let Y be a scheme, $\mathcal{A}$ a quasicoherent $\mathcal{O}_Y$-algebra and $f : \mathbf{Spec}(\mathcal{A}) \to Y$ the global spec. Then, $f_*\mathcal{O}_{\mathbf{Spec}(\mathcal{A})} \cong \mathcal{A}$.*

Proof. In the proof, we showed that for any open affine $V \subseteq Y$, we have $f_*\mathcal{O}_{\mathbf{Spec}(\mathcal{A})|f^{-1}(V)} \cong \pi_{V*}\mathcal{O}_{\mathrm{Spec}(\mathcal{A}(V))} \cong \mathcal{A}_{|V}$ and this isomorphism is compatible with restrictions. Consequently, we have an isomorphism between $f_*\mathcal{O}_{\mathbf{Spec}(\mathcal{A})}$ and $\mathcal{A}$ over a base, which gives the required isomorphism as sheaves over Y. □

It is immediate to see by above theorem that global spec is always affine over the base.

Corollary 1.12.5.7. *Let Y be a scheme and $\mathcal{A}$ a quasicoherent $\mathcal{O}_Y$-algebra. Then the morphism*

$$f : \mathbf{Spec}(\mathcal{A}) \to Y$$

is affine. □

We now prove the converse of the above corollary.

Proposition 1.12.5.8. *Let $f : X \to Y$ be an affine morphism. Then,*

1. *$f_*\mathcal{O}_X$ is a quasicoherent $\mathcal{O}_Y$-algebra,*

2. *there is an isomorphism*

$$X \cong \mathbf{Spec}(f_*\mathcal{O}_X).$$

Proof. 1. This is immediate from the fact that the morphism f is quasicompact and separated by Lemma 1.12.5.4 (Lemma 1.9.1.17).

2. Let $\{V_\alpha\}$ be a basis consisting of open affines of Y. Then, $\{f^{-1}(V_\alpha)\}$ is an open affine basis of X by Proposition 1.12.5.3. Then, we have a canonical isomorphism $f^{-1}(V_\alpha) \cong \operatorname{Spec}\left(\mathcal{O}_X(f^{-1}(V_\alpha))\right)$. Moreover, for $V_\alpha \hookrightarrow V_\beta$, we have $f^{-1}(V_\alpha) \hookrightarrow f^{-1}(V_\beta)$ obtained by restriction $\rho : \mathcal{O}_X(f^{-1}(V_\beta)) \to \mathcal{O}_X(f^{-1}(V_\alpha))$. Hence by uniqueness of Theorem 1.12.5.5, we conclude the proof. $\square$

We may sum this up in the following bijection.

Corollary 1.12.5.9. [56] *Let Y be a scheme. We have the following bijection*

$$\left\{\text{Affine morphisms } X \xrightarrow{f} Y\right\} \cong \left\{\begin{matrix}\text{Quasicoherent} & \mathcal{O}_Y\text{-} \\ \text{algebras } \mathcal{A} & \end{matrix}\right\}$$

established by $f \mapsto f_*\mathcal{O}_X$ and $\mathbf{Spec}(\mathcal{A}) \leftarrow\!\shortmid \mathcal{A}$. $\square$

We see that any closed immersion is an affine map.

Proposition 1.12.5.10. *Let $i : Z \to X$ be a closed immersion. Then i is an affine map.*

Proof. By Proposition 1.4.4.11, there is an open affine cover $\{V_k\}$ of X such that $i : Z \cap V_k \hookrightarrow V_k$ is a closed immersion. Denote $V_k = \operatorname{Spec}(A_k)$. Thus, $Z \cap V_k = \operatorname{Spec}(A_k/I_k)$ and hence $i^{-1}(V_k) = \operatorname{Spec}(A_k/I_k)$, as required. Alternatively, it follows from the fact that any closed immersion is a finite map (Proposition 1.12.3.9). $\square$

Example 1.12.5.11 (A non-affine map)**.** Consider the map $\mathbb{A}^2 \setminus \{0\} \hookrightarrow \mathbb{A}^2$. We claim that this is not an affine map. Indeed, assuming to the contrary, there is a basic open affine $U = D(f)$ for $f \in k[x,y]$ of $\mathbb{A}^2$ containing 0 such that $U \setminus 0$ is open affine. But as can be checked, the coordinate rings of U and $U \setminus 0$ are isomorphic. It follows at once that $U \cong \operatorname{Spec}(k[x,y]_f) \cong U \setminus 0$, a contradiction to the fact that $U \ncong U \setminus 0$.

Remark 1.12.5.12. There is a canonical way of constructing quasi-coherent algebras out of quasi-coherent modules, that is, by using symmetric algebra. Thus for any quasi-coherent $\mathcal{O}_X$-module $\mathcal{E}$, we get an algebra $\operatorname{Sym}(\mathcal{E})$. Denote $\mathbb{V}(\mathcal{E}) := \mathbf{Spec}(\operatorname{Sym}(\mathcal{E}))$. Thus by Theorem 1.12.5.5, we have an affine map

$$f : \mathbb{V}(\mathcal{E}) \to X$$

such that $f_*\mathcal{O}_{\mathbb{V}(\mathcal{E})} = \operatorname{Sym}(\mathcal{E})$, a graded $\mathcal{O}_X$-algebra, where the first graded piece is $\mathcal{E}$.

Here is the universal property of **Spec**.

Theorem 1.12.5.13. *Let X be a scheme and $\mathcal{A}$ be a quasi-coherent $\mathcal{O}_X$-algebra. For any X-scheme $f : T \to X$, there is a natural isomorphism:*

$$\operatorname{Hom}_{\mathbf{Sch}_{/X}}(T, \mathbf{Spec}(\mathcal{A})) \longrightarrow \operatorname{Hom}_{\mathbf{QCohAlg}(\mathcal{O}_X)}(\mathcal{A}, f_*\mathcal{O}_T).$$

[56] Exercise II.5.17, d) of Hartshorne.

1.12.6 Vector bundles

We study vector bundles over schemes.

Definition 1.12.6.1 (Geometric vector bundles). Let X be a scheme. A geometric vector bundle of rank n over X is a map $p : E \to X$ such that there is a cover U_i of X and isomorphisms $\varphi_i : p^{-1}(U_i) \to \mathbb{A}^n_{\mathbb{Z}} \times U_i$ such that

$$\begin{tikzcd} p^{-1}(U_i) \arrow[r, "\varphi_i"] \arrow[d, "p"'] & U_i \times \mathbb{A}^n_{\mathbb{Z}} \arrow[dl, "\pi_1"] \\ U_i & \end{tikzcd}$$

commutes and for any open affine $V = \mathrm{Spec}\,(A) \subseteq U_i \cap U_j$, the composite

$$\mathbb{A}^n_A \cong V \times \mathbb{A}^n_{\mathbb{Z}} \xleftarrow{\varphi_i} p^{-1}(V) \xrightarrow{\varphi_j} V \times \mathbb{A}^n_{\mathbb{Z}} \cong \mathbb{A}^n_A$$

is a linear isomorphism of $\mathbb{A}^n_A$, i.e. $\varphi_j \circ \varphi_i^{-1} : \mathbb{A}^n_A \to \mathbb{A}^n_A$ is given by $\theta : A[x_1, \dots, x_n] \to A[x_1, \dots, x_n]$ which is A-linear and $\theta(x_i) = \sum_j a_{ij}x_j$ for some $a_{ij} \in A$.

If $p : E \to X$ and $p' : E' \to X$ are two vector bundles of rank n and m over X, then a map of vector bundles is an X-morphism $f : E \to E'$ such that if $\varphi : p^{-1}(U) \to U \times \mathbb{A}^n_{\mathbb{Z}}$ and $\psi : p'^{-1}(U') \to U' \times \mathbb{A}^m_{\mathbb{Z}}$ are local trivializations of E and E', then the horizontal composite

$$(U \cap U') \times \mathbb{A}^n_{\mathbb{Z}} \xleftarrow[\cong]{\varphi} p^{-1}(U \cap U') \xrightarrow{f} p'^{-1}(U \cap U') \xrightarrow[\cong]{\psi} (U \cap U') \times \mathbb{A}^m_{\mathbb{Z}}$$

is a linear map of affine spaces $\mathbb{A}^n_{U\cap U'} \to \mathbb{A}^m_{U\cap U'}$. We denote the category of geometric vector bundles on X by $\mathbf{VB}(X)$.

The following is the main theorem. Denote the category of locally free modules of finite rank on a scheme X by $\mathbf{LocFree}(X)$.

Theorem 1.12.6.2. *Let X be a scheme. There is an equivalence of categories*

$$\mathbf{LocFree}(X) \xrightarrow{\equiv} \mathbf{VB}(X).$$

We construct functors which are essential inverse of each other as follows.

Construction 1.12.6.3 (Vector bundle from a locally free module)**.** Consider the assignment

$$\begin{aligned} V : \mathbf{LocFree}(X) &\longrightarrow \mathbf{VB}(X) \\ \mathcal{E} &\longmapsto \mathbf{Spec}(\mathrm{Sym}(\mathcal{E})). \end{aligned}$$

If $\mathcal{E}$ is locally free of rank n, then we now show that $V(\mathcal{E})$ is a indeed a geometric vector bundle of rank n. Let $U \subseteq X$ be an open such that $\mathcal{E}_{|U}$ is a free $\mathcal{O}_{X|U}$-module of rank n. We wish to show that $V(\mathcal{E})$ on U is trivial. Indeed,

Construction 1.12.6.4 (Locally free module from a vector bundle)**.**

1.12.7 Proper

This and the next section brings us closer to detecting when a scheme is projective (i.e. is a subscheme of projective scheme). Proper maps corresponds roughly to the intuition that the scheme $X \to Y$ should not have any *missing points*.

Definition 1.12.7.1. (**Universally closed and proper maps**) A map $f : X \to Y$ is said to be universally closed if f is closed and for any base extension $Y' \to Y$, the base extension of X, denoted $f' : X' \to Y'$ is also closed as in the diagram below:

$$\begin{tikzcd} X' \arrow[r] \arrow[d, "f'"'] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "f"] \\ Y' \arrow[r] & Y \end{tikzcd}.$$

Consequently, f is said to be proper if it is separated, finite type and is universally closed.

The main result here is again a valuative criterion which allows a lot of properties of such maps to be derived quite easily.

Theorem 1.12.7.2. *(Valuative criterion of properness) Let $f : X \to Y$ be a finite type map of schemes where X is noetherian. Then the following are equivalent.*

1. *f is proper.*
2. *Pick any field K and any valuation ring R with fraction field K (see Section 16.10). Let i : $\mathrm{Spec}\,(K) \to \mathrm{Spec}\,(R)$ be the map corresponding to $R \hookrightarrow K$. For all $g : \mathrm{Spec}\,(R) \to Y$ and $h : \mathrm{Spec}\,(K) \to X$ such that the square commutes, there exists a unique lift of g along f as to make the following diagram commute:*

$$\begin{tikzcd} \mathrm{Spec}\,(K) \arrow[r, "h"] \arrow[d, "i"'] & X \arrow[d, "f"] \\ \mathrm{Spec}\,(R) \arrow[r, "g"'] \arrow[ur, dashed, "!"] & Y \end{tikzcd}.$$

Note that whereas in Theorem 1.12.4.4 we had that there exists *atmost one* lift, here we have that there exists *unique* lift (it exists and there is only one).

Corollary 1.12.7.3. *Let us work in the category of noetherian schemes. Then,*

1. *if $X \to Y \to Z$ is proper and $Y \to Z$ is separated, then $X \to Y$ is proper,*
2. *closed immersion are proper,*
3. *proper maps are stable under base extensions,*
4. *composite of proper maps is proper,*
5. *for two proper schemes $X \to S, Y \to S$ in* **Sch**$/S$*, their product $X \times_S Y \to S$ is proper,*
6. *a map $f : X \to Y$ is proper if and only if there exists an open cover V_i of Y such that the restriction $f|_{f^{-1}(V_i)} : f^{-1}(V_i) \to V_i$ is proper.*

Proof. **TODO : From notebook.** □

Remark 1.12.7.4 (Serre's GAGA-1)**.** Let $f : X \to Y$ be a map between $\mathbb{C}$-varieties. Then Serre proved the following equivalence:

1. $f : X \to Y$ is proper,
2. $f : X(\mathbb{C}) \to Y(\mathbb{C})$ is a proper map of topological spaces,

where recall that a map of spaces is proper if inverse image of any compact is compact.

1.12.8 Projective

We now define maps of schemes which factors through a projective space over the target. This will be fundamental, as the most natural type of schemes we find in nature are projective varieties appearing as closed subschemes of the projective space over a field. Though we will work more generally, but this will pay off in some of the later discussions. See Definition 1.8.2.14 for projective spaces over a scheme.

Definition 1.12.8.1. (**Projective and quasi-projective maps**) Let $f : X \to Y$ be a map of schemes. We say f is projective if there exists an $n \in \mathbb{N}$ such that f factors as a closed immersion $X \to \mathbb{P}^n_Y$ followed by the struture map $\mathbb{P}^n_Y \to Y$ as in

$$\begin{tikzcd} X \arrow[r, "\text{cl. imm.}"] \arrow[dr, "f"'] & \mathbb{P}^n_Y \arrow[d] \\ & Y \end{tikzcd}.$$

Further, a map $f : X \to Y$ is said to be *quasi-projective* if f factors first into an open immersion $X \to X'$ and then a projective map $X' \to Y$ as in

$$\begin{tikzcd} X' \arrow[r, "\text{cl. imm.}"] & \mathbb{P}^n_Y \arrow[d] \\ X \arrow[u, "\text{op. imm.}"] \arrow[r, "f"'] & Y \end{tikzcd}.$$

Thus quasi-projective maps corresponds to the usual notion of quasi-projective varieties (open subsets of projective varieties in a projective n-space).

The important point to keep in mind about projective maps is that they are proper.

Theorem 1.12.8.2. *Let X and Y be noetherian schemes.*

1. *If $f : X \to Y$ is projective, then f is proper.*
2. *If $f : X \to Y$ is quasi-projective, then f is finite type and separated.*

Proof. 1. Since closed immersions are proper and proper maps are stable under base change (Corollary 1.12.7.3), we may reduce to showing that for each $n \in \mathbb{N}$, the scheme $\mathbb{P}^n_{\mathbb{Z}} \to \operatorname{Spec}(\mathbb{Z})$ is proper. It is clear that $\mathbb{P}^n_{\mathbb{Z}}$ is finite type $\mathbb{Z}$-scheme which is furthermore separated by Corollary 1.12.4.7.

In order to show that $\mathbb{P}^n_{\mathbb{Z}}$ is proper, we will proceed by induction over n. For $n = 0$, we have $\mathbb{P}^0_{\mathbb{Z}} \cong \operatorname{Spec}(\mathbb{Z})$, which is trivially proper over $\operatorname{Spec}(\mathbb{Z})$. Now suppose $\mathbb{P}^{n-1}_{\mathbb{Z}}$ is proper over $\operatorname{Spec}(\mathbb{Z})$. We wish to show that $\mathbb{P}^n_{\mathbb{Z}}$ is proper. We will use valuative criterion for this (Theorem 1.12.7.2). Consider a valuation ring R with fraction field K such that we have maps g, h making the following commute:

$$\begin{tikzcd} \operatorname{Spec}(K) \arrow[r, "h"] \arrow[d] & \mathbb{P}^n_{\mathbb{Z}} \arrow[d] \\ \operatorname{Spec}(R) \arrow[r, "g"'] & \operatorname{Spec}(\mathbb{Z}) \end{tikzcd}.$$

Consequently, we wish to define a unique map Spec $(R) \to \mathbb{P}^n_{\mathbb{Z}}$ which makes everything commute.

Denote Spec $(K) = \{\star\}$ and $\xi = h(\star) \in \mathbb{P}^n_{\mathbb{Z}}$. We now observe that if $\xi \in V(x_{i_0})$ for any $i_0 = 0, \dots, n$, then by the natural isomorphism $V(x_{i_0}) \cong \mathbb{P}^{n-1}_{\mathbb{Z}}$ and obvious restrictions, we get the following commutative diagram:

$$\begin{tikzcd} \mathrm{Spec}\,(K) \arrow[r, "h"] \arrow[d] & \mathbb{P}^{n-1}_{\mathbb{Z}} \arrow[d] \\ \mathrm{Spec}\,(R) \arrow[r, "g"'] & \mathrm{Spec}\,(\mathbb{Z}) \end{tikzcd}.$$

Consequently, by inductive hypothesis, we have a unique lift Spec $(R) \to \mathbb{P}^{n-1}_{\mathbb{Z}}$ and thus a map Spec $(R) \to \mathbb{P}^n_{\mathbb{Z}}$ making the diagram commutative. This is sufficient by the fact that $\mathbb{P}^n_{\mathbb{Z}}$ is separated (Lemma 1.12.4.7) and by valuative criterion (Theorem 1.12.4.4).

We next need to cover the case when ξ is not in any hyperplane $V(x_i)$, that is, when $\xi \in \bigcap_{i=0}^n D_+(x_i)$. We will construct a map Spec $(R) \to \mathbb{P}^n_{\mathbb{Z}}$ which makes everything commute and we will be done by separatedness of $\mathbb{P}^n_{\mathbb{Z}}$ (Lemma 1.12.4.7). In this case, we obtain that $\mathcal{O}_{\mathbb{P}^n_{\mathbb{Z}},\xi} \cong \mathbb{Z}[x_0/x_i, \dots, \widehat{x_i/x_i}, \dots, x_n/x_i]$ for all $i = 0, \dots, n$ as $D_+(x_i) \cong$ Spec $\left(\mathbb{Z}[x_0, \dots, x_n]_{(x_i)}\right)$, (Lemma 1.8.2.4). Consequently, $x_i/x_j \in \mathcal{O}_{\mathbb{P}^n_{\mathbb{Z}},\xi}$ is invertible for all $i, j = 0, \dots, n$, hence $x_i/x_j \notin \mathfrak{m}_{\mathbb{P}^n_{\mathbb{Z}},\xi}$. Denote further $f_{ij} \in \kappa(\xi)$ to be the image of x_i/x_j under the map $\mathcal{O}_{\mathbb{P}^n_{\mathbb{Z}},\xi} \to \mathcal{O}_{\mathbb{P}^n_{\mathbb{Z}},\xi}/\mathfrak{m}_{\mathbb{P}^n_{\mathbb{Z}},\xi} = \kappa(\xi)$.

The map $h :$ Spec $(K) \to \mathbb{P}^n_{\mathbb{Z}}$ is equivalent to the data of the point $\xi \in \mathbb{P}^n_{\mathbb{Z}}$ and $\kappa(\xi) \hookrightarrow K$ (Lemma 1.6.1.1). Thus we have $f_{ij} \in K$ for all $i, j = 0, \dots, n$. In order to define the map $j :$ Spec $(R) \to \mathbb{P}^n_{\mathbb{Z}}$ in this case, it is sufficient to obtain a map $\mathbb{Z}[x_0/x_i, \dots, \widehat{x_i/x_i}, \dots, x_n/x_i] \to R$ such that the following commutes:

$$\begin{tikzcd} K & Z[x_0/x_i, \dots, \widehat{x_i/x_i}, \dots, x_n/x_i] \arrow[l] \arrow[dl] \\ R \arrow[u] & \mathbb{Z} \arrow[l] \arrow[u] \end{tikzcd}.$$

We will now construct such a map. Let $v : K \to G$ be the valuation corresponding to the valuation ring R (so that R is the value ring of v), where G is a totally ordered abelian group. Consider the collection of elements $f_{10}, \dots, f_{n0} \in K$ and denote $g_i = v(f_{i0}) \in G$. Let $g_m = \min_i g_i$. Consequently, for each $i = 0, \dots, n$ we obtain $0 \le g_i - g_m = v(f_{i0}) - v(f_{m0}) = v(f_{i0} f_{0m}) = v(f_{im})$. Thus, $f_{im} \in R$. Hence, we can construct the following map:

$$\begin{aligned} Z[x_0/x_i, \dots, \widehat{x_i/x_i}, \dots, x_n/x_i] &\longrightarrow R \\ \frac{x_i}{x_m} &\longmapsto f_{im}. \end{aligned}$$

It is immediate that the above map makes the above diagram commute.
2. Since open immersions are separated (Corollary 1.12.4.5) and an open immersion $X \to X'$ where X is noetherian is immediately quasicompact, so by Proposition 1.12.2.5, 2, the result follows. $\square$

1.12.9 Flat

Look at the following MO post for more clarifications. Flat maps of schemes capture the notion of a "continuous family of schemes parameterized by points of a base scheme". However, the notion of flatness is very algebraic, as we shall soon see. We collect the properties of flat modules in the Special Topics, Chapter 16.

1.12.10 Rational

We discuss rationality questions for varieties. We fix an algebraically closed field k for this section and all schemes are over k.

Definition 1.12.10.1 (Rational & birational maps). Let X, Y be varieties and $U, V \subseteq X$ be open subsets. Note they are dense and so is $U \cap V$. Two maps $f_U : U \to Y$ and $f_V : V \to Y$ are *equivalent* if $f_U|_W = f_V|_W$ for some open $W \subseteq U \cap V$. A rational map $f : X \dashrightarrow Y$ is an equivalence class of maps $f_U : U \to Y$, denoted $\langle f_U, U \rangle$, as defined above. Composition of two rational maps $f : X \dashrightarrow Y$ and $g : Y \dashrightarrow Z$ is given as follows: one has $f_U : U \to Y$ and $g_V : V \to Z$ where $U \subseteq X$ and $V \subseteq Y$ are open subsets, then $g_V \circ f_U : U \to Z$ is the required map which gives the rational composite $g \circ f : X \dashrightarrow Z$. One easily checks that this is well-defined. A rational map $f : X \dashrightarrow Y$ is *birational* if there is an inverse rational map $g : Y \dashrightarrow X$, that is, $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$ as rational maps. We get the category of varieties with rational maps, denoted $\mathbf{Var}_k^{\mathrm{R}}$. A rational map $f : X \dashrightarrow Y$ is *dominant* if some representative $f_U : U \to Y$ is dominant, that is, has dense image. It follows that every representative is dominant and composite of dominant rational maps is again rational. We denote the category of varieties and dominant rational maps as $\mathbf{Var}_k^{\mathrm{DR}}$. Clearly, there is an inclusion $\mathbf{Var}_k^{\mathrm{DR}} \hookrightarrow \mathbf{Var}_k^{\mathrm{R}}$.

Remark 1.12.10.2 (Faithful embedding of varieties)**.** For any map $f : X \to Y$ of varieties, we get a rational map $f : X \dashrightarrow Y$, which is given by the class $\langle f, X \rangle$. One immediately checks that this gives a functor

$$\langle -, - \rangle : \mathbf{Var}_k \longrightarrow \mathbf{Var}_k^{\mathrm{R}}.$$

We claim that the above functor is faithful. Indeed, if $\langle f, X \rangle = \langle g, X \rangle$ for two map of varieties $f, g : X \to Y$, then $f|_U = g|_U$ for some open dense $U \subseteq X$. We wish to show that $f = g$. This follows from the following result (Lemma 1.12.10.3). Hence, varieties embed faithfully into varieties with rational maps.

Lemma 1.12.10.3. *Let $f, g : X \to Y$ be two map of schemes such that there exists a dense open $U \subseteq X$ on which $f|_U = g|_U$. If X is reduced and Y is separated, then $f = g$.*

Proof. Denote $C = \{x \in X \mid f(x) = g(x)\}$. Observe that $C = h^{-1}(\Delta(Y))$ where $h = (f, g) : X \to Y \times Y$ mapping $x \mapsto (f(x), g(x))$ and $\Delta : Y \to Y \times Y$ is the diagonal map. As Y is separated, thus $\Delta(Y)$ is closed and hence $C \subseteq X$ is closed and contains U. As $\bar{U} = X$, therefore $C = X$ as sets. We wish to show that this equality is of schemes. To this end, we need only show that the only closed subscheme of a reduced scheme containing a dense open is X itself. Indeed, let $\operatorname{Spec}(A) \subseteq X$ be any affine open. We wish to show that the closed subscheme of $\operatorname{Spec}(A)$ given by $C \cap \operatorname{Spec}(A)$ is $\operatorname{Spec}(A)$ itself. Let $C \cap \operatorname{Spec}(A) = \operatorname{Spec}(A/I)$. Suffices to show that $I = 0$. As $\operatorname{Spec}(A/I) = \operatorname{Spec}(A)$ since $C = X$ as sets, therefore $I \subseteq \mathfrak{n}$, the nilradical of A. As A is reduced, therefore $\mathfrak{n} = 0$ and hene $I = 0$, as required. $\square$

Remark 1.12.10.4. We really do require reduced domain in the above lemma, as the following example shows. For the two maps $\varphi, \psi : k \to k[\epsilon]$ mapping $\varphi : cc$ and $\psi : c \mapsto c\epsilon$, we get two maps $f, g : \operatorname{Spec}(k[\epsilon]) \to \operatorname{Spec}(k)$. Both maps are equal on the whole space (both sets are singleton). Yet the maps are not the same, as they are induced by different functions. The conclusion is, maps on an unreduced scheme are not determined by their mapping on points.

To state the main theorem of this section, we have to promote the construction of function fields to a functor.

Construction 1.12.10.5. Let $f : X \dashrightarrow Y$ be a dominant rational map of varieties. Take any element $(V, \varphi) \in K(Y)$. We define an element of $K(X)$ as follows. By shrinking $U = f^{-1}(V)$ which is possible as f is dominant, we may assume that $f_U : U \to Y$ is a representative of the rational map f. Then we define the class $(U, f_V^\flat(\varphi))$ in $K(X)$. One checks easily that this is a well-defined homomorphism of fields, denoted

$$K(f) : K(Y) \longrightarrow K(X).$$

Hence, K is a contravariant functor from category of varieties and dominant rational maps to category of fields over k.

The main theorem here is the following.

Theorem 1.12.10.6. *The function field functor is an equivalence between category of varieties over k with dominant rational map and category of finitely generated field extensions of k:*

$$K : \mathbf{Var}_k^{\mathrm{DR}} \longrightarrow \mathbf{Fld}_k^{\mathrm{fg}}.$$

One reason to discuss rational maps is that they give a geometric meaning to function fields of varieties.

Proposition 1.12.10.7. *Let X be a variety. Then there is a natural isomorphism*

$$K(X) \cong \mathrm{Hom}_{\mathbf{Var}_k^{\mathrm{DR}}}\left(X, \mathbb{P}^1\right).$$

Proof. By Theorem 1.12.10.6, we have a natural isomorphism

$$\mathrm{Hom}_{\mathbf{Var}_k^{\mathrm{DR}}}\left(X, \mathbb{P}^1\right) \cong \mathrm{Hom}_{\mathbf{Fld}_k^{\mathrm{fg}}}\left(K(\mathbb{P}^1), K(X)\right).$$

As $K(\mathbb{P}^1) = k(T)$ where T is the coordinate of an open affine patch $\mathrm{Spec}\,(k[T])$ and since any k-linear field homomorphism $\alpha : k(T) \to K(X)$ equivalently is determined by any non-constant function in $K(X)$, hence we have the natural isomorphism

$$\mathrm{Hom}_{\mathbf{Fld}_k^{\mathrm{fg}}}\left(k(T), K(X)\right) \cong K(X).$$

This completes the proof. $\square$

1.12.11 Smooth

1.12.12 Unramified

1.12.13 Étale

Étale maps is the place from where one enters the land of algebraic topology via algebraic geometry. Indeed, the fundamental goal here is to capture the notion of local isomorphism but in an algebraic context. The simplest place where one can understand them is a restricted version of this called finite étale maps. This is where we begin from as we shall need this in our discussion of Galois theory of schemes.

Finite étale

We refer to Algebra, Chapter 16 for background on separable algebras, in particular, to Definition 16.22.2.2 for free separable algebras.

We now define finite étale maps.

Definition 1.12.13.1. (**Finite étale scheme**) Let X be a base scheme. An X-scheme $p : Y \to X$ is said to be finite étale if there is an open affine covering of X given by $\{\operatorname{Spec}(A_i)\}_{i\in I}$ such that $p^{-1}(\operatorname{Spec}(A_i))$ is an open affine subscheme of Y given by $\operatorname{Spec}(B_i)$ such that the induced map $A_i \to B_i$ makes B_i a free separable A_i-algebra, for all $i \in I$. In such a situation, one calls Y a finite étale covering of X. Denote the category $\mathbf{Et}_{\text{fin}}(X)$ to be the full subcategory of $\mathbf{Sch}/X$ consisting of finite étale coverings of X.

Let us now give an example of finite étale scheme.

Example 1.12.13.2.

1.13 Coherent and quasicoherent sheaf cohomology

All schemes in this section are Noetherian. Cohomology will serve as an important tool to derive invariants on a given scheme. We would need the cohomology of abelian sheaves over a space (Chapter 20) and the notion of Noetherian schemes (Section 1.4) in this section. *Apart from Mumford, you may like to give Hida a visit.*

We refer to Topics in Sheaf Theory, Chapter 20 for classical Čech cohomology, derived functor cohomology and relations between them on topological spaces.

Since we are only dealing with noetherian schemes and the most important such schemes would be those which are closed subvarieties of projective space, so finite dimensional, therefore the following theorem of Grothendieck is of particular importance.

Theorem 1.13.0.1 (Grothendieck)**.** *Let X be a noetherian topological space of dimension n and $\mathcal{F}$ be an abelian sheaf over X. Then*

$$H^i(X;\mathcal{F}) = 0$$

for all $i > n$.

We now show some basic theorems in cohomology of sheaves over schemes which allows us to use Čech-cohomology for calculations instead of derived functor cohomology, because both becomes isomorphic.

1.13.1 Quasicoherent sheaf cohomology

Do from Hartshorne and Bruzzo.

As a corollary of Theorem 1.13.0.1, we have the following.

Corollary 1.13.1.1. *Let X be a noetherian scheme of dimension n. Then for any $\mathcal{O}_X$-module $\mathcal{M}$, $H^i(X;\mathcal{M}) = 0$ for $i > n$.* □

Theorem 1.13.1.2 (Serre)**.** *Let X be a projective scheme over a noetherian ring A. Then for any coherent $\mathcal{O}_X$-module $\mathcal{M}$, the $H^i(X;\mathcal{M})$ is a finitely generated A-module.*

Remark 1.13.1.3 (Euler characteristic)**.** If X is a projective variety over a noetherian ring A and $\mathcal{M}$ is any coherent $\mathcal{O}_X$-module $\mathcal{M}$, we have that $H^i(X;\mathcal{M})$ are finitely generated A-modules concentrated in $0 \leq i \leq n$. We may thus define the Euler-characteristic of $\mathcal{M}$ if $A = k$ as

$$\chi(\mathcal{M}) = \sum_{i\geq 0}(-1)^i \dim_k H^i(X;\mathcal{M}).$$

Theorem 1.13.1.4 (Serre)**.** *If X is a noetherian scheme then the following are equivalent:*
1. *X is affine.*
2. *$H^i(X;\mathcal{F}) = 0$ for all $i \geq 1$ and all quasicoherent $\mathcal{O}_X$-modules $\mathcal{F}$.*
3. *$H^i(X;\mathcal{F}) = 0$ for all $i \geq 1$ and all coherent $\mathcal{O}_X$-modules $\mathcal{F}$.*

Proposition 1.13.1.5 (Künneth)**.** *Let X be a noetherian separated scheme over k and L/k be an extension. Then for any coherent sheaf $\mathcal{F}$ over X*

$$H^i(X \times_k L;\mathcal{F}_L) \cong H^i(X;\mathcal{F}) \otimes_k L.$$

Corollary 1.13.1.6. *The arithmetic genus of a curve is invariant under change of base field.* □

Remark 1.13.1.7. To see the importance of arithmetic genus, one may look at Falting's theorem; any curve of arithmetic genus $g_a \geq 2$ has finitely many rational points.

GAGA

Serre proved the following famous equivalence in the mid 1950s.

Theorem 1.13.1.8 (Serre's GAGA)**.** *Let X be a projective scheme over $\mathbb{C}$. Then there is an equivalence of categories*

$$\mathbf{Coh}(X) \equiv \mathbf{Coh}(X(\mathbb{C})).$$

1.13.2 Application : Serre-Grothendieck duality

Do from Hida and Hartshorne.

1.13.3 Application : Riemann-Roch theorem for curves

Do from Hida and Hartshorne.

We denote by $\ell(D) := \dim_k H^0(X; \mathcal{L}(D))$ for a Cartier divisor D on X.

Theorem 1.13.3.1. *Let X be a regular, projective, geometrically integral variety of pure dimension 1. If D be a Cartier divisor on X, then*

$$\ell(D) + \ell(K \setminus D) = \deg(D) + 1 - g_a$$

where K is the canonical divisor and $g_a = \dim_k H^1(X; \mathcal{O}_X)$ is the arithmetic genus of X.

Part II

The Arithmetic Viewpoint

Chapter 2

Foundational Arithmetic

Contents

In this foundational chapter, we will discuss some topics in classical number theory.

2.1 Fundamental properties of $\mathbb{Z}$

We wish to see the following list of properties of integers, all of which are immediate, but good to keep in mind.

Theorem 2.1.0.1. *Consider the ring of integers $\mathbb{Z}$. Then, $\mathbb{Z}$ is*

1. *an Euclidean domain,*
2. *a gcd domain,*
3. *a principal ideal domain,*
4. *a unique factorization domain,*
5. *a noetherian ring,*
6. *a normal domain,*
7. *a dimension 1 ring.*

2.2 Algebraic number fields

The main objects of study in number theory are number fields. We will define them and will discuss some of their basic properties, before doing a more involved study of real and imaginary quadratic number fields.

Definition 2.2.0.1 (**Algebraic number fields and ring of integers**)**.** A field K is an algebraic number field if it is a finite extension of $\mathbb{Q}$. The ring of integers or the integral ring of an algebraic number field K is the integral closure of the inclusion $\mathbb{Z} \hookrightarrow K$, and is denoted by $\mathcal{O}_K$.

Remark 2.2.0.2. The Proposition 16.7.1.10 guarantees that $\mathcal{O}_K$ is indeed a ring. Observe that $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$ as $\mathbb{Z}$ is a normal domain. Indeed, this can be generalized.

It is clear that every algebraic number field is of characteristic 0, as it contains $\mathbb{Q}$. Some of the first (but important) properties of $\mathcal{O}_K$ are as follows.

Proposition 2.2.0.3. *Let K be an algebraic number field.*

1. *The integral ring $\mathcal{O}_K$ is a domain.*
2. *Let R be a normal domain and K be its fraction field. Then $\mathcal{O}_K$ is a subring of R.*
3. *The fraction field of $\mathcal{O}_K$ is K.*
4. *We have $K = \mathbb{Q} \cdot \mathcal{O}_K$.*
5. *The integral ring $\mathcal{O}_K$ is a normal domain.*
6. *The integral ring $\mathcal{O}_K$ is of dimension 1.*
7. *The integral ring $\mathcal{O}_K$ is noetherian.*

Proof. 1. As $\mathcal{O}_K \subseteq K$, therefore it has no zero-divisors.
2. As R is a normal domain, therefore its integral closure in K is R itself. Thus, $\mathcal{O}_K \subseteq R$.
3. Pick any $s \in K$. As K is algebraic over $\mathbb{Q}$, therefore there exists $p(x) \in \mathbb{Q}[x]$ such that $p(s) = 0$ in K. Multiplying by common denominators of coefficients of p, it follows that

$$d_n s^n + d_{n-1} s^{n-1} + \cdots + d_1 s + d_0 = 0$$

in K where $d_i \in \mathbb{Z}$. Multiplying this by d_n^{n-1} and writing $t = sd_n$, we get

$$t^n + c_{n-1} t^{n-1} + \cdots + c_1 t + c_0 = 0$$

where $c_i \in \mathbb{Z}$. It follows that $t \in \mathcal{O}_K$ and thus $s = a/m$ where $a \in \mathcal{O}_K$ and $m \in \mathbb{Z}$. Thus, $s \in L$ where L is the fraction field of $\mathcal{O}_K$.
4. Follows from the proof of item **3**.
5. Let $C \subseteq K$ be normalization of $\mathcal{O}_K$ in K. Thus we have $\mathbb{Z} \hookrightarrow \mathcal{O}_K \hookrightarrow C$ where both are integral maps. It follows from Lemma 16.7.1.12 that the composite $\mathbb{Z} \hookrightarrow C$ is integral. Consequently, $C \subseteq \mathcal{O}_K$, yielding that $\mathcal{O}_K$ is a normal domain.
6. This follows from the corollary of Cohen-Seidenberg theorems (Corollary **??**) and that $\mathcal{O}_K$ is integral over $\mathbb{Z}$.
7. **TODO.** □

A direct corollary of this is that $\mathcal{O}_K$ is a very special type of ring.

Corollary 2.2.0.4. *Let K be an algebraic number field and $\mathcal{O}_K$ be its integral ring. Then $\mathcal{O}_K$ is a Dedekind domain (see Definition 16.11.0.1).*

Proof. From Proposition 2.2.0.3, 5,6,7, the result follows. □

The following is an important example of ring of integers.

Theorem 2.2.0.5. *Let $L = \mathbb{Q}(\sqrt{d})$ where $d \in \mathbb{Z}$ is a square-free integer (i.e. product of distinct primes). Then, $\mathcal{O}_L$ is given as follows:*

$$\mathcal{O}_L = \begin{cases} \mathbb{Z}[\sqrt{d}] & \text{if } d = 2, 3 \mod 4 \\ \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right] & \text{if } d = 1 \mod 4. \end{cases}$$

Proof. □

We now wish to study units and irreducibles in $\mathcal{O}_L$. To this end we require norm and trace of a finite extension as discussed in §**??**.

2.2.1 Quadratic number fields

We study imaginary quadratic number fields obtained by taking square root of some integer. A key tool for studying the relation between arithmetic and algebra of the situation is that of the norm and trace. We will also state for which quadratic number fields is the ring of integers a UFD, thus solving the fundamental problem of algebraic number theory in this restricted case; when is a ring of integers of a number field a UFD?

Part III

The Topological Viewpoint

Chapter 3

Foundational Geometry

Contents

Complete this chapter from Wedhorn's manifolds, sheaves and cohomology, and by Bredon's topology and geometry.
We will define the notion of a real and complex manifold. Some foundational constructions are made on them. We will take a rather modern viewpoint on the matter. We will further discuss ....

3.1 Locally ringed spaces and manifolds

We will make very fluid use of sheaves (see Chapter 20). Let us begin by the foundational structure in all of geometry, a (locally)ringed space.

Definition 3.1.0.1. (**Ringed and locally ringed spaces**) A ringed space is a pair $(X, \mathcal{O}_X)$ where X is a topological space and $\mathcal{O}_X$ is a sheaf of commutative R-algebras. The space $(X, \mathcal{O}_X)$ is locally ringed if the stalk $\mathcal{O}_{X,x}$ at each point $x \in X$ is a local ring. The sheaf $\mathcal{O}_X$ is called the structure sheaf of X.

In order to understand the relation between two such spaces, we next have to understand the morphism of (locally)ringed spaces. For a motivation, see Example 1.2.2.1.

Definition 3.1.0.2. (**Morphism of ringed and locally ringed spaces**) Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be two ringed spaces. A morphism $(f, f^\sharp) : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is given by a continuous map $f : X \to Y$ and a map of sheaves over X denoted $f^\sharp : f^{-1}\mathcal{O}_Y \to \mathcal{O}_X$. If $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ are locally ringed, then for $(f, f^\sharp)$ to be morphism of locally ringed spaces has to satisfy an additional condition that the induced map on stalks is a map of local rings. That is, for each $x \in X$, the induced map on stalks

$$f_x^\sharp : \mathcal{O}_{Y,f(x)} \longrightarrow \mathcal{O}_{X,x}$$

is such that $(f_x^\sharp)^{-1}(\mathfrak{m}_{X,x}) = \mathfrak{m}_{Y,f(x)}$ (see Special Topics, Remark 20.5.0.6). We call this map the comorphism at $x \in X$. In particular, this map is given by the unique map obtained by universality of direct limits under question: consider any open $V \ni f(x)$ in Y, we then obtain the following diagram:

$$\begin{tikzcd} \mathcal{O}_{X,x} & & \mathcal{O}_{Y,f(x)} \arrow[ll, dashed, "=:f_x^\sharp"'] \\ \mathcal{O}_X(f^{-1}(V)) \arrow[u, "\iota_{f^{-1}(V)}"] & & \mathcal{O}_Y(V) \arrow[ll, "f_V^\flat"] \arrow[u, "\iota_V"'] \arrow[llu, "\iota_{f^{-1}(V)} \circ f_V^\flat"] \end{tikzcd} .$$

In most of our purposes, the map $f^\flat$ will be given on sections by composing with f. In such situations, the map on stalks being local corresponds to the geometric intuition that all non-invertible functions around some open subset of $f(x)$ comes from non-invertible maps around x. This in some sense makes sure that the local data around $f(x)$ is completely available via f.

Definition 3.1.0.3. (**Composition**) Composition of two maps of locally ringed spaces is defined in the obvious manner. For $X \xrightarrow{g} Y \xrightarrow{f} Z$, we get maps $g^\flat : \mathcal{O}_Y \to g_*\mathcal{O}_X$ and $f^\sharp : f^{-1}\mathcal{O}_Z \to \mathcal{O}_Y$. Then, the map $f \circ g : X \to Z$ is defined on space level by just the composite $f \circ g$ of the continuous maps and on the sheaf level as the corresponding flat and sharp maps of $f \circ g : X \to Z$:

$$h^\flat : \mathcal{O}_Z \longrightarrow (f \circ g)_*\mathcal{O}_X$$
$$h^\sharp : (f \circ g)^{-1}\mathcal{O}_Z \longrightarrow \mathcal{O}_X.$$

In particular, for an open set $U \subseteq Z$, the corresponding map on local sections $h_U^\flat$ is given by the following composite:

$$\begin{tikzcd} \mathcal{O}_Z(U) \arrow[r, dashed, "h_U^\flat"] \arrow[d, "f_U^\flat"] & (f_*g_*\mathcal{O}_X)(U) \arrow[r, equal] & \mathcal{O}_X(g^{-1}f^{-1}(U)) \\ (f_*\mathcal{O}_Y)(U) \arrow[rr, equal] & & \mathcal{O}_Y(f^{-1}(U)) \arrow[u, "g^\flat_{f^{-1}(U)}"'] \end{tikzcd} .$$

Similarly, the corresponding morphism of stalks given by $h_x^\sharp$ is given by the usual

$$h_x^\sharp : (g^{-1}f^{-1}\mathcal{O}_Z)_x \cong \mathcal{O}_{Z,f(g(x))} \longrightarrow \mathcal{O}_{X,x}$$

which is the composite

$$\begin{tikzcd}
\mathcal{O}_{Z,h(x)} \arrow[d, "f^\sharp_{g(x)}"'] & \mathcal{O}_Z(W) \arrow[l] \arrow[d, "f^\flat_W"] \\
\mathcal{O}_{Y,g(x)} \arrow[d, "g^\sharp_x"'] & \mathcal{O}_Y(f^{-1}(W)) \arrow[l] \arrow[d, "g^\flat_{f^{-1}(W)}"] \\
\mathcal{O}_{X,x} & \mathcal{O}_X(g^{-1}(f^{-1}(W))) \arrow[l]
\end{tikzcd}.$$

Lemma 3.1.0.4. *Let $h : X \xrightarrow{g} Y \xrightarrow{f} Z$ be a morphism of ringed spaces. Consider the base change functors corresponding to maps g and f:*

$$g^{-1} : \mathbf{Sh}(Y) \longrightarrow \mathbf{Sh}(X)$$
$$f_* : \mathbf{Sh}(Y) \longrightarrow \mathbf{Sh}(Z).$$

and consider the following composite in $\mathbf{Sh}(Y)$

$$f^{-1}\mathcal{O}_Z \xrightarrow{f^\sharp} \mathcal{O}_Y \xrightarrow{g^\flat} g_*\mathcal{O}_X \ .$$

Then,
1. *$g^{-1}(g^\flat \circ f^\sharp) \cong h^\sharp$,*
2. *$f_*(g^\flat \circ f^\sharp) \cong h^\flat$.*

Proof. These are cumbersome but straightforward identities. For example, one has to observe that $f_*(f^\sharp) \cong f^\flat$ and that for an open set $U \subseteq Z$, we have $(f_*(g^\flat))_U = g^\flat_{f^{-1}(U)}$. □

We have a simple lemma for isomorphism of ringed spaces.

Lemma 3.1.0.5. *Let $f : X \to Y$ be a morphism of ringed spaces. Then, f is an isomorphism if and only if $f : X \to Y$ is a homeomorphism and $f^\flat : \mathcal{O}_Y \to f_*\mathcal{O}_X$ is an isomorphism.*

Proof. (L ⇒ R) Use Theorem 20.3.0.6, 3 and 4.
(R ⇒ L) One can explicitly construct a map of sheaves in the other direction in a straightforward manner. □

An open subspace of a ringed space also inherits the structure of a ringed space.

Definition 3.1.0.6. (**Open subspace and embedding**) Let $(X, \mathcal{O}_X)$ be a (locally) ringed space. An open subspace of $(X, \mathcal{O}_X)$ is an open subset $i : U \hookrightarrow X$ together with the inverse image sheaf $i^{-1}\mathcal{O}_X = \mathcal{O}_{X|U}$[1]. The pair $(U, \mathcal{O}_{X|U})$ is called an open subspace, $(U, \mathcal{O}_{X|U}) \hookrightarrow (X, \mathcal{O}_X)$. A map $(j, j^\sharp) : (Z, \mathcal{O}_Z) \to (X, \mathcal{O}_X)$ is an open embedding if $U := j(Z) \hookrightarrow X$ is open and $(j, j^\sharp) : (Z, \mathcal{O}_Z) \to (U, \mathcal{O}_{X|U})$ is an isomorphism of ringed spaces.

An important concept is of local isomorphism of ringed spaces, which will prove it's worth while defining manifolds.

Definition 3.1.0.7. (**Local isomorphism**) Let $f : X \to Y$ be a morphism of ringed spaces. One calls f to be a local isomorphism if there exists an open cover $\{U_i\}_{i \in I}$ of X such that $f|_{U_i} : U_i \to Y$ is an open embedding for all $i \in I$.

[1]It's a trivial matter to observe that inverse image of a sheaf to an open inclusion will be the restriction sheaf (see Lemma 20.5.0.3).

3.1.1 Local models and manifolds

Before we proceed further, we have to clearly state some of our local model spaces that we are going to use while defining the manifolds. Therefore the following example of ringed spaces are foundational.

Example 3.1.1.1. (*Sheaf of C^α-maps*) Let $X \subseteq \mathbb{R}^n$ be an open set and $\alpha \in \mathbb{N}^\infty$. One defines the following presheaf

$$\mathcal{C}^\alpha_{X;\mathbb{R}^m} := \{f : X \to \mathbb{R}^m \mid f \text{ is } C^\alpha\}$$

where the restriction maps are usual functional restrictions. Then, $\mathcal{C}^\alpha_{X;\mathbb{R}^m}$ forms a sheaf, called the sheaf of C^α maps on X. This sheaf has stalks as local rings which can be seen quite easily (set of all functions defined in *some* neighborhood of $x \in X$ has a ring structure with maximal ideal being all those functions taking value 0 at x). Hence, $(X, \mathcal{C}^\alpha_{\mathbb{R}^m})$ is a locally ringed space, where we dropped the subscript X for notational convenience.

Example 3.1.1.2. (*Sheaf of holomorphic maps*) Let $X \subseteq \mathbb{C}^n$ be an open set. One defines the following presheaf

$$\mathcal{C}^{\mathrm{hol}}_{X;\mathbb{C}^m} := \{f : X \to \mathbb{C}^m \mid f \text{ is holomorphic}\}$$

where the restriction maps are the usual functional restriction. This is easily seen to be a sheaf, called the sheaf of holomorphic functions over X. This endows $(X, \mathcal{C}^{\mathrm{hol}}_{\mathbb{C}^m})$ with the structure of a locally ringed space.

With these two examples, we can come to the notion of real and complex manifolds as follows.

Definition 3.1.1.3. (**Real and complex manifolds**) Let X be a Hausdorff and second-countable topological space. Then,

1. A locally $\mathbb{R}$-ringed space $(X, \mathcal{O}_X)$ is a real C^α-manifold if there exists an open covering $\{U_i\}_{i\in I}$ of X and for each $i \in I$, there exists a positive integer $n_i \in \mathbb{N}$ and an isomorphism of locally $\mathbb{R}$-ringed spaces $\varphi_i : (U_i, \mathcal{O}_{X|U_i}) \xrightarrow{\cong} (Y_i, \mathcal{C}^\alpha_{\mathbb{R}})$ for some open $Y_i \subseteq \mathbb{R}^{n_i}$. Hence a real C^α-manifold structure on X is the following tuple of data:
$$\left(X, \mathcal{O}_X, \{U_i\}_{i\in I}, \{Y_i \subseteq \mathbb{R}^{n_i}\}_{i\in I}, \{\varphi_i : (U_i, \mathcal{O}_{X|U_i}) \xrightarrow{\cong} (Y_i, \mathcal{C}^\alpha_{\mathbb{R}})\}_{i\in I}\right)$$
2. A locally $\mathbb{C}$-ringed space $(X, \mathcal{O}_X)$ is a complex manifold if there exists an open covering $\{U_i\}_{i\in I}$ of X and for each $i \in I$ there exists $n_i \in \mathbb{N}$ and an isomorphism of locally $\mathbb{C}$-ringed spaces $\varphi_i : (U_i, \mathcal{O}_{X|U_i}) \xrightarrow{\cong} (Y_i, \mathcal{C}^{\mathrm{hol}}_{\mathbb{C}})$ for some open $Y_i \subseteq \mathbb{C}^{n_i}$. Hence a complex manifold structure on X is the following tuple of data:
$$\left(X, \mathcal{O}_X, \{U_i\}_{i\in I}, \{Y_i \subseteq \mathbb{C}^{n_i}\}_{i\in I}, \{\varphi_i : (U_i, \mathcal{O}_{X|U_i}) \xrightarrow{\cong} (Y_i, \mathcal{C}^{\mathrm{hol}}_{\mathbb{C}})\}_{i\in I}\right)$$

In both of these, the isomorphisms $\{\varphi_i\}$ are called *charts* of the manifold and the sheaf $\mathcal{O}_X$ the structure sheaf of the manifold. Also, we can rather consider $\{\varphi_i\}_{i\in I}$ to be open embeddings. A map of manifolds is just defined to be a map of locally ringed spaces. Let $\mathbf{Mfd}^{\mathbb{R}}_\alpha$ and $\mathbf{Mfd}^{\mathbb{C}}$ denote the category of real C^α and complex manifolds respectively. A map of manifolds are just locally ringed maps between them. Isomorphisms in them are called C^α-diffeomorphism and biholomorphic maps respectively.

Let us now dwell into some of the immediate observations and remarks coming out of this definition. Let us first ease some notations. Let $(X, \mathcal{O}_X)$ be a real or complex manifold. The local chart (U_i, φ_i) is usually denoted by (U_i, x) where $x : U_i \to \mathbb{R}^n$ is a local embedding of locally ($\mathbb{R}$ or $\mathbb{C}$)-ringed spaces, where n depends on U_i. We usually suppress all the sheaves and their morphisms unless necessary (we will soon see why that's the case). For a local chart (U_i, x), the n component maps $\pi_j \circ x : U_i \to \mathbb{R}$ are denoted by x^j. Moreover, since $x : U \to x(U)$ is an isomorphism, therefore we denote $x^{-1} : x(U) \to U$ to be its inverse. All this will come in handy when we will start doing geometry over $(X, \mathcal{O}_X)$.

Let $(X, \mathcal{O}_X)$ be a real or complex manifold. We call an open subspace $(U, \mathcal{O}_{X|U}) \hookrightarrow (X, \mathcal{O}_X)$ an open submanifold.

One now sees that any morphism of manifolds as locally ringed spaces is completely determined by what happens at the level of points. In-fact, the sheaf allowed on X is also restricted if its a manifold. This is why we usually completely suppress the map of sheaves from our notation as that will be vacuous as long as we are working with map of manifolds. Let $(M, \mathcal{O}_M), (N, \mathcal{O}_N)$ be two manifolds ($\mathbb{R}$ or $\mathbb{C}$, but both of same type). We can define a sheaf $\mathcal{O}_{M;N}$ on M given by following sections: for some open $U \subseteq M$, we have a sheaf

$$\mathcal{O}_{M;N}(U) := \{f : (U, \mathcal{O}_{X|U}) \to (N, \mathcal{O}_N) \mid f \text{ is a map of manifolds}\}.$$

Now we show a foundational result which says that the notion of morphism of locally ringed spaces are nothing new in the classical world of $\mathbb{R}^n$ or $\mathbb{C}^n$. We place high importance on the following result as it becomes our point of departure (and thus a point of motivation) as to why the notion of a morphism of locally ringed spaces is defined as what it is; because it is the right notion of a "geometric map" in more abstract situations.

Theorem 3.1.1.4. *Let K be either $\mathbb{R}$ or $\mathbb{C}$, $X \subseteq K^n$ and $Y \subseteq K^m$ be two open subsets of the standard spaces. If $f : (X, \mathcal{C}^\alpha_X) \to (Y, \mathcal{C}^\alpha_Y)$ is a map of locally ringed spaces, then*

1. *$f^\flat : \mathcal{C}^\alpha_Y \to f_*\mathcal{C}^\alpha_X$ is given on an open set $V \subseteq Y$ by the standard composition map*

$$\begin{aligned} f^\flat_V : \mathcal{C}^\alpha_Y(V) &\longrightarrow \mathcal{C}^\alpha_X(f^{-1}(V)) \\ V \xrightarrow{t} K &\longmapsto f^{-1}(V) \xrightarrow{f} V \xrightarrow{t} K, \end{aligned}$$

2. *f is a C^α-map.*

Remark 3.1.1.5. As a slogan, we may remember the above theorem as the following principle:

In $\mathbb{R}^n$ or $\mathbb{C}^n$, locally ringed maps are exactly real C^α or holomorphic maps.

As a consequence of this, whenever we would like to consider C^α maps from, say $\mathbb{R}^n$ to $\mathbb{R}^m$, we might as well ask to produce a map of locally ringed spaces $(\mathbb{R}^n, \mathcal{C}^\alpha_{\mathbb{R}^n})$ to $(\mathbb{R}^m, \mathcal{C}^\alpha_{\mathbb{R}^m})$, which again shows how much geometric information is hidden in the notion of sheaves.

Proof of Theorem 3.1.1.4. [2] Pick any open $V \subseteq Y$ and any $t \in \mathcal{C}^\alpha_Y(V)$. We wish to show that $f^\flat_V(t) = t \circ f$ as a map $f^{-1}(V) \to K$. Consequently, pick any point $p \in f^{-1}(V)$. We wish to show that

[2] First proof in my new creator of meaning!

$f_V^\flat(t)(p) = t(f(p))$. To this end, we consider the evaluation homomorphism which are available at stalks. Observe that we have the following commutative square of K-algebras:

$$\begin{tikzcd} \mathcal{C}^\alpha_Y(V) \arrow[r, "f_V^\flat"] \arrow[d] & \mathcal{C}^\alpha_X(f^{-1}(V)) \arrow[d] \\ \mathcal{C}^\alpha_{Y,f(p)} \arrow[r, "f_p^\sharp"'] & \mathcal{C}^\alpha_{X,p} \end{tikzcd}.$$

In order to show $f_V^\flat(t)(p) = t(f(p))$, it is sufficient to show that the following triangle commutes:

$$\begin{tikzcd} \mathcal{C}^\alpha_{Y,f(p)} \arrow[rr, "f_p^\sharp"] \arrow[dr, "\mathrm{ev}_{f(p)}"'] & & \mathcal{C}^\alpha_{X,p} \arrow[dl, "\mathrm{ev}_p"] \\ & K & \end{tikzcd}.$$

But this is immediate from the fact that the K-algebra homomorphism $f_p^\sharp$ is a local ring homomorphism and the kernels of the evaluation maps are exactly the corresponding unique maximal ideals, so by quotienting by the maximal ideals, we obtain a K-algebra homomorphism $K \to K$ which necessarily is identity as it is a K-algebra homomorphism. Hence the triangle indeed commutes.

In order to show that the map f is a C^α-map, we need only show that the m projection maps $\pi_i : K^m \to K$ when composed with f yields C^α maps given by $X \to K$, but that is immediate from 1. $\square$

Using the above result, one can show that any manifold essentially has a unique structure sheaf of the form $\mathcal{O}_{X;\mathbb{R}}$ or $\mathcal{O}_{X;\mathbb{C}}$.

Proposition 3.1.1.6. *Let $(X, \mathcal{O}_X)$ be a locally ringed space. If $(X, \mathcal{O}_X)$ is a real or complex manifold, then $\mathcal{O}_X \cong \mathcal{O}_{X;\mathbb{R}}$ or $\mathcal{O}_X \cong \mathcal{O}_{X;\mathbb{C}}$.*

Proof. We wish to show that there is an isomorphism of sheaves $\varphi : \mathcal{O}_{X;\mathbb{R}} \to \mathcal{O}_X$. For an open set $U \subseteq X$, we define φ_U as follows:

$$\begin{aligned} \varphi_U : \mathcal{O}_{X;\mathbb{R}}(U) &\longrightarrow \mathcal{O}_X(U) \\ t : (U, \mathcal{O}_{X|U}) \to (\mathbb{R}, \mathcal{C}^\alpha_\mathbb{R}) &\longmapsto t^\flat_\mathbb{R}(\mathrm{id}_\mathbb{R}). \end{aligned}$$

We claim that this map of sheaves is an isomorphism. We need only show that the map on stalks $\varphi_x : \mathcal{O}_{X;\mathbb{R},x} \to \mathcal{O}_{X,x}$ is an isomorphism. So we may assume that X has a global chart $\eta : (X, \mathcal{O}_X) \cong (W, \mathcal{C}^\alpha_{W;\mathbb{R}})$ where $W \subseteq \mathbb{R}^n$ is an open subset. Consequently, we have $\eta_x^\sharp : \mathcal{C}^\alpha_{W;\mathbb{R},\eta(x)} \cong \mathcal{O}_{X,x}$. Furthermore, $\mathcal{O}_{X;\mathbb{R},x} \cong \mathcal{C}^\alpha_{W;\mathbb{R},\eta(x)}$. Consequently, we wish to show that $\varphi_x : \mathcal{C}^\alpha_{W;\mathbb{R},\eta(x)} \to \mathcal{C}^\alpha_{W;\mathbb{R},\eta(x)}$ given by $(W, t : W \to \mathbb{R})_{\eta(x)} \mapsto (W, t^\flat_\mathbb{R}(\mathrm{id}_\mathbb{R}))_{\eta(x)}$ is an isomorphism. Since by Theorem 3.1.1.4, 1, the map $t^\flat_\mathbb{R}$ is given by precomposition by t, therefore $t^\flat_\mathbb{R}(\mathrm{id}_\mathbb{R})$ is just t. Consequently, φ_x is identity, which proves the result. $\square$

Remark 3.1.1.7. By virtue of Proposition 3.1.1.6, we can assume that any C^α-manifold is a locally ringed space of the form $(X, \mathcal{O}_{X;\mathbb{R}})$ (similarly for $\mathbb{C}$-manifolds).

3.1.2 Sheaves & atlases

We have defined a manifold to be a space with an open covering by a model locally ringed spaces. There is a traditional definition, whereas, which is used heavily in traditional geometry because we really care about the charts (which is usually not done in algebraic geometry). This elucidates how one has to undertake a different viewpoint of geometry in algebraic geometry.

We wish to show that giving a manifold structure on a second countable Hausdorff space X as defined above is equivalent to giving an atlas in the classical sense. Indeed, for each atlas on X, we first define a sheaf on X.

Definition 3.1.2.1 (**Atlas sheaf).** Let X be a second countable Hausdorff space and $\mathcal{A} = (U_i, x_i)_{i\in I}$ be a C^α-atlas on X where $x_i : U_i \to \mathbb{C}^{n_i}$ is an open embedding. Consider the following assignment for each open $V \subseteq X$:

$$\mathcal{O}_{\mathcal{A}}(V) := \{f : V \to K \mid f \circ x_i^{-1} : x_i(U_i \cap V) \to K \text{ is } C^\alpha\text{-map}\}.$$

Then $\mathcal{O}_{\mathcal{A}}$ is a sheaf of $\mathbb{R}$-algebras, called the sheaf of atlas $\mathcal{A}$. Similarly for the holomorphic case.

We first observe that equivalent atlases give same atlas sheaves.

Lemma 3.1.2.2. *Let X be a second-countable Hausdorff space with $\mathcal{A} = (U_i, x_i)_i$ and $\mathcal{B} = (V_i, y_i)_i$ being two equivalent C^α or holomorphic atlases on X. Then the atlas sheaves $\mathcal{O}_{\mathcal{A}}$ and $\mathcal{O}_{\mathcal{B}}$ are isomorphic.*

Proof. Indeed, for each open $W \subseteq X$, define the map

$$\begin{aligned} \varphi_W : \mathcal{O}_{\mathcal{A}}(W) &\longrightarrow \mathcal{O}_{\mathcal{B}}(W) \\ f : W \to K &\longmapsto f : W \to K. \end{aligned}$$

To show that this is well-defined, we have to show that $f \in \mathcal{O}_{\mathcal{B}}(W)$. Indeed, pick any chart $y_i : V_i \to K$ of $\mathcal{B}$. We wish to show that $f \circ y_i^{-1} : y_i(V_i \cap W) \to K$ is C^α or holomorphic. As either condition is local on domain, so pick any point in $y_i(V_i \cap W)$. Pick a chart $x_i : U_i \to x_i(U_i)$ containing that point. Note that it is sufficent to show $f \circ y_i^{-1} : y_i(V_i \cap U_i \cap W) \to K$ is C^α or holomorphic. Indeed, we can write this as

$$f \circ y_i^{-1} = (f \circ x_i^{-1}) \circ (x_i \circ y_i^{-1}) : y_i(U_i \cap V_i \cap W) \to K.$$

Since $\mathcal{A}$ and $\mathcal{B}$ are equivalent and $f \in \mathcal{O}_{\mathcal{A}}$, it follows repsectively that $(x_i \circ y_i^{-1})$ and $(f \circ x_i^{-1})$ are C^α or holomorphic, as required.

Thus $\varphi : \mathcal{O}_{\mathcal{A}} \to \mathcal{O}_{\mathcal{B}}$ is a sheaf map, which is identity, hence both sheaves are same. □

We next see that a C^α or holomorphic atlas sheaf on a space X gives a C^α or $\mathbb{C}$ manifold structure on X.

Proposition 3.1.2.3. *Let $(X, \mathcal{O}_{X;\mathbb{C}})$ be a locally ringed space and $Y \subseteq \mathbb{C}^n$ be open. If $\varphi : (X, \mathcal{O}_{X;\mathbb{C}}) \to (Y, \mathcal{C}^{\mathrm{hol}}_{Y;\mathbb{C}})$ is a map of locally ringed spaces, then $\varphi^\flat$ on open $V \subseteq Y$ is given by*

$$\begin{aligned} \varphi^\flat_V : \mathcal{C}^{\mathrm{hol}}_{Y;\mathbb{C}}(V) &\longrightarrow \mathcal{O}_{X;\mathbb{C}}(\varphi^{-1}(V)) \\ t : V \to \mathbb{C} &\longmapsto t \circ \varphi : \varphi^{-1}(V) \to \mathbb{C}. \end{aligned}$$

Moreover, the following are equivalent:

1. *$\varphi : (X, \mathcal{O}_{X;\mathbb{C}}) \to (Y, \mathcal{C}^{\mathrm{hol}}_{Y;\mathbb{C}})$ is an isomorphism of locally ringed spaces.*
2. *$\varphi : X \to Y$ is a homeomorphism such that for any open $U \subseteq X$ and any $f : U \to \mathbb{C}$ in $\mathcal{O}_X(U)$, $f \circ \varphi^{-1} : \varphi(U) \to \mathbb{C}$ is a holomorphic map.*

The same conclusions hold true for C^α-manifolds as well.

Proof. The proof of the first statement is exactly same as that of Theorem 3.1.1.4, hence is omitted. We now show the equivalence of items 1 and 2.

(1. $\Rightarrow$ 2.) This is immediate as the map $\varphi^\flat$ is an isomorphism, so in particular a bijection on sections.

(2. $\Rightarrow$ 1.) Pick any open $V \subseteq Y$. Then $\varphi^\flat_V$ is injective as φ is an isomorphism. It is also surjective by the given hypothesis and homeomorphism φ. This shows that $\varphi^\flat$ is an isomorphism. □

Theorem 3.1.2.4. *Let X be a second-countable Hausdorff space and $(X, \mathcal{O}_X)$ be a locally ringed space. Then the following are equivalent.*

1. *$(X, \mathcal{O}_X)$ is a C^α/complex manifold.*
2. *$\mathcal{O}_X$ is a C^α/complex atlas sheaf.*

To avoid repetitions, we will do the complex case only, as there is no change in the proof for the real case.

Proof. (1. $\Rightarrow$ 2.) By Proposition 3.1.1.6, we may assume that $\mathcal{O}_X$ is just $\mathcal{O}_{X;\mathbb{C}}$, the sheaf of locally ringed maps from X to $\mathbb{C}$. We have an open cover $\{U_i\}_{i\in I}$ of X and isomorphisms of locally ringed spaces $\varphi_i : (U_i, \mathcal{O}_{U_i;\mathbb{C}}) \to (Y_i, \mathcal{C}^{\mathrm{hol}}_{\mathbb{C}})$. This makes (U_i, φ_i) into an usual atlas as follows. For any i, j such that $U_i \cap U_j \neq \emptyset$, we obtain that the map

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j).$$

This is holomorphic since $\varphi_j : U_i \cap U_j \to \mathbb{C}$ is a map of locally ringed spaces in $\mathcal{O}_{X;\mathbb{C}}(U_i \cap U_j)$. Now, $\varphi_i : (U_i, \mathcal{O}_{U_i;\mathbb{C}}) \to (Y_i, \mathcal{C}^{\mathrm{hol}}_{Y_i;\mathbb{C}})$ is an isomorphism, therefore by Proposition 3.1.2.3, it follows that $\varphi_j \circ \varphi_i^{-1}$ is a holomorphic map, as required.

We claim that this makes $\mathcal{O}_X$ into an atlas sheaf. Indeed, observe that $f \in \mathcal{O}_X(V)$ is a locally ringed map $f : (V, \mathcal{O}_{V;\mathbb{C}}) \to (Y, \mathcal{C}^{\mathrm{hol}}_{\mathbb{C}})$. We claim that the data of f is equivalent to saying that $f \circ \varphi_i^{-1} : \varphi_i(V \cap U_i) \to \mathbb{C}$ is holomorphic. Indeed, this is the content of Proposition 3.1.2.3.

(2. $\Rightarrow$ 1.) Let $\mathcal{A} = (U_i, \varphi_i)$ be a complex atlas where $\varphi_i : U_i \to Y_i$ for open $Y_i \subseteq \mathbb{C}^{n_i}$ is a homeomorphism with holomorphic transitions. We need only show the item 2 of Proposition 3.1.2.3 for φ_i as then it would follow that $\varphi_i : (U_i, \mathcal{O}_{U_i;\mathbb{C}}) \to (Y_i, \mathcal{C}^{\mathrm{hol}}_{Y_i;\mathbb{C}})$ is an isomorphism of locally ringed spaces, completing the proof. Indeed, pick any open $U \subseteq X$ and any $f : U \to \mathbb{C}$ in $\mathcal{O}_X(U)$. As $\mathcal{O}_X$ is the atlas sheaf of $\mathcal{A}$, therefore for φ_i in particular, we have that $f \circ \varphi_i^{-1} : \varphi_i(U \cap U_i) \to \mathbb{C}$ is a holomorphic map, as required. This completes the proof. □

3.2 Linearization

Write till fib
ucts, Chapt

3.3 Constructions on manifolds

3.4 Lie groups

3.5 Global algebra

Let $(X, \mathcal{O}_X)$ be a locally ringed space. We will discuss here the operations on and properties of $\mathbf{Mod}(\mathcal{O}_X)$, the category of $\mathcal{O}_X$-modules [3]. An $\mathcal{O}_X$-module is a sheaf $\mathcal{M}$ on X such that $\mathcal{M}(U)$ is an $\mathcal{O}_X(U)$-module and the restriction maps of $\mathcal{M}$ are given as module homomorphism w.r.t the corresponding restriction map of $\mathcal{O}_X$ (more precisely below). There are several important constructions and properties that one can make with these. In-fact, just like one understands a ring R by understanding R-modules, one can understand $\mathcal{O}_X$ by understanding $\mathcal{O}_X$-modules. The similarity runs deeper as we can also define in certain cases the very same constructions we do in module, but in the case of $\mathcal{O}_X$-modules, and these constructions and operations becomes indispensable in doing geometry over locally ringed spaces of special kind, like schemes. A lot of such phenomenon is merely due to the fact that $\mathbf{Mod}(\mathcal{O}_X)$ is an abelian category. In-fact, notice that for each singleton space $X = \{\text{pt.}\}$, a ring R can be seen as the structure sheaf $\mathcal{O}_X$ over X and any R-module as a $\mathcal{O}_X$-module. Hence one may also think of the concept of $\mathcal{O}_X$-modules as the global version of classical commutative algebra.

Needless to say, this is an indispensable section for the purposes of geometry in general.

Let us first observe that over any topological space X, the product of two sheaves $\mathcal{F}, \mathcal{G}$ over X defined by $(\mathcal{F} \times \mathcal{G})(U) = \mathcal{F}(U) \times \mathcal{G}(U)$ is indeed a sheaf with restriction maps as products of the restrictions. This allows us to define $\mathcal{O}_X$-modules very naturally.

For the rest of this section, we fix a ringed space $(X, \mathcal{O}_X)$.

Definition 3.5.0.1. ($\mathcal{O}_X$**-modules**) An abelian sheaf $\mathcal{F}$ over X is an $\mathcal{O}_X$-module if there is a map of sheaves

$$\begin{aligned} \mathcal{O}_X \times \mathcal{F} &\longrightarrow \mathcal{F} \\ (c, s) &\longmapsto cs \end{aligned}$$

where $c \in \mathcal{O}_X(U), s \in \mathcal{F}(U)$ for all open $U \subseteq X$ which endows $\mathcal{F}(U)$ an $\mathcal{O}_X(U)$-module structure.

An $\mathcal{O}_X$-linear map of $\mathcal{O}_X$-modules is defined as a sheaf map $\varphi : \mathcal{F} \to \mathcal{G}$ between $\mathcal{O}_X$-modules such that for each open $U \subseteq X$, the map $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is an $\mathcal{O}_X(U)$-linear map and that the restrictions preserves the respective module structures.

The above definition, when unravelled, yields that the scalar multiplication of each $\mathcal{O}_X(U)$-module $\mathcal{F}(U)$ commutes with restrictions; for $c \in \mathcal{O}_X(U)$, $s \in \mathcal{F}(U)$ and an open subset $V \subseteq U$, we have $(c \cdot s)|_V = c|_V \cdot s|_V$.

Remark 3.5.0.2. For an $\mathcal{O}_X$-module $\mathcal{F}$ we have the following easy observations:

1. $\mathcal{F}_x$ is an $\mathcal{O}_{X,x}$-module for all $x \in X$. Indeed, this follows from the universal property of direct limits and the fact that direct limits commutes with product; we have the following

[3] we will give some general constructions for arbitrary sheaves over a topological case at times, before specializing to $\mathcal{O}_X$-module case.

diagram

$$\begin{tikzcd}
\mathcal{O}_{X,x} \times \mathcal{F}_x \arrow[ddrr, dashed, "\tilde{m}"] & & \\
\varinjlim_{x \in V} \mathcal{O}_X(V) \times \mathcal{F}(V) \arrow[u, "\cong"] & & \\
\mathcal{O}_X(U) \times \mathcal{F}(U) \arrow[u] \arrow[r, "m_U"] & \mathcal{F}(U) \arrow[r] & \mathcal{F}_x
\end{tikzcd}.$$

Explicitly, the $\mathcal{O}_{X,x}$-module structure on $\mathcal{F}_x$ is given by

$$\begin{aligned} \mathcal{O}_{X,x} \times \mathcal{F}_x &\longrightarrow \mathcal{F}_x \\ ((U,c)_x, (U,s)_x) &\longmapsto (U, c \cdot s)_x \end{aligned}$$

where we may assume c and s are defined on same open neighborhood of x by appropriately restricting.

2. For a homomorphism $f : \mathcal{F} \to \mathcal{G}$ of $\mathcal{O}_X$-modules, we get a $\mathcal{O}_{X,x}$-module homomorphism $f_x : \mathcal{F}_x \to \mathcal{G}_x$ mapping as $(U,s)_x \mapsto (U, f_U(s))_x$ for each $x \in X$,
3. Let X be locally ringed space. Then, $\mathcal{F}_x/\mathfrak{m}_{X,x}\mathcal{F}_x \cong \mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{O}_{X,x}/\mathfrak{m}_{X,x} \cong \mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \kappa(x)$ is a $\kappa(x)$-vector space. This is called the *fiber of module $\mathcal{F}$ over x*, denoted by $\mathcal{F}(x)$. Recall this is how the fiber of a module over a prime ideal of the ring is defined.

We first give few basic constructions, which is useful to keep in mind.

Definition 3.5.0.3. (**Support of a sheaf**) Let X be a topological space and $\mathcal{F}$ be an abelian sheaf over X. Let $U \subseteq X$ be an open set. For $s \in \mathcal{F}(U)$, we define the support of s as the subset

$$\text{Supp}\,(s) := \{x \in U \mid (U,s)_x \neq 0 \text{ in } \mathcal{F}_x\}.$$

We further define the support of the sheaf as

$$\text{Supp}\,(\mathcal{F}) := \{x \in X \mid \mathcal{F}_x \neq 0\}.$$

Support of a section is always a closed subset, but the support of a sheaf may not be closed.

Lemma 3.5.0.4. [4] *Let X be a space and $\mathcal{F}$ be a sheaf over X with $s \in \mathcal{F}(U)$ for an open set $U \subseteq X$. Then* $\text{Supp}\,(s) \subseteq U$ *is a closed subset of U.*

Proof. Take any point $y \in U \setminus \text{Supp}\,(s)$. We will find an open set $W \subseteq U \setminus \text{Supp}\,(s)$ with $W \ni y$. Indeed, as $(U,s)_y = 0$, therefore we get a $W \subseteq U$ with $s|_W = 0$. For any $z \in W$, one further checks that $(U,s)_z = (W, s|_W)_z = 0$. Thus, $z \notin \text{Supp}\,(s)$ and consequently, $W \subseteq U \setminus \text{Supp}\,(s)$. □

Do skyscraper and subsheaf with support (Exercises 1.17 and 1.20 in Hartshorne.)

[4] Exercise II.1.14 of Hartshorne.

3.5.1 Global algebra : The algebra of $\mathcal{O}_X$-modules

In our quest to do geometry over schemes, we will make heavy use of the algebra of sheaves, especially that of exact sequences, so we give a lot of constructions that we may have to make out in the wild. We will make heavy use of sheafification (Theorem 20.2.0.1) in the sequel. An important question that arises is whether sheafification of an algebraic construction over collection of $\mathcal{O}_X$-modules actually is again an $\mathcal{O}_X$-module or not? The answer is yes, as can be easily checked by explicitly looking at sections of sheafification directly (see Remark 20.2.0.4 to observe that its not difficult, anyways we will show the explicit checks consistently).

Caution 3.5.1.1. The following pages might seem to be filled with *unnecessary details* about checking whether a given construction on $\mathcal{O}_X$-modules results in an $\mathcal{O}_X$-module or not. While for some this might be unnecessary, but working this out in experience has been satisfying and tends to give a deeper understanding of the various module structures that gets associated with an $\mathcal{O}_X$-module $\mathcal{F}$ and how they interrelate. Indeed, we will see that with more elaborate constructions, we get more and more module structures to handle with. Thus it is necessary to work some details out of this. At any rate, we will be using notions presented in the sequel quite frequently in algebraic geometry and in particular while doing cohomology (Cěch cohomology in particular!) so we need a good knowledge of the $\mathcal{O}_X$-modules and their internal technicalities.

Remark 3.5.1.2. Since there are a lot of constructions in the sequel, so to have a sense of mental clarity, let us list them here:

- Submodules and ideals of $\mathcal{O}_X$.✓
- Quotient of modules.✓
- Image and kernel modules.✓
- Exact sequences of modules.✓
- The $\Gamma(\mathcal{O}_X, X)$-module $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$.✓
- $\mathcal{H}om_{\mathcal{O}_X}$ module.✓
- Direct sum of modules.✓
- Direct product of modules.✓
- Tensor product of modules.✓
- Free, locally free & finite locally free $\mathcal{O}_X$-modules.✓
- Invertible modules and the Picard group.✓
- Direct and inverse image modules.✓
- Sums & intersections of submodules.
- Modules generated by sections.
- Inverse limit.
- Direct limit.
- Tensor, symmetric & exterior algebras.
- $\mathcal{E}xt$ module.
- $\mathcal{T}or$ module.

Remark 3.5.1.3. Let $\mathbf{V}$ be the category of abelian groups and X be a locally ringed space. Consider a functor $F : \mathbf{V} \times \cdots \times \mathbf{V} \to \mathbf{V}$. Given abelian sheaves $\mathcal{F}_1, \dots, \mathcal{F}_k$ over X, we obtain a sheaf $\mathcal{F}_F := F(\mathcal{F}_1, \dots, \mathcal{F}_k)$ by the following procedure: first define the presheaf $\mathcal{F}_F^-$ on X given by $U \mapsto F(\mathcal{F}_1(U), \dots, \mathcal{F}_k(U))$, then define the sheaf $\mathcal{F}_F = (\mathcal{F}_F^-)^{++}$ to be the sheafification of $\mathcal{F}_F^-$. We will follow this general strategy in all the constructions in the following.

Submodules and ideals of $\mathcal{O}_X$

Definition 3.5.1.4. (**Submodules and ideals**) Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. A *submodule* of $\mathcal{F}$ is an $\mathcal{O}_X$-module which is a subsheaf $\mathcal{G} \subseteq \mathcal{F}$ such that for all open $U \subseteq X$, the inclusion

$$\mathcal{G}(U) \hookrightarrow \mathcal{F}(U)$$

is an $\mathcal{O}_X(U)$-module homomorphism. Since $\mathcal{O}_X$ is an $\mathcal{O}_X$-module, thus, to be in line with usual terminology, we define submodules of $\mathcal{O}_X$ as *ideals* of $\mathcal{O}_X$.

Remark 3.5.1.5. Note that for any $\mathcal{O}_X$ submodule $\mathcal{G} \subseteq \mathcal{F}$, we get a submodule $\mathcal{G}_x \subseteq \mathcal{F}_x$ of the $\mathcal{O}_{X,x}$-module $\mathcal{F}_x$.

Quotient of modules

Definition 3.5.1.6. (**Quotient modules**) Let $\mathcal{F}$ be an $\mathcal{O}_X$-module and $\mathcal{G}$ be a submodule of $\mathcal{F}$. The *quotient module* is the sheafification of the presheaf $U \mapsto \mathcal{F}(U)/\mathcal{G}(U)$, denoted by $\mathcal{F}/\mathcal{G}$ (see Definition 20.3.0.4). Indeed, $\mathcal{F}/\mathcal{G}$ is an $\mathcal{O}_X$-module by the following lemma.

Lemma 3.5.1.7. *$\mathcal{F}/\mathcal{G}$ is an $\mathcal{O}_X$-module.*

Proof. We will use the definition of sheafification as given in Remark 20.2.0.4. For each open set $U \subseteq X$, consider the following map:

$$\begin{aligned} \eta_U : \mathcal{O}_X(U) \times (\mathcal{F}/\mathcal{G})(U) &\longrightarrow (\mathcal{F}/\mathcal{G})(U) \\ (c, s) &\longmapsto \eta_U(c, s) : U \to \amalg_{x \in U} \mathcal{F}_x/\mathcal{G}_x \end{aligned}$$

where $\eta_U(c, s)(x) := c_x \cdot s(x)$ where $c_x \in \mathcal{O}_{X,x}$ and $s(x) \in \mathcal{F}_x/\mathcal{G}_x$ and the multiplication $c_x \cdot s(x)$ is coming from the $\mathcal{O}_{X,x}$-module structure that $\mathcal{F}_x/\mathcal{G}_x$ has. We now need to show following two statements:

1. $\eta_U(c, s)$ is indeed in $(\mathcal{F}/\mathcal{G})(U)$,
2. $\eta : \mathcal{O}_X \times \mathcal{F}/\mathcal{G} \to \mathcal{F}/\mathcal{G}$ is a sheaf map.

For statement 1, we need to show that for each $x \in U$, there exists an open set $x \in V \subseteq U$ and there exists $r \in \mathcal{F}(U)/\mathcal{G}(U)$ such that for all $y \in V$ we have the equality $c_y \cdot s(y) = r_y$ in $\mathcal{F}_y/\mathcal{G}_y$. Indeed, this can easily be seen via the fact that $s \in (\mathcal{F}/\mathcal{G})(U)$. Statement 2 is immediate after drawing the relevant square whose commutativity is under investigation. □

Remark 3.5.1.8. Note further that we get a natural map

$$\mathcal{F} \to \mathcal{F}/\mathcal{G}$$

which factors through the inclusion of the presheaf $U \mapsto \mathcal{F}(U)/\mathcal{G}(U)$ into the sheaf $\mathcal{F}/\mathcal{G}$.

Image and kernel modules

Definition 3.5.1.9. (**Image and kernel modules**) Let $f : \mathcal{F} \to \mathcal{G}$ be a $\mathcal{O}_X$-module homomorphism. We then get the image sheaf $\mathrm{Im}\,(f)$ and the kernel sheaf $\mathrm{Ker}\,(f)$ by Definition 20.3.0.5. Indeed, both of these are $\mathcal{O}_X$-modules as the following lemma shows.

Lemma 3.5.1.10. $\mathrm{Im}\,(f)$ *and* $\mathrm{Ker}\,(f)$ *are* $\mathcal{O}_X$*-modules.*

Proof. $\mathrm{Ker}\,(f)$ is straightforward. For $\mathrm{Im}\,(f)$, we first observe that if we denote $\mathrm{Im}\,(f) = (\mathrm{im}\,(f))^{++}$, then $(\mathrm{im}\,(f))_x = f_x(\mathcal{F}_x)$. We thus define the $\mathcal{O}_X$-module structure on $\mathrm{Im}\,(f)$ as follows:

$$\begin{aligned}\eta_U : \mathcal{O}_X(U) \times \mathrm{Im}\,(f)(U) &\longrightarrow \mathrm{Im}\,(f)(U)\\ (c, s : U \to \amalg_{x\in U} f_x(\mathcal{F}_x)) &\longmapsto \eta_U(c,s)\end{aligned}$$

where $\eta_U(c,s)(x) = c_x \cdot s(x)$ where $s(x) \in f_x(\mathcal{F}_x) \subseteq \mathcal{G}_x$. One checks like for quotient modules that this defines an $\mathcal{O}_X$-module structure on $\mathrm{Im}\,(f)$. Further, it is clear that $\mathrm{Im}\,(f) \subseteq \mathcal{G}$. □

Corollary 3.5.1.11. *For a $\mathcal{O}_X$-module homomorphism $f : \mathcal{F} \to \mathcal{G}$, we have $\mathrm{Ker}\,(f) \leq \mathcal{F}$ and $\mathrm{Im}\,(f) \leq \mathcal{G}$ are submodules.*

Proof. Use Remark 20.2.0.4 to get this immediately. □

We have a "first isomorphism theorem" for modules then.

Lemma 3.5.1.12. *For a map $f : \mathcal{F} \to \mathcal{G}$ of $\mathcal{O}_X$-modules, we obtain an isomorphism*

$$\mathcal{F}/\mathrm{Ker}\,(f) \cong \mathrm{Im}\,(f).$$

Proof. For each $x \in X$ let $\varphi_x : \mathcal{F}_x/\ker f_x \xrightarrow{\cong} \mathrm{im}\,(f_x)$. Then we define the following for any $U \subseteq X$ open

$$\begin{aligned}(\mathcal{F}/\mathrm{Ker}\,(f))(U) &\longrightarrow \mathrm{Im}\,(f)(U)\\ s : U \to \amalg_{x\in U}\mathcal{F}_x/\ker f_x &\mapsto \varphi \circ s\end{aligned}$$

where $(\varphi \circ s)(x) = \varphi_x(s(x))$. This is immediately an isomorphism by going to stalks (Theorem 20.3.0.6, 3). □

Exact sequences of modules

Definition 3.5.1.13. (**Exact sequences**) A sequence of $\mathcal{O}_X$-modules

$$\mathcal{F}' \xrightarrow{f} \mathcal{F} \xrightarrow{g} \mathcal{F}''$$

is said to be *exact* if $\mathrm{Ker}\,(g) = \mathrm{Im}\,(f)$.

Remark 3.5.1.14. By Lemma 20.3.0.8, $\mathcal{F}' \xrightarrow{f} \mathcal{F} \xrightarrow{g} \mathcal{F}''$ is exact if and only if $\mathrm{Ker}\,(g_x) = \mathrm{Im}\,(f_x)$ at all points $x \in X$.

The $\Gamma(\mathcal{O}_X, X)$-module $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$

We now consider the set of all $\mathcal{O}_X$-module homomorphisms $f : \mathcal{F} \to \mathcal{G}$ and observe very easily that it has a $\Gamma(\mathcal{O}_X, X)$-module structure. This generalizes the fact that under point-wise addition and scalar multiplication, the set $\mathrm{Hom}_R(M, N)$ for two R-modules M, N is again an R-module.

Definition 3.5.1.15. (**$\Gamma(\mathcal{O}_X, X)$-module** $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$) Let $\mathcal{F}, \mathcal{G}$ be two $\mathcal{O}_X$-modules. Then the collection of all $\mathcal{O}_X$-module homomorphisms $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ is a $\Gamma(X, \mathcal{O}_X)$-module. Indeed, for two $f, g \in \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ and $c \in \Gamma(\mathcal{O}_X, X)$, we define $f + g : \mathcal{F} \to \mathcal{G}$ by $s \mapsto f(s) + g(s)$ and we define $c \cdot f : \mathcal{F} \to \mathcal{G}$ by $s \mapsto \rho_{X,U}(s) \cdot f(s)$ for any open set $U \subseteq X$ and $s \in \mathcal{F}(U)$.

We will now globalize the construction of $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ to obtain an $\mathcal{O}_X$-module out of it.

$\mathcal{H}om_{\mathcal{O}_X}$ module

Definition 3.5.1.16. (**Hom module** $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$) Let $\mathcal{F}, \mathcal{G}$ be two $\mathcal{O}_X$-modules. Then the following presheaf

$$U \mapsto \mathcal{H}om_{\mathcal{O}_{X|U}}(\mathcal{F}_{|U}, \mathcal{G}_{|U})$$

with restriction given by restriction of sheaf maps, is an $\mathcal{O}_X$-module denoted by $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$, as the following lemma shows.

Lemma 3.5.1.17. *$\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ is an $\mathcal{O}_X$-module*

Proof. The fact that $\mathcal{H}om(\mathcal{F}, \mathcal{G})$ is a sheaf can be seen immediately. The $\mathcal{O}_X$-module structure is defined as follows: pick any open $U \subseteq X$

$$\begin{aligned} \eta_U : \mathcal{O}_X(U) \times \mathrm{Hom}_{\mathcal{O}_U}\left(\mathcal{F}_{|U}, \mathcal{G}_{|U}\right) &\longrightarrow \mathrm{Hom}_{\mathcal{O}_U}\left(\mathcal{F}_{|U}, \mathcal{G}_{|U}\right) \\ (c, f) &\longmapsto cf \end{aligned}$$

where $cf : \mathcal{F}_{|U} \to \mathcal{G}_{|U}$ is given on an open set $V \subseteq U$ by

$$\begin{aligned} (cf)_V : \mathcal{F}(V) &\longmapsto \mathcal{G}(V) \\ s &\longmapsto \rho_{U,V}(c) \cdot f_V(s). \end{aligned}$$

One easily check that η is a well-defined natural map of sheaves, thus making $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ into an $\mathcal{O}_X$-module. □

Remark 3.5.1.18. It is in general NOT true that $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})_x \cong \mathrm{Hom}_{\mathcal{O}_{X,x}}(\mathcal{F}_x, \mathcal{G}_x)$.

We now define the dual of a module in the obvious manner.

Definition 3.5.1.19. (**Dual module**) Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. The dual of $\mathcal{F}$ is defined to be the module $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{O}_X)$. We denote the dual by $\mathcal{F}^\vee$.

There are some isomorphisms regarding $\mathcal{H}om$ that is akin to their usual algebraic counterparts. We outline them in the following lemma.

Lemma 3.5.1.20. *Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. Then,*
1. *$\mathcal{H}om(\mathcal{O}_X^n, \mathcal{F}) \cong \mathcal{H}om(\mathcal{O}_X, \mathcal{F})^n$,*
2. *$\mathcal{H}om(\mathcal{O}_X, \mathcal{F}) \cong \mathcal{F}$.*

Proof. In both cases we construct a map and its inverses and it is straightforward to see that they are well-defined, natural and indeed inverses of each other.
1. Consider the map

$$\mathcal{H}om(\mathcal{O}_X^n, \mathcal{F}) \longrightarrow \mathcal{H}om(\mathcal{O}_X, \mathcal{F})^n$$

which on an open set $U \subseteq X$ maps as

$$\begin{aligned} \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}^n, \mathcal{F}_{|U}\right) &\longrightarrow \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}, \mathcal{F}_{|U}\right)^n \\ f : \mathcal{O}_{X|U}^n \to \mathcal{F}_{|U} &\longmapsto (f_i)_{i=1,\dots,n} \end{aligned}$$

where for $V \subseteq U$, we have that $f_{i,V} : \mathcal{O}_X(V) \to \mathcal{F}(V)$ maps as $s \mapsto s \cdot f_V(e_i) = f_V(s \cdot e_i)$ where e_i is i^{th} standard vector in $\mathcal{O}_X(V)^n$. Conversely, define the map

$$\mathcal{H}om(\mathcal{O}_X, \mathcal{F})^n \longrightarrow \mathcal{H}om(\mathcal{O}_X^n, \mathcal{F})$$

which on $U \subseteq X$ open maps as

$$\left(g_i : \mathcal{O}_{X|U} \to \mathcal{F}_{|U}\right)_{i=1,\dots,n} \longmapsto g : \mathcal{O}_{X|U}^n \to \mathcal{F}_{|U}$$

where on $V \subseteq U$ open, we define $g_V : \mathcal{O}_X(V)^n \to \mathcal{F}(V)$ as $(s_1, \dots, s_n) \mapsto \sum_{i=1}^n g_{i,V}(s_i) = \sum_{i=1}^n s_i \cdot g_{i,V}(e_i)$.
2. Define the map

$$\mathcal{H}om(\mathcal{O}_X, \mathcal{F}) \longrightarrow \mathcal{F}$$

on open $U \subseteq X$ by

$$\begin{aligned} \operatorname{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}, \mathcal{F}_{|U}\right) &\longrightarrow \mathcal{F}(U) \\ f : \mathcal{O}_{X|U} \to \mathcal{F}_{|U} &\longmapsto f_U(1). \end{aligned}$$

Define the inverse

$$\mathcal{F} \longrightarrow \mathcal{H}om(\mathcal{O}_X, \mathcal{F})$$

on an open set $U \subseteq X$ by

$$\begin{aligned} \mathcal{F}(U) &\longrightarrow \operatorname{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}, \mathcal{F}_{|U}\right) \\ s &\longmapsto f : \mathcal{O}_{X|U} \to \mathcal{F}_{|U} \end{aligned}$$

where for an open set $V \subseteq U$, we define $f_V(t) = f_V(t \cdot 1) := t \cdot s$. □

The following the usual adjunction from algebra.

Proposition 3.5.1.21 (⊗-hom adjunction)**.** *For any $\mathcal{O}_X$-modules $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$, we have*

$$\operatorname{Hom}_{\mathcal{O}_X}(\mathcal{E} \otimes \mathcal{F}, \mathcal{G}) \cong \operatorname{Hom}_{\mathcal{O}_X}\left(\mathcal{F}, \mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{G})\right).$$

Proof. Let $R = \Gamma(\mathcal{O}_X, X)$. We construct an R-linear map $\varphi : \operatorname{Hom}_{\mathcal{O}_X}(\mathcal{E} \otimes \mathcal{F}, \mathcal{G}) \to \operatorname{Hom}_{\mathcal{O}_X}\left(\mathcal{F}, \mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{G})\right)$ as follows: for any $f \in \operatorname{Hom}_{\mathcal{O}_X}(\mathcal{E} \otimes \mathcal{F}, \mathcal{G})$, define $\varphi(f) : \mathcal{F} \to \mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{G})$ by the following on open $U \subseteq X$:

$$\varphi(f) : s \mapsto \varphi(f)(s) : t \mapsto f(s \otimes t).$$

This is R-linear by construction. To show its a bijection, we construct an inverse as follows: for any $f : \mathcal{F} \to \mathcal{H}om(\mathcal{E}, \mathcal{G})$, define $\theta(f) : \mathcal{E} \otimes \mathcal{F} \to \mathcal{G}$ as the unique map corresponding to the map on presheaves:

$$\begin{aligned} (\mathcal{E} \otimes \mathcal{F})^- &\longrightarrow \mathcal{G} \\ s \otimes t &\mapsto f(t)(s). \end{aligned}$$

It is easy to see that this is an inverse of φ, as required. □

Direct sum of modules

Definition 3.5.1.22. (**Direct sum of modules**) Let $\{\mathcal{F}_i\}_{i\in I}$ be a family of $\mathcal{O}_X$-modules. The direct sum of $\mathcal{F}_i$ is the sheafification of the presheaf

$$U \mapsto \bigoplus_{i\in I} \mathcal{F}_i(U)$$

whose restriction is the direct sum of the corresponding restrictions. We denote this sheaf by $\bigoplus_{i\in I} \mathcal{F}_i$ and it is an $\mathcal{O}_X$-module by the following lemma. If for all $i \in I$, we have $\mathcal{F}_i = \mathcal{F}$, then we write

$$\bigoplus_{i\in I} \mathcal{F} = \mathcal{F}^{\oplus I} = \mathcal{F}^{(I)}$$

as usually is done in algebra.

Lemma 3.5.1.23. *$\bigoplus_{i\in I} \mathcal{F}_i$ is an $\mathcal{O}_X$-module and $\left(\bigoplus_{i\in I} \mathcal{F}_i\right)_x \cong \bigoplus_{i\in I} \mathcal{F}_{i,x}$ for all $x \in X$.*

Proof. Since stalks functor is left adjoint (to skyscraper, we didn't covered this but this is a basic known fact), therefore it preserves all colimits and thus $\left(\bigoplus_{i\in I} \mathcal{F}_i\right)_x \cong \bigoplus_{i\in I} \mathcal{F}_{i,x}$. Now, the $\mathcal{O}_X$-module structure over $\bigoplus_{i\in I} \mathcal{F}_i$ is obtained as follows: pick any $U \subseteq X$ open and consider the map

$$\begin{aligned} \eta_U : \mathcal{O}_X(U) \times \left(\bigoplus_{i\in I} \mathcal{F}_i\right)(U) &\longrightarrow \left(\bigoplus_{i\in I} \mathcal{F}_i\right)(U) \\ (c, s : U \to \amalg_{x\in U} \oplus_{i\in I} \mathcal{F}_{i,x}) &\longmapsto cs \end{aligned}$$

where $cs(x) = c_x \cdot s(x)$ where $s(x) \in \bigoplus_{i\in I} \mathcal{F}_{i,x}$ and $\bigoplus_{i\in I} \mathcal{F}_{i,x}$ is an $\mathcal{O}_{X,x}$-module. By exactly same techniques employed in proving them in earlier cases, it can be observed that the above defines a map $\eta : \mathcal{O}_X \times \oplus_{i\in I}\mathcal{F}_i \to \oplus_{i\in I}\mathcal{F}_i$ which is a sheaf map. $\square$

We now cover the other construction we know from algebra.

Direct product of modules

Definition 3.5.1.24. (**Direct product of modules**) Let $\{\mathcal{F}\}_{i\in I}$ be a family of $\mathcal{O}_X$-modules. The direct product of them is defined to be the sheaf

$$U \mapsto \prod_{i\in I} \mathcal{F}_i(U)$$

with product of restrictions as its restriction. Indeed, it is immediate it is a sheaf and that the canonical map $\eta_U : \mathcal{O}_X(U) \times \prod_{i\in I} \mathcal{F}_i(U) \to \prod_{i\in I} \mathcal{F}_i(U)$ mapping as $(c, (s_i)_{i\in I}) \mapsto (c \cdot s_i)_{i\in I}$ makes $\prod_{i\in I} \mathcal{F}_i$ an $\mathcal{O}_X$-module. If $\mathcal{F}_i = \mathcal{F}$ for all $i \in I$, then we denote

$$\prod_{i\in I} \mathcal{F} = \mathcal{F}^{\prod I} = \mathcal{F}^I$$

as is usually done in algebra.

We now define tensor product of two $\mathcal{O}_X$-modules.

Tensor product of modules

Definition 3.5.1.25. (**Tensor product of modules**) Let $\mathcal{F}, \mathcal{G}$ be two $\mathcal{O}_X$-modules. The tensor product of $\mathcal{F}$ and $\mathcal{G}$ is given by the sheafification of the presheaf

$$U \mapsto \mathcal{F}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{G}(U),$$

denoted by $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G}$, as the following lemma shows.

Lemma 3.5.1.26. *$\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G}$ is an $\mathcal{O}_X$-module and $(\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G})_x \cong \mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{G}_x$ for each $x \in X$.*

Proof. The second statement is immediate from Lemma 16.5.1.2. The $\mathcal{O}_X$-module structure is the obvious one: pick any open $U \subseteq X$ and then consider the map

$$\begin{aligned} \eta_U : \mathcal{O}_X(U) \times (\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G})(U) &\longrightarrow (\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G})(U) \\ (a, s : U \to \amalg_{x \in U} \mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{G}_x) &\longmapsto as \end{aligned}$$

where $as(x) = a_x s(x)$. One easily checks that this defines a well-defined natural sheaf map. $\square$

A simple observation also yields the usual identity we know from modules.

Lemma 3.5.1.27. *Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. Then,*

$$\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X \cong \mathcal{F}.$$

Proof. Consider the map

$$\eta : \mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X \longrightarrow \mathcal{F}$$

given on an open $U \subseteq X$ by the map corresponding to the following natural isomorphism (Theorem 20.2.0.1)

$$\eta_U : \mathcal{F}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{O}_X(U) \xrightarrow{\cong} \mathcal{F}(U).$$

This yields the similar isomorphic map on stalks via Lemma 3.5.1.26 to yield the result via Theorem 20.3.0.6, 3. $\square$

Tensor product of modules is obviously commutative.

Lemma 3.5.1.28. *Let $\mathcal{F}, \mathcal{G}$ be two $\mathcal{O}_X$-modules. Then, $\mathcal{F} \otimes \mathcal{G} \cong \mathcal{G} \otimes \mathcal{F}$.*

Proof. Construct the map $\tilde{\eta} : \mathcal{F} \otimes \mathcal{G} \to \mathcal{G} \otimes \mathcal{F}$ as the unique map corresponding to the following

$$\begin{tikzcd} \mathcal{F}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{G}(U) \arrow[r, "\eta_U", "\cong"'] \arrow[dr, "j_U \eta_U"'] & \mathcal{G}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{F}(U) \arrow[d, "j_U"] \\ & (\mathcal{G} \otimes \mathcal{F})(U) \end{tikzcd}.$$

This map on the stalks gives the usual twist isomorphism $\mathcal{F}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{G}_x \cong \mathcal{G}_x \otimes_{\mathcal{O}_{X,x}} \mathcal{F}_x$. $\square$

Free, locally free & finite locally free $\mathcal{O}_X$-modules

Definition 3.5.1.29. (**Free, locally free and finite locally free modules**) Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. Then,

1. $\mathcal{F}$ is called *free* if $\mathcal{F} \cong \mathcal{O}_X^{(I)}$ for some index set I,
2. $\mathcal{F}$ is called *locally free* if for all $x \in X$, there exists open $U \ni x$ such that $\mathcal{F}_{|U} \cong \mathcal{O}_{X|U}^{(I_x)}$ where I_x is an indexing set depending on x,
3. $\mathcal{F}$ is called *finite locally free* if $\mathcal{F}$ is locally free and the indexing set I_x is finite for each $x \in X$. If $I_x = I$ and I has size n, then we say that $\mathcal{F}$ is *locally free of rank n*.

We now observe that the hom sheaf of two locally free modules of finite rank is again locally free of finite rank.

Lemma 3.5.1.30. *Let $\mathcal{F}, \mathcal{E}$ be two locally free $\mathcal{O}_X$-modules of ranks n and m respectively. Then $\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{E})$ and $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{E}$ are both locally free module of rank nm.*

Proof. For each $x \in X$, there exists an open set $U \ni x$ such that $\mathcal{F}_{|U} \cong \mathcal{O}_{X|U}^n$ and $\mathcal{E}_{|U} \cong \mathcal{O}_{X|U}^m$. We then observe the following

$$\mathcal{H}om_{\mathcal{O}_X}(\mathcal{F}, \mathcal{E})(U) = \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{F}_{|U}, \mathcal{E}_{|U}\right) \cong \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}^n, \mathcal{O}_{X|U}^m\right) \cong \mathcal{O}_{X|U}^{nm}$$

where the last isomorphism can be established easily by reducing to the usual module case ($\mathrm{Hom}_R(R^n, R^m) \cong R^{nm}$).

For tensor, we proceed similarly as above. By replacing X by U, we need only show that $\mathcal{O}_X^n \otimes \mathcal{O}_X^m \cong \mathcal{O}_X^{nm}$. Indeed, by universal property of sheafification, it is sufficient to describe a map of presheaves $(\mathcal{O}_X^n \otimes \mathcal{O}_X^m)^- \to \mathcal{O}_X^{nm}$ which is an isomorphism on stalks. The usual isomorphism $R^n \otimes R^m \to R^{nm}$ gives such a map of presheaves, as required. $\square$

An important corollary of the above lemma is as follows.

Corollary 3.5.1.31. *Let $\mathcal{F}$ be be a locally free module of rank n. Then the dual $\mathcal{F}^\vee$ is locally free of rank n.*

Proof. By Lemma 3.5.1.30, $\mathcal{F}^\vee$ is locally free of rank n. $\square$

One may think of finite locally free modules as those modules which are locally free in the usual sense. Consequently, these modules satisfy global version of the properties enjoyed by the usual notion of free modules, as the following result shows.

Proposition 3.5.1.32. [5] *Let $\mathcal{E}$ be a finite locally free of rank n. Then,*

1. *$\mathcal{E}^{\vee\vee} \cong \mathcal{E}$.*
2. *For any $\mathcal{O}_X$-module $\mathcal{F}$, we have*

$$\mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{F}) \cong \mathcal{E}^\vee \otimes_{\mathcal{O}_X} \mathcal{F}.$$

Proof. As $\mathcal{E}$ is locally of free of rank n, therefore there is an open cover $\{U_i\}$ of X such that $\mathcal{E}_{|U_i} \cong \mathcal{O}_{X|U_i}^n$. Let $\{B_j\}$ be a basis of X where each B_j is in some U_i. Consequently, we reduce to constructing an isomorphism in each case only as sheaves over the basis $\{B_j\}$.

[5] Exercise II.5.1 of Hartshorne.

1. Indeed, as each B_j is in some U_i, therefore $\mathcal{E}_{|B_j} \cong \mathcal{O}^n_{X|B_j}$. Consequently, we get the following isomorphisms for any $U \in \{B_j\}$

$$\begin{aligned}
\mathcal{E}^{\vee\vee}(U) &= \mathrm{Hom}_{\mathcal{O}_{X|U}}(\mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{O}_X)\big|_U, \mathcal{O}_{X|U}) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}(\mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}^n_X, \mathcal{O}_X)\big|_U, \mathcal{O}_{X|U}) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}((\mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X, \mathcal{O}_X))^n\big|_U, \mathcal{O}_{X|U}) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}(\mathcal{O}^n_{X|U}, \mathcal{O}_{X|U}) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}, \mathcal{O}_{X|U}\right)^n \\
&\cong \mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X, \mathcal{O}_X)(U)^n \\
&\cong \mathcal{O}_X(U)^n \\
&\cong \mathcal{E}(U),
\end{aligned}$$

and its naturality with resepect to restrictions is evident.

2. Pick any $U \in \{B_j\}$. We then have

$$\begin{aligned}
\mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{F})(U) &\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{E}_{|U}, \mathcal{F}_{|U}\right) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}^n_{X|U}, \mathcal{F}_{|U}\right) \\
&\cong \mathrm{Hom}_{\mathcal{O}_{X|U}}\left(\mathcal{O}_{X|U}, \mathcal{F}_{|U}\right)^n \\
&\cong \mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X, \mathcal{F})(U)^n \\
&\cong \mathcal{F}(U)^n \\
&\cong (\mathcal{O}^n_X \otimes_{\mathcal{O}_X \mathcal{F}})(U)
\end{aligned}$$

by Lemma 3.5.1.27. The fact that this isomorphism is natural with respect to restrictions is immediate. □

Invertible modules and the Picard group

Definition 3.5.1.33. (**Invertible modules**) An $\mathcal{O}_X$-module $\mathcal{L}$ is said to be invertible if it is locally free of rank 1.

The name is justified by the fact that the set of all invertible modules upto isomorphism forms a group under tensor product and is one of the important invariants of a (ringed) space amongst many others. We now show that indeed this forms a group. We will drop the subscript $\mathcal{O}_X$ from the tensor product, for clarity, in the following.

Proposition 3.5.1.34. *Let $\mathcal{L}, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ be invertible $\mathcal{O}_X$-modules. Then,*

1. *$\mathcal{L}_1 \otimes \mathcal{L}_2$ is invertible,*
2. *$(\mathcal{L}_1 \otimes \mathcal{L}_2) \otimes \mathcal{L}_3 \cong \mathcal{L}_1 \otimes (\mathcal{L}_2 \otimes \mathcal{L}_3)$,*
3. *$\mathcal{L}^\vee \otimes \mathcal{L} \cong \mathcal{O}_X$.*

Proof. 1. This is a local question, so pick $x \in X$ and an open set $U \ni x$ such that $\mathcal{L}_{1|U} \cong \mathcal{O}_{X|U} \cong \mathcal{L}_{2|U}$. We wish to construct a natural map $(\mathcal{L}_1 \otimes_{\mathcal{O}_X} \mathcal{L}_2)(U) \to \mathcal{O}_X(U)$ which is an isomorphism. By

Theorem 20.2.0.1, it suffices to show a natural isomorphism $\mathcal{L}_1(U) \otimes_{\mathcal{O}_X(U)} \mathcal{L}_2(U) \to \mathcal{O}_X(U)$. This is constructed quite easily as $\mathcal{L}_1(U) \otimes_{\mathcal{O}_X(U)} \mathcal{L}_2(U) \cong \mathcal{O}_X(U) \otimes_{\mathcal{O}_X(U)} \mathcal{O}_X(U) \cong \mathcal{O}_X(U)$. Thus we just need to consider $\mathrm{id}_{\mathcal{O}_X(U)}$.
2. This is again a local question, which can be answered by establishing an isomorphism (by using Theorem 20.2.0.1)

$$(\mathcal{L}_1(U) \otimes_{\mathcal{O}_X(U)} \mathcal{L}_2(U)) \otimes_{\mathcal{O}_X(U)} \mathcal{L}_3(U) \cong \mathcal{L}_1(U) \otimes_{\mathcal{O}_X(U)} (\mathcal{L}_2(U) \otimes_{\mathcal{O}_X(U)} \mathcal{L}_3(U))$$

for any open $U \subseteq X$, but that is an immediate observation from algebra.
3. By Corollary 3.5.1.31, we have that $\mathcal{L}^\vee$ is invertible. By Theorem 20.3.0.6, 3, the result would follow if we can show that there is a natural $\mathcal{O}_X$-linear map $\varphi : \mathcal{L}^\vee \otimes \mathcal{L} \to \mathcal{O}_X$ such that for each point $x \in X$ there exists an open set $x \in U \subseteq X$ such that on U, φ yields an $\mathcal{O}_X(U)$-linear isomorphism $(\mathcal{L}^\vee \otimes \mathcal{L})(U) \cong \mathcal{O}_X(U)$. We may take U small enough so that $\mathcal{L}^\vee_{|U} \cong \mathcal{O}_{X|U} \cong \mathcal{L}_{|U}$. Thus, after replacing X by U, we may assume $\mathcal{L} = \mathcal{O}_X = \mathcal{L}^\vee$. By Lemmas 3.5.1.20 and 3.5.1.27, we obtain the following isomorphisms

$$\mathcal{L}^\vee \otimes \mathcal{L} = \mathcal{H}om(\mathcal{L}, \mathcal{O}_X) \otimes \mathcal{O}_X \cong \mathcal{H}om(\mathcal{O}_X, \mathcal{O}_X) \otimes \mathcal{O}_X \cong \mathcal{O}_X \otimes \mathcal{O}_X \cong \mathcal{O}_X.$$

This can easily be promoted to a sheaf map. $\square$

Definition 3.5.1.35. (**Picard group of X**) The Picard group of X is defined to be the set of all isomorphism classes of invertible modules with the operation of tensor product. We denote this by

$$\mathrm{Pic}(X)$$

The Proposition 3.5.1.34 and Lemma 3.5.1.28 shows that $\mathrm{Pic}(X)$ is indeed an abelian group.

Direct and inverse image modules

In this and the next sections, we show how the modules behave under map of ringed spaces.

Definition 3.5.1.36. (**Direct image**) Let $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ be a map of ringed spaces and let $\mathcal{F}$ be an $\mathcal{O}_X$-module. Then the direct image of $\mathcal{F}$ under f is the direct image sheaf $f_*\mathcal{F}$ which is again an $\mathcal{O}_Y$-module given by the following composition

$$\mathcal{O}_Y \times f_*\mathcal{F} \xrightarrow{f^\flat \times \mathrm{id}} f_*\mathcal{O}_X \times f_*\mathcal{F} \xrightarrow{f_* m} f_*\mathcal{F}$$

where $m : \mathcal{O}_X \times \mathcal{F} \longrightarrow \mathcal{F}$ is the $\mathcal{O}_X$-module structure on $\mathcal{F}$. Note that f_* commutes with products as f_* is a right-adjoint.

The inverse image of a module, on the other hand, is an involved construction.

Definition 3.5.1.37. (**Inverse image**) Let $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ be a map of ringed spaces and let $\mathcal{G}$ be an $\mathcal{O}_Y$-module. The inverse image of $\mathcal{G}$ is defined to be the map

$$f^*\mathcal{G} := \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}\mathcal{G}$$

which is indeed an $\mathcal{O}_X$-module as the following lemma shows.

Lemma 3.5.1.38. *The sheaf $f^*\mathcal{G}$ is an $\mathcal{O}_Y$-module.*

Proof. We need to show three statements:

1. $\mathcal{O}_X$ is an $f^{-1}\mathcal{O}_Y$-module.
2. $f^{-1}\mathcal{G}$ is an $f^{-1}\mathcal{O}_Y$-module.
3. $f^*\mathcal{G}$ is an $\mathcal{O}_X$-module.

Statement 1 follows from the following composition

$$f^{-1}\mathcal{O}_Y \times \mathcal{O}_X \xrightarrow{f^\sharp \times \mathrm{id}} \mathcal{O}_X \times \mathcal{O}_X \longrightarrow \mathcal{O}_X$$

where the latter is just the multiplication structure on $\mathcal{O}_X$. Statement 2 follows from $\mathcal{O}_Y$-module structure on $\mathcal{G}$ and the fact that $f^{-1}(\mathcal{G} \times \mathcal{G}') = f^{-1}\mathcal{G} \times f^{-1}\mathcal{G}'$ for two sheaves $\mathcal{G}, \mathcal{G}'$ over Y. Indeed, the latter follows from the fact that $f^+(\mathcal{G} \times \mathcal{G}') = f^+\mathcal{G} \times f^+\mathcal{G}'$, which in turn follows from the fact that filtered colimit commutes with finite limits. Statement 3 now follows immediately. $\square$

We now state an important result, that is $f_* \vdash f^*$.

Proposition 3.5.1.39. *Let $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ be a map of ringed spaces. Then,*

$$\begin{tikzcd} \mathbf{Mod}(\mathcal{O}_Y) \arrow[r, shift left=2, "f^*"] \arrow[r, phantom, "\perp"] & \mathbf{Mod}(\mathcal{O}_X) \arrow[l, shift left=2, "f_*"] \end{tikzcd} .$$

In other words, we have a natural isomorphism of groups

$$\mathrm{Hom}_{\mathcal{O}_X}(f^*\mathcal{G}, \mathcal{F}) \cong \mathrm{Hom}_{\mathcal{O}_Y}(\mathcal{G}, f_*\mathcal{F}).$$

Proof. Omitted. $\square$

Sums & intersections of submodules

Modules generated by sections

Inverse limit

Do Hartshorne Exercise 1.12 as well.

Direct limit

Do Hartshorne Exercise 1.11 as well.

Tensor, symmetric & exterior powers

We now define $T(\mathcal{F})$, $S(\mathcal{F})$ and $\wedge(\mathcal{F})$ for a module $\mathcal{F}$.

Definition 3.5.1.40 ($T(\mathcal{F}), \mathrm{Sym}(\mathcal{F})$ **and** $\wedge(\mathcal{F})$)**.** Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. The sheafification of presheaf $U \mapsto T(\mathcal{F}(U))$ or $\mathrm{Sym}(\mathcal{F}(U))$ or $\wedge(\mathcal{F}(U))$ is denoted to be $T(\mathcal{F})$ or $S(\mathcal{F})$ or $\wedge(\mathcal{F})$ called the tensor or symmetric or exterior algebra, respectively. This is an $\mathcal{O}_X$-algebra, i.e. a sheaf of rings which is an $\mathcal{O}_X$-module. Moreover, we have

$$T(\mathcal{F}) = \bigoplus_{n\geq 0} T^n(\mathcal{F})$$

where $T^n(\mathcal{F})$ is the sheafification of $U \mapsto T^n(\mathcal{F}(U))$. Note that this makes sense as sheafification is a left adjoint, so it commutes with all colimits. We call $T^n(\mathcal{F})$ the n^{th}-tensor power of $\mathcal{F}$. Similarly, we define $\mathrm{Sym}^n(\mathcal{F})$ and $\wedge^n(\mathcal{F})$.

Lemma 3.5.1.41. *If $\mathcal{F} = \mathcal{O}_X^n$ is a free $\mathcal{O}_X$-module of rank n, then*
1. $T^r(\mathcal{F}) \cong \mathcal{O}_X^{n^r}$,
2. $\mathrm{Sym}^r(\mathcal{F}) \cong \mathcal{O}_X^{{}^{n+r-1}C_{n-1}}$,
3. $\wedge^r(\mathcal{F}) \cong \mathcal{O}_X^{{}^nC_r}$.

Proof. All three isomorphisms are obtained by defining a corresponding map of presheaves which is an isomorphism on stalks, where this map is induced from the usual map in algebra:

$$\begin{aligned} R^n \otimes R^m &\cong R^{nm} \\ \mathrm{Sym}^r(R^n) &\cong R^{{}^{n+r-1}C_r} \\ \wedge^r(R^n) &\cong R^{{}^nC_r}. \end{aligned}$$

Then the corresponding map on sheaves induced by universal property of sheafification is an isomorphism as it is so on stalks. □

We now indulge in generalizing some local properties of tensor algebra to this global case. We first have the standard observation of instantiating these definitions on the finite locally free case, which generalizes the usual tensor calculations of free modules.

Lemma 3.5.1.42. [6] *Let $\mathcal{F}$ be a finite locally free $\mathcal{O}_X$-module of rank n. Then, $T^r(\mathcal{F}), \mathrm{Sym}^r(\mathcal{F})$ and $\wedge^r(\mathcal{F})$ is a finite locally free $\mathcal{O}_X$-module of rank n^r, ${}^{n+r-1}C_{n-1}$ and ${}^nC_r$ respectively.*

Proof. Let $\{U_\alpha\}$ be an open cover of X where $\mathcal{F}$ is $\mathcal{O}^n_{X|U_\alpha}$ for each α. Let $\mathcal{B}$ be a basis of X such that for any $B \in \mathcal{B}$, we have $B \subseteq U_\alpha$ for some α. Observe that $\mathcal{F}_{|B} \cong \mathcal{O}^n_{X|B}$. Hence, we may replace X by B to assume that $\mathcal{F}$ is free of rank n. The result now follows from Lemma 3.5.1.41. □

Another global phenomenon that is borrowed by tensor calculation of free modules is the perfect pairing of wedge product.

[6] Exercise II.5.16 of Hartshorne.

$\mathcal{E}xt$ module

$\mathcal{T}or$ module

3.5.2 The abelian category of $\mathcal{O}_X$-modules

We now show an important result that category of $\mathcal{O}_X$-modules over any ringed space is an abelian category (thus we can do whole of homological algebra over it!). We have essentially done everything, but we write it here for clear reference.

Theorem 3.5.2.1. *Let $(X, \mathcal{O}_X)$ be a ringed space. Then the category $\mathbf{Mod}(\mathcal{O}_X)$ of $\mathcal{O}_X$-modules is an abelian category.*

Proof. For any two $\mathcal{O}_X$-modules $\mathcal{F}, \mathcal{G}$, we have $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$ is an abelian group where for any two $f, g \in \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$, the sum $h = f + g$ is defined to as follows: pick any open $U \subseteq X$ and define $h_U = f_U + g_U$. This is an $\mathcal{O}_X$-linear sheaf map because f and g are. Hence $\mathbf{Mod}(\mathcal{O}_X)$ is preadditive. Moreover $\mathbf{Mod}(\mathcal{O}_X)$ is additive. This is what we did in the preceding section while defining finite products of $\mathcal{O}_X$-modules. The preceding section also shows that $\mathbf{Mod}(\mathcal{O}_X)$ has all kernels and cokernels. Consequently, we need only show that the for any $f : \mathcal{F} \to \mathcal{G}$ in $\mathbf{Mod}(\mathcal{O}_X)$, $\mathrm{CoIm}\,(f) \cong \mathrm{Im}\,(f)$. Indeed, this is a local question and can be thus immediately seen by first isomorphism theorem. More precisely, we need only construct this isomorphism on a basis of X, where the canonical map $\mathrm{CoIm}\,(f) \to \mathrm{Im}\,(f)$ is an isomorphism by first isomorphism theorem. This completes the proof. □

Theorem 3.5.2.2. *Let $(X, \mathcal{O}_X)$ be a ringed space. Then the abelian category $\mathbf{Mod}(\mathcal{O}_X)$ has enough injectives.*

Proof. Let $\mathcal{F}$ be an $\mathcal{O}_X$-module. We wish to find an injective $\mathcal{O}_X$-module $\mathcal{I}$ such that $\mathcal{F} \hookrightarrow \mathcal{I}$. First note that for each $x \in X$, we have an injective $\mathcal{O}_{X,x}$-module I_x such that $\mathcal{F}_x \hookrightarrow I_x$ by Theorem 19.2.2.7. Observe that I_x is a sheaf over $i : \{x\} \hookrightarrow X$. Let $\mathcal{I} = \prod_{x \in X} i_* I_x$ be the corresponing $\mathcal{O}_X$-module. We claim that $\mathcal{I}$ is an injective $\mathcal{O}_X$-module and there is an injective map $\mathcal{F} \hookrightarrow \mathcal{I}$.

To see that there is an injective map $\mathcal{F} \hookrightarrow \mathcal{I}$, we claim the following three isomorphisms

$$\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{I}) \cong \prod_{x \in X} \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, i_* I_x) \cong \prod_{x \in X} \mathrm{Hom}_{\mathcal{O}_{X,x}}(\mathcal{F}_x, I_x).$$

The first isomorphism is immediate from limit preserving property of covariant hom. The second isomorphism is obtained by the following isomorphism

$$\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{F}, i_* I_x) \cong \mathrm{Hom}_{\mathcal{O}_{X,x}}(\mathcal{F}_x, I_x) \tag{$*$}$$

for each $x \in X$. Indeed, this follows from the maps $f \mapsto f_x$ and $(\tilde{\kappa} : \mathcal{F} \to i_* I_x) \leftarrow\!\shortmid (\kappa : \mathcal{F}_x \to I_x)$ where $\tilde{\kappa}$ is defined on an open set $U \subseteq X$ as $\tilde{\kappa}_U : \mathcal{F}(U) \to I_x$ mapping as $s \mapsto \kappa((U, s)_x)$. These are clearly inverses of each other. It then follows that a map $\mathcal{F} \to \mathcal{I}$ is equivalent to a collection of maps $\mathcal{F}_x \to I_x$ and since we have $\mathcal{F}_x \hookrightarrow I_x$, therefore we obtain a unique injective map $\mathcal{F} \hookrightarrow \mathcal{I}$.

Finally, we claim that $\mathrm{Hom}_{\mathcal{O}_X}(-, \mathcal{I})$ is exact as a functor into the category of abelian groups. To this end, by left exactness of hom, we need only show that this is right exact. This immediately follows from isomorphism $(*)$ and I_x being injective and that product of surjective homomorphisms is surjective. This completes the proof. □

3.6 Torsors and 1^{st}-Čech cohomology group

Once we have understood the constructions of the last section, we can now start doing some serious geometry over our manifolds. Indeed, this is what we start laying out in this section.

3.7 Bundles

We give here the general theory of fiber, principal and vector bundles. When the need arises, we will instantiate this into different areas (like in the chapter on differential geometry). The material in previous chapter will allow a very united way of looking at the notion of bundles, and will start portraying the intimate connection that bundles and cohomology has.

3.7.1 Generalities on twisting atlases

Let $p : E \to B$ be a map of topological spaces/manifolds together with a specified subsheaf of groups $\mathcal{G} \subseteq \mathcal{A}_B(E) \in \mathbf{Sh}(B)$ where $\mathcal{A}_B(E)$ is the sheaf of homeomorphisms/isomorphisms over B; for any open $U \subseteq B$, the group $\mathcal{A}_B(E)(U)$ consists of all homeomorphisms/isomorphisms $\varphi : p^{-1}(U) \to p^{-1}(U)$ such that $p \circ \varphi = p$.

The tuple $(p : E \to B, \mathcal{G})$ is the pre-datum for defining $(p, \mathcal{G})$-twisting atlas for a map $\pi : X \to B$.

Definition 3.7.1.1 (**$(p, \mathcal{G})$-twisting atlas for a map**)**.** Let $p : E \to B$ be a map and $\mathcal{G}$ be a subsheaf of groups $\mathcal{G} \subseteq \mathcal{A}_B(E)$. Let $\pi : X \to B$ be a map. Then, a $(p, \mathcal{G})$-twisting atlas for π is a family $(U_i, h_i)_{i\in I}$ where $\{U_i\}_{i\in I}$ is an open cover of B and $h_i : \pi^{-1}(U_i) \xrightarrow{\cong} p^{-1}(U_i)$ is an isomorphism over U_i such that for any $i, j \in I$, denoting $U_{ij} = U_i \cap U_j$, we have

$$\begin{tikzcd} p^{-1}(U_{ij}) \arrow[dr, "p"'] \arrow[rr, shift right, "h_j^{-1}\big|_{p^{-1}(U_{ij})}"'] & & \pi^{-1}(U_{ij}) \arrow[ll, shift right, "h_i|_{\pi^{-1}(U_{ij})}"'] \arrow[dl, "\pi"] \\ & U_{ij} & \end{tikzcd}$$

and from which we require that

$$h_{ij} = h_i|_{\pi^{-1}(U_{ij})} \circ h_j^{-1}\Big|_{p^{-1}(U_{ij})}$$

is a section in $\mathcal{G}(U_{ij})$. We then call $\pi : X \to B$ together with (U_i, h_i) *a twist of $p : E \to B$ with structure sheaf $\mathcal{G}$.*

Using this, we may define a general notion of a bundle.

Definition 3.7.1.2 (**Bundles**)**.** Let $\pi : X \to B$ be a map, F a space/manifold and $p : B \times F \to B$ be the projection map onto first coordinate. Then π is a bundle with fiber F if there is a $(p, \mathcal{A}_B(B\times F))$-twisting atlas for π. Equivalently, π is a bundle with fiber F if it is a twist of $p : B \times F \to B$ with full structure sheaf $\mathcal{A}_B(B \times F)$.

Remark 3.7.1.3. Let $\pi : X \to B$ be a bundle with fiber F. Consequently we have a $\mathcal{A}_B(B \times F)$-twisting atlas of $p : B \times F \to B$ denoted (U_i, h_i), where $h_i : \pi^{-1}(U_i) \to p^{-1}(U_i)$ is an isomorphism over U_i such that the transition maps $h_{ij} : p^{-1}(U_{ij}) = U_{ij} \times F \to U_{ij} \times F = p^{-1}(U_{ij})$ is just an isomorphism over U_{ij} (i.e. $h_{ij} \in \mathcal{A}_B(B \times F)(U_{ij})$).

3.8 Differential forms and de-Rham cohomology

Do this from Section 8.6 and Section 10.4 of Wedhorn, via sheaf cohomology. Add motivation from courses.

3.8.1 Differential forms on $\mathbb{R}^n$

We first discuss differential forms on $\mathbb{R}^n$ as this provides clear and sufficient motivation for the abstract treatment of differential forms in all other places where it is used. We begin by defining the main ingredients. The material of Section 16.5 is used in the following.

Definition 3.8.1.1. (**Coordinate forms on** $\mathbb{R}^n$) Fix $n \in \mathbb{N}$. Let $V = \mathbb{R}^n$ be the n-dimensional $\mathbb{R}$-module. The functional

$$\begin{aligned} dx_i : V &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longmapsto x_i \end{aligned}$$

is called the i^{th}-coordinate form on V, for each $i = 1, \dots, n$. Note that dx_i is a 1-form/1-tensor, i.e. $dx_i \in M^1(V) = V^*$. Observe that dx_i is the dual basis of V^* corresponding to standard basis e_i of V.

Next, we define a multilinear map which for each choices of axes, gives the volume of the parallelopiped obtained by the projection along those axes, given a parallelopiped spanned by some vectors.

Definition 3.8.1.2. (**Projection forms on** $\mathbb{R}^n$) Fix $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $V = \mathbb{R}^n$ be the n-dimensional $\mathbb{R}$-module. Let $I = (i_1, \dots, i_k)$ be an ordered k-tuple where $1 \leq i_j \leq n$ for each $j = 1, \dots, k$. Then, we define the I-projection form as

$$dx_I := \pi_k(dx_{i_1} \otimes \ldots \otimes dx_{i_k}) = D_I$$

which is an alternating k-form on V, that is $dx_I \in \Lambda^k(V)$ (see Example 16.5.3.11). More explicitly, it is given by the following k-linear form on V

$$\begin{aligned} dx_I : V \times \cdots \times V &\longrightarrow \mathbb{R} \\ (v_1, \dots, v_k) &\longmapsto \det \begin{bmatrix} dx_{i_1}(v_1) & dx_{i_2}(v_1) & \dots & dx_{i_k}(v_1) \\ dx_{i_1}(v_2) & dx_{i_2}(v_2) & \dots & dx_{i_k}(v_2) \\ \vdots & \vdots & \dots & \vdots \\ dx_{i_1}(v_k) & dx_{i_2}(v_k) & \dots & dx_{i_k}(v_k) \end{bmatrix}. \end{aligned}$$

Remark 3.8.1.3. Recall from Theorem 16.5.3.14 that $\Lambda^k(V)$ has basis given by dx_I for distinct increasing k-tuples from $1, \dots, n$. Thus, $\{dx_I\}_I$ forms an $\mathbb{R}$-basis of $\Lambda^k(V)$ of size ${}^nC_k$.

Remark 3.8.1.4. Recall that wedge product of forms is given by the following (where one defines them only on the basis elements)

$$\begin{aligned} \Lambda^k(V) \times \Lambda^l(V) &\longrightarrow \Lambda^{k+l}(V) \\ (dx_I, dx_J) &\longmapsto dx_I \wedge dx_J := dx_{(I,J)} \end{aligned}$$

where recall that $dx_{(I,J)}$ will be zero if there is any index common in I and J (see Definition 16.5.4.1), where I, J are increasing tuples of indices from $\{1, \dots, n\}$ of lengths k and l respectively. From the above, we see that for any alternating k-form $\omega = \sum_I a_I dx_I$ and alternating l-form $\eta = \sum_J b_j dx_J$, their wedge product is defined as

$$\omega \wedge \eta = \sum_J \sum_I a_I b_J (dx_I \wedge dx_J).$$

Remark 3.8.1.5. Let $U \subseteq \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$. Observe that $\mathcal{C}^\infty(U)$, the ring of smooth $\mathbb{R}$-valued functions on U, is an $\mathbb{R}$-algebra. In the same vein, we know that alternating k-forms $\Lambda^k(\mathbb{R}^n)$ forms an $\mathbb{R}$-vector space of dimension ${}^nC_k$ (see Theorem 16.5.3.14).

Definition 3.8.1.6. (**Differential k-forms**) Let $U \subseteq \mathbb{R}^n$ be an open set and $0 \le k \le n$. The module of differential k-forms is defined to be the following $\mathbb{R}$-vector space

$$\Omega_U^k = \Lambda^k(\mathbb{R}^n) \otimes_{\mathbb{R}} \mathcal{C}^\infty(U).$$

As $\Lambda^k(\mathbb{R}^n)$ is a free $\mathbb{R}$-module with rank ${}^nC_k$, therefore Ω_U^k is a free $\mathcal{C}^\infty(U)$-module of rank ${}^nC_k$.

Remark 3.8.1.7. Observe that $\{\Omega_U^k\}$ obtains the wedge product structure from the wedge product on $\{\Lambda^k(\mathbb{R}^n)\}$ as we may define for $\omega = \sum_I f_I dx_I \in \Lambda^k(\mathbb{R}^n)$ and $\eta = \sum_J g_J dx_J$ the following

$$\begin{aligned} \omega \wedge \eta &:= \left(\sum_I f_I dx_I\right) \wedge \left(\sum_J g_J dx_J\right) \\ &= \sum_I \sum_J f_I g_J dx_I \wedge dx_J. \end{aligned}$$

Thus, $\bigoplus_{k \ge 0} \Omega_U^k$ forms a graded $\mathcal{C}^\infty(U)$-algebra.

Remark 3.8.1.8. An arbitrary element $\omega \in \Omega_U^k$ is called a differential k-form over U and is written as

$$\omega = \sum_{I \in X_k} f_I(x_1, \dots, x_n) dx_I$$

where X_k is the set of size ${}^nC_k$ of all k-combinations in increasing order of $\{1, \dots, n\}$ and $f_I \in \mathcal{C}^\infty(U)$ is a smooth function. Observe that $\Omega_U^0 = \mathcal{C}^\infty(U)$.

We now construct the exterior derivative which will be a differential over the chain complex Ω_U^k, as we will see soon.

Definition 3.8.1.9. (**Exterior derivative**) Let $U \subseteq \mathbb{R}^n$ be an open subset and $\{\Omega_U^k\}_{k \in \mathbb{N}}$ be the modules of differential k-forms. For each $k \in \mathbb{N} \cup \{0\}$, we define a map $d : \Omega_U^k \to \Omega^{k+1_U}$ as follows. Define for $k = 0$ the following

$$\begin{aligned} d : \Omega_U^0 &\longrightarrow \Omega_U^1 \\ f &\longmapsto \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i \end{aligned}$$

where since $f \in \mathcal{C}^\infty(U)$ is smooth, therefore so is $\partial f/\partial x_i$. Further, since $dx_i \in \Lambda^1(\mathbb{R}^n)$, therefore the above is well-defined. For $k \geq 1$, we define d as follows

$$\begin{gathered} d : \Omega_U^k \longrightarrow \Omega_U^{k+1} \\ \omega = \sum_{I \in X_k} f_I dx_I \longmapsto d\omega = \sum_{I \in X_k} df_I \wedge dx_I \end{gathered}$$

where $dx_I \in \lambda^k(\mathbb{R}^n)$. Observe that $df_I \in \Omega_U^1$, thus indeed $df_I \wedge dx_I \in \Omega_U^{k+1}$. This map d is called the exterior derivative of differential forms.

The following are immediate but important properties of exterior derivative. **TODO**.

Chapter 4

Foundational Differential Geometry

Contents

The material of previous chapter has already gotten us very close to topics in geometry over C^∞-manifolds.

4.1 Bundles in differential geometry and applications

4.2 Cohomological methods

4.3 Covariant derivative, connections, classes and curvatures

Chapter 5

Foundational Homotopy Theory

Contents

We introduce basic players of homotopy theory.

Let us first engage in a discussion of the type of spaces we would like to work with, that is, compactly generated space.

Definition 5.0.0.1 (Compactly generated spaces). A space X is said to be compactly generated if it satisfies

1. (*weak Hausdorff*) for any compact Hausdorff space K and a map $g : K \to X$, the image $g(K)$ is closed,
2. (*k-space*) for any $A \subseteq X$, if $g^{-1}(A)$ is closed in K for any $g : K \to X$ where K is a compact Hausdorff space[1], then A is closed in X.

The following are some immediate observations.

Proposition 5.0.0.2. *Let X be a compactly generated space. Then,*

1. *Every compact subspace of X is closed.*
2. *If K is compact Hausdorff and $g : K \to X$ is a map, then $g(K) \subseteq X$ is compact Hausdorff.*
3. *If X is compactly generated and $f : X \to Y$ is a function, then f is continuous if and only if $f|_K$ is continuous for all compact subspaces $K \subseteq X$.*
4. *Any closed subspace of a compactly generated space is compactly generated.*

Proof. **TODO.** □

Example 5.0.0.3. Following are some examples of compactly generated spaces.
1. Any compact Hausdorff space is compactly generated. Indeed, for any compact Hausdorff K and a map $g : K \to X$, we have $g(K)$ is compact in X which is Hausdorff, so closed. Furthermore, if $A \subseteq X$ and $g^{-1}(A)$ is closed in K for any such g, then letting $K = X$ and $g = \mathrm{id}$, we immediately deduce that A is closed, as required.

[1] we then call A to be *compactly closed*

2. Any Hausdorff space X which is locally compact is compactly generated. Indeed, for any compact Hausdorff K and a map $g : K \to X$, we have $g(K)$ is compact in X which is Hausdorff, so closed. Furthermore, if $A \subseteq X$ and $g^{-1}(A)$ is closed in K for any such g, then letting $\tilde{X}$ denote the 1-pt. compactification of X, we see that $\tilde{X}$ is compact Hausdorff. Consequently we may consider the map id : $\tilde{X} \to \tilde{X}$. As any compact Hausdorff space is compactly generated as shown above, therefore $\mathrm{id}^{-1}(A) = A$ is closed by hypothesis, as needed.

3. Hence, every CW-complex is a compactly generated space.

Remark 5.0.0.4. The above example in particular shows that any real or complex manifold is a compactly generated space.

Construction 5.0.0.5. (*k-ification*) Let X be a weak-Hausdorff space. Then, X can be made into a compactly generated space. Define kX to have the same set as X but a finer topology obtained by deeming any compactly closed subspace to be closed in kX. It then follows that

1. kX is compactly generated,
2. the function id : $kX \to X$ is continuous,
3. X and kX have same compact subsets,
4. for weak Hausdorff spaces X and Y, we have $k(X \times Y) = kX \times kY$.

We now show why we restrict our gaze to only these spaces. In part because the category of compactly generated spaces is well-behaved.
TODO *Category* $\mathbf{Top}^{cg}$ *has limits, colimits and exponential objects (all after* k*-ification) and that the dual notion of homotopy as a path in function space is same as that of the usual notion.*

Remark 5.0.0.6. From now on in this chapter, we only work with the category of compactly generated spaces, $\mathbf{Top}^{cg}$. Moreover, any construction on spaces that we do is assumed to be k-ified, i.e. functor k is applied to it to always end up with the category of compactly generated spaces.

Next, we introduce constructions that one can do on based spaces. We denote $\mathbf{Top}_*^{cg}$ to be the category of based compactly generated spaces and based maps between them.

Construction 5.0.0.7 (*Based constructions*)**.** Let X and Y be two based spaces. Then, we denote by

1. $[X, Y]$ the based homotopy classes of based maps from X to Y. This is a based set itself, the basepoint being the homotopy class of $c_* : X \to Y$ mapping $x \mapsto *$. If $X \simeq X'$ and $Y \simeq Y'$, then there is a base point preserving bijection $[X, Y] \cong [X', Y']$.
2. $X \wedge Y$ the smash product given by $X \times Y / X \vee Y$ where $X \vee Y = \{*\} \times Y \cup X \times \{*\}$. This is a based space, the base point being the point corresponding to the subspace $X \vee Y$.
3. $\mathrm{Map}_*(X, Y)$ the collection of based maps from X to Y. This is again a based space in compact-open topology where the basepoint is c_*.
4. X_+ the based space obtained by adjoining a distinct point $*$ to X.
5. $X \wedge I_+$ the reduced cylinder of X where X is based. For any based X and unbased Y the based space $X \wedge Y_+$ is naturally homeomorphic to $X \times Y / \{*\} \times Y$.

There is a natural "$\otimes$-Hom" adjunction in $\mathbf{Top}_*^{cg}$.

Theorem 5.0.0.8. *Let* X, Y, Z *be based spaces in* $\mathbf{Top}_*^{cg}$. *Then we have a natural isomorphism*

$$\mathrm{Map}_*(X \wedge Y, Z) \cong \mathrm{Map}_*(X, \mathrm{Map}_*(Y, Z)).$$

Proof. (*Sketch*) Let $f : X \wedge Y \to Z$. Then by universal property of quotients, we get a map $\bar{f} : X \times Y \to Z$ which is constant on $X \vee Y$. Now construct

$$\begin{aligned} \tilde{f} : X &\longrightarrow \mathrm{Map}_*(Y, Z) \\ x &\longmapsto y \mapsto \bar{f}(x, y). \end{aligned}$$

The fact that this is based follows from $\bar{f}$ being constant on $X \vee Y$.

Let $g : X \to \mathrm{Map}_*(Y, Z)$ a based map. Then we get

$$\begin{aligned} \bar{g} : X \times Y &\longrightarrow Z \\ (x, y) &\longmapsto g(x)(y). \end{aligned}$$

This is based immediately. Further, on $X \vee Y$, we see that $\bar{g}$ is constant. By universal property of quotients, we get the required $\tilde{g} : X \wedge Y \to Z$. □

This theorem shows the duality between smash products and mapping space constructions.

Construction 5.0.0.9 (*More based constructions*)**.** We now give two constructions each for smash product and mapping space which complement each other.

1. CX the cone of X obtained by $X \wedge I$ where 1 is the basepoint of I.
2. ΣX the suspension of X obtained by $X \wedge S^1$.
3. PX the path space of X obtained by $\mathrm{Map}_*(I, X)$.
4. ΩX the loop space of X obtained by $\mathrm{Map}_*(S^1, X)$.

It follows from Theorem 5.0.0.8 that we have following natural isomorphisms

$$\mathrm{Map}_*(CX, Y) \cong \mathrm{Map}_*(X, PY)$$

and

$$\mathrm{Map}_*(\Sigma X, Y) \cong \mathrm{Map}_*(X, \Omega Y),$$

the latter being the famous *suspension-loop space* adjunction.

In the next few items, we give results which are simple to see but important as technical tools.

Proposition 5.0.0.10. *Let X, Y be based spaces in $\mathbf{Top}_*^{cg}$. Then*

$$\pi_0(\mathrm{Map}_*(X, Y)) \cong [X, Y].$$

In particular, we have

$$[\Sigma X, Y] \cong [X, \Omega Y].$$

Proof. (*Sketch*) In $\mathbf{Top}^{cg}$, both left and right notions of homotopy are equivalent. Consequently, a path-component in $\mathrm{Map}_*(X, Y)$ is equivalently the set of based maps $X \to Y$ which are homotopic, as required. □

Every space can be *pointified*.

Definition 5.0.0.11 (**Pointification**)**.** The functor $(-)_+ : \mathbf{Top} \to \mathbf{Top}_*$ given by $X \mapsto X_+$ and $f : X \to Y$ mapping to $f_+ : X_+ \to Y_+$ is called the pointification functor.

There are important relationships between based and unbased constructions. We first have the following simple observation.

Lemma 5.0.0.12. *Let X be a based space. We have the following bijection*

$$\begin{Bmatrix} \textit{Based homotopies } h : \\ X \times I \to Y \end{Bmatrix} \cong \mathrm{Map}_*(X \wedge I_+, Y).$$

□

Remark 5.0.0.13. Let X be an unbased space. All the construction of Construction 5.0.0.9 have an unbased counterpart where smash products are replaced by Cartesian product and Map_* are replaced by Map. In particular,

1. CX the unreduced cone of X obtained by $X \times I/X \times \{1\}$.
2. ΣX the unreduced suspension of X obtained by $X \times S^1/X \times \{1\}$.
3. PX the unbased path space of X obtained by $\mathrm{Map}(I, X)$.
4. ΩX the unbased loop space of X obtained by $\mathrm{Map}(S^1, X)$.

We also call them by same name, if it is clearly understood that the space in question is unbased.

The following is an important observation about pointification and cones.

Lemma 5.0.0.14. *Let X be an unbased space. Then, the unreduced cone of X is isomorphic to the reduced cone on X_+. That is,*

$$CX \cong CX_+.$$

□

5.1 Fundamental group and covering maps

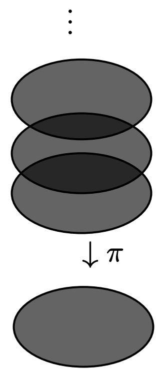
5.1.1 Covering spaces

We will now study a very important concept which is used everywhere in algebraic topology, the concept of covering spaces. This concept captures the notion of when does another space *covers* another space. Even though at this time it may seem completely unrelated to what we've been doing, but we will soon see that using this simple idea we would be able to calculate first homotopy group of S^1. So let us first give the definition of a covering space:

Definition 5.1.1.1. (**Covering space**) Let X be a topological space and suppose $\pi : \tilde{X} \to X$ is a continuous map such that for all $x \in X$, there exists open neighborhood $U_x \ni x$ such that:

1. $\pi^{-1}(U_x) = \coprod_{\alpha \in J_x} V_\alpha$ where V_α's are disjoint open sets in $\tilde{X}$,
2. $\pi|_{V_\alpha} : V_\alpha \to U_x$ is a homeomorphism.

Then, $\pi : \tilde{X} \to X$ is said to be a **covering map** and $\tilde{X}$ is said to be a covering space over X. In this case, the open neighborhood $U_x \subseteq X$ containing x is said to be the **evenly-covered neighborhood** of $x \in X$.



Let us begin with an important example.

Example 5.1.1.2. Well, clearly, the easiest way to get a covering space out of any space is to simply consider that map $X \amalg X \to X$. But that's not interesting.
The most important example of covering spaces that we will consider in this course is the exponential map:

$$\begin{aligned}\exp : \mathbb{R} &\longrightarrow S^1\\ \theta &\longmapsto e^{2\pi i\theta}.\end{aligned}$$

Let us make sure that this is indeed a covering map. Take any point $e^{2\pi i\theta} \in S^1$ where $0 < \theta \leq 1$. Now consider an open set U of S^1, formed by $B_\epsilon(e^{2\pi i\theta}) \cap S^1$ where $0 < \epsilon < 2$. Denote $U =: e^{2\pi i(\theta-\delta,\theta+\delta)}$ where clearly $0 < \delta < 1/2$. Consider now $\pi^{-1}(U) \subseteq \mathbb{R}$. We will have

$$\pi^{-1}(U) = \coprod_{n\in\mathbb{Z}} (\theta + 2\pi n - \delta, \theta + 2\pi n + \delta).$$

Denote $V_n := (\theta + 2\pi n - \delta, \theta + 2\pi n + \delta)$. Moreover, it is clear that

$$\pi|_{V_n} : V_n \longrightarrow U$$

is a homeomorphism. So indeed π is a covering map of S^1. This is a very famous covering map as well. You should think of it as an infinite spiral (homeomorphic to $\mathbb{R}$) which covers the S^1 in the sense that when you view the spiral from the top, you will see only S^1.

We will use this covering map $\exp : \mathbb{R} \to S^1$ to find the first homotopy group of S^1. The main idea there will be *resolve* complicated loops in S^1 to $\mathbb{R}$, where each loop is homotopic to constant loop at the starting/ending point of the loop(!)

Remark 5.1.1.3. It is clear that every covering map is surjective.

The following is an important example of a covering map.

Lemma 5.1.1.4. *The map $\varphi : S^1 \to S^1$ given by $z \mapsto z^n$ is a covering map.*

Proof. Pick any $z_0 = e^{i\theta_0} \in S^1$. We wish to show that there exists an open set $U_0 \ni z_0$ in S^1 such that

$$\varphi^{-1}(U_0) = \coprod_{k=0}^{n-1} V_k$$

where V_k are open in S^1 and $\varphi|_{V_k} : V_k \to U_0$ is a homeomorphism.

Denote by $\gamma : \mathbb{R} \to S^1$ the continuous surjective map given by $t \mapsto e^{it}$. Thus, $z_0 = \gamma(\theta_0)$. Consider the interval $I_0 = \left(\theta_0 - \frac{\pi}{n}, \theta_0 + \frac{\pi}{n}\right)$. As the map $\gamma : \mathbb{R} \to S^1$ is an open map, therefore we have $U_0 = \gamma(I_0)$ which is an open set of S^1 containing z_0. We claim that U_0 is an evenly covered neighborhood for z_0. Indeed, we see that

$$\begin{aligned}
\varphi^{-1}(U_0) &= \{z \in S^1 \mid z^n \in U_0\} \\
&= \left\{e^{i\theta} \in S^1 \mid e^{ni\theta} \in \gamma(I_0)\right\} \\
&= \left\{e^{i\theta} \in S^1 \mid \exists \kappa \in I_0 \text{ s.t. } \gamma(\kappa) = e^{i\kappa} = e^{ni\theta}\right\} \\
&= \left\{e^{i\theta} \in S^1 \mid \exists \kappa \in I_0 \text{ s.t. } n\theta = \kappa + 2k\pi, \text{ for some } k \in \mathbb{Z}\right\} \\
&= \left\{e^{i\theta} \in S^1 \mid \exists \kappa \in I_0 \text{ s.t. } \theta = \frac{\kappa}{n} + \frac{2\pi k}{n}, \text{ for some } k \in \mathbb{Z}\right\} \\
&= \left\{e^{i\theta} \in S^1 \mid \theta \in \coprod_{k\in\mathbb{Z}} \left(\frac{\theta_0}{n} - \frac{\pi}{n^2} + \frac{2\pi k}{n}, \frac{\theta_0}{n} + \frac{\pi}{n^2} + \frac{2\pi k}{n}\right)\right\} \\
&= \coprod_{k=0}^{n-1} \gamma\left(\left(\frac{\theta_0}{n} - \frac{\pi}{n^2} + \frac{2\pi k}{n}, \frac{\theta_0}{n} + \frac{\pi}{n^2} + \frac{2\pi k}{n}\right)\right).
\end{aligned}$$

This completes the proof. □

We next discuss the notion of mapping torus of a map and how van Kampen can be used to compute its fundamental group.

Definition 5.1.1.5 (Mapping torus). For any map $f : X \to X$ the mapping torus of f is $T_f := X \times I/\sim$ where $(x, 0) \sim (f(x), 1)$.

Example 5.1.1.6. For $\mathrm{id} : X \to X$, one can check that $T_{\mathrm{id}} = X \times S^1$.

We have the following basic, but useful lemma.

Lemma 5.1.1.7. *Let $\pi : \tilde{X} \to X$ be a covering map. Then, for all $x \in X$ the fiber $\pi^{-1}(x) \subseteq \tilde{X}$ is a discrete subspace of $\tilde{X}$, that is, each $\tilde{x} \in \pi^{-1}(x)$ is both open and closed.*

Proof. To see this, take any $\tilde{x} \in \pi^{-1}(x)$ and an evenly covered neighborhood $U_x \subseteq X$ of x. Since $\pi^{-1}(U_x) = \coprod_{\alpha\in J_x} V_\alpha$, where each V_α is homeomorphic to U_x under $\pi|_{V_\alpha}$. Thus, the unique $\tilde{x}_\alpha \in V_\alpha$ such that $\pi(\tilde{x}_\alpha) = x$ is an element of $\pi^{-1}(x)$, one for each $\alpha \in J_x$. Now an open set of $\pi^{-1}(x)$ is of the form $V \cap \pi^{-1}(x)$ where $V \subseteq \tilde{V}$ is open, therefore $V_\alpha \cap \pi^{-1}(x)$ is open in $\pi^{-1}(x)$. But $V_\alpha \cap \pi^{-1}(x) = \{\tilde{x}_\alpha\}$ because each V_α are disjoint. Therefore $\{\tilde{x}_\alpha\}$ is open in $\pi^{-1}(x)$. Similarly, it is closed in $\pi^{-1}(x)$ by considering the complement of $\cup_{\beta\neq\alpha} V_\beta$ in $\pi^{-1}(x)$. Hence $\pi^{-1}(x)$ is a discrete subspace of $\tilde{X}$. □

5.1.2 Path lifting

Covering maps are important in algebraic topology because they come equipped with a lot of unique lifting properties. We will first spell out the unique path lifting property of covering spaces, which is a baby version of unique homotopy lifting property. Before that, we need some specific property of a path in space X which is covered by a covering space $\tilde{X}$.

Lemma 5.1.2.1. *Let $\gamma : I \to X$ be a path in X and $\pi : \tilde{X} \to X$ be a covering map. Then there exists a partition $0 = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = 1$ of unit interval I such that for all $i = 0, \dots, k-1$, the image $\gamma([t_i, t_{i+1}]) \subseteq X$ is contained in an evenly-covered neighborhood of X.*

Proof. So first, for all $t \in I$, there exists an evenly-covered neighborhood $U_t \subseteq X$ of $\gamma(t) \in X$. Thus, by continuity of γ, we get that there exists $(a_t, b_t) \subseteq I$ containing $t \in I$ such that $\gamma((a_t, b_t)) \subseteq U_t$. Since each open interval contains a compact interval, therefore we can assume (a_t, b_t) to be $[a_t, b_t]$. So we have a family of closed subintervals $\{[a_t, b_t]\}_{t\in I}$ of I. By compactness of I, we get that there exists a finite subcover $[a_{t_1}, b_{t_1}], \dots, [a_{t_n}, b_{t_n}]$ of I. Now suppose $[a_{t_i}, b_{t_i}]$ and $[a_{t_j}, b_{t_j}]$ intersect, then we can break down $[a_{t_i}, b_{t_i}] \cup [a_{t_j}, b_{t_j}]$ into three disjoint closed intervals $[a_{t_i}, a_{t_j}] \cup [a_{t_j}, b_{t_i}] \cup [b_{t_i}, b_{t_j}]$. Furthermore note that each of the above three have their images contained inside an evenly-covered neighborhood. Since there are only finitely many such intersections, therefore we have a finite disjoint cover of I by closed intervals, each of which has image under γ contained in an evenly covered neighborhood. $\square$

Theorem 5.1.2.2. *(Unique path lifting of covering maps) Let $\pi : \tilde{X} \to X$ be a covering map. Suppose there is a path $\gamma : I \to X$ and a prescribed point $\tilde{\gamma}_0 : \{0\} \to \tilde{X}$ such that $\pi(\tilde{\gamma}_0) = \gamma(0)$, then there exists a unique path $\tilde{\gamma} : I \to \tilde{X}$ such that $\pi \circ \tilde{\gamma} = \gamma$ and $\tilde{\gamma}(0) = \tilde{\gamma}_0$. That is, the following lifting problem is uniquely filled:*

$$\begin{tikzcd}
\{0\} \arrow[r, hook, "\tilde{\gamma}_0"] \arrow[d, hook] & \tilde{X} \arrow[d, "\pi"] \\
I \arrow[r, "\gamma"'] \arrow[ur, dashed, "\tilde{\gamma}"] & X
\end{tikzcd}.$$

Proof. Let us first construct such a path lift. By Lemma 5.1.2.1, we have a partition of I into $I = \cup_{i=0}^{k-1}[t_i, t_{i+1}]$ of disjoint closed intervals where $\gamma([t_i, t_{i+1}]) \subset U_i \subset X$ and U_i is evenly-covered in X. Now to construct the said $\tilde{\gamma}$, we will have to do it for each $[t_i, t_{t+1}]$, starting from $i = 0$, making use of $\tilde{\gamma}_0 \in \tilde{X}$ that has been already given to us. Now, let us first denote $\pi^{-1}(U_i) = \coprod_{\alpha\in J_i} V_\alpha^i$ for all $i = 0, \dots, k-1$ where $V_\alpha^i \cong U_i$, which is given by the fact that π is a covering map. Also keep in note that $\forall t \in [t_i, t_{i+1}]$, $\gamma(t) \in U_i \subseteq X$ which is evenly-covered.

So let us first define $\tilde{\gamma}$ for $[t_0, t_1] = [0, t_1]$. Since $\pi(\tilde{\gamma}_0) = \gamma(0) \in U_0$, therefore $\tilde{\gamma}_0 \in \pi^{-1}(U_0)$ and hence there is unique $\alpha_0 \in J_0$ such that $\tilde{\gamma}_0 \in V_{\alpha_0}^0$.

$$\begin{aligned}
\tilde{\gamma}|_{[t_0,t_1]} : [t_0, t_1] &\longrightarrow \tilde{X} \\
t &\longmapsto \left(\pi|_{V_{\alpha_0}^0}\right)^{-1}(\gamma(t)),
\end{aligned}$$

where $\pi|_{V^0_{\alpha_0}} : V^0_{\alpha_0} \to U_0$ is a homeomorphism and we are using it's inverse map in the above definition. Ok, so we first observe that $\tilde{\gamma}|_{[t_0,t_1]}(0) = \left(\pi|_{V^0_{\alpha_0}}\right)^{-1}(\gamma(0)) = \left(\pi|_{V^0_{\alpha_0}}\right)^{-1}(\pi(\tilde{\gamma}_0)) = \tilde{\gamma}_0$. That is, the starting point of path $\tilde{\gamma}$ is indeed $\tilde{\gamma}_0$. So we have constructed a path in $\tilde{X}$ from $\tilde{\gamma}_0$ to $\tilde{\gamma}|_{[t_0,t_1]}(t_1)$. Moreover, this path satsfies that $\pi \circ \tilde{\gamma}|_{[t_0,t_1]} = \gamma|_{[t_0,t_1]}$, which is exactly what we wanted.

Next, let us continue defining $\tilde{\gamma}$ for $[t_1, t_2]$ by using where we left off at $[t_0, t_1]$. This in turn will suggest us how to completely define the whole path $\tilde{\gamma}$. So we first note that $\gamma(t_1) \in U_0 \cap U_1$, therefore the end point of path $\tilde{\gamma}|_{[t_0,t_1]}$ at t_1, takes value in $\pi^{-1}(U_1)$ as well, so let $\tilde{\gamma}|_{[t_0,t_1]}(t_1) \in V^1_{\alpha_1}$. It should be clear by now what we are about to do; now define:

$$\begin{aligned} \tilde{\gamma}|_{[t_1,t_2]} : [t_1, t_2] &\longrightarrow \tilde{X} \\ t &\longmapsto \left(\pi|_{V^1_{\alpha_1}}\right)^{-1}(\gamma(t)). \end{aligned}$$

As usual, we again observe that $\tilde{\gamma}|_{[t_1,t_2]}(t_1) = \tilde{\gamma}|_{[t_0,t_1]}(t_1)$ because we have

$$\begin{aligned} \left(\pi|_{V^1_{\alpha_1}}\right)^{-1}(\gamma(t_1)) &= \left(\pi|_{V^1_{\alpha_1}}\right)^{-1}(\pi(\tilde{\gamma}|_{[t_0,t_1]}(t_1))) \\ &= \tilde{\gamma}|_{[t_0,t_1]}(t_1) \end{aligned}$$

where we conclude second line from first as $\gamma(t_1) \in U_0 \cap U_1$, where $\left(\pi|_{V^1_{\alpha_1}}\right)^{-1}$ is indeed defined. So we have indeed define a path $\tilde{\gamma}|_{[t_1,t_2]}$ whose starting point is same as the ending point of $\tilde{\gamma}|_{[t_0,t_1]}$, so we have defined the $\tilde{\gamma}$ upto $[t_0, t_2]$.

Having done the above, we now give general procedure of continuing the definition of path $\tilde{\gamma}$ till $[t_{k-1}, t_k]$. Suppose $2 \le j \le k-1$ and suppose we have constructed $\tilde{\gamma}|_{[t_{j-1},t_j]} : [t_{j-1}, t_j] \to \tilde{X}$ as of yet. So we know the point $\tilde{\gamma}|_{[t_{j-1},t_j]}(t_j) \in V^{j-1}_{\alpha_{j-1}}$ where $\gamma(t_j) \in U_{j-1} \cap U_j$. We now construct with this information the next piece of path $\tilde{\gamma}|_{[t_j,t_{j+1}]} : [t_j, t_{j+1}] \to \tilde{X}$. Well, the following definition shouldn't be a surprise:

$$\begin{aligned} \tilde{\gamma}|_{[t_j,t_{j+1}]} : [t_j, t_{j+1}] &\longrightarrow \tilde{X} \\ t &\longmapsto \left(\pi|_{V^j_{\alpha_j}}\right)^{-1}(\gamma(t)) \end{aligned}$$

where we again observe that the starting point of the above path is same as $\tilde{\gamma}|_{[t_{j-1},t_j]}(t_j)$. Moreover, it is easy to observe that $\pi \circ \tilde{\gamma}|_{[t_j,t_{j+1}]} = \gamma|_{[t_j,t_{j+1}]}$.

Finally, since there are only finitely many $[t_j, t_{j+1}]$s, therefore we have constructed a path $\tilde{\gamma}$ in $\tilde{X}$ such that it starts from $\tilde{\gamma}_0$ ($\tilde{\gamma}_0 = \tilde{\gamma}(0)$) and when projected back to X under π, we obtain the path γ back ($\pi \circ \tilde{\gamma} = \gamma$). In particular, the end point $\tilde{\gamma}(1) \in \pi^{-1}(\gamma(1))$. The uniqueness of $\tilde{\gamma}$ follows by construction. $\square$

A simple yet useful observation about higher homotopy groups of universal covers is the following.

Lemma 5.1.2.3. *Let (X, x_0) be a path-connected, locally path-connected and semi-locally simply connected space and denote $p : \tilde{X} \to X$ be its universal cover. Then,*

$$p_* : \pi_k(\tilde{X}) \to \pi_k(X)$$

is an isomorphism for all $k \geq 2$.

Proof. We have a homomorphism $p_* : \pi_k(\tilde{X}) \to \pi_k(X)$ for all $k \geq 2$. We shall show that this homomorphism has an inverse. Indeed, we have a map

$$\begin{aligned} \psi : \pi_k(X) &\longrightarrow \pi_k(\tilde{X}) \\ [\gamma] &\longmapsto [\tilde{\gamma}] \end{aligned}$$

where $\tilde{\gamma}$ is the unique lift of γ which exists as S^k and $\tilde{X}$ are simply connected for $k \geq 2$. It follows immediately that $p_* \circ \psi = \text{id}$ and by uniqueness of lifts that $\psi \circ p_* = \text{id}$. Hence p_* is a bijection, as required. □

5.1.3 Homotopy lifting

The Theorem 5.1.2.2 will be the building block for it's generalization, which is the homotopy lifting of covering maps. Let us first define what does it mean for a map to have homotopy lifting property.

Definition 5.1.3.1. (**Homotopy lifting property**) Let $p : E \to B$ be a continuous map. The map p is said to have homotopy lifting property if for any homotopy $H : Y \times I \to B$ and any map $\tilde{H}_0 : Y \times \{0\} \to E$ such that $p \circ \tilde{H}_0 = H(-, 0)$, there exists a homotopy $\tilde{H} : Y \times I \to E$ such that $\tilde{H}(-, 0) = \tilde{H}_0$ and $p \circ \tilde{H} = H$. That is, the following lifting problem is filled:

$$\begin{tikzcd} Y \times 0 \arrow[r, "\tilde{H}_0"] \arrow[d, hook] & E \arrow[d, "p"] \\ Y \times I \arrow[r, "H"'] \arrow[ur, dashed, "\tilde{H}"] & B \end{tikzcd} .$$

Remark 5.1.3.2. It is clear that path lifting property is obtained from homotopy lifting property by setting $Y = \{0\}$ in the diagram of homotopy lifting problem above.

We then have the following theorem.

Theorem 5.1.3.3. *(Unique homotopy lifting of covering maps) Let $\pi : \tilde{X} \to X$ be a covering map. Then π satisfies unique homotopy lifting property. That is, given any homotopy $H : Y \times I \to X$ and a map $\tilde{H}_0 : Y \to \tilde{X}$ such that $\pi \circ \tilde{H}_0 = H(-, 0)$, there exists a unique homotopy $\tilde{H} : Y \times I \to \tilde{X}$ such that $\tilde{H}(-, 0) = \tilde{H}_0$ and $\pi \circ \tilde{H} = H$. In other words, the following lifting problem is uniquely filled:*

$$\begin{tikzcd} Y \times \{0\} \arrow[r, "\tilde{H}_0"] \arrow[d, hook] & \tilde{X} \arrow[d, "\pi"] \\ Y \times I \arrow[r, "H"'] \arrow[ur, dashed, "\tilde{H}"] & X \end{tikzcd} .$$

Proof. **[TODO] Proof is quite long and detailed so I will do it when I will get time.**. □

5.1.4 $\pi_1(S^1) \cong \mathbb{Z}$

We now prove using the covering map $\exp : \mathbb{R} \to S^1$ that the first homotopy group of S^1 is $\mathbb{Z}$.

Theorem 5.1.4.1. $\pi_1(S^1) \cong \mathbb{Z}$.

Proof. Consider the following map which is quite intuitive to define:

$$\begin{aligned} \varphi : \mathbb{Z} &\longrightarrow \pi_1(S^1) \\ n &\longmapsto [\gamma_n] \end{aligned}$$

where $\gamma_n : I \to S^1$ is the loop $\theta \longmapsto e^{2\pi i n \theta}$, that is, γ_n is the loop corresponding to travelling around n-times on the circle S^1. Let us first show that it is indeed a group homomorphism. We see that

$$\begin{aligned} \varphi(n+m) &= [\gamma_{n+m}] \\ &= [\gamma_n * \gamma_m] \\ &= [\gamma_n] * [\gamma_m] \\ &= \varphi(n) * \varphi(m), \end{aligned}$$

so no qualms there.

The major hurdle starts when we try to prove the injectivity and surjectivity. This is where we will need to use the path and homotopy lifting properties of the covering map $\exp : \mathbb{R} \to S^1$ where we indeed verified that exp is a covering map in the example below the definition of covering spaces.

Let us first show surjectivity. So take any $[\gamma] \in \pi_1(S^1)$. We need to show that $\exists n \in \mathbb{Z}$ such that $[\gamma_n] = [\gamma]$. So we have that $\exp(\tilde{x}) = \gamma_n(0)$, which in diagrammatic form is

$$\begin{tikzcd} \{0\} \arrow[r, hook, "\tilde{x}"] \arrow[d, hook] & \mathbb{R} \arrow[d, "\exp"] \\ I \arrow[r, "\gamma"'] & S^1 \end{tikzcd} .$$

Since $\exp : \mathbb{R} \to S^1$ is a covering map, therefore using the unique path lifting property of covering maps (Theorem 5.1.2.2), we get that there is a unique $\tilde{\gamma} : I \to \mathbb{R}$ such that the above lifting problem is filled and then we get $\exp \circ \tilde{\gamma} = \gamma$ and $\tilde{\gamma}(0) = \tilde{x} \in (\exp)^{-1}(1)$. Now, we also have that $\tilde{\gamma}(1) \in (\exp)^{-1}(1)$. Therefore $\tilde{\gamma}(1) - \tilde{\gamma}(0) =$ total number of times the loop γ crosses $1 = n$, say. So $\tilde{\gamma}$ is homotopic to the straight line joining $\tilde{\gamma}(0)$ and $\tilde{\gamma}(1)$, that is $\kappa(t) = (1-t)\tilde{\gamma}(0) + t\tilde{\gamma}(1)$. Let this homotopy between κ and $\tilde{\gamma}$ be denoted by $H : I \times I \to \mathbb{R}$. Then $\exp \circ H$ is a homotopy between $\exp \circ \kappa$ and $\exp \circ \tilde{\gamma}$ where the former is the γ_n and the latter is γ. We thus have a homotopy between them and therefore $[\gamma] = [\gamma_n]$.

Let us next show injectivity. So suppose $\varphi(n) = [\gamma_n] = [c_1] = [\gamma_0]$ where $c_1 = \gamma_0 : I \to S^1$ is the constant loop at $1 \in S^1$. We need to show that this implies $n = 0$. We will use homotopy lifting to prove this, that is, we will lift the homotopy which makes γ_n homotopic to c_1 to a homotopy in $\mathbb{R}$ between the lift of γ_n to a constant path. More precisely, consider the homotopy

$$H : I \times I \longrightarrow S^1$$

establishing a homotopy between $H(-,0) = \gamma_n$ and $H(-,1) = \gamma_0$ and moreover $H(0,-) = H(1,-) = 1$. Also consider the map $\tilde{\gamma}_n : I \longrightarrow \mathbb{R}$ given by $t \longmapsto nt$. This is the other map which the lifted homotopy will give a homotopy from to some other map (which we have to figure out). We then observe that $\tilde{\gamma}_n$ is the right map to define here because $\exp \circ \tilde{\gamma}_n(s) = e^{2\pi i n s} = \gamma_n(s) = H(s,0)$. Ok so now we lift. Using Theorem 5.1.3.3, the following lifting problem is uniquely solved:

$$\begin{tikzcd} I \times \{0\} \arrow[r, "\tilde{\gamma}_n"] \arrow[d, hook] & \mathbb{R} \arrow[d, "\exp"] \\ I \times I \arrow[r, "H"'] \arrow[ur, dashed, "\tilde{H}"] & S^1 \end{tikzcd}.$$

So we have a homotopy $\tilde{H} : I \times I \longrightarrow \mathbb{R}$ such that $\tilde{H}(s,0) = \tilde{\gamma}_n(s)$ and, more importantly, $\exp \circ \tilde{H} = H$. Thus, $\exp(\tilde{H}(s,1)) = H(s,1) = 1$, that is, $\operatorname{Im}\left(\tilde{H}(-,1)\right) \subseteq (\exp)^{-1}(1)$. Since fibres of a covering map are necessarily discrete (Lemma 5.1.1.7) and $\tilde{H}(-,1)$ is a continuous map from a connected set I, so it's image has to be connected as well and hence $\operatorname{Im}\left(\tilde{H}(-,1)\right)$ has to be a point inside $(\exp)^{-1}(1)$. What this means is that $\tilde{H}(-,1)$ is a constant map, to a point in $\mathbb{R}$, which we denote as $a \in \mathbb{R}$ such that $\exp(a) = 1$. So $\tilde{H}$ is a homotopy between $\tilde{\gamma}_n$ and c_a (the constant path at a). Moreover, we also have that $\tilde{H}(0,t) = \tilde{H}(1,t)$ for all $t \in I$ because $\tilde{H}$ is a based homotopy. So we get that the map $\tilde{H}(1,t) = \tilde{H}(0,t) = a \in (\exp)^{-1}(1)$ for all $t \in I$ as it is a for $t = 1$. So this **forces** $H(\tilde{s,0}) = \tilde{\gamma}_n(s)$ to have starting point and ending point same, equal to a. But this can only happen when $n = 0$ (see definition of $\tilde{\gamma}_n$). We are done. □

5.1.5 Couple of properties of covering spaces

Covering maps are quite nice maps as is shown by Theorem 5.1.3.3. We will consider a couple of important properties that covering spaces hold in this section. The first one being that all fibers of a covering map of a path-connected space (which is discrete, Lemma 5.1.1.7) are bijective (so have same *size*).

Lemma 5.1.5.1. *Let $\pi : \tilde{X} \to X$ be a covering map and let X be a path-connected space*[2]*. Let $x_0, x_1 \in X$ be two points, then there is a set bijection*

$$\pi^{-1}(x_0) \cong \pi^{-1}(x_1).$$

Proof. **[TODO]**. □

[2] or work over path-components.

Another use of covering spaces is that if $\pi : \tilde{X} \to X$ is a covering map where both the spaces are path-connected, then the fundamental group of $\tilde{X}$ is naturally embedded inside the fundamental group of X.

Proposition 5.1.5.2. *Let $\pi : \tilde{X} \to X$ be a covering map where both X and $\tilde{X}$ are path-connected. Then the map*

$$\pi_1(\pi) : \pi_1(\tilde{X}, \tilde{x}_0) \longrightarrow \pi_1(X, \pi(\tilde{x}_0))$$

is injective.

5.1.6 Fun applications of $\pi_1(S^1) \cong \mathbb{Z}$

We first have the famous Brouwer's fixed point theorem.

Proposition 5.1.6.1. *(Brouwer's fixed point theorem) For any continuous $f : D^2 \to D^2$, there exists a point $x \in D^2$ such that $f(x) = x$.*

Next is something we know very well but didn't knew that it can be done from the methods we have developed till now:

Proposition 5.1.6.2. *(Fundamental theorem of algebra) Let $p(x) \in \mathbb{C}[x]$. Then there exists a $c \in \mathbb{C}$ such that $x - c$ divides $p(x)$. That is, every complex polynomial has a root in $\mathbb{C}$ (and thus have all roots in $\mathbb{C}$).*

The last one is something we saw in the departmental seminar a week ago, using which we saw that one can prove very non-trivial combinatorial results.

Proposition 5.1.6.3. *(Borsuk-Ulam theorem) If $f : S^2 \to \mathbb{R}^2$ is a continuous map, then there exists a pair of anti-podal points which are mapped to same point under f.*

5.1.7 Covering spaces, group actions and Galois theory of covers

So in this second phase of the course, we will be seeing some more fancy theorems, but the main goal will be to go to some calculative things, like computing homology groups and all that. In any case, we covered covering spaces, but it would be rather incomplete if we don't say something about universal covering and more theorems in that direction. The first theorem we therefore discuss, tells us how a certain type of G-space naturally enriches the quotient map with the structure of a covering space. We first define the type of G-space we wish to look out for.

Definition 5.1.7.1. (**Properly discontinuous action**) Let G be a group and X be a space with a continuous action[3] of G. The action of G is said to be properly discontinuous if for all $x \in X$, there exists an open set $U_x \subseteq X$ containing x such that $gU_x \cap U_x = \emptyset$ for all $g \in G$.

There is another type of action:

Definition 5.1.7.2. (**Free action**) Let G be a group acting continuously on space X. The the action is said to be free if for all $x \in X$, the stabilizer subgroup is trivial, that is, $S_G(x) = \{e\}$.

[3] this means that the action map $G \times X \to X$ is a continuous map where G is given the discrete topology.

There are some consequences of the above definition which we collectively state in the the following lemma:

Lemma 5.1.7.3. *Let G be a group and X be a space with continuous G-action.*

1. *If the action is properly discontinuous, then it is free.*
2. *If G is finite and X is locally finite*[4]*, then the action is free if and only if it is properly discontinuous.*

Proof. 1. Take any $x \in X$. Let U_x be the open set containing x obtained from properly discontinuous action of G. If $g \in S_G(x)$, then $gU_x \cap U_x \neq \emptyset$. Thus $g = e$.

2. R $\Rightarrow$ L is simple. For L $\Rightarrow$ R, we go by contradiction. So suppose the action is free but not properly discontinuous. Take any point $x \in X$. So for any open $U \ni x$ and for any $g \in G$, $gU \cap U \neq \emptyset$. Now, we have a sequence of open sets each containing x, U_n, such that $\cap_n U_n = \{x\}$. Since $gU_n \cap U_n \neq \emptyset$ for each n, therefore we get a sequence $\{x_n\}$ where $x_n \in U_n$ such that $\varprojlim x_n = x$ and $\varprojlim gx_n = x$. Since $g \in G$ can be treated as $g : X \to X$ a homeomorphism, therefore $g(\varprojlim x_n) = x$ that is $gx = x$, a contradiction to the fact that G acts freely[5]. $\square$

Let us now state the theorem of interest.

Theorem 5.1.7.4. *Let G be a group and X be a space with continuous G-action. If the action is properly discontinuous, then the quotient map*

$$q : X \twoheadrightarrow X/G$$

is a covering map.

Before stating the proof, we would like to give some example uses of this theorem.

Example 5.1.7.5. Consider $G = \mathbb{Z}^n$ and $X = \mathbb{R}^n$. There is a canonical action we can define on $\mathbb{R}^n$ using $\mathbb{Z}^n$ given by

$$\begin{aligned} G \times X &\longrightarrow X \\ ((m_1, \dots, m_n), (x_1, \dots, x_n)) &\longmapsto (m_1 + x_1, \dots, m_n + x_n). \end{aligned}$$

The fact that this is a continuous action is trivial to check. We first claim that this action is properly discontinuous. It is simple to see why that's the case; for an $x \in X$ simply take any $0 < a < 1/2$ and define $U = \prod (x_i - a, x_i + a)$. This U is open and for any $m := (m_1, \dots, m_n) \in \mathbb{Z}^n$, $(m + U) \cap U = \emptyset$ for any $m \neq 0$. So indeed the action is properly discontinuous.

Next, we observe that $X/G = \mathbb{R}^n/\mathbb{Z}^n$ is simply homeomorphic to $[0,1]^n/G$ and which is in turn homeomorphic to $([0,1]/0 \sim 1)^n$ and which is just $(S^1)^n$. So that is why the questions regarding $\mathbb{R}/\mathbb{Z}$ are so innumerable in literature, as they quickly form spaces which are quite weird to imagine.

[4]this means that for all $x \in X$, there exists a sequence of open sets U_n containing x such that $\bigcap_n U_n = \{x\}$.
[5]this is in-line with what the wonderful man *I.P. Freely* had to say.

Example 5.1.7.6. (*Configuration space of k-points in space X*) Let X be a space. The configuration space of k points in X, denoted $F_k(X)$, is intuitively the set of all possible positions that k particles moving in X can inhabit. More precisely, we define:

$$F_k(X) = \{(x_1, \dots, x_k) \in \prod_{i=1}^{k} X \mid \forall\, i \neq j = 1, \dots, k,\ x_i \neq x_j\}.$$

This space has an action of S_k, the symmetry group of k letters, given by:

$$\begin{aligned} S_k \times F_k(X) &\longrightarrow F_k(X) \\ (\sigma, x_1, \dots, x_k) &\longmapsto (x_{\sigma(1)}, \dots, x_{\sigma(k)}). \end{aligned}$$

In other words, we just permute the k points which we find in some position in X. For $k = 2$, we get that since $S_2 = \mathbb{Z}_2$, so the only action possible is

$$\begin{aligned} \mathbb{Z}_2 \times F_2(X) &\longrightarrow F_2(X) \\ (0, x_1, x_2) &\longmapsto (x_1, x_2) \\ (1, x_1, x_2) &\longmapsto (x_2, x_1). \end{aligned}$$

In other words, we swap the two points. Then, orbits of the action of $\mathbb{Z}_2$ over $F_2(X)$ will consist of just the point itself and it's swapped counterpart. Hence,

$$F_2(X)/\mathbb{Z}_2 \cong (X \times X)/\sim$$

where $(x_1, x_2) \sim (y_1, y_2)$ iff $x_1 = y_2$ and $x_2 = y_1$. To better understand the situation, suppose $X = S^1$. Then, $F_2(S^1) = S^1 \times S^1/\sim$. Since $S^1 \times S^2 = T^2$, therefore we get $F^2(S^1) = T^2 \setminus \Delta(S^1)$, where $\Delta(S^1)$ is the diagonal subspace of $S^1 \times S^2$. But $T^2 \setminus \Delta(S^1)$ will look like quotient of $I \times I \setminus \Delta(I)$ which looks like two disjoint right triangles together. Now, we can obtain $F_2(S^1)/\sim$ by identifying the two triangles and doing the ensuing identifications of $I \times I$ to reach some weird object.

Example 5.1.7.7. The next example that we do is known for it's weirdness. It is the construction of *lens space*. Consider the odd sphere $S^{2k+1} \subset \mathbb{C}^{k+1}$ for $k \in \mathbb{N}$. Consider the cyclic group $\mathbb{Z}_d$ where we take the following presentation of it: $\mathbb{Z}_d = \langle \xi \rangle$ where ξ is the d^{th} root of unity. We then have the following action of $\mathbb{Z}_d$ on S^{2k+1}:

$$\begin{aligned} \mathbb{Z}_d \times S^{2k+1} &\longrightarrow S^{2k+1} \\ (\xi, z_1, \dots, z_k) &\longmapsto (\xi z_1, \dots, \xi z_k). \end{aligned}$$

This is indeed a valid action. In particular, we claim that this is a free action so that by Lemma 5.1.7.3, 2, this action becomes properly discontinuous and we can then use Theorem 5.1.7.4 to get that S^{2k+1} is a cover of this so-called lens space. To see that it is free, take any $(z_1, \dots, z_k) \in S^{2k+1}$. We see that if $(\xi^n z_1, \dots, \xi^n z_k) = (z_1, \dots, z_k)$, then $\xi^n = 1$. So each stabilizer subgroup is trivial. Hence the action is free. Then, the lens space is defined to be the quotient $S^{2k+1}/\mathbb{Z}_d$. Whatever that may look like, it has a structure of a $2k + 1$ dimension manifold, as we have a cover by Theorem 5.1.7.4.

With all these examples out of the way, let us now get to the proof of the theorem at hand.

Proof of Theorem 5.1.7.4. Since the action of G is properly discontinuous, therefore for each $x \in X$, there exists open $U_x \subseteq X$ such that $gU_x \cap U_x = \emptyset$ for all $g \in G$. We claim that for any $[x] \in X/G$, the set $V_x := q(U_x)$ is evenly covered open neighborhood of $[x]$. In order to show this, we first claim the following

$$q^{-1}(V_x) = \coprod_{g \in G} gU_x.$$

Now, since $g : X \to X$ is a homeomorphism, thus $gU_x = g(U_x) \subseteq X$ is open in X. Hence, $q^{-1}(V_x)$ is open in X, if the above claim is true. So in order to see the claim, we see that

$$\begin{aligned} q^{-1}(V_x) &= \{y \in X \mid q(y) \in V_x = q(U_x)\} \\ &= \{y \in X \mid \exists z \in U_x \text{ s.t. } q(y) = q(z)\} \\ &= \{y \in X \mid \exists z \in U_x \text{ s.t. } y = gz \text{ for some } g \in G\} \\ &= \bigcup_{g \in G} gU_x. \end{aligned}$$

So we need only show that $gU_x \cap hU_x =$. This is simple because if it is not the case, then for some $y, z \in U_x$, we get $gy = hz$, so $y = g^{-1}hz$, a contradiction to $U_x \cap g^{-1}hU_x = \emptyset$ by properly discontinuous action of G on X. So indeed the claim is true.

We need only show now that for any $g \in G$, the restriction

$$q|_{gU_x} : gU_x \longrightarrow V_x$$

is a homeomorphism. Firstly, it is rather easy to see that $q(gU_x) = q(U_x)$, after all, q kills all orbits so that $q(gy) = q(y)$. Next, since $q(U_x) =: V_x$, so the above map is well defined. We now only need to show that it is a homeomorphism. For that, we can consider the following inverse:

$$\begin{aligned} w : V_x := q(U_x) &\longrightarrow gU_x \\ q(y) &\longmapsto gy. \end{aligned}$$

This is indeed well-defined. To see this, take any $z \in U_x$ such that $q(y) = q(z)$. Thus there is an $h \in G$ such that $y = hz$. Since $y, z \in U_x$ and U_x is such that $kU_x \cap U_x = \emptyset \ \forall k \in G$, thus, if $q(y) = q(z)$, then $y = z$, hence $gy = gz$. It is now easy to see that w is a continuous inverse of $q|_{gU_x}$, as $gy \mapsto q(gy) \mapsto w(q(gy)) = gy$ and conversely $q(y) \mapsto gy \mapsto q(gy) = y$. This completes the proof. □

5.1.8 Category of covering maps

Let (X, x_0) be a based space. It is easy to see that knowing information about all covers of (X, x_0), would be pretty handy. But how can one do that? Well, we will try to do exactly that in this section. Since we want to handle all covers of X, so it is better we start giving this collection of all covers of (X, x_0) some structure. One structure that it has is that it forms a category.

Definition 5.1.8.1. (**The category Cov** (X, x_0)) Let (X, x_0) be a based map. The category of covering maps of (X, x_0) and homomorphisms between them is defined by:

1. **Objects**: An object of $\mathbf{Cov}(X, x_0)$ is a covering map $p : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$.
2. **Arrows**: An arrow in $\mathbf{Cov}(X, x_0)$ is a continuous based map $f : (\tilde{X}_1, \tilde{x}_1) \to (\tilde{X}_2, \tilde{x}_2)$ such that the following commutes:

$$\begin{tikzcd} (\tilde{X}_1, \tilde{x}_1) \arrow[rr, "f"] \arrow[dr, "p_1"'] & & (\tilde{X}_2, \tilde{x}_2) \arrow[dl, "p_2"] \\ & (X, x_0) & \end{tikzcd}.$$

It is clear that $\mathbf{Cov}(X, x_0)$ is a sub-category of the category $\mathbf{Top}_*$ over (X, x_0), that is, $\mathbf{Cov}(X, x_0) \subseteq \mathbf{Top}_*/(X, x_0)$.

We will see in this and the following sections that the main ingredient of our goal to understand a covering space will be, just like in Galois theory, the automorphism group of $(\tilde{X}, \tilde{x}_0)$ in the category $\mathbf{Cov}(X, x_0)$. We denote the set of all **automorphisms** of $(\tilde{X}, \tilde{x}_0)$ by $\mathsf{Deck}(\tilde{X}, x_0)$. Note that in the unbased setting, we will denote the automorphism group of $\tilde{X} \in \mathbf{Cov}(X, x_0)$ as just $\mathsf{Deck}(X)$.

From now, we will abbreviate a based space (X, x_0) by just X. Similarly for the covering spaces.

For our purposes, we see the following result.

Proposition 5.1.8.2. *Let X be a path connected and locally path connected based space and consider $(\tilde{X}_1, p_1)$ and $(\tilde{X}_2, p_2)$ to be two path-connected covers in $\mathbf{Cov}(X)$. Let $\varphi : (\tilde{X}_1, p_1) \to (\tilde{X}_2, p_2)$ be a map of covering spaces. Then, φ is a covering map over $(\tilde{X}_2, p_2)$.*

Proof. We break the proof into following steps.

Act 1 : *The map φ is surjective.*

Take any point $y \in \tilde{X}_2$. Since $\tilde{X}_2$ is path connected, so there is a path $\eta : I \to \tilde{X}_2$ with $\eta(0) = \tilde{x}_2$ and $\eta(1) = y$. Then we have $z := p_2(y) \in X$. Since X is path-connected, we thus have a path $\gamma : I \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = z$. By Theorem 5.1.2.2 on $\tilde{X}_2$, it can be easily seen that η is the unique lift of γ. Now, by Theorem 5.1.2.2 for covering space $\tilde{X}_1$ with starting point $\tilde{x}_1$, we get a path $\tilde{\gamma}_1 : I \to \tilde{X}_1$ such that $\tilde{\gamma}_1(0) = \tilde{x}_1$ and $p_1 \circ \tilde{\gamma}_1 = \gamma$. Moreover, it is unique w.r.t. these properties. Now denote $x := \tilde{\gamma}_1(1) \in \tilde{X}_1$. Now, we have another path $\tilde{\gamma}_2 := \varphi \circ \tilde{\gamma}_1 : I \to \tilde{X}_2$ such that $\tilde{\gamma}_2(0) = \tilde{x}_2$. Moreover, by the fact that $p_2 \circ \varphi = p_1$, we get that $p_2 \circ \tilde{\gamma}_2 = \gamma$. So if we apply Theorem 5.1.2.2 on $\tilde{X}_2$, then the path that we must get should exactly be $\tilde{\gamma}_2$ because it satisfies the conditions that makes the path coming from the theorem unique. But then, $\eta = \tilde{\gamma}_2$. Hence $\tilde{\gamma}_2(1) = \eta(1) = y$. Hence $\varphi(x) = y$. This completes Act 1.

Act 2 : *Each point of $\tilde{X}_2$ has an evenly covered neighborhood.*

Take any point $y \in \tilde{X}_2$. To get an evenly covered neighborhood of y, we begin with $z := p_2(y) \in X$. Since both $\tilde{X}_1, \tilde{X}_2$ are covering X, therefore there are evenly covered neighborhoods $U_1, U_2 \subseteq X$ containing z. Then $V := U_1 \cap U_2$ is an open set which is an evenly covered neighborhood for both the covers. Now, $(p_2)^{-1}(V) \ni y$. Since $(p_2)^{-1}(V) = \coprod_{i \in J_z} V_i$. Let $y \in V_{i_y}$. We claim that this

V_{i_y} will be an evenly covered neighborhood of $y \in \tilde{X}_2$ for φ. Clearly, $(\varphi)^{-1}(V_{i_y}) \cong (p)^{-1}(V) \cong \coprod_{i \in I_z} W_i$ where $p_1|_{W_i} : W_i \to V$ which is a homeomorphism. This concludes Act 2.

This concludes the proof. $\square$

We now define universal covering space of a based space.

Definition 5.1.8.3. (**Universal covering**) Let (X, x_0) be a path-connected and locally path-connected space. A simply connected covering space $(\tilde{X}, \tilde{x}_0)$ is called a universal covering space of (X, x_0).

The justification of the name will come soon, but for the time being, let us develop some more theory of covering spaces, which we would need in order to prove Theorem **??**, which classifies coverings of a space upto isomorphism!

More properties of covering spaces & classification

Let us discuss few more properties of morphisms of covering spaces. It is good to remind ourselves that a space is path-connected and locally path-connected if and only if it is connected and locally path-connected.

Remark 5.1.8.4. It is clear by the definition of covering maps that if X is a locally path-connected space, then any covering space $\tilde{X}$ is also a locally path-connected space. But it is in general not true that if X is connected then $\tilde{X}$ is connected, a simple example is the trivial covering $X \amalg X \to X$. In conclusion, if X is connected and locally path-connected, then $\tilde{X}$ may not be connected but is locally path-connected.

The following lemma shows that to check equality of two maps in $\mathbf{Cov}(X)$ of connected covering spaces, we may check only at one point(!)

Lemma 5.1.8.5. *Let X be a path-connected and locally path-connected space. If $\varphi_0, \varphi_1 : (\tilde{X}_1, p_1) \rightrightarrows (\tilde{X}_2, p_2)$ are two maps of covering spaces in $\mathbf{Cov}(X)$ between connected covers $\tilde{X}_1$ and $\tilde{X}_2$, such that there exists a point $x_1 \in \tilde{X}_1$ for which $\varphi_0(x_1) = \varphi_1(x_1)$, then $\varphi_0 = \varphi_1$.*

Proof. Let $x \in \tilde{X}_1$. We wish to show that $\varphi_0(x) = \varphi_1(x)$. For this, we first denote $z := p_1(x) = p_2 \circ \varphi_0(x) = p_2 \circ \varphi_1(x)$. Hence it is clear that $y_0 := \varphi_0(x), y_1 := \varphi_1(x) \in (p_2)^{-1}(z)$, i.e. $y_0, y_1 \in \tilde{X}_2$ are in the same fiber. We now need to show that the points $y_0, y_1 \in p^{-1}(z)$ are literally the same. Suppose to the contrary that $y_0 \neq y_1$. Let $z \in U \subseteq X$ be an evenly covered neighborhood of z. Now, $(p_2)^{-1}(U) = \coprod_{i \in J} V_i$ where $p_2|_{V_i} : V_i \to U$ is an homeomorphism. Since $y_0 \neq y_1$, therefore, say $y_0 \in V_0$ and $y_1 \in V_1$ where V_0 and V_1 are disjoint in $\tilde{X}_2$. Since φ_0 and φ_1 are continuous, therefore there are open sets $W_0, W_1 \subseteq \tilde{X}_1$ containing x such that $\varphi_0(W_0) \subseteq V_0$ and $\varphi_1(W_1) \subseteq V_1$. Now, denote $W = W_0 \cap W_1$, so we have $\varphi_0(W) \subseteq V_0$ and $\varphi_1(W) \subseteq V_1$. So for each $x \in \tilde{X}_1$, we have an open set $x \in W_x \subseteq \tilde{X}_1$ such that $\varphi_0(W_x) \cap \varphi_1(W_x) = \emptyset$. This contradicts the fact that $x_1 \in \tilde{X}_1$ is not such a point. $\square$

Remark 5.1.8.6. Hence, for any $\varphi \in \mathsf{Deck}(\tilde{X})$ where $\tilde{X}$ is connected, φ doesn't have any fixed points.

The next result is an important one for our purposes, for it generalizes the unique path lifting property of covering maps to that of any path-connected and locally path-connected space, by comparing it's fundamental group.

Theorem 5.1.8.7 (*Unique lifting property*)**.** *Let (X, x_0) be a path-connected and locally path-connected space and let $p : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ be a covering map. Let (Y, y_0) be a path-connected and locally path-connected space. If $\varphi : (Y, y_0) \to (X, x_0)$ is a based map, then there exists a unique lift $\tilde{\varphi} : (Y, y_0) \to (\tilde{X}, \tilde{x}_0)$ if and only if $\varphi_*(\pi_1(Y, y_0)) \le p_*(\pi_1(\tilde{X}, \tilde{x}_0))$.*

More diagrammatically, the following lifting problem is uniquely solved if and only if $\varphi_*(\pi_1(Y, y_0)) \le p_*(\pi_1(\tilde{X}, \tilde{x}_0))$:

$$\begin{tikzcd} & (\tilde{X}, \tilde{x}_0) \arrow[d, "p"] \\ (Y, y_0) \arrow[r, "\varphi"'] \arrow[ur, dashed, "\exists!\, \tilde{\varphi}"] & (X, x_0) \end{tikzcd} .$$

Proof. (L $\Rightarrow$ R) Since $p \circ \tilde{\varphi} = \varphi$, therefore $\varphi_*(\pi_1(Y, y_0)) = (p \circ \tilde{\varphi})_*(\pi_1(Y, y_0)) = p_*(\tilde{\varphi}_*(\pi_1(Y, y_0))) \le p_*(\pi_1(\tilde{X}, \tilde{x}_0))$.

(R $\Rightarrow$ L) We define the following candidate for the lift: for each point $y \in Y$, we join it to y_0 using $\gamma_y : I \to Y$ where $\gamma_y(0) = y_0$ and $\gamma_y(1) = y$, and then lift (Theorem 5.1.2.2) $\varphi \circ \gamma_y$ to a path $\tilde{\gamma}_y$ in $\tilde{X}$ from $\tilde{x}_0$ to $\tilde{\gamma}_y(1) \in p^{-1}(\varphi(y))$. This process gives the following map

$$\begin{aligned} \tilde{\varphi} : Y &\longrightarrow \tilde{X} \\ y &\longmapsto \tilde{\gamma}_y(1). \end{aligned}$$

We complete the rest of the proof in the following acts.

Act 1 : *The map $\tilde{\varphi}$ is well-defined.*

The plan is to use both homotopy an path liftings for this. So what we need to show is that for any other choice $\eta : I \to Y$ with $\eta(0) = y_0$ and $\eta(1) = y$, we get that $\tilde{\eta}_y(1) = \tilde{\gamma}_y(1)$. In order to do this, we first note that we get a loop $\gamma_y * \bar{\eta}_y$ at y_0 in Y, so that we have an element $[\gamma_y * \bar{\eta}_y] \in \pi_1(Y, y_0)$. Now, $\varphi_*([\gamma_y * \bar{\eta}_y]) = [\varphi \circ \gamma_y * \varphi \circ \bar{\eta}_y]$. Now since $\varphi_*(\pi_1(Y, y_0)) \le p_*(\pi_1(\tilde{X}, \tilde{x}_0))$, therefore there exists a loop $[\xi] \in \pi_1(\tilde{X}, \tilde{x}_0)$ such that $[p \circ \xi] = [\varphi \circ \gamma_y * \varphi \circ \bar{\eta}_y]$. Let us denote $[\varphi \circ \gamma_y * \varphi \circ \bar{\eta}_y] =: [\chi]$. So we have $p \circ \xi \simeq \chi$. Now, by Theorem 5.1.3.3, we get that ξ is homotopic to a loop at $\tilde{x}_0$, denoted τ such that $p \circ \tau = \chi$. Now note that $\tilde{\gamma}_y$ joins $\tilde{x}_0$ to a point, say $\omega \in \tilde{X}$ such that $p(\omega) = \varphi(y)$. Since we have a path $\varphi \circ \bar{\eta}_y$ which connects $\varphi(y)$ to x_0 in X, therefore if we lift (Theorem 5.1.2.2) $\varphi \circ \bar{\eta}_y$ to a path $\tilde{\bar{\eta}}_y$ beginning from ω and ending to a point in $p^{-1}(x_0)$ in $\tilde{X}$, we get that we get a unique path $\tilde{\gamma}_y * \tilde{\bar{\eta}}_y$ from $\tilde{x}_0$ to a point in $p^{-1}(x_0)$ in $\tilde{X}$ which is unique w.r.t the property that $p \circ (\tilde{\gamma}_y * \tilde{\bar{\eta}}_y) = \chi$. But, τ is also a path beginning from $\tilde{x}_0$ such that $p \circ \tau = \chi$, hence $\tilde{\gamma}_y * \tilde{\bar{\eta}}_y = \tau$, and thus the lift of $\bar{\eta}_y$ in $\tilde{X}$ starts at ω and ends at $\tilde{x}_0$. So now if we lift η_y in $\tilde{X}$, we get the path $\bar{\tilde{\bar{\eta}}}_y$ because of uniqueness of path lifts. Hence $\tilde{\eta}_y$ is a path from $\tilde{x}_0$ to $\omega =: \tilde{\gamma}_y(1)$. Hence well-definedness of $\tilde{\varphi}$ follows.

Act 2 : *The map $\tilde{\varphi}$ is continuous.*

It is at this point that we will use the hypotheses imposed on Y. We will show that $\tilde{\varphi}$ is locally a continuous map. Take any point $y \in Y$ and let $\varphi(y) \in X$. There is an evenly covered neighborhood of $\varphi(y)$, which we denote by $U \ni \varphi(y)$ so that $p^{-1}(U) = \coprod_{i \in I} V_i$. Denote $\tilde{\varphi}(y) \in V_0$. We also have an open set $\varphi^{-1}(U)$ of Y. Since Y is locally path-connected, let $W \subseteq \varphi^{-1}(U)$ be a path-connected subset of Y containing y. We now claim that $\tilde{\varphi}|_W = \left(p|_{V_0}\right)^{-1} \circ \varphi|_W$. For this, take any point

$z \in W$, and since W is path-connected, therefore there exists ξ joining $y \to z$. Since γ_y already joins $y_0 \to y$, therefore we have that $\gamma_y * \xi$ joins $y_0 \to z$. By Act 1, we get

$$\begin{aligned}\tilde{\varphi}(z) &= (\varphi \circ \widetilde{(\gamma_y * \xi)})(1) \\ &= (\varphi \tilde{\circ} \gamma_y) * (\varphi \tilde{\circ} \xi)(1).\end{aligned}$$

Now, since $p|_{V_0}$ is a homeomorphism of V_0 to U and since $\varphi(y), \varphi(z) \in U$ are connected by a path $\varphi \circ \xi$, so V_0 also has a path connecting $\tilde{\varphi}(y)$ and $\tilde{\varphi}(z)$. Hence, by uniqueness of path lifts (Theorem 5.1.2.2), we get $(\varphi \tilde{\circ} \gamma_y) * (\varphi \tilde{\circ} \xi)(1) = \left(p|_{V_0}\right)^{-1} (\varphi(z))$. We are now gladly done.

Act 3 : *The map $\tilde{\varphi}$ is unique.*

Essentially by construction. If the reader is not convinced, just start doing the brute force verification and you will see why that's the case.

This proof is now complete. $\square$

This theorem is an extremely important result as it will allow us to classify all connected covers of a connected and path-connected space upto isomorphism, as we will soon see. We will in the following few results see the beginnings of the Galois theory of covering spaces.

Lemma 5.1.8.8. *Let (X, x_0) be a path-connected and locally path-connected space and consider* **Cov** (X, x_0). *If $(\tilde{X^H}_1, \tilde{x}_1, p_1)$ and $(\tilde{X^H}_2, \tilde{x}_2, p_2)$ are two connected covering spaces over (X, x_0) such that*

$$p_{1*}(\pi_1(\tilde{X^H}_1, \tilde{x}_1)) = p_{2*}(\pi_1(\tilde{X^H}_2, \tilde{x}_2)) = H \leq \pi_1(X, x_0),$$

then there exists a unique homeomorphism $\varphi : (\tilde{X^H}_1, \tilde{x}_1, p_1) \to (\tilde{X^H}_2, \tilde{x}_2, p_2)$, that is, $(\tilde{X^H}_1, \tilde{x}_1, p_1)$ and $(\tilde{X^H}_2, \tilde{x}_2, p_2)$ are equivalent. In diagrammatic terms,

$$\begin{tikzcd} (\tilde{X^H}_1, \tilde{x}_1) \arrow[rr, dashed, "\exists!\, \varphi", "\cong"'] \arrow[dr, "p_1"'] & & (\tilde{X^H}_2, \tilde{x}_2) \arrow[dl, "p_2"] \\ & (X, x_0) & \end{tikzcd}.$$

Proof. We will use Theorem 5.1.8.7 for this purpose. By the said theorem, where, in the notation of the theorem, we let $Y = \tilde{X^H}_1$ and $\varphi = p_1$, we get that there is a unique map $\varphi : \tilde{X^H}_1 \to \tilde{X^H}_1$ such that $p_2 \circ \varphi = p_1$. This follows because the condition of the theorem is trivially satisfied. We now need only show that it has an inverse. This is also easy because of the equality of the image subgroups; since $H = p_{2*}(\pi_1(\tilde{X^H}_2, \tilde{x}_2)) \subseteq p_{1*}(\pi_1(\tilde{X^H}_1, \tilde{x}_1)) = H$, therefore another application of Theorem 5.1.8.7 yields a unique map $\varpi : (\tilde{X^H}_2, \tilde{x}_2) \to (\tilde{X^H}_1, \tilde{x}_1)$ such that $p_1 \circ \varpi = p_2$. To show that φ and ϖ are inverses of each other, consider the composite $\varphi \circ \varpi : (\tilde{X^H}_2, \tilde{x}_2) \to (\tilde{X^H}_2, \tilde{x}_2)$. Since $\varphi \circ \varpi$ is a unique map w.r.t. the property that $p_2 \circ (\varphi \circ \varpi) = (p_2 \circ \varphi) \circ \varpi = p_1 \circ \varpi = p_2$, but since so is $\mathrm{id}(\tilde{X^H}_2, \tilde{x}_2)$, therefore $\varphi \circ \varpi = \mathrm{id}(\tilde{X^H}_2, \tilde{x}_2)$. Similarly, $\varpi \circ \varphi = \mathrm{id}(\tilde{X^H}_1, \tilde{x}_1)$. This completes the proof. $\square$

Remark 5.1.8.9. Let $\tilde{X}$ be a connected cover of a p.c., l.p.c. space $(X, \tilde{x}_0)$. Then, we would like to know whether for any two choice of $\tilde{x}_1, \tilde{x}_2 \in \tilde{X}$, we get an element $\varphi \in \mathsf{Deck}(\tilde{X})$ such that

$\varphi(\tilde{x}_1) = \tilde{x}_2$ and $\varphi(\tilde{x}_2) = \tilde{x}_1$. In such a case, we can say that the cover $\tilde{X}$ will be the one with *maximal symmetry*. Now with the result above, we can partly answer that, for if $p_{1*}(\pi_1(\tilde{X}, \tilde{x}_1)) = p_{2*}(\pi_1(\tilde{X}, \tilde{x}_2))$ in $\pi_1(X, x_0)$, then there is a *unique* deck transformation $\varphi \in \mathsf{Deck}(\tilde{X})$ such that $\varphi(\tilde{x}_1) = \tilde{x}_2$ as $p_2 \circ \varphi = p_1$, where $p_i : (\tilde{X}, \tilde{x}_i) \to (X, x_0)$. But the question for the converse remains open and we see how to resolve it in the next big theorem.

We now state one of the major theorems of this course.

Theorem 5.1.8.10. *(Classification of coverings) Let (X, x_0) be a path-connected and locally path-connected space. Then,*

1. *(Based version) Two connected covers $(\tilde{X}_1, \tilde{x}_1, p_1)$ and $(\tilde{X}_2, \tilde{x}_2, p_2)$ are equivalent if and only if*
$$p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) = p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2)) \text{ in } \pi_1(X, x_0).$$
2. *(Unbased version) Two connected covers $(\tilde{X}_1, p_1)$ and $(\tilde{X}_2, p_2)$ are equivalent if and only if for any $\tilde{x}_1 \in p_1^{-1}(x_0)$ and $\tilde{x}_2 \in p_2^{-1}(x_0)$, we have that*
$$p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) \;\&\; p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2)) \text{ are conjugate subgroups of } \pi_1(X, x_0).$$

Proof. 1. (R $\Rightarrow$ L) This is exactly the Lemma 5.1.8.8 above.
(L $\Rightarrow$ R) Suppose the two covers are equivalent. Then there is a homeomorphism $\varphi : (\tilde{X}_1, \tilde{x}_1) \to (\tilde{X}_2, \tilde{x}_2)$ such that $p_2 \circ \varphi = p_1$. Let its inverse be $\varpi : (\tilde{X}_2, \tilde{x}_2) \to (\tilde{X}_1, \tilde{x}_1)$, which satisfies $p_1 \circ \varpi = p_2$. The former gives us $p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) = p_{2*} \circ \varphi_*(\pi_1(\tilde{X}_1, \tilde{x}_1)) \le p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))$. Similarly, the latter gives us $p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2)) = p_{1*} \circ \varpi_*(\pi_1(\tilde{X}_2, \tilde{x}_2)) \le p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1))$. Hence we get the equality.

2. (L $\Rightarrow$ R) Choose $\tilde{x}_i \in p_i^{-1}(x_0)$. We know that there is a homeomorphism $\varphi : \tilde{X}_1 \to \tilde{X}_2$ such that $p_2 \circ \varphi = p_1$. Hence $\varphi(\tilde{x}_1) \in p_2^{-1}(x_0)$ and may not be equal to $\tilde{x}_2$. So we have two based covers $(\tilde{X}_2, \tilde{x}_2)$ and $(\tilde{X}_2, \varphi(\tilde{x}_1))$ with the same projection map p_2. Now since $(\tilde{X}_1, \tilde{x}_1)$ and $(\tilde{X}_2, \varphi(\tilde{x}_1))$ are equivalent, then by 1. above, they induce the same subgroups of $\pi_1(X, x_0)$. So if we can show that the subgroups induced by $(\tilde{X}_2, \varphi(\tilde{x}_1))$ and $(\tilde{X}_2, \tilde{x}_2)$ are conjugates, then we would be done. So we reduce to showing that $p_{2*}(\pi_1(\tilde{X}_2, \varphi(\tilde{x}_1)))$ and $p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))$ are conjugates. Since $\tilde{X}_2$ is path-connected, therefore we have a path $\gamma : I \to \tilde{X}_2$ such that $\gamma(0) = \varphi(\tilde{x}_1)$ and $\gamma(1) = \tilde{x}_2$. Now recall from proof of Lemma **??** that the following establishes an isomorphism of groups:

$$\begin{aligned} \Phi : \pi_1(\tilde{X}_2, \tilde{x}_2) &\longrightarrow \pi_1(\tilde{X}_2, \varphi(\tilde{x}_1)) \\ [\xi] &\longmapsto [\gamma * \xi * \bar{\gamma}]. \end{aligned}$$

So, applying p_{2*} on the above map Φ yields

$$\begin{aligned} p_{2*}(\Phi) : p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2)) &\longrightarrow p_{2*}(\pi_1(\tilde{X}_2, \varphi(\tilde{x}_1))) \\ [p_2 \circ \xi] &\longmapsto [(p_2 \circ \gamma) * (p_2 \circ \xi) * (p_2 \circ \bar{\gamma})], \end{aligned}$$

which is also an isomorphism. But this tells us more, that each element of $p_{2*}(\pi_1(\tilde{X}_2, \varphi(\tilde{x}_1)))$ can be written as a conjugate of an element of $p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))$ by a fixed element $[p_2 \circ \gamma]$, conditioned on the fact that we somehow show that $[\overline{p_2 \circ \gamma}] = [p_2 \circ \bar{\gamma}]$, but that's a tautology. Hence we are done.

(R $\Rightarrow$ L) We are given that there exists $[\gamma] \in \pi_1(X, x_0)$ for any choice of $\tilde{x}_1$ and $\tilde{x}_2$ such that

$$p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) = [\bar{\gamma}] p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))[\gamma].$$

In order to get a homeomorphism $\varphi : (\tilde{X}_1, \tilde{x}_1, p_1) \to (\tilde{X}_2, \tilde{x}_2, p_2)$, we will use statement 1. above. Since we need a homeomorphism φ such that $p_2 \circ \varphi = p_1$, therefore we may show that $p_{1*}(\pi_1(\tilde{X}_1, \tilde{y}_1)) = p_{2*}(\pi_1(\tilde{X}_2, \tilde{y}_2))$ for any $\tilde{y}_i \in p_i^{-1}(x_0)$ and then use 1. to conclude the existence of such φ. To show this, we first lift the loop γ in X to a unique path $\tilde{\gamma}$ in $\tilde{X}_2$ where we start the lift at $\tilde{x}_2$ (Theorem 5.1.2.2). Hence we have a path $\tilde{\gamma} : I \to \tilde{X}_2$ where $\tilde{\gamma}(0) = \tilde{x}_2$ and denote $z := \tilde{\gamma}(1) \in p_2^{-1}(x_0)$. Now, if $[p_2 \circ \xi] \in p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))$, then $[\bar{\gamma} * (p_2 \circ \xi) * \gamma]$ is equal to $[(p_2 \circ \bar{\tilde{\gamma}}) * (p_2 \circ \xi) * (p_2 \circ \tilde{\gamma})]$ because $p_2 \circ \tilde{\gamma} = \gamma$, and then we further get that it is equal to $[p_{2*} \circ (\bar{\tilde{\gamma}} * \xi * \tilde{\gamma})]$ where $[\bar{\tilde{\gamma}} * \xi * \tilde{\gamma}] \in \pi_1(\tilde{X}_2, z)$. Conversely, for any $[p_2 \circ \eta] \in p_{2*}(\pi_1(\tilde{X}_2, z))$, we get the loop $[\alpha] := [\tilde{\gamma} * \eta * \bar{\tilde{\gamma}}] \in \pi_1(\tilde{X}_2, \tilde{x}_2)$ which is such that $[\bar{\gamma} * (p_2 \circ \alpha) * \gamma] = [p_2 \circ \eta]$. Hence indeed, we get that $[\bar{\gamma}] p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))[\gamma] = p_{2*}(\pi_1(\tilde{X}_2, z))$. Since $p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) = [\bar{\gamma}] p_{2*}(\pi_1(\tilde{X}_2, \tilde{x}_2))[\gamma]$, therefore we get $p_{1*}(\pi_1(\tilde{X}_1, \tilde{x}_1)) = p_{2*}(\pi_1(\tilde{X}_2, z))$, so we are done as now we can take $\tilde{y}_1 := \tilde{x}_1$ and $\tilde{y}_2 := z$. □

Construction of universal cover

We will show some striking results about the group of deck transformations of the universal cover and the fundamental group of the base space. Before that, let us define a class of connected covers which have in some sense maximal symmetry.

Definition 5.1.8.11. (**Normal covers**) Let (X, x_0) be a path-connected and locally path-connected space. A connected cover $p : \tilde{X} \to X$ is said to be normal if for any two $\tilde{x}_1, \tilde{x}_2 \in p^{-1}(x_0)$ there exists a $\varphi \in \mathrm{Deck}(\tilde{X})$ such that $\varphi(\tilde{x}_1) = \tilde{x}_2$.

Clearly, this induces the following map when $\tilde{X}$ is normal:

$$\begin{aligned} \mathrm{Deck}(\tilde{X}) &\longrightarrow S_{p^{-1}(x_0)} \\ \varphi &\longmapsto \varphi|_{p^{-1}(x_0)} . \end{aligned}$$

We will use this map later. The following gives a characterization of normal covers.

Lemma 5.1.8.12. *Let (X, x_0) be a path-connected and locally path-connected space. Then, a connected cover $p : \tilde{X} \to X$ is normal if and only if for all $\tilde{x}_0 \in p^{-1}(x_0)$, we have that $p_*(\pi_1(\tilde{X}, \tilde{x}_0))$ is a normal subgroup of $\pi_1(X, x_0)$.*

Proof. (L ⇒ R) Take any $[\gamma] \in \pi_1(X, x_0)$ and let $\tilde{\gamma}$ be the unique lift of γ in $\tilde{X}$ starting from $\tilde{x}_0 \in \tilde{X}$ (Theorem 5.1.2.2). Denote $\tilde{x}_1 := \tilde{\gamma}(1) \in p^{-1}(x_0)$ as γ is a lift of a loop so both endpoints are in $p^{-1}(x_0)$. Now, since $\tilde{X}$ is normal, therefore there exists $\varphi \in \mathrm{Deck}(\tilde{X})$ such that $\varphi(\tilde{x}_0) = \tilde{x}_1$. Hence $(\tilde{X}, \tilde{x}_0)$ and $(\tilde{X}, \tilde{x}_1)$ are equivalent connected based covers. Therefore by Theorem 5.1.8.10, 1, we get that $H_i := p_*(\pi_1(\tilde{X}, \tilde{x}_i))$[6], $i = 0, 1$, are exactly equal. Now, $[\bar{\gamma}] H_0 [\gamma] = [\bar{\gamma}] p_*(\pi_1(\tilde{X}, \tilde{x}_0))[\gamma] = p_*([\bar{\tilde{\gamma}}] \pi_1(\tilde{X}, \tilde{x}_0)[\tilde{\gamma}]) = p_*(\pi_1(\tilde{X}, \tilde{x}_1)) = H_1 = H_0$ where the third to last equality follows from proof of Lemma **??**. Hence H_0 is a normal subgroup.

(R ⇒ L) Take any two points $\tilde{x}_1, \tilde{x}_2 \in p^{-1}(x_0)$. To find the required deck transformation φ, we see that since $(\tilde{X}, \tilde{x}_1)$ and $(\tilde{X}, \tilde{x}_2)$ are two covers such that $H_1 := p_*(\pi_1(\tilde{X}, \tilde{x}_1))$ and $H_2 := p_*(\pi_1(\tilde{X}, \tilde{x}_2))$ are normal subgroups of $\pi_1(X, x_0)$. Now since $\tilde{X}$ is path-connected, therefore there is a path joining $\tilde{x}_1$ to $\tilde{x}_2$ and let us denote it by $\gamma : I \to \tilde{X}$. Now, we get a loop $\xi := p \circ \gamma : I \to X$, based at x_0, and hence $[\xi] \in \pi_1(X, x_0)$. By uniqueness of path lifts (Theorem 5.1.2.2), we see

[6] Should have made this notation earlier?

that the lift of ξ (started at $\tilde{x}_1$) indeed has to be γ. We thus get $[\bar{\xi}]H_1[\xi] = [\bar{\xi}]p_*(\pi_1(\tilde{X},\tilde{x}_1))[\xi] = p_*([\bar{\gamma}]\pi_1(\tilde{X},\tilde{x}_1)[\gamma]) = p_*(\pi_1(\tilde{X},\tilde{x}_2)) = H_2$, where second to last equality follows from proof of Lemma **??**. Thus, H_1 and H_2 are conjugate, but both are normal, therefore $H_1 = H_2$ and by Theorem 5.1.8.10, 1, we are done. □

Let us now briefly outline the construction of universal covering space. Let (X, x_0) be a path-connected, locally path-connected and semi-locally simply connected space[7]. For such a space, the universal cover exists and is unique upto isomorphism (in $\mathbf{Cov}\,(X, x_0)$). We construct the universal cover by quotienting out $\mathrm{Path}_*\,(X, x_0)$, the space of all paths starting at x_0, by an equivalence relation given by the following:

$$\gamma \sim \eta \iff [\gamma\bar{\eta}] = [c_{x_0}] \in \pi_1(X, x_0).$$

This is a loaded relation, so let us explain. First, γ and η are two elements of $\mathrm{Path}_*\,(X, x_0)$, so they are paths both starting from x_0. The fact that we are demanding $[\gamma\bar{\eta}] = [c_{x_0}]$ tells us that we are demanding two things: 1) that γ and $\bar{\eta}$ be joinable, that is both γ and η have same end points, and 2) $\gamma\bar{\eta}$ is homotopy equivalent to constant loop x_0. This is indeed an equivalence relation on $\mathrm{Path}_*\,(X, x_0)$. Hence, by quotienting $\mathrm{Path}_*\,(X, x_0)$ by this relation we obtain a quotient, denoted:

$$\tilde{X} := \mathrm{Path}_*\,(X, x_0)/\sim.$$

This inherits a topology from compact-open topology of $\mathrm{Path}_*\,(X, x_0)$. Let us only state what is a basis of that topology, because verifying that indeed is so will unnecessarily deviate us from our goal. A basis of $\tilde{X}$ is given by subsets of the following form: for each path-connected, locally path-connected and semi-locally simply connected open subset $U \subseteq X$ and any $[\gamma] \in \tilde{X}$ whose endpoint lies in U, define

$$U_{[\gamma]} := \{[\gamma\alpha] \in \mathrm{Path}_*\,(X, x_0) \mid \alpha \text{ is contained in } U\}.$$

Such sets $U_{[\gamma]}$ forms a basis of $\tilde{X}$. A basic fact that can be checked about this basis is the following:

$$U_{[\gamma]} \cap U_{[\eta]} \neq \emptyset \implies U_{[\gamma]} = U_{[\eta]}.$$

This is because if $[\gamma\alpha] = [\eta\beta]$, then for any $[\gamma\delta] \in U_{[\gamma]}$, we have $[\gamma\delta] = [\eta\beta\bar{\alpha}\delta] \in U_{[\eta]}$, similarly the converse. We then have the following natural map:

$$\begin{aligned} p : \tilde{X} &\longrightarrow X \\ [\gamma] &\longmapsto \gamma(1). \end{aligned}$$

This is indeed well-defined. Moreover, it's a covering map as for any $x = \gamma(1) \in X$ for some path γ and any p.c., l.p.c., s.l.s.c. open set $U \ni x$, we get $p^{-1}(U) = \coprod_{[\alpha] \in \pi_1(X, x_0)} U_{[\alpha\gamma]}$. Finally, note that $\tilde{X}$ is simply-connected.

[7] This means that for all $x \in X$, there exists an open set $U \ni x$ which also contains x_0 such that $\iota_*(\pi_1(U, x_0)) = \{0\} \leq \pi_1(X, x_0)$. Note that this doesn't necessarily means that $\pi_1(U, x_0) = \{0\}$(!)

Construction of a connected cover from a subgroup

Construction 5.1.8.13. Let (X, x_0) be a connected, path-connected and semi-locally simply connected space. Let $H \leq \pi_1(X, x_0)$ be a subgroup. We will construct a connected cover $(X_H, \tilde{x}_0)$ of X such that $p_*(\pi_1(X_H, \tilde{x}_0)) = H$. This is obtained as follows.

Consider the following map:

$$\begin{aligned} H \times \text{Path}_* (X, x_0)/\sim &\longrightarrow \text{Path}_* (X, x_0)/\sim \\ ([\alpha], [\gamma]) &\longmapsto [\alpha * \gamma]. \end{aligned}$$

This is well-defined because if $([\alpha], [\gamma]) = ([\beta], [\eta])$, then $[\alpha*\gamma] = [\beta*\eta]$ in $\text{Path}_* (X, x_0)/\sim$ obtained by concatenating the two homotopies. Moreover, we have the following

$$\begin{aligned} ([c_{x_0}], [\gamma]) &\mapsto [\gamma] \\ ([\alpha], [\beta\gamma]) &\mapsto [\alpha\beta\gamma]. \end{aligned}$$

So we have that the group H acts on the universal covering space $\text{Path}_* (X, x_0)/\sim= \tilde{X}$. Now, consider the quotient $\tilde{X}/H$. Explicitly, this is the quotient of $\tilde{X}$ obtained by the relation

$$[\gamma] \sim_H [\eta] \iff \exists [\alpha] \in H \text{ s.t. } [\gamma] = [\alpha\eta].$$

The above holds if and only if $\gamma(1) = \eta(1)$, hence $\gamma\bar{\eta}$ is a loop of X based at x_0. The relation above can thus be read as:

$$[\gamma] \sim_H [\eta] \iff [\gamma\bar{\eta}] \in H.$$

Now, note that the quotient space $X_H := \tilde{X}/H$ will identify certain decks of the cover. Let us explain. Let $\gamma(1) = x \in X$ for some path γ in X and $U \subseteq X$ be an evenly covered neighborhood of x. Therefore

$$p^{-1}(U) = \coprod_{[\alpha]\in\pi_1(X,x_0)} U_{[\alpha\gamma]}.$$

That is, the cardinality of decks is exactly the order of $\pi_1(X, x_0)$. Now, when we apply the quotient map $q : \tilde{X} \twoheadrightarrow \tilde{X}/H$, we get that

$q(U_{[\xi]})$ and $q(U_{[\eta]})$ are identified if and only if $[\xi] = [\alpha\eta]$ for some $[\alpha] \in \pi_1(X, x_0)$

Hence, applying q on $p^{-1}(U)$ will give us

$$\begin{aligned} q(p^{-1}(U)) &= q\left(\coprod_{[\alpha]\in\pi_1(X,x_0)} U_{[\alpha\gamma]}\right) \\ &= \bigcup_{[\alpha]\in\pi_1(X,x_0)} q(U_{[\alpha\gamma]}) \\ &= \coprod_{[\alpha]\in H} q(U_{[\alpha\gamma]}). \end{aligned}$$

Now since q is a quotient map and $p : \tilde{X} \to X$ is map such that p identifies all elements of an equivalence class of $\tilde{X}/H$, therefore we have a unique map $p_H : X_H \to X$, which is the required covering map corresponding to subgroup H. Moreover, one can show that $p_{H*}(\pi_1(X_H, \tilde{x}_0)) = H$.

5.1.9 Covers of $\mathbb{R}P^2 \times \mathbb{R}P^2$

We will classify all covers of this space, and in the process will portray the power of tools developed so far. We first begin with a section on background calculations. The reader interested only in the classification result may safely jump on to Theorem 5.1.9.4 and may refer back to results in the following section whenever it is used in the proof.

Background calculations

Let us begin by trying to understand the structure of $\pi_1(\mathbb{R}P^2)$.

Lemma 5.1.9.1. *The antipodal action of $\mathbb{Z}_2$ on S^n is a free action. This induces a covering map $p : S^n \to \mathbb{R}P^n$.*

Proof. The action is defined by

$$\begin{aligned} \mathbb{Z}_2 \times S^n &\longrightarrow S^n \\ (0, x) &\longmapsto x \\ (1, x) &\longmapsto -x. \end{aligned}$$

So if $x \in S^n$ is any point, then for any $g \in S_{\mathbb{Z}_2}(x)$, we get $g \cdot x = x$. This implies that either $g = 0$ or $x = -x$. Since there is no point in S^n such that $x = -x$, therefore $g = 0$. So the action is free. Now, since $\mathbb{Z}_2$ is finite and S^n is locally finite, therefore by Lemma 5.1.7.3, 2, we get that this action is properly discontinuous. Now, using Theorem 5.1.7.4, we get that the quotient map $p : S^n \to S^n/\mathbb{Z}_2$ is a covering map. But since $S^n/\mathbb{Z}_2$ is exactly how $\mathbb{R}P^n$ constructed, therefore we have S^n as a cover of $\mathbb{R}P^n$.

ALITER : One can show that we get a covering map $p : S^n \to \mathbb{R}P^n$ by the $\mathbb{Z}_2$ action without using Theorem 5.1.7.4. For this, take any point $[x] \in \mathbb{R}P^n$ where we identify $\mathbb{R}P^n$ as the quotient of S^n by $\mathbb{Z}_2$, so each element of $\mathbb{R}P^n$ represents an equivalence class of two points which are antipodal. To find the required evenly covered neighborhood of $[x]$, we first notice that we get an open subset of S^n, denoted U and it's antipodal version $-U$ such that $x \in U$ and $-x \in -U$ and, most importantly, $U \cap -U = \emptyset$. *This last fact follows most importantly from the fact that the action of $\mathbb{Z}_2$ on S^n is* ***properly discontinuous***. Defining p to be the quotient map $p : S^n \to S^n/\mathbb{Z}_2$, we get that $p^{-1}(U) = U$. So we have that p is a 2-sheeted covering of $\mathbb{R}P^n$. This explicit proof shows the importance of the action of the finite group $\mathbb{Z}_2$ being free on S^n. □

Next we calculate the fundamental group of $\mathbb{R}P^2$ and as a result, gets pleasantly surprised in the process.[8]

Lemma 5.1.9.2. $\pi_1(\mathbb{R}P^n) = \mathbb{Z}_2$ *for* $n > 1$.

Proof. Take any $n > 1$. The Lemma 5.1.9.1 tells us that $p : S^n \to \mathbb{R}P^n$ is a covering map for $\mathbb{R}P^n$. We take it as a fact that $\pi_1(S^n) = 0$. Thus, S^n is a simply, path and locally path-connected space where $\mathbb{R}P^n$ is also semi-locally simply connected. Hence by the corollary of **main theorem of**

[8] You see, the fact that $\mathbb{R}P^n$ are such weird manifolds to imagine and also the fact that they are not embeddable in $\mathbb{R}^n$ (for $n > 1$) entices and invites one to think that their fundamental group is quite bad and complicated. But it is not so!

universal covering of a space, we get that $\pi_1(\mathbb{R}P^n) \cong \mathsf{Deck}(S^n)$. It is clear that $\mathsf{Deck}(S^n)$ is just $\mathbb{Z}_2$, as S^n is a 2-sheeted cover of $\mathbb{R}P^n$ (**by Galois equivalence for connected covers**). □

Lemma 5.1.9.3. *$\mathbb{R}P^2$ is connected, locally path-connected and semi-locally simply connected.*

Proof. Since $\mathbb{R}^3$ is satisfies all of the three properties and the quotient map $q : \mathbb{R}^3 \twoheadrightarrow \mathbb{R}P^2$ is continuous, so $\mathbb{R}P^2$ is connected. To show that $\mathbb{R}P^2$ is also locally path-connected, take any point $[x] \in \mathbb{R}P^2$, then $l_x := q^{-1}([x]) \subseteq \mathbb{R}^3$ is a line passing through origin in $\mathbb{R}^3$. For any open set $V \ni [x]$ in $\mathbb{R}P^2$, we have $U := q^{-1}(V)$ is open in $\mathbb{R}^3$, containing the line l_x. Choose an $\epsilon > 0$ small enough so that $l_x \times B_\epsilon \subseteq U$. Clearly, $l_x \times B_\epsilon$ is path-connected (it's a solid infinite cylinder with open boundary). Now, since q is a quotient map so $q(l_x \times B_\epsilon)$ is an open set inside V which is path-connected (as it is a continuous image of a path-connected set). Hence $\mathbb{R}P^2$ is both connected and locally path-connected.
Since $\mathbb{R}P^n$ is an n-dimensional manifold, so for each point there is an open neighborhood U which is homeomorphic to an open ball of $\mathbb{R}^n$, which is contractible. Hence $\mathbb{R}P^n$ is semi-locally simply connected. □

The classification theorem

Theorem 5.1.9.4. *(Classification of covers of $\mathbb{R}P^2 \times \mathbb{R}P^2$) Each connected cover of $\mathbb{R}P^2 \times \mathbb{R}P^2$ belongs to equivalence class of one of the following:*
1. *$\mathbb{R}P^2 \times \mathbb{R}P^2$,*
2. *$\mathbb{R}P^2 \times S^2$,*
3. *$S^2 \times \mathbb{R}P^2$,*
4. *$S^2 \times S^2$,*
5. *$S^2 \times S^2/\sim$ where $\sim$ is generated by $(x, y) \sim (-x, -y)$.*

Proof. In Lemma 5.1.9.2, we obtained $\pi_1(\mathbb{R}P^2) = \mathbb{Z}_2$. By Lemma **??**, we get that $\pi_1(\mathbb{R}P^2 \times \mathbb{R}P^2) = \mathbb{Z}_2 \times \mathbb{Z}_2$. Now, there are the following five subgroups of $\mathbb{Z}_2 \times \mathbb{Z}_2$:
1. $H_1 = \{(0, 0)\} = \{e\}$,
2. $H_2 = \{(0, 0), (0, 1)\}$,
3. $H_3 = \{(0, 0), (1, 0)\}$,
4. $H_4 = \{(0, 0), (0, 1), (1, 0), (1, 1)\} = \mathbb{Z}_2 \times \mathbb{Z}_2$.
5. $H_5 = \{(0, 0), (1, 1)\}$.

Now, note that $\mathbb{Z}_2 \times \mathbb{Z}_2$ is an abelian group, therefore, each subgroup of $\mathbb{Z}_2 \times \mathbb{Z}_2$ is normal. We know the following equivalence:

$$\begin{tikzcd} \{\text{Connected covers of } (X, x_0)\}/\text{equivalence} \arrow[r, bend left, "(\tilde{X},p) \longmapsto p_*(\pi_1(\tilde{X},\tilde{x}_0))"] & \{\text{Subgroups of } \pi_1(X, x_0)\}/\text{conjugacy} \arrow[l, bend left, "X_H \longleftarrow H"] \end{tikzcd}$$

for a path-connected, locally-path connected and semi-locally simply connected space X. Now, remember that X_H for some $H \leq \pi_1(X, x_0)$ is made via quotienting the universal cover of X by the action of H that is obtained by restricting the global action of $\pi_1(X, x_0)$ on $\tilde{X}$, via the deck transformations (we have $\pi_1(X, x_0) \cong \mathsf{Deck}(\tilde{X})$). Hence, X_H will be obtained by identifying the sheets of the universal cover $\tilde{X}$. In our case, $\mathsf{Deck}(\mathbb{R}P^2) = \mathbb{Z}_2 \times \mathbb{Z}_2$ and $(1, 0)$ acts on $(x, y) \in S^2$

via $(1,0) \cdot (x,y) \mapsto (-x,y)$, similarly for $(0,1),(1,1)$. This gives the required identifications on the sheets and hence the five classes of connected covers as stated above. $\square$

5.2 Cofibrations and cofiber sequences

Most of the long exact sequences appearing in algebraic topology are derived from the topics that we will cover in this chapter. These should rather be seen as an important conceptual tool in order to do computations. We will begin with cofibrations, closed subspaces from whose homotopies can be extended to the whole space, and then fibrations, which can be thought of as generalizations of covering spaces (more generally, fiber bundles) which one studies in a first course in algebraic topology.

Cofibrations can be treated as an intermediary tool for developing more sophisticated concepts in algebraic topology. In particular, we will be using this to derive an exact sequence of groups out of a map of based spaces.

Note that there is little to no difference in based or unbased cofibrations, so we will prove something for unbased context and will use it as it has been proved for based context as well. We will give some remarks towards the end.

5.2.1 Definition and first properties

Definition 5.2.1.1. (**Cofibrations**) A map $i : A \to X$ is a cofibration if it satisfies the homotopy extension property; if $f : X \to Y$ is a continuous map such that there is a homotopy $h : A \times I \to Y$ where $h(-, 0) = f \circ i$, then that homotopy can be lifted to $\tilde{h} : X \times I \to Y$ where $\tilde{h}(-, 0) = f$. More abstractly, if $h \circ i_0 = f \circ i$ in the following diagram, then there exists $\tilde{h}$ such that the following diagram commutes:

$$\begin{tikzcd}
A \arrow[r, "i_0"] \arrow[d, "i"'] & A \times I \arrow[d, "h"'] \arrow[dr, "i\times\mathrm{id}"] & \\
X \arrow[r, "f"] \arrow[rr, bend right, "i_0"'] & Y & X \times I \arrow[l, dashed, "\tilde{h}"]
\end{tikzcd}.$$

One sees that pushout of a cofibration along any map is a cofibration.

Lemma 5.2.1.2. *Let $i : A \to X$ be a cofibration and $f : A \to B$ be any other map. Then, the pushout $j : B \to B \cup_f X$ is a cofibration.*

Proof. Take any map $g : B \cup_f X \to Y$ and a homotopy $h : B \times I \to Y$ where $h \circ i_0 = g \circ j$. We have the following diagram:

$$\begin{tikzcd}
A \arrow[dr, "f"] \arrow[dddd, "i\text{ (cofibration)}"'] & & & \\
& B \arrow[rr, "i_0"] \arrow[dd, "j"'] & & B\times I \arrow[dl, "j\times\mathrm{id}"'] \arrow[dd, "h"] \\
\arrow[r, phantom, "\text{p.o.}"] & {} & (B\cup_f X)\times I & \\
& B\cup_f X \arrow[ur, "i_0"] \arrow[rr, "g"'] & & Y \\
X \arrow[ur] & & &
\end{tikzcd} .$$

We wish to show that there is a map $\tilde{h} : (B \cup_f X) \times I \to Y$ which commutes with the diagram shown above. Since we have the following pushout square:

$$\begin{tikzcd}
B\cup_f X & X \arrow[l] \\
B \arrow[u, "j"] \arrow[ur, phantom, "\text{p.o.}"] & A \arrow[l, "f"] \arrow[u, "i"']
\end{tikzcd} ,$$

therefore after applying functor $- \times I$, which has a right adjoint, so is colimit preserving (we are working in the category of compactly generated spaces which is cartesian closed), we get the following pushout square which is closer to what we have in the first diagram:

$$\begin{tikzcd}
(B\cup_f X)\times I & X\times I \arrow[l] \\
B\times I \arrow[u, "j\times\mathrm{id}"] \arrow[ur, phantom, "\text{p.o.}"] & A\times I \arrow[l, "f\times\mathrm{id}"] \arrow[u, "i\times\mathrm{id}"']
\end{tikzcd} .$$

Now, we get a map h' as below by the virtue of i being a cofibration:

$$\begin{tikzcd}
A \arrow[rrrr, "i_0"] \arrow[dr, "f"] \arrow[dddd, "i\text{ (cofibration)}"'] & & & & A\times I \arrow[dl, "f\times\mathrm{id}"'] \arrow[dddd, "i\times\mathrm{id}"] \\
& B \arrow[rr, "i_0"] \arrow[dd, "j"'] & & B\times I \arrow[dl, "i\times\mathrm{id}"'] \arrow[dd, "h"] & \\
\arrow[r, phantom, "\text{p.o.}"] & {} & (B\cup_f X)\times I & & \\
& B\cup_f X \arrow[ur, "i_0"] \arrow[rr, "g"'] & & Y & \\
X \arrow[ur] \arrow[rrrr, "i_0"'] & & & & X\times I \arrow[ul, dashed, "h'"']
\end{tikzcd} .$$

Next, by the universal property of pushout $(B \cup_f X) \times I$, we get a map $\tilde{h}$

$$\begin{tikzcd}
A \arrow[rrrr, "i_0"] \arrow[dr, "f"] \arrow[dddd, "i\text{ (cofibration)}"'] & & & & A \times I \arrow[dl, "f\times\mathrm{id}"'] \arrow[dddd, "i\times\mathrm{id}"] \\
& B \arrow[rr, "i_0"] \arrow[dd, "j"'] \arrow[dd, phantom, "\text{p.o.}", shift right=6ex] & & B \times I \arrow[dl, "i\times\mathrm{id}"'] \arrow[dd, "h"] & \\
& & (B \cup_f X) \times I \arrow[dr, dashed, "\tilde{h}"'] & & \\
& B \cup_f X \arrow[ur, "i_0"] \arrow[rr, "g"'] & & Y & \\
X \arrow[ur] \arrow[rrrr, "i_0"'] & & & & X \times I \arrow[ul, dashed, "h'"']
\end{tikzcd},$$

which satisfies the required commutativity. $\square$

To check that a map $i : A \to X$ is a cofibration, we can reduce to checking the homotopy extension property to the map $X \to Mi$ where Mi is the *mapping cylinder*.

Definition 5.2.1.3 (Mapping cylinder). Let $f : X \to Y$ be a map. Then the mapping cylinder of f is the following pushout space

$$\begin{tikzcd}
Mf \arrow[dr, phantom, "\ulcorner", very near start] & X \times I \arrow[l] \\
Y \arrow[u] & X \arrow[l, "f"] \arrow[u, "i_0"']
\end{tikzcd}.$$

More explicitly, it is $((X \times I) \amalg Y)/\sim$ where $(x, 0) \sim f(x)$ for all $x \in X$.

Let $f : X \to Y$ be a map. More pictorially, Mf is formed by gluing cylinder $X \times I$ to Y along f. In mind, one pictures a cylinder "popping out" of Y from where $f(X)$ lived in Y, as shown in the following diagram: A based version of mapping cylinder is as follows.

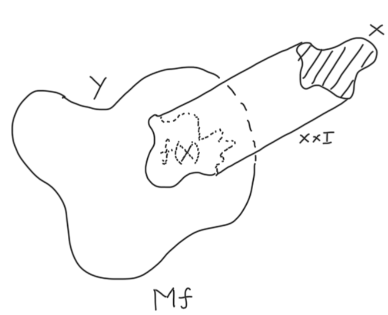


Figure 5.1: Schematic representation of mapping cylinder for $f : X \to Y$.

Definition 5.2.1.4 (Based mapping cylinder). Let $f : X \to Y$ be a based map. The based mapping cylinder M_*f is the pushout of reduced cylinder about f:

$$\begin{tikzcd}
M_*f \arrow[dr, phantom, "\lrcorner", very near start] & X \wedge I_+ \arrow[l] \\
Y \arrow[u] & X \arrow[l, "f"] \arrow[u, "i_0"']
\end{tikzcd}.$$

Indeed, we have the following result:

Proposition 5.2.1.5. *Let $i : A \to X$ be a map. Then the following are equivalent:*

1. *i is a cofibration.*
2. *i satisfies homotopy extension property for any $f : X \to Y$ and for any Y.*
3. *i satisfies homotopy extension property for the natural map $X \to Mi$ and the homotopy $h : A \times I \to Mi$ obtained from pushout.*

Proof. The only non-trivial part is to show $3 \Rightarrow 2$. Take any map $f : X \to Y$ and any homotopy $h : A \times I \to Y$. Consider

$$\begin{tikzcd}
A \arrow[rrr, "i_0"] \arrow[ddd, "i"'] & & & A \times I \arrow[ddll] \arrow[dl, "i \times \mathrm{id}"] \arrow[ddd, "h"] \\
& Mi \arrow[ddrr, dashed, "g"] & X \times I \arrow[l, dashed, "h_1"'] & \\
& & & \\
X \arrow[uur] \arrow[uurr, "i_0"] \arrow[rrr, "f"'] & & & Y
\end{tikzcd}.$$

The map h_1 is formed by homotopy extension property of i for $X \to Mi$ and g is formed by universal property of pushout which is Mi. The map $gh_1 : X \times I \to Y$ follows the required commutativity relations. □

Consequently, we have the following result.

Proposition 5.2.1.6. *Any cofibration $i : A \to X$ is an inclusion with closed image.*

Proof. Consider the natural maps $j : X \to Mi$ and $h : A \times I \to Mi$ obtained by the pushout square. Since $hi_0 = ji$, therefore by Proposition 5.2.1.5, 3, we obtain a map $\tilde{h} : X \times I \to Mi$ fitting in the following commutative diagram

$$\begin{tikzcd}
A \arrow[r, "i_0"] \arrow[d, "i"'] & A \times I \arrow[d, "h"] \arrow[dr] & \\
X \arrow[r, "j"] \arrow[rr, bend right, "i_0"'] & Mi & X \times I \arrow[l, dashed, "\tilde{h}"']
\end{tikzcd}.$$

Let $k : Mi \to X \times I$ be obtained by the following diagram

$$\begin{tikzcd}
A \arrow[r, "i_0"] \arrow[d, "i"'] & A \times I \arrow[d, "h"'] \arrow[ddr, "i\times\mathrm{id}"] & \\
X \arrow[r, "j"] \arrow[drr, "i_0"'] & Mi \arrow[ul, phantom, "\lrcorner", very near start] \arrow[dr, dashed, "k"] & \\
 & & X \times I
\end{tikzcd}.$$

It follows that $\tilde{h} \circ k : Mi \to Mi$ is id, that is, Mi is a retract of $X \times I$. Consequently, restricting onto $i(A)$, we see that $i(A)$ is a retract of $X \times I$, hence closed as $X \times I$ is compactly generated. It also follows from $\tilde{h} \circ k = \mathrm{id}$ that i is injective. $\square$

We see the following from the proof of the above result.

Corollary 5.2.1.7. *Let $i : A \to X$ be a map. Then the following are equivalent:*

1. *Map $i : A \to X$ is a cofibration.*
2. *Mapping cylinder Mi is a retract of $X \times I$.*

Proof. 1. $\Rightarrow$ 2. is immediate from the proof. For 2. $\Rightarrow$ 1. we seee that if $Mi \hookrightarrow X \times I \twoheadrightarrow Mi$ is a retract, then letting $\tilde{h} : X \times I \twoheadrightarrow Mi$, we have $\tilde{h} \circ i_0 = \mathrm{id}_X$ and $\tilde{h}|_{A\times I} = h$, as needed. $\square$

Let $f : X \to Y$ be an arbitrary map of spaces. We can replace f by a cofibration followed by a homotopy equivalence.

Construction 5.2.1.8 (Replacement by a cofibration and a homotopy equivalence)**.** Let $f : X \to Y$ be a map of spaces. Consider the following commutative triangle:

$$\begin{tikzcd}
 & Mf \arrow[dr, "r"] & \\
X \arrow[ur, "j"] \arrow[rr, "f"'] & & Y
\end{tikzcd}$$

where $Mf = Y \cup_f (X \times I)$ is the mapping cylinder and the other two maps are given as follows:

1. Map $j : X \to Mf$ is given by $x \mapsto (x, 1)$. We claim that j is a cofibration. Indeed, if $g : Mf \to Z$ is any map and we have a diagram as in Definition 5.2.1.1, then we can form the required homotopy $\tilde{h} : Mf \times I \to Z$ by defining
$$\tilde{h}([(x,s)],t) := \begin{cases} g(x) & \text{if } x \in Y \\ h(x,st) & \text{if } [(x,s)] \in X \times I. \end{cases}$$
We then see that $\tilde{h}(j \times \mathrm{id})(x,t) = \tilde{h}([(x,1)],t) = h(x,t)$ and that $\tilde{h}i_0([(x,s)]) = \tilde{h}([(x,s)],0) = h(x,0) = g(x)$. So we have the required extension and hence $j : X \to Mf$ is a cofibration.
2. Map $r : Mf \to Y$ is given by $r|_Y = \mathrm{id}_Y$ and $r|_{X\times I}(x,t) = f(x)$ for $t > 0$. We claim that r is a homotopy equivalence. For this, we have a map $i : Y \to Mf$ taking $y \mapsto [y]$. We then see that $ri = \mathrm{id}_Y$ and $ir \simeq \mathrm{id}_{Mf}$. The former is simple and the latter is established by the following homotopy $h : Mf \times I \to Mf$ mapping as $([(x,s)],t) \mapsto [(x,(1-t)s)]$ on $X \times I$ and $(y,t) \mapsto y$ on Y. This is indeed a homotopy from ir to id_{Mf}. Thus, $r : Mf \to Y$ establishes that Y is a deformation retract of the mapping cylinder Mf.

Hence, one can replace a map of spaces $f : X \to Y$ by a cofibration $j : X \to Mf$ followed by a homotopy equivalence $r : Mf \to Y$.

We now discuss an important characterization of cofibrations. For this we define first the following notion.

Definition 5.2.1.9 (Neighborhood deformation retract). A pair (X, A) where $A \subseteq X$ is a neighborhood deformation retract (NDR) if there exists a map $u : X \to I$ such that $u^{-1}(0) = A$ and a homotopy $h : X \times I \to X$ such that $h(x, 0) = \mathrm{id}_X(x) = x$, $h(a, t) = a$ for all $a \in A$ and all $t \in I$ and $h(x, 1) \in A$ if $u(x) < 1$.

Remark 5.2.1.10. Let (X, A) be an NDR-pair. If $u(X) \subseteq [0, 1)$, then $A \hookrightarrow X$ is a closed subspace which is a deformation retract of X.

Theorem 5.2.1.11. *Let A be a closed subsapce of X. Then the following are equivalent:*

1. *(X, A) is an NDR-pair.*
2. *$i : A \to X$ is a cofibration.*

We now define the notion of homotopy equivalence under a space. This will come in handy later. Recall that if $\mathbf{C}$ is a category $c \in \mathbf{C}$ is an object, then $\mathbf{C}_{c/}$ denotes the under category at c, i.e., where objects are $i : c \to a$ and maps are commutative triangles

$$\begin{tikzcd} & c \arrow[dl, "i"'] \arrow[dr, "j"] & \\ a \arrow[rr, "f"'] & & b \end{tikzcd} .$$

Definition 5.2.1.12 (Relative homotopy). Let $i : A \to X$ and $j : A \to Y$ be in $\mathbf{Top}^{cg}_{A/}$. Let $f, g : X \rightrightarrows Y$ be maps in $\mathbf{Top}^{cg}_{A/}$. Then $h : X \times I \to Y$ is a homotopy rel A between f and g if $h(x, 0) = f(x)$, $h(x, 1) = g(x)$ and $h(i(a), t) = j(a)$ for all $a \in A$ and $t \in I$.

The notion of homotopy equivalence rel A is special as the Theorem 5.2.1.14 shows, hence we give it the following name.

Definition 5.2.1.13 (Cofiber homotopy equivalence). Let $i : A \to X$ and $j : A \to Y$ be two spaces under A in $\mathbf{Top}^{cg}_{A/}$. If i and j homotopy equivalent under A, then X and Y are said to be cofiber homotopy equivalent.

Theorem 5.2.1.14. *Let $i : A \to X$ and $j : A \to Y$ be two cofibrations under A and $f : X \to Y$ be a map under A. If f is a homotopy equivalence, then f is a cofiber homotopy equivalence.*

Example 5.2.1.15. Let $i : A \to X$ be a cofibration. Then by Construction 5.2.1.8, we have

$$\begin{tikzcd} & Mi \arrow[dr, "r"] & \\ A \arrow[ur, "j"] \arrow[rr, "i"'] & & X \end{tikzcd}$$

where j is a cofibration and r is a homotopy equivalence. Since r is a homotopy equivalence under A, therefore by Theorem 5.2.1.14, r is a cofiber homotopy equivalence. Consequently, there is a homotopy inverse $\kappa : X \to Mi$ of r under A.

The following is a mild generalization of Theorem 5.2.1.14 in the sense that we allow mapping between two cofibration pairs now.

Proposition 5.2.1.16. *Let (X, A) and (Y, B) be two cofibration pairs and let $f : X \to Y$ and $d : A \to B$ be maps such that $f|_A = d$. If f and d are homotopy equivalences, then the map of pairs $(f, d) : (X, A) \to (Y, B)$ is a homotopy equivalence of pairs*[9].

We next portray how a cofibration pair (X, A) in some cases behaves homotopically same as the quotient X/A.

Proposition 5.2.1.17. *Let $i : A \to X$ be a cofibration and A be contractible. Then the quotient map $p : X \twoheadrightarrow X/A$ is a homotopy equivalence.*

Proof. As A is contractible, therefore for some $x_0 \in A$, we have a homotopy $h : A \times I \to A$ such that $h_0 = \mathrm{id}_A$ and $h_1 = c_{x_0}$. Consequently, we obtain $\tilde{h}$ as in the commutative square

$$\begin{tikzcd} A \arrow[r, "i_0"] \arrow[d, "i"'] & A \times I \arrow[d, "h"] \arrow[dr, "i \times \mathrm{id}"] & \\ X \arrow[r, "\mathrm{id}"] \arrow[rr, bend right, "i_0"'] & X & X \times I \arrow[l, dashed, "\tilde{h}"'] \end{tikzcd}$$

where we have $\tilde{h}_0 = \mathrm{id}_X$, $\tilde{h}_t(A) \subseteq A$ for all $t \in I$ and $\tilde{h}_1(A) = \{x_0\} \in A$. Consequently, $\tilde{h}_1$ fits in the following diagram

$$\begin{tikzcd} X \arrow[d, two heads, "p"'] \arrow[dr, "\tilde{h}_1"] & \\ X/A \arrow[r, "g"'] & X \end{tikzcd}$$

where $g : X/A \to X$ comes from the universal property of quotients. We claim that g is the required homotopy inverse of p. Indeed, by definition $\tilde{h} : \mathrm{id}_X \simeq g \circ p$. Consequently, we need only show that $\mathrm{id}_{X/A} \simeq p \circ g$. We derive this homotopy from $\tilde{h}$ as well. Indeed, for any $t \in I$, we obtain $\tilde{q}_t$ by universal property of quotients as in

$$\begin{tikzcd} X \arrow[r, "\tilde{h}_t"] \arrow[d, two heads, "p"'] & X \arrow[d, two heads, "p"] \\ X/A \arrow[r, "\tilde{q}_t"'] & X/A \end{tikzcd} .$$

It follows that the homotopy $\tilde{q} : X/A \times I \to X/A$ is such that $\tilde{q}_0 = \mathrm{id}_{X/A}$ and $\tilde{q}_1 = p \circ g$, as needed. $\square$

Let us end this section by discussing how we will tell the same story in the based setting.

Remark 5.2.1.18 (*Based cofibration*)**.** A based map $i : A \to X$ is a based cofibration if it satisfies the based version of homotopy extension property. The following are few remarks which are easily verifiable of the situation in the based case.

[9] as defined in Definition 5.4.1.1.

1. If a based map $i : A \to X$ is an unbased cofibration, then it is a based cofibration.
2. If $A \subseteq X$ is a closed subspace such that $* \to A$ and $* \to X$ are cofibrations and $i : A \to X$ is a based cofibration, then $i : A \to X$ is an unbased cofibration.
3. A based map $i : A \to X$ is a based cofibration if and only if $M_* i$ is a retract of $X \wedge I_+$.

We see the following example of above remark.

Lemma 5.2.1.19. *Let X be a based space. Then the inclusion $X \hookrightarrow CX$ to the base of the cone*
1. *is a deformation retract,*
2. *is a cofibration.*

Proof. The inclusion map is $x \mapsto [x, 0]$. The fact that X is deformation retract is immediate by the based homotopy $h : CX \times I \to CX$ given by $([x, t], s) \mapsto [x, t(1 - s)]$. We will use Remark 5.2.1.18, 3 for showing $i : X \hookrightarrow CX$ is a cofibration. Indeed, consider the map $CX \wedge I_+ \to M_* i$ given by $[[x, t], s] \mapsto [x, s + t]$. The inclusion $M_* i \to Y \wedge I_+$ is the map which on CX is $[x, t] \mapsto [[x, t], 0]$ and on $X \wedge I_+$ is $[x, t] \mapsto [[x, 0], t]$. One checks that this makes $M_* i$ a retract of $CX \wedge I_+$. □

5.2.2 Based cofiber sequences

The main point of cofiber sequences is to obtain an exact sequence of groups, which will prove to be helpful later. All cofibrations in this section are based cofibrations. We first observe that $[\Sigma X, Y]$ is a group.

Proposition 5.2.2.1. *Let X, Y be based spaces. Then*
1. *$[\Sigma X, Y]$ is a group under concatenation,*
2. *$[\Sigma^2 X, Y]$ is an abelian group under the same operation.*

Proof. The concatenation operation here is as follows : for $f, g \in \mathrm{Map}_*(\Sigma X, Y)$, define $f + g$ as

$$(f + g)([(x, t)]) := \begin{cases} f([(x, 2t)]) & \text{if } 0 \le t \le 1/2 \\ g([x, 2t - 1]) & \text{if } 1/2 \le t \le 1. \end{cases}$$

This tells us that $[\Sigma X, Y] \cong [X, \Omega Y]$ is a group. The second statement uses Theorem 5.0.0.8 to observe that a map $\Sigma^2 X \to Y$ is a map $S^1 \wedge S^1 = S^2 \to \mathrm{Map}_*(X, Y)$. Hence we reduce to showing that $[S^2, X]$ is an abelian group, this is well-known. □

Definition 5.2.2.2. (**Homotopy cofiber/Mapping cone**) Let $f : X \to Y$ be a based map and let $j : X \to M_* f$, $x \mapsto (x, 1)$ be it's cofibrant replacement. The homotopy cofiber Cf of f is defined to be the quotient of the based mapping cylinder $M_* f$ of f by the image of the map j taking $x \mapsto (x, 1)$. That is,

$$Cf := M_* f / j(X).$$

Alternatively, it is the pushout $Cf = Y \cup_f CX$.

There is a relationship between unbased cofiber and based cofiber.

Lemma 5.2.2.3. *Let X be an unbased space. Then the unreduced cone of X is isomorphic to the reduced cone of pointification of X. That is,*

$$CX \cong CX_+.$$

Proof. We have

$$CX_+ = X_+ \wedge I = \frac{X_+ \times I}{\{\text{pt.}\} \times I \amalg X \times \{1\}} = \frac{X \times I \amalg \{\text{pt.}\} \times I}{\{\text{pt.}\} \times I \amalg X \times \{1\}}$$
$$\cong \frac{X \times I}{X \times \{1\}} = CX,$$

as needed. □

This is an important observation, as it says that unreduced homotopy cofiber is isomorphic to the homotopy cofiber of the poinitification.

Proposition 5.2.2.4. *Let X, Y be unbased spaces and $f : X \to Y$ be an unbased map. Then the unreduced homotopy cofiber of f is isomorphic to the homotopy cofiber of $f_+ : X_+ \to Y_+$. That is,*

$$Cf \cong Cf_+.$$

Proof. By Lemma 5.2.2.3, we can write

$$Cf_+ = Y_+ \cup_{f_+} CX_+ \cong Y_+ \cup_{f_+} CX$$

where $X_+ \to CX$ is the map which takes pt. $\mapsto [x, 1]$ as the basepoint of CX is $[x, 1]$. Consequently, $Y_+ \cup_{f_+} CX$ is isomorphic to $Y \cup_f CX$. □

Remark 5.2.2.5. It follows from Proposition 5.2.2.4 that there is really no difference between reduced and unreduced cofiber as unreduced cofiber is really a special case of reduced cofiber by pointification.

The following result shows that the homotopy cofiber of a based cofibration is is of the same homotopy type as X/A. This is an important property of cofibrations.

Proposition 5.2.2.6. *Let $i : A \to X$ be a based cofibration between based spaces. Then,*

1. *$Ci/CA \cong X/A$,*
2. *$\pi : Ci \to Ci/CA$ is a based homotopy equivalence.*

Proof. **TODO.** □

Pictorially, one sees that the mapping cone Cf of $f : X \to Y$ is obtained by gluing Y to the cone of X at it's base. We are now ready to construct cofiber sequence of a based map $f : X \to Y$.

Construction 5.2.2.7 (Cofiber sequence)**.** Let $f : X \to Y$ be a based map and denote Cf to be the mapping cone of f. We have a natural map $i : Y \to Cf$ which is the inclusion of Y into the mapping cone. This is a cofibration because it is the pushout (Lemma 5.2.1.2) of the inclusion $X \to CX$ of X into the 0-th level of the cone CX and this inclusion is a cofibration (Lemma 5.2.1.19). The sequence $X \to Y \to Cf$ is called the *short cofiber sequence of f*.

Consider also the map $-\Sigma f : \Sigma X \to \Sigma Y$ which maps $[(x, t)] \mapsto [(f(x), 1 - t)]$. We have another natural map from the mapping cone to its quotient by Y given by $\pi : Cf \to Cf/Y \cong \Sigma X$. We then get the following sequence of based maps, called the *long cofiber sequence of map f*:

$$\begin{tikzcd}
X \arrow[r, "f"] & Y \arrow[r, tail, "i"] & Cf \arrow[dll, two heads, "\pi"'] \\
\Sigma X \arrow[r, "-\Sigma f"'] & \Sigma Y \arrow[r, "-\Sigma i"] & \Sigma Cf \arrow[dll, "-\Sigma \pi"'] \\
\Sigma^2 X \arrow[r, "\Sigma^2 f"'] & \Sigma^2 Y \arrow[r, "\Sigma^2 i"] & \Sigma^2 Cf \arrow[dll, dash] \\
{} & &
\end{tikzcd}$$

The main theorem that will be used continuously elsewhere is that cofiber sequence of a map gives a long exact sequence in homotopy sets. First, recall that for any based space Z, we have the homotopy classes of maps $[X, Z]$. Moreover, $[-, Z]$ is contravariantly functorial as for any based map $f : X \to Y$, we get

$$\begin{aligned}
[f, Z] : [Y, Z] &\longrightarrow [X, Z] \\
g &\longmapsto g \circ f.
\end{aligned}$$

We are now ready to state the main theorem.

Theorem 5.2.2.8 (Main theorem of cofiber sequences)**.** *Let $f : X \to Y$ be a based map and Z be a based space in $\mathbf{Top}_*^{cg}$. Then the functor $[-, Z]$ applied on the long cofiber sequence of f yields a long exact sequence of based sets:*

$$\begin{tikzcd}
& & {} \\
{[\Sigma^2 Cf, Z]} \arrow[urr, dash] & {[\Sigma^2 Y, Z]} \arrow[l] & {[\Sigma^2 X, Z]} \arrow[l] \\
{[\Sigma Cf, Z]} \arrow[urr] & {[\Sigma Y, Z]} \arrow[l] & {[\Sigma X, Z]} \arrow[l] \\
{[Cf, Z]} \arrow[urr, "\pi_*"] & {[Y, Z]} \arrow[l, "i^*"] & {[X, Z]} \arrow[l, "f^*"]
\end{tikzcd}$$

The proof of this theorem relies on the following fundamental observation.

Proposition 5.2.2.9. *Let $f : X \to Y$ be a based map and Z be a based space. Consider the short cofiber sequence*

$$X \xrightarrow{f} Y \xrightarrow{i} Cf.$$

Then the sequence of based sets

$$[Cf, Z] \longrightarrow [Y, Z] \longrightarrow [X, Z]$$

is exact.

Proof. Let $g \in [Y, Z]$ such that $gf \simeq c_*$ in $[X, Z]$. We wish to show that there is a map $k \in [Cf, Z]$ such that $ki \simeq g$ in $[Y, Z]$. We first have a based homotopy $h : X \times I \to Z$ between gf and c_*. As h is constant on $X \vee I$, therefore we obtain a map $\bar{h} : CX \to Z$. Note that the following pushout diagram commutes so to give a unique map $k : Cf \to Z$

$$\begin{tikzcd}
Z & & \\
& Cf \arrow[ul, dashed, "k"'] \arrow[r, phantom, ""] & CX \arrow[l] \arrow[ull, bend right, "\bar{h}"'] \\
& Y \arrow[u, "i"] \arrow[uul, bend left, "g"] \arrow[ur, phantom, "\ulcorner"] & X \arrow[l, "f"] \arrow[u, "i_0"']
\end{tikzcd}$$

Hence we have that $ki = g$, hence we don't even need to construct a homotopy between ki and g. ☐

We will now show that each term in the cofiber sequence is obtained by taking cofiber of the previous map. For that, we would need the following small result.

Lemma 5.2.2.10. *Let $f : X \to Y$ be a based map. Then,*

1. *We have a natural based homeomorphism $\Sigma Cf \cong C\Sigma f$.*
2. *The suspension functor takes the short cofiber sequence*

$$X \xrightarrow{f} Y \xrightarrow{i} Cf$$

to a short cofiber sequence

$$\Sigma X \xrightarrow{\Sigma f} \Sigma Y \xrightarrow{\Sigma i} \Sigma Cf.$$

Proof. The first one follows from Σ being a left adjoint. The second statement follows from first statement as we have the following isomorphism

$$\begin{tikzcd}
\Sigma X \arrow[r, "\Sigma f"] & \Sigma Y \arrow[r, "\Sigma i"] \arrow[dr, hook] & \Sigma Cf \arrow[d, "\cong"] \\
& & C\Sigma f
\end{tikzcd} .$$

This completes the proof. ☐

Proposition 5.2.2.11. *Let $f : X \to Y$ be a based map. Then each consecutive pair of maps in the long cofiber sequence of f is a short cofiber sequence.*

Proof. Note that the following square commutes

$$\begin{tikzcd}
\Sigma Cf \arrow[r, "\Sigma\pi"] \arrow[d, "\cong"'] & \Sigma^2 X \arrow[r, "-\Sigma^2 f"] \arrow[d, "\tau"] & \Sigma^2 Y \arrow[d, equal] \\
C\Sigma f \arrow[r, "\pi'"'] & \Sigma^2 X \arrow[r, "\Sigma^2 f"'] & \Sigma^2 Y
\end{tikzcd}$$

where $\tau([x,t,s]) = [x,s,t]$ is a homeomorphism and $\pi' : C\Sigma f \to C\Sigma f/\Sigma Y$ is the quotient map. We claim that τ is homotopic to $-\mathrm{id}$, where $(-\mathrm{id})([x,t,s]) = [x,t,1-s]$. With this claim and Lemma 5.2.2.10, we would reduce to showing that $Y \to Cf \to \Sigma X$ and $Cf \to \Sigma X \to \Sigma Y$ in the cofiber sequence of f are short cofiber sequences.

To see a based homotopy between τ and $-\mathrm{id}$ as based maps $\Sigma^2 X \to \Sigma^2 X$, we see that the following map will work

$$\begin{aligned} h : \Sigma^2 X \times I &\longrightarrow \Sigma^2 X \\ ([x,t,s],r) &\longmapsto [x,(1-r)s+rt,(1-r)t+r(1-s)]. \end{aligned}$$

We now wish to show that the two pairs are short cofiber sequenes. The fact that $Y \to Cf \to \Sigma X$ is a short cofiber sequence is immediate from Proposition 5.2.2.6 as it will yield the following diagram

$$\begin{tikzcd} & & Ci \arrow[d, "\simeq"] \\ Y \arrow[r, "i"'] & Cf \arrow[ur, "\pi'"] \arrow[r, "\pi"'] & \Sigma X \end{tikzcd}.$$

The fact that $Cf \to \Sigma X \to \Sigma Y$ is also a short cofiber sequence follows from the following diagram which can be seen to be commutative, albeit requires a lot of work:

$$\begin{tikzcd} Cf \arrow[r, "\pi"] \arrow[d, equal] & \Sigma X \arrow[r, "-\Sigma f"] \arrow[d, "\simeq"] & \Sigma Y \arrow[d, "\cong"] \\ Cf \arrow[r, "\pi'"'] & Ci \arrow[r, "\pi''"'] & Ci/Cf \end{tikzcd}.$$

This completes the proof. □

5.3 Fibrations and fiber sequences

We now study fibrations, which is a generalization of covering spaces. Indeed, recall that covering spaces satisfies homotopy lifting property. That *becomes* the definition of a fibration. Indeed, one can have a fruitful time reading about fibrations by keeping the basic results about covering spaces in mind. We'll see that familiar objects from geometry are fibrations (fiber bundles, for example).

5.3.1 Definition and first properties

Definition 5.3.1.1 (Fibrations). A surjective map $p : E \to B$ is a fibration if it satisfies homotopy lifting property. That if, for any map $f : Y \to E$ and any homotopy $h : Y \times I \to B$ such that $p \circ f = h \circ i_0$, there exists $\tilde{h} : Y \times I \to E$ such that the following commutes

$$\begin{tikzcd} Y \arrow[r, "f"] \arrow[d, "i_0"'] & E \arrow[d, "p"] \\ Y \times I \arrow[r, "h"'] \arrow[ur, dashed, "\tilde{h}"] & B \end{tikzcd}.$$

Just as pushouts of cofibrations along any map is a cofibration, we have pullback of a fibration along any map is a fibration.

Lemma 5.3.1.2. *Let $p : E \to B$ be a fibration and $g : A \to B$ be any map. Then the pullback of p along g given by $p' : E \times_B A \to A$ is a fibration.*

Proof. Consider the following diagram

$$\begin{tikzcd}
Y \arrow[r, "f"] \arrow[d, "i_0"'] & E \times_B A \arrow[r, "\pi"] \arrow[d, "p'"] & E \arrow[d, "p"] \\
Y \times I \arrow[r, "h"'] & A \arrow[r, "g"'] & B
\end{tikzcd}.$$

As p is a fibration, we yield a homotopy $\tilde{h}_1 : Y \times I \to E$ as in

$$\begin{tikzcd}
Y \arrow[r, "\pi f"] \arrow[d, "i_0"'] & E \arrow[d, "p"] \\
Y \times I \arrow[r, "gh"'] \arrow[ur, dashed, "\tilde{h}_1"] & B
\end{tikzcd}.$$

Consequently, we get a pullback diagram

$$\begin{tikzcd}
Y \times I \arrow[r, dashed, "!\tilde{h}"] \arrow[rr, bend left, "\tilde{h}_1"] \arrow[dr, "h"'] & E \times_B A \arrow[r, "\pi"] \arrow[d, "p'"'] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\
& A \arrow[r, "g"'] & B
\end{tikzcd}$$

which yields $\tilde{h} : Y \times I \to E \times_B A$. We claim that this is the required homotopy extension. We immediately have $p'\tilde{h} = h$ from the above diagram. We need only show that $\tilde{h}i_0 = f$. To this end, consider the following pullback square

$$\begin{tikzcd}
Y \arrow[r, dashed, "!\kappa"] \arrow[rr, bend left, "\pi f"] \arrow[dr, "hi_0"'] & E \times_B A \arrow[r, "\pi"] \arrow[d, "p'"'] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\
& A \arrow[r, "g"'] & B
\end{tikzcd}$$

which yields a unique $\kappa : Y \to E \times_B A$. It follows that both f and $\tilde{h}i_0$ satisfies the same commutation properties as κ. It follows from uniqueness of κ w.r.t. these properties that $\tilde{h}i_0 = f$, as required. $\square$

We now introduce a sort of intermediary space for further studying fibrations.

Definition 5.3.1.3 (Mapping path space). Let $f : X \to Y$ be a map. The mapping path space Nf is defined to be the following pullback

$$\begin{tikzcd} Nf := X \times_Y Y^I \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "f"] \\ Y^I \arrow[r, "p_0"'] & Y \end{tikzcd}$$

where $p_0 : Y^I \to Y$ takes $\gamma \mapsto \gamma(0)$.

Remark 5.3.1.4. Consequently, the mapping path space $Nf = \{(x, \gamma) \in X \times Y^I \mid f(x) = \gamma(0)\}$. Hence a point in Nf is the data of a point $x \in X$ *upstairs* and a path $\gamma \in Y^I$ starting *downstairs* at the image of x under f.

With regards to mapping path spaces, one important type of function Nf is that of a *path lifters*.

Definition 5.3.1.5 (Path lifters). Let $f : X \to Y$ be a map. Let $k : X^I \to Nf$ be the unique map obtained by the following pullback diagram

$$\begin{tikzcd} X^I \arrow[r, dashed, "k"] \arrow[rr, bend left, "p_0"] \arrow[dr, "f^I"'] & Nf \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "f"] \\ & Y^I \arrow[r, "p_0"'] & Y \end{tikzcd}.$$

A path lifter $s : Nf \to X^I$ is a global section of k, i.e. $k \circ s = \mathrm{id}_{Nf}$.

Remark 5.3.1.6. The main content of a path lifter $s : Nf \to X^I$ is the fact that its a global section of k. That is, if we let $\tilde{\gamma} = s(x, \gamma) \in X^I$, then $k(\tilde{\gamma}) = (p_0(\tilde{\gamma}), f \circ \tilde{\gamma}) = (x, \gamma)$. It follows that $s(x, \gamma) = \tilde{\gamma}$ is a lift of the path $\gamma \in Y^I$ starting at $f(x)$ to a path $\tilde{\gamma} \in X^I$ starting at x. We may keep the following picture in mind (Figure 5.2).

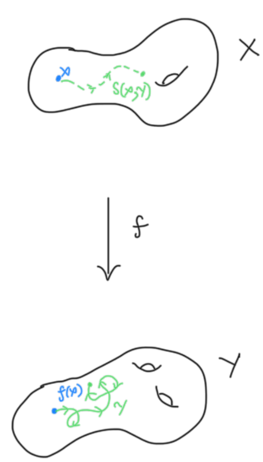


Figure 5.2: Path lifter s taking (x, γ) downstairs to a lift $s(x, \gamma)$ in X upstairs.

Remark 5.3.1.7. (*Covering maps have a unique path lifter*). Recall that a covering space $p : E \to B$ has *unique* homotopy lifting property, hence in particular it is a cofibration. Furthermore recall that a covering space also has *unique* path lifting property, hence in particular it has a unique path-lifter.

We have the following reduction of fibration criterion to mapping path space.

Proposition 5.3.1.8. *Let $p : E \to B$ be a surjective map. Then the following are equivalent:*
1. *p is a fibration.*
2. *p satisfies homotopy lifting property for the natural projection map $Np \to E$.*

Proof. 1. $\Rightarrow$ 2. is definition. For 2. $\Rightarrow$ 1. we proceed as follows. Consider the following diagram

$$\begin{tikzcd} Y \arrow[r, "f"] \arrow[d, "i_0"'] & E \arrow[d, "p"] & Np \arrow[l, "\pi"'] \arrow[d, "\eta"] \arrow[dl, phantom, "\llcorner", very near start] \\ Y \times I \arrow[r, "h"'] & B & B^I \arrow[l, "p_0"] \end{tikzcd} .$$

We may write $h : Y \times I \to B$ as $h^T : Y \to B^I$. Observe that $p_0 h^T = pf$, leading to the following unique map $\kappa : Y \to Np$ as below

$$\begin{tikzcd} Y \arrow[r, dashed, "\kappa"] \arrow[rr, bend left, "f"] \arrow[dr, "h^T"'] & Np \arrow[r, "\pi"] \arrow[d, "\eta"] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\ & B^I \arrow[r, "p_0"'] & B \end{tikzcd} .$$

Similar to h^T, we also have $\eta^T : Np \times I \to B$. It is immediate from $\eta\kappa = h^T$ that $\eta^T(\kappa \times \mathrm{id}) = h : Y \times I \to B$. Consequently, we have the following commutative diagram

$$\begin{tikzcd} Y \arrow[r, "\kappa"] \arrow[rr, bend left, "f"] \arrow[d, "i_0"'] & Np \arrow[r, "\pi"] \arrow[d, "i_0"'] & E \arrow[d, "p"] \\ Y \times I \arrow[r, "\kappa\times\mathrm{id}"'] \arrow[rr, bend right, "h"'] & Np \times I \arrow[r, "\eta^T"'] \arrow[ur, dashed, "\tilde{\eta}^T"'] & B \end{tikzcd}$$

and composing $\tilde{\eta}^T$ with $\kappa \times \mathrm{id}$ yields the required lift of h. $\square$

Proposition 5.3.1.9. *Let $p : E \to B$ be a map. Then the following are equivalent:*
1. *$p : E \to B$ is a fibration.*
2. *There exists a path lifter $s : Np \to E^I$.*

Proof. The forward direction is immediate from dualizing the homotopy lifting property into mappings into path space. For the converse, use Proposition 5.3.1.8. $\square$

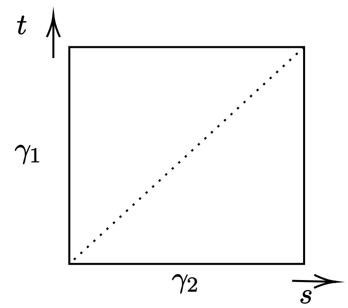
We see that map that the canonical maps $p_0, p_1 : Y^I \to Y$ is a fibration.

Lemma 5.3.1.10. *Let Y be a space. The map*

$$\begin{aligned} p_0 : Y^I &\longrightarrow Y \\ \gamma &\longmapsto \gamma(0) \end{aligned}$$

is a fibration.

Proof. By Proposition 5.3.1.9, it suffices to show that there is a path lifter $s : Np_0 \to Y^{I \times I}$, i.e. a global section of $k : Y^{I\times I} \to Np_0$ mapping $h(s,t) \mapsto (h(s,0), h(0,t))$. Indeed, we define $s((\gamma_1, \gamma_2))$ for $\gamma_i \in Y^I$ such that $\gamma_1(0) = \gamma_2(0)$ by the following homotopy square:



This gives us a map $h \in Y^{I\times I}$ such that $h(0,t) = \gamma_1$ and $h(s,0) = \gamma_2$. This completes the proof. $\square$

Let $f : X \to Y$ be an arbitrary map of spaces. We can replace f by a homotopy equivalence followed by a fibration.

Construction 5.3.1.11 (Replacement by a homotopy equivalence and a fibration)**.** Let $f : X \to Y$ be a map. Consider the following commutative triangle

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "\nu"'] & & Y \\ & Nf \arrow[ur, two heads, "\rho"'] & \end{tikzcd}$$

where

$$\begin{aligned} \nu : X &\longrightarrow Nf \\ x &\longmapsto (x, c_{f(x)}) \end{aligned}$$

and

$$\begin{aligned} \rho : Nf &\longrightarrow Y \\ (x,\gamma) &\longmapsto \gamma(1). \end{aligned}$$

We now make the following claims:

1. Map ν is a homotopy equivalence. Indeed, consider the natural projection map $\pi : Nf \to X$ given by $(x,\gamma) \mapsto x$. We claim that π is a homotopy inverse of ν. Indeed, $\pi\nu = \mathrm{id}_X$ is immediate. We claim $\nu\pi \simeq \mathrm{id}_{Nf}$. Indeed, we may consider the homotopy

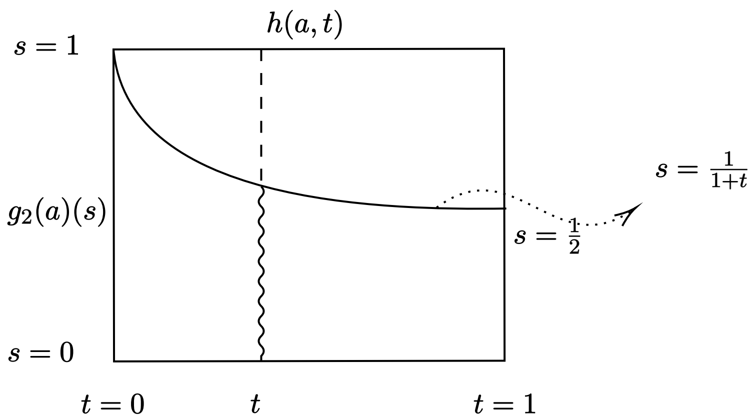
$$\begin{aligned} h : Nf \times I &\longrightarrow Nf \\ ((x,\gamma),t) &\longmapsto (x, \gamma_t) \end{aligned}$$

where $\gamma_t(s) = \gamma((1-t)s)$.

2. Map ρ is a fibration. Let $g : A \to Nf$ be a map such that the following square commutes

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, "i_0"'] & Nf \arrow[d, "\rho"] \\ A \times I \arrow[r, "h"'] & Y \end{tikzcd} .$$

We wish to construct $\tilde{h} : A \times I \to Nf$ which would lift h. Indeed, let $g(a) = (g_1(a), g_2(a))$ where $g_1 : A \to X$ and $g_2 : A \to Y^I$ are the component functions. In order to construct $\tilde{h}$, we need only construct $\alpha : A \times I \to Y^I$ and $\beta : A \times I \to X$ such that the following holds (these are obtained by unravelling $\rho\tilde{h} = h$, $\tilde{h}i_0 = g$ and the respective pullback square):
(a) $f\beta = p_0\alpha$,
(b) $\beta(a, 0) = g_1(a)$,
(c) $\alpha(a, 0) = g_2(a)$,
(d) $\alpha(a, t)(1) = h(a, t)$.
We may immediately set $\beta(a, t) = g_1(a)$. For $\alpha : A \times I \to Y^I$, we may dually write α as $\alpha : A \times I \times I \to Y$ (recall we are in compactly generated spaces, where the dual notion of homotopy is equivalent to the usual one). We construct this α as follows. Fix $a \in A$. We then define the following homotopy



which more explicitly is given by

$$\alpha(a, t, s) = \begin{cases} g_2(a)(s \cdot (1+t)) & \text{if } 0 \leq s \leq \frac{1}{1+t} \\ h(a, s \cdot (1+t) - 1) & \text{if } \frac{1}{1+t} \leq s \leq 1. \end{cases}$$

One can then observe that this α satisfies conditions (a), (c) and (d) mentioned above.

5.3.2 Bundles and change of fibers

We now see that, under some mild hypothesis, fibration is a local property on base. As a consequence, we will show that under some mild hypothesis any bundle (Definition 3.7.1.2) is a fibration.

An open cover $\{U_\alpha\}$ of B will be called *numerable* if for each α, there is a map $f_\alpha : B \to I$ such that $f_\alpha^{-1}((0, 1]) = U_\alpha$ and $\{U_\alpha\}$ is a locally finite cover.

Theorem 5.3.2.1. *Let $p : E \to B$ be a map and $\{U_\alpha\}$ be a numerable open cover of B. Then the following are equivalent:*

1. *$p : E \to B$ is a fibration.*
2. *$p : p^{-1}(U_\alpha) \to U_\alpha$ is a fibration for each α.*

Proof. 1. $\Rightarrow$ 2. is immediate from Lemma 5.3.1.2.

(2. $\Rightarrow$ 1.) The main idea is to patch up the lifts of a homotopy that we obtain by virtue of each $p|_{p^{-1}(U_\alpha)}$ being a fibration. **TODO**. $\square$

We claimed in the beginning that fibrations are upto homotopy generalizations of covering spaces/certain bundles. We know that such objects have homeomorphic fibres (say, when base is path-connected). This fact can be generalized to fibrations which would yield that fibres of a fibration may not be homeomorphic, but will be of same homotopy type!

Construction 5.3.2.2. (*Homotopy invariance of path-lifting for fibrations*). We now show that a path γ in the base gives a map of fibers which is invariant under homotopy class of γ.

In particular, let $p : E \to B$ be a fibration and $\gamma : I \to B$ be a path from b to b' in B. Let E_b and $E_{b'}$ be fibers at b and b' respectively under p. We claim that we get a map $\tilde{\gamma} : E_b \to E_{b'}$ whose homotopy class is independent of the path γ upto homotopy.

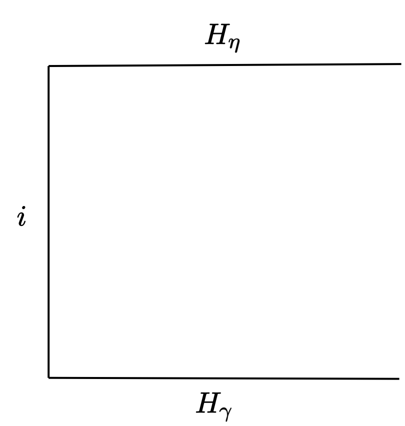
We first construct $\tilde{\gamma} : E_b \to E_{b'}$. Indeed, we have the following diagram

$$\begin{tikzcd}
E_b \arrow[rr, "i"] \arrow[d, "i_0"'] & & E \arrow[d, "p"] \\
E_b \times I \arrow[r, "\pi_2"'] \arrow[urr, dashed, "H_\gamma"] & I \arrow[r, "\gamma"'] & B
\end{tikzcd}$$

by virtue of fibration p. Observe that $H_{\gamma,1}(e) = H_\gamma(e,1)$ is such that $pH_\gamma(e,1) = \gamma(1) = b'$ for all $e \in E_b$. Consequently, $\tilde{\gamma} = H_{\gamma,1} : E_b \to E_{b'}$ is the required map. This shows the construction of $\tilde{\gamma}$. We now show that its homotopy class is invariant of homotopy class of γ.

Let $\gamma, \eta \in B^I$ be two paths joining b and b' together with a homotopy $h : I \times I \to B$ rel $\{0,1\}$ such that $h_0 = \gamma$ and $h_1 = \eta$, that is h is a homotopy between γ and η through paths joining b and b'. We wish to show that $\tilde{\gamma}$ and $\tilde{\eta}$ are homotopy equivalent as well. To this end, we need to construct a homotopy $\tilde{h} : E_b \times I \to E_{b'}$ satisfying $\tilde{h}_0 = \tilde{\gamma} = H_{\gamma,1}$ and $\tilde{h}_1 = \tilde{\eta} = H_{\eta,1}$.

Fix an $e \in E_b$. Our goal is to fill the right side of this square continuously with $e \in E_b$



where $i : E_b \hookrightarrow E$ the inclusion. To this end, we first observe that there is a homeomorphism of pairs

$$(I \times I, S) \xrightarrow{\alpha} (I \times I, I \times 0)$$

where S is the union of three sides of the square as shown above; $S = I \times \{0,1\} \cup \{0\} \times I$. Using this homeomorphism, we obtain the following square

$$\begin{tikzcd} E_b \times S \arrow[rr, "f"] \arrow[d, "k"'] & & E \arrow[d, "p"] \\ E_b \times I \times I \arrow[r, "\kappa"'] \arrow[urr, dashed, "l"] & I \times I \arrow[r, "h"'] & B \end{tikzcd}$$

where $k = \iota(\mathrm{id} \times \alpha)$ where $\iota : E_b \times (I \times 0) \hookrightarrow E_b \times (I \times I)$ and $\kappa(e,t,s) = \alpha^{-1}(t,s)$. Moreover, $f : E_b \times S \to E$ is defined as in the incomplete square above; on $I \times \{0\}$, f is given by H_γ, on $I \times \{1\}$, f is given by H_η and on $0 \times I$, f is given by i. .Observe that $\kappa k(e,t,s) = (t,s)$. The fact that this is a commutative square is immediate. It follows from p being a fibration that there is a lift $l : E_b \times I \times I \to E$ which fits in the above commutative square. Consequently, we have $pl = h\pi_2$ and $lk = f$. By appropriately composing l with α and replacing l with this composition, we get that $l : E_b \times I \times I \to E_{b'}$ which is given by following schematic homotopy cube, which we leave the reader to draw. Consequently, we get the following map $\tilde{h} : E_b \times I \to E_{b'}$ where

$$\tilde{h}(-,s) := l(-,1,s) : E_b \times I \to E_{b'}$$

where $l(e,1,s) \in E_{b'}$ because $h(1,s) \in b'$ (h is a homotopy through paths joining b and b'). Moreover, $\tilde{h}(e,0) = l(e,1,0) = H_{\gamma,1}(e) = \tilde{\gamma}(e)$ and $\tilde{h}(e,1) = H_{\eta,1}(e) = \tilde{\eta}(e)$. Thus, $\tilde{h}$ is the required homotopy between $\tilde{\gamma}$ and $\tilde{\eta}$.

5.3.3 Based fiber sequences

Just as for cofibrations, we had a long cofiber sequence, similarly we have a long fiber sequence for a map of based spaces. As is customary, for based case, we change the definition of mapping path space of $f : X \to Y$, to $Nf = \{(x,\gamma) \mid f(x) = \gamma(1)\}$. We thus define homotopy fiber of a map and construct the short and long fiber sequences of a map.

Definition 5.3.3.1 (**Homotopy fiber/Mapping path space**)**.** Let $f : X \to Y$ be a based map of based spaces. The homotopy fiber of f, denoted Ff, is the following pullback space:

$$\begin{tikzcd} Ff \arrow[r, "\pi"] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & X \arrow[d, "f"] \\ PY \arrow[r, "p_1"'] & Y \end{tikzcd}.$$

Remark 5.3.3.2 (Homotopy fiber is the fiber of mapping path space)**.** Let $f : X \to Y$ be a based map. Denote $Nf = X \times_Y Y^I = \{(x,\gamma) \in X \times Y^I \mid f(x) = \gamma(1)\}$ to be the mapping path space of Y. Then, we have a map

$$\begin{aligned} q : Nf &\longrightarrow Y \\ (x,\gamma) &\longmapsto \gamma(0). \end{aligned}$$

As the base point of Nf is $(*, c_*)$ which is mapped to $*$ under q, thus, q is based as well. Moreover, the fiber of q is

$$q^{-1}(*) = \{(x, \gamma) \mid f(x) = \gamma(1) \;\&\; \gamma(0) = *\}.$$

Hence, $q^{-1}(*) = Ff$, as required.

We first see why this is called homotopy fiber.

Lemma 5.3.3.3. *Let $f : X \to Y$ be a based map of based spaces.*

1. *The map $\pi : Ff \to X$ is a fibration.*
2. *If $\rho : Nf \to Y$ is the fibration replacement of f (Construction 5.3.1.11) where Nf is the mapping path space of f, then*

$$\rho^{-1}(*) = Ff.$$

Proof. For item 1, consider the fibration $p_1 : PY \to Y$ (Lemma 5.3.1.10). By Lemma 5.3.1.2, we see that $\pi : Ff \to X$ as above is a fibration. For item 2, recall that $\rho(x, \gamma) = \gamma(0)$. Thus, we have $\rho^{-1}(*) = \{(x, \gamma) \in Nf \mid \gamma(0) = *, \ \gamma(1) = f(x)\}$. But this is exactly the fiber Ff as PY is the based path space. □

We expect the fiber of a fibration to be homotopy equivalent to the homotopy fiber. Indeed it is true.

Proposition 5.3.3.4. *Let $p : E \to B$ be a based fibration. Then the fiber $F := p^{-1}(*)$ is based homotopy equivalent to homotopy fiber Fp.*

Proof. Let $F = p^{-1}(*)$. Consider the map

$$\begin{aligned} \phi : F &\longrightarrow Fp \\ e &\longmapsto (e, c_*). \end{aligned}$$

Indeed as $p_1(c_*) = * = p(e)$, so $(e, c_*) \in Fp$. To construct a homotopy inverse, we will begin from the mapping path space of p. Recall from Remark 5.3.3.2 that Fp is the fiber of mapping path space $q : Np \to B$, $(e, \gamma) \mapsto \gamma(0)$. Consider the following homotopy

$$\begin{aligned} H : Np \times I &\longrightarrow B \\ ((e, \gamma), t) &\longmapsto \gamma(1 - t). \end{aligned}$$

Observe that the following map commutes where the top horizontal map is $(e, \gamma) \mapsto e$, so that we get $\tilde{H}$ as shown:

$$\begin{tikzcd} Np \arrow[r] \arrow[d, "i_0"'] & E \arrow[d, "p"] \\ Np \times I \arrow[r, "H"'] \arrow[ur, dashed, "\tilde{H}"] & B \end{tikzcd} .$$

Define the following homotopy using $\tilde{H}$:

$$\begin{aligned} G : Fp \times I &\longrightarrow Fp \\ ((e, \gamma), t) &\longmapsto \left(\tilde{H}((e, \gamma), t), \, \gamma|_{[0,1-t]}\right). \end{aligned}$$

Indeed, as $p(\tilde{H}((e,\gamma),t) = H((e,\gamma),t) = \gamma(1-t) = p_1(\gamma|_{[0,1-t]})$, thus G is well-defined. Let $g : Fp \times I \to E$ given by $((e,\gamma),t) \mapsto \tilde{H}((e,\gamma),t)$, that is the first coordinate of homotopy G. Then consider the map

$$\begin{aligned} \psi : Fp &\longrightarrow F \\ (e,\gamma) &\longmapsto g((e,\gamma),1). \end{aligned}$$

Indeed, as $p(\tilde{H}((e,\gamma),1)) = H((e,\gamma),1) = \gamma(1-1) = \gamma(0) = *$ as $(e,\gamma) \in Fp$, thus ψ is well-defined. We claim that ψ is the homotopy inverse of ϕ. Indeed, we have

$$\begin{aligned} \phi \circ \psi : Fp &\longrightarrow Fp \\ (e,\gamma) &\longmapsto (g((e,\gamma),1), c_*). \end{aligned}$$

Observe that $G_1(e,\gamma) = (g((e,\gamma),1), c_*)$ and $G_0 = \mathrm{id}_{Fp}$, so that G forms a homotopy between $\phi \circ \psi$ and id. Conversely, we have

$$\begin{aligned} \psi \circ \phi : F &\longrightarrow F \\ e &\longmapsto g((e,c_*),1) = \tilde{H}((e,c_*),1). \end{aligned}$$

Consider the restriction of G onto the subspace T of elements $((e,c_*),t) \in Fp \times I$. Note that G maps onto T as well. Thus we have $G : T \times I \to T$ and $G_1(e,c_*) = \tilde{H}((e,c_*),1)$ and $G_0 = \mathrm{id}_T$. Moreover, observe that $F \to T$, $e \mapsto (e,c_*)$ is a homeomorphism. Hence the above restriction of G is a homotopy from $\psi \circ \phi$ to id_F. This completes the proof. □

Construction 5.3.3.5 (Fiber sequence)**.** Let $f : X \to Y$ be a based map of based spaces. Consider the following three maps

$$\begin{aligned} \pi : Ff &\longrightarrow X \\ (x,\gamma) &\longmapsto x \\ \iota : \Omega Y &\longrightarrow Ff \\ \gamma &\longmapsto (*,\gamma). \end{aligned}$$

The sequence

$$Ff \xrightarrow{\pi} X \xrightarrow{f} Y$$

is called the *short fiber sequence*.

We can continue the above short fiber sequence into a *long fiber sequence* as follows. Consider the functor $-\Omega : \mathbf{Top}_*^{cg} \to \mathbf{Top}_*^{cg}$ taking X to ΩX and $f : X \to Y$ to $-\Omega f : -\Omega X \to -\Omega Y$ given by $\gamma(t) \mapsto f \circ \gamma(1-t)$. Thus, we get the following sequence of maps

$$\begin{tikzcd}
 & & \arrow[dll] \\
\Omega^2 Ff \arrow[r, "\Omega^2\pi"'] & \Omega^2 X \arrow[r, "\Omega^2 f"] & \Omega^2 Y \arrow[dll, "-\Omega\iota"'] \\
\Omega Ff \arrow[r, "-\Omega\pi"'] & \Omega X \arrow[r, "-\Omega f"] & \Omega Y \arrow[dll, "\iota"'] \\
Ff \arrow[r, two heads, "\pi"'] & X \arrow[r, "f"'] & Y
\end{tikzcd}$$

which we call the long fiber sequence of $f : X \to Y$.

The main theorem is the following, which associates an exact sequence of based sets to the long fiber sequence.

Theorem 5.3.3.6 (Main theorem of fiber sequences)**.** *Let $f : X \to Y$ be a based continuous map of based spaces and Z be a based space. Then, the long cofiber sequence of f induces a long exact sequence of based homotopy sets:*

$$
\begin{tikzcd}
{[Z, \Omega^2 Ff]} \arrow[r] & {[Z, \Omega^2 X]} \arrow[r] & {[Z, \Omega^2 Y]} \arrow[dll] \\
{[Z, \Omega Ff]} \arrow[r] & {[Z, \Omega X]} \arrow[r] & {[Z, \Omega Y]} \arrow[dll] \\
{[Z, Ff]} \arrow[r, "\pi_*"'] & {[Z, X]} \arrow[r, "f_*"'] & {[Z, Y]}
\end{tikzcd}
$$

Taking $Z = S^0$ and recalling the suspension-loop space adjunction (Proposition 5.0.0.10), we immediately get the following long exact sequence of homotopy groups.

Corollary 5.3.3.7 (Homotopy L.E.S.-1)**.** *Let $f : X \to Y$ be a based map of based space. Then the fiber sequence of f induces the following long exact sequence of homotopy groups (basepoint suppressed):*

$$
\begin{tikzcd}
\pi_2(Ff) \arrow[r, "\pi_*"'] & \pi_2(X) \arrow[r, "f_*"] & \pi_2(Y) \arrow[dll, "\partial"'] \\
\pi_1(Ff) \arrow[r, "\pi_*"'] & \pi_1(X) \arrow[r, "f_*"] & \pi_1(Y) \arrow[dll, "\partial"'] \\
\pi_0(Ff) \arrow[r, "\pi_*"'] & \pi_0(X) \arrow[r, "f_*"] & \pi_0(Y)
\end{tikzcd}
.
$$

□

5.3.4 Serre spectral sequence

For any fibration (more generally, for Serre fibration) $p : E \to B$, there is a spectral sequence converging to homology of the total space E.

Theorem 5.3.4.1. *Let $F \xrightarrow{i} E \xrightarrow{\pi} B$ be a Serre fibration with fiber F. If B is simply connected, then there is a first quadrant homology spectral seqeunce converging to homology of E:*

$$E^2_{pq} = H_p(B; H_q(F)) \Rightarrow H_{p+q}(E).$$

See cite[HopSSeq] for a proof. We will see some applications of the above spectral sequence below.

Theorem 5.3.4.2 (Loop fibration). *Let $\Omega B \to PB \xrightarrow{\pi} B$ be the loop space fibration where $\pi(\gamma) = \gamma(1)$ (see Lemma 5.3.1.10). Then,*

1. *$H_1(\Omega B; \mathbb{Z}) \cong H_2(B; \mathbb{Z})$,*
2. *there is an exact sequence*

$$H_4(B) \to H_2(B) \otimes H_2(B) \to H_2(\Omega B) \to H_3(B) \to 0.$$

Theorem 5.3.4.3 (Fibrations over S^n/Wang sequence). *Let $F \xrightarrow{i} E \xrightarrow{\pi} S^n$ be a fibration for $n \geq 2$. Then there is a long exact sequence*

$$\begin{tikzcd}
H_{q-n+1}(F) \arrow[r, "d^{n-1}"'] & H_q(F) \arrow[r, "i_*"] & H_q(E) \arrow[dll] \\
H_{q-n}(F) \arrow[r, "d^n"'] & H_{q-1}(F) \arrow[r, "i_*"] & H_{q-1}(E)
\end{tikzcd}.$$

Theorem 5.3.4.4 (Sphere fibrations/Gysin sequence). *Let $S^n \xrightarrow{i} E \xrightarrow{\pi} B$ be a fibration for $n \geq 1$ and B be simply connected. Then there is a long exact sequence*

$$\begin{tikzcd}
H_{p-n}(B) \arrow[r] & H_p(E) \arrow[r, "\pi_*"] & H_p(B) \arrow[dll, "d^{n+1}"] \\
H_{p-n-1}(B) \arrow[r] & H_{p-1}(E) \arrow[r, "\pi_*"] & H_{p-1}(B)
\end{tikzcd}.$$

We discuss some more general properties now.

Useful properties of Serre spectral sequence

Acyclic fiber theorem

Theorem 5.3.4.5 (Acyclic fiber)**.** *Let $f : X \to Y$ be a based map between connected CW-complexes. Then the following are equivalent:*

1. *For all $k \geq 0$, we have*

$$f_* : H_k(X; M) \xrightarrow{\cong} H_k(Y; M)$$

   for every $\pi_1(Y)$-module M[10]*.*

2. *The homotopy fiber Ff of f is acyclic*[11]*.*

Proof. □

[10] That is, M is a left $\mathbb{Z}[\pi_1(Y)]$-module.
[11] that is, Ff has homology of a point.

5.4 Homology theories

We will begin by introducing (co)homology from an axiomatic point of view and will derive few properties off of it. This will come in handy for discussing the main properties of differential manifolds in (co)homological language, especially characteristic classes and orientations and what not. The main thing that we wish to do is the Hurewicz theorem, which will allow us to connect homotopy groups and homology groups on the one hand, and will allow us to prove the uniqueness of homology theories for CW complexes on the other hand.

All spaces X are assumed to be compactly generated (Definition 5.0.0.1).

We will use the theory of cofibrations and fibrations as developed above quite freely.

5.4.1 Homology theories

We begin with the category of pairs on which homology theories are defined.

Definition 5.4.1.1 ($\mathbf{Top}_2$). The $\mathbf{Top}_2$ denotes the category of pairs (X, A) of spaces where $A \hookrightarrow X$ and maps $(X, A) \to (Y, B)$ which consists of the pair $f : X \to Y$ and $g : A \to B$ such that $g = f|_A$. A map of pairs $(f, d) : (X, A) \to (Y, B)$ is said to be a homotopy equivalence if there is a map of pairs $(g, e) : (Y, B) \to (X, A)$ and there are homotopies $H : g \circ f \simeq \mathrm{id}_X$ and $K : f \circ g \simeq \mathrm{id}_Y$ which extends the homotopies $h : e \circ d \simeq \mathrm{id}_A$ and $k : d \circ e \simeq \mathrm{id}_B$ respectively.

Definition 5.4.1.2. (**Homology theory**) A homology theory for an abelian group π is a sequence of functors

$$H_q(-,-;\pi) : \mathbf{Top}_2 \longrightarrow \mathbf{AbGrp}$$

for $q \in \mathbb{Z}$ equipped with natural transformations

$$\partial : H_q(-,-;\pi) \longrightarrow H_{q-1}(-,-;\pi)$$

whose component at (X, A) is given by $\partial : H_q(X, A;\pi) \to H_{q-1}(A, \emptyset;\pi)$. Denote $H_q(X;\pi) := H_q(X, \emptyset;\pi)$. This data must satisfy the following axioms:

1. *(Homology of a point)* : If $X = \{\text{pt.}\}$, then homology must be concentrated at degree 0:

$$H_q(\{\text{pt.}\};\pi) = \begin{cases} \pi & \text{if } q = 0, \\ 0 & \text{if } q \neq 0. \end{cases}$$

2. (*Homology long exact sequence*) : The trivial inclusions $A \hookrightarrow X$ and $(X, \emptyset) \hookrightarrow (X, A)$ induces the following long exact sequence:

$$\begin{tikzcd}
& & \cdots \arrow[dll, dashed, "\partial"'] \\
H_q(A;\pi) \arrow[r] & H_q(X;\pi) \arrow[r] & H_q(X,A;\pi) \arrow[dll, dashed, "\partial"'] \\
H_{q-1}(A;\pi) \arrow[r] & H_{q-1}(X;\pi) \arrow[r] & H_{q-1}(X,A;\pi) \arrow[dll, dashed, "\partial"'] \\
\cdots & &
\end{tikzcd}.$$

3. *(Excision invariance)* : For an excisive triple (X, A, B), that is $A, B \hookrightarrow X$ and $X = A^\circ \cup B^\circ$, the inclusion $(A, A \cap B) \hookrightarrow (X, B)$ induces an isomorphism at all degree $q \in \mathbb{Z}$:

$$H_q(A, A \cap B; \pi) \xrightarrow{\cong} H_q(X, B; \pi) \ .$$

4. *(Coproduct preserving)* : If (X_i, A_i) is an arbitrary collection of objects in $\mathbf{Top_2}$, then the homology in any degree of their disjoint union is the sum of the corresponding homology groups:

$$\bigoplus_i H_q(X_i, A_i; \pi) \xrightarrow{\cong} H_q\left(\coprod_i (X_i, A_i); \pi\right)$$

where the maps are induced by the inclusions $(X_{i_0}, A_{i_0}) \hookrightarrow \coprod_i (X_i, A_i)$.
5. *(π_*-insensitivity)* : If $f : (X, A) \longrightarrow (Y, B)$ is a weak equivalence, then in all degrees the corresponding homology groups are isomorphic:

$$f_p : H_q(X, A; \pi) \xrightarrow{\cong} H_q(Y, B; \pi).$$

Remark 5.4.1.3. In nature, there are some homology theories which satisfy all of the above axioms except the dimension axiom, that is, the group that they assign to a point is not concentrated in degree 0 (axiom 1. above). A famous example of this is K-theory via the Bott-periodicity theorem. One calls such a homology theory to be a *generalized homology theory*. All results that we will derive here will hold true for a generalized homology theory E_q.

General properties

We now discuss some general properties of homology theories that one can deduce from the axioms.

Proposition 5.4.1.4. *Let π be a group and E_q be a generalized homology theory. Let X be a space.*

1. *If $A \hookrightarrow X \xrightarrow{r} A$ is a retract of X, then the following natural maps form a short-exact sequence of E-homology groups:*

$$0 \to E_q(A) \to E_q(X) \to E_q(X, A) \to 0.$$

2. *$E_q(X, X) \cong 0$.*

Proof. 1. The fact that $E_q(A) \to E_q(X)$ is injective follows from a set theoretic observation; any factorization of identity is a monic followed by an epic. By homology long-exact sequence, we then have that all boundary maps ∂ are trivial. It follows that maps $E_q(X) \to E_q(X, A)$ is surjective. The exactness at middle is given by the homology long-exact sequence.
2. Since X is always a retract of itself, therefore from item 1, it follows that $E_q(X, X) \cong E_q(X)/E_q(X) \cong 0$. □

The following is a long exact sequence in homology that one obtains from a *triplet* (X, A, B) where $X \supseteq A \supseteq B$.

Proposition 5.4.1.5 (Triplet long-exact sequence)**.** *Let (X, A, B) be a triplet and denote $i : (A, B) \hookrightarrow (X, B)$ and $j : (X, B) \hookrightarrow (X, A)$ to be inclusions. Also denote $\partial' : E_q(X, A) \to E_{q-1}(A, B)$ to be the composite $E_q(X, A) \xrightarrow{\partial} E_{q-1}(A) \to E_{q-1}(A, B)$. Then there is a long exact sequence*

$$
\begin{tikzcd}
E_q(A,B) \arrow[r, "i_*"] & E_q(X,B) \arrow[r, "j_*"] & E_q(X,A) \arrow[dll, "\partial'"'] \\
E_{q-1}(A,B) \arrow[r, "i_*"] & E_{q-1}(X,B) \arrow[r, "j_*"] & E_{q-1}(X,A)
\end{tikzcd}
.
$$

Proof. This follows from a fairly long diagram chase involving the homology long-exact sequence corresponding to each of the pairs (A, B), (X, B) and (X, A) which one has to expand for degrees q and $q-1$. From that big diagram, the chase is straightforwad after some reductions and hence is omitted. $\square$

There is an equivalent form of excision which is also quite useful.

Lemma 5.4.1.6 (Excision-II)**.** *Let $(X, A) \in \mathbf{Top}_2$ be a pair and E_q be a homology theory. If $B \subseteq A$ is a subspace such that $\bar{B} \subseteq A^\circ$, then B can be excised, that is, the inclusion*

$$(X - B, A - B) \hookrightarrow (X, A)$$

induces an isomorphism in homology:

$$E_q(X - B, A - B; \pi) \cong E_q(X, A; \pi).$$

Proof. Consider the triple $(X, A, X - B)$. This is an excisive triple since $A^\circ \cup (X - B)^\circ = X$ since $(X - B)^\circ = X - \bar{B}$. Thus by excision axiom, the inclusion

$$j : (X - B, A \cap X - B) \hookrightarrow (X, A)$$

induces isomorphism in E_q. As $A \cap (X - B) = A - B$, we get the desired result. $\square$

5.4.2 Reduced homology

For each homology theory $E_q(-, -)$, we can construct a based version of the theory denoted $\tilde{E}_q(-, \mathrm{pt.})$. For a based space $(X, \mathrm{pt.})$, define the following

$$\tilde{E}_q(X) := E_q(X, \mathrm{pt.}).$$

This tends to remove the effect of the defining group of the homology theory, so to normalize the theory in the sense of Lemma 5.4.2.1, 1. In particular, if E_q satisfies dimension axiom, it follows that $E_0(\mathrm{pt.}) = \pi$. Thus this lemma will tell that $\tilde{E}_0(X) = \tilde{E}_0(X) \oplus \pi$.

Let us spell out some basic relations of this reduced homology $\tilde{E}_q$ to that of original homology E_q.

Proposition 5.4.2.1. *Let π be a group and E_q be a generalized homology theory. Let $(X, \text{pt.})$ be a based space and $(A, \text{pt.}) \hookrightarrow (X, \text{pt.})$ be a based subspace.*

1. *$E_q(X) = \tilde{E}_q(X) \oplus E_q(\text{pt.})$ and the map $\iota_* : E_q(A) \to E_q(X)$ restricted on $E_q(\text{pt.})$ is the identity map $\iota_* : E_q(\text{pt.}) \to E_q(\text{pt.})$.*
2. *There is a long exact sequence*

$$\begin{tikzcd}
 & & \dots \arrow[dll, "\partial"'] \\
\tilde{E}_q(A) \arrow[r] & \tilde{E}_q(X) \arrow[r] & E_q(X, A) \arrow[dll, "\partial"'] \\
\tilde{E}_{q-1}(A) \arrow[r] & \tilde{E}_{q-1}(X) \arrow[r] & E_{q-1}(X, A) \arrow[dll, "\partial"'] \\
\dots & &
\end{tikzcd}.$$

3. *If E_q is an ordinary homology theory, then for any $q \geq 2$, we have*

$$\tilde{E}_q(X) \cong E_q(X).$$

Proof. 1. The following is split exact on the left as the map $\text{pt.} \hookrightarrow X$ is a retract (Proposition 5.4.1.4):

$$0 \to E_q(\text{pt.}) \to E_q(X) \to E_q(X, \text{pt.}) \to 0.$$

Note that the left map here is split by the retraction $r_* : E_q(X) \to E_q(\text{pt.})$. The latter statement follows from the fact that $E_q(-, \emptyset)$ is a functor and thus takes $\text{id}_{\text{pt.}}$ to $\text{id} : E_q(\text{pt.}) \to E_q(\text{pt.})$.
2. Consider $i : A \hookrightarrow X$. Then $E_q(A) \to E_q(X)$ takes $E_q(\text{pt.})$ to $E_q(\text{pt.})$ isomorphically as in item 1. Hence we may quotient it out under the exactness to get the desired sequence.
3. This is immediate from long exact sequence of the pair $(X, \text{pt.})$. □

In-fact, one can obtain the unreduced homology back by reduced homology via a simple use of coproduct preservation axiom.

Lemma 5.4.2.2. *Let X be a space and denote X_+ to be the based space obtained by disjoint union of X with a point* pt.. *For any generalized homology theory E_q, we have*

$$E_q(X) \cong \tilde{E}_q(X_+).$$

Proof. As $X_+ = X \amalg \{\text{pt.}\}$, therefore by additivity of homology theories, we obtain

$$\begin{aligned}
\tilde{E}_q(X_+) &= E_q(X \amalg \{\text{pt.}\}, \text{pt.}) = E_q((X, \text{pt.}) \amalg (\text{pt.}, \text{pt.})) \cong E_q(X, \text{pt.}) \oplus E_q(\text{pt.}, \text{pt.}) \\
&\cong \tilde{E}_q(X) \oplus E_q(\text{pt.}) \cong E_q(X)
\end{aligned}$$

where the second-to-last isomorphism follows from Proposition 5.4.2.1, 1 and the last from 4. □

5.4.3 Mayer-Vietoris sequence in homology

We now cover an important calculational tool for generaized homology theories, which relates the homology groups of X with those of A, B and $A \cap B$ where (X, A, B) forms an excisive triad.

Theorem 5.4.3.1 (Mayer-Vietoris for homology)**.** *Let (X, A, B) be an excisive triple and denote $i : A \cap B \hookrightarrow A$, $j : A \cap B \hookrightarrow B$, $k : A \hookrightarrow X$ and $l : B \hookrightarrow X$. Then there is a long exact sequence*

$$\begin{tikzcd}
E_q(A \cap B) \arrow[r, "\begin{bmatrix} i_* \\ j_* \end{bmatrix}"] & E_q(A) \oplus E_q(B) \arrow[r, "\begin{bmatrix} k_* - l_* \end{bmatrix}"] & E_q(X) \arrow[dll, "\overline{\partial}"] \\
E_{q-1}(A \cap B) \arrow[r] & E_{q-1}(A) \oplus E_{q-1}(B) \arrow[r] & E_{q-1}(X) \arrow[dll, dashed] \\
{} & &
\end{tikzcd}$$

where $\overline{\partial}$ is obtained as the following composite

$$\begin{tikzcd}
E_q(X) \arrow[r] \arrow[d, dashed, "\overline{\partial}"'] & E_q(X, B) \arrow[d, "\cong"] \\
E_{q-1}(A \cap B) & E_q(A, A \cap B) \arrow[l, "\partial"]
\end{tikzcd}$$

where top horizontal arrow is corresponds to $(X, \emptyset) \hookrightarrow (X, B)$, the right vertical is exicision isomorphism and the bottom horizontal is the boundary map of homology long exact sequence of the pair $(A, A \cap B)$.

Proof. The proof will follow from excision and long exact sequence for homology quite easily. **TODO.** □

5.4.4 Relative homology of cofibrations and suspension isomorphism

There are two important results for homology. The first affirms our intuition that the homology of pair (X, A) ought to behave as homology of X/A, but it works out only when $A \hookrightarrow X$ is a cofibration. The second gives a suspension isomorphism type result akin to that of homotopy groups.

Relative homology of cofibrations

Theorem 5.4.4.1. *Let $i : A \hookrightarrow X$ be a cofibration and E_q a generalized homology theory. Then the quotient map $p : (X, A) \twoheadrightarrow (X/A, \mathrm{pt}.$ induces an isomorphism*

$$p_* : E_q(X, A) \xrightarrow{\cong} E_q(X/A).$$

Suspension isomorphism

Theorem 5.4.4.2. *Let (X, x_0) be a non-degenerately based space, that is, the inclusion $\{x_0\} \hookrightarrow X$ is a cofibration. Let E_q be a generalized homology theory. Then, there is a natural isomorphism*

$$\tilde{E}_q(\Sigma X) \cong \tilde{E}_{q-1}(X).$$

5.4.5 Fundamental theorem of homology theories

We will now see that reduced homology and unreduced homology theories are equivalent. To this end, we first axiomatize reduced homology theory. The category $\mathbf{Top}_*$ denotes the category of well-pointed spaces.

Definition 5.4.5.1 (**Reduced homology theory**). A reduced homology theory for an abelian group π is a sequence of functors

$$\tilde{H}_q(-;\pi) : \mathbf{Top}_* \longrightarrow \mathbf{AbGrp}$$

for $q \in \mathbb{Z}$ which satisfies the following axioms (we suppress π):

1. *(Cofibration exactness)* If $i : A \hookrightarrow X$ is a cofibration, then

$$\tilde{H}_q(A) \to \tilde{H}_q(X) \to \tilde{H}_q(X/A)$$

is exact.

2. *(Suspension isomorphism)* For all $q \geq 0$, we have a natural isomorphism

$$\Sigma : \tilde{H}_q(X) \stackrel{\cong}{\to} \tilde{H}_{q+1}(\Sigma X).$$

3. *(Additivity)* If $X = \bigvee_{i\in I} X_i$ where each X_i is well-pointed, then the natural inclusions $\iota_i : X_i \hookrightarrow X$ induces an isomorphism

$$\bigoplus_{i\in I} \tilde{H}_q(X_i) \cong \tilde{H}_q(X)$$

4. *(Weak equivalence)* If $f : X \to Y$ is a weak equivalence, then

$$f_* : \tilde{H}_q(X) \to \tilde{H}_q(Y)$$

is an isomorphism.

5.4.6 Singular homology & applications

We define the usual singular homology groups and will mention that it is a homology theory. Once that's set-up, then with the explicit description of chain complexes in singular homology and the ES-axioms and all the surrounding results, we will have a good toolbox to compute homology groups of very many spaces. In-fact, these applications are important to really showcase that if in any situation we have an invariant of any class of objects which is a homology theory, then we can immediately make this invariant very palpatable to calculations, which is very important in aspects where the objects are abstract entities like rings or schemes.

For this section, we may assume that our spaces are not compactly generated.

Definition 5.4.6.1 (**Singular homology**). Let X be a space and fix a field F. Let $S_i(X)$ be the free F-vector space generated by the set of all i-simplices $\{f : \Delta^i \to X \mid f \text{ is continuous}\}$. An element of $S_i(X)$ is called *singular i-chain*. Consider the map $\partial : S_i(X) \to S_{i-1}(X)$ which on an i-simplex σ is given by $\sigma \mapsto \sum_{j=0}^{i}(-1)^j \partial_j \sigma$ where $\partial_j\sigma$ is the σ restricted to the face opposite to j^{th}-vertex. It follows that $\partial^2 = 0$. Thus, we have a chain complex $(S_i(X), \partial)$, called the singular chain complex. The homology of this chain complex is defined to be the singular homology of X, denoted $H_i(X;\mathbb{Z})$ or simply $H_i(X)$. A map $f : X \to Y$ on spaces yields a map on singular complex $f_\sharp : S_\bullet(X) \to S_\bullet(Y)$. As map of complexes induces map on homology, we get $f_* : H_\bullet(X) \to H_\bullet(Y)$.

Let (X,A) be a pair. We define the relative singular i-chains to be

$$S_\bullet(X,A) := S_\bullet(X)/S_\bullet(A).$$

The boundary map of $S_\bullet(X)$ descends to a boundary map on $S_\bullet(X,A)$ by properties of quotients and thus we define the singular homology of a pair (X,A) to be homology of the complex $S_\bullet(X,A)$ denoted $H_i(X,A;\mathbb{Z})$.

In the following result, we state some important first properties of singular homology.

Theorem 5.4.6.2 (Singular homology is a homology theory)**.** *Let X be a space.*

1. *If $\{X_k\}$ is the collection of path-components of X, then*
$$H_i(X;\mathbb{Z}) \cong \bigoplus_k H_i(X_k\, \mathbb{Z}).$$
2. *Singular homology satisfies dimension axiom:*
$$H_i(\{\text{pt.}\};\mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{if } i = 0 \\ 0 & \text{else.} \end{cases}$$
3. *X is path-connected if and only if*
$$H_0(X;\mathbb{Z}) \cong \mathbb{Z}.$$
4. *Singular homology has long exact sequence of pairs, that is, if (X,A) is a pair, then there is a long exact sequence obtained by inclusions $A \hookrightarrow X$ and $(X,\emptyset) \hookrightarrow (X,A)$ as follows:*
$$\begin{tikzcd} & & \cdots \arrow[dll, dashed] \\ H_q(A;\pi) \arrow[r] & H_q(X;\pi) \arrow[r] & H_q(X,A;\pi) \arrow[dll, dashed, "\partial"'] \\ H_{q-1}(A;\pi) \arrow[r] & H_{q-1}(X;\pi) \arrow[r] & H_{q-1}(X,A;\pi) \arrow[dll, dashed, "\partial"'] \\ \cdots & & \end{tikzcd}.$$
5. *Singular homology is excision invariant; for an excisive triple (X,A,B), that is $A,B \hookrightarrow X$ and $X = A^\circ \cup B^\circ$, the inclusion $(A, A\cap B) \hookrightarrow (X,B)$ induces an isomorphism at all degree $q \in \mathbb{Z}$:*
$$H_q(A, A\cap B;\mathbb{Z}) \xrightarrow{\cong} H_q(X,B;\mathbb{Z}) \ .$$
An equivalent restatement is that if $A \supseteq B$ such that $\bar{B} \subseteq A^\circ$, then the inclusion $(X-B, A-B) \hookrightarrow (X,A)$ induces isomorphism in homology
$$H_q(X-B, A-B;\mathbb{Z}) \xrightarrow{\cong} H_q(X,A;\mathbb{Z}).$$

6. *Singular homology preserves coproducts, that is, if* $\{(X_i, A_i)\}_{i\in I}$ *is a collection of pairs of spaces, then*

$$\bigoplus_i H_q(X_i, A_i; \pi) \xrightarrow{\cong} H_q\left(\coprod_i (X_i, A_i); \pi\right)$$

where the maps are induced by the inclusions $(X_{i_0}, A_{i_0}) \hookrightarrow \coprod_i (X_i, A_i)$.

7. *Singular homology satisfies strong* π_**-insensitivity, that is, if* $f, g : X \to Y$ *are two homotopic maps, then* $f_* = g_* : H_i(X; \mathbb{Z}) \to H_i(Y; \mathbb{Z})$.

Proof. 1. Observe that $S_i(X) = \bigoplus_k S_i(X_k)$ by path-connectedness of each X_k. Moreover, $Z_i(X) = \bigoplus_k Z_i(X_k)$ and $B_i(X) = \bigoplus_k B_i(X_k)$. The result follows.

2. First observe that every $S_i(X)$ is isomorphic to $\mathbb{Z}$ as there is only one i-simplex, namely $c_{\text{pt.}}$, the constant map. We have for $c_{\text{pt.}} \in Z_{i+1}(X)$ its boundary as

$$\partial(c_{pt}) = \sum_{j=0}^{i+1} (-1)^j \partial_j(c_{\text{pt.}})$$

where note that the j^{th}-boundary of $c_{\text{pt.}}$ is still $c_{\text{pt.}}$. Thus, if $i+2$ is even, then $\partial : S_{i+1}(X) \to S_i(X)$ is zero and if $i+2$ is odd, then $\partial : S_{i+1}(X) \to S_i(X)$ is an isomorphism. Hence, we get that

$$d_p : S_p(X) \to S_{p-1}(X)$$

is 0 if p is odd and an isomorphism if p is even. From this, it immediately follows that $H_p(\text{pt.}; \mathbb{Z}) = 0$ if $p > 0$ and $H_0(\text{pt.}; \mathbb{Z}) \cong \mathbb{Z}$.

3. (L $\Rightarrow$ R) Let X be a path-connected space. Recall that $H_0(X; \mathbb{Z}) = S_0(X)/\text{Im}\,(\partial_1)$. Consider the following map

$$\begin{aligned} \epsilon : S_0(X) &\longrightarrow \mathbb{Z} \\ \sum_j n_j x_j &\longmapsto \sum_j n_j. \end{aligned}$$

Clearly this is surjective. We claim that $\text{Ker}\,(\epsilon) = \text{Im}\,(\partial_1)$. Suppose $\sum_j n_j x_j \in S_0(X)$ and each x_j is distinct with $\sum_j n_j = 0$. We wish to find a 1-chain $\sigma = \sum_j m_j \sigma_j$ such that $\partial_1 \sigma = \sum_j n_j x_j$. Fix $x_0 \in X$ a point different from x_j and let $\gamma_j : I \to X$ be a path joining x_0 to x_j. Consider $\sigma = \sum_j n_j \gamma_j$. We claim that $\partial\sigma = \sum_j n_j x_j$. Indeed, we have

$$\partial\sigma = \sum_j n_j(\gamma_j(1) - \gamma_j(0)) = \sum_j n_j(x_j - x_0) = \sum_j n_j x_j - \left(\sum_j n_j\right) x_0 = \sum_j n_j x_j,$$

as required.

TODO □

Corollary 5.4.6.3. *The construction of the sequence of functors* $H_k(-,-;\mathbb{Z}) : \mathbf{Top_2} \to \mathbf{AbGrp}$ *is a homology theory.* □

Remark 5.4.6.4 (Mayer-Vietoris sequence for singular homology)**.** Consider a space X and an excisive triple (X, A, B). Then since singular homology is a homology theory, hence we have the Mayer-Vietoris sequence as in Theorem 5.4.3.1. After long exact sequence for pairs, this is the second most important long exact sequence in homology:

$$\begin{tikzcd}
H_q(A \cap B) \arrow[r, "\begin{bmatrix} i_* \\ j_* \end{bmatrix}"] & H_q(A) \oplus H_q(B) \arrow[r, "\begin{bmatrix} k_* - l_* \end{bmatrix}"] & H_q(X) \arrow[dll, "\partial"'] \\
H_{q-1}(A \cap B) \arrow[r] & H_{q-1}(A) \oplus H_{q-1}(B) \arrow[r] & H_{q-1}(X) \arrow[dll, dashed]
\end{tikzcd} .$$

This also holds for reduced homology.

Remark 5.4.6.5 (Triplet long exact sequence for singular homology)**.** Consider a triplet (X, A, B) where $X \supseteq A \supseteq B$. Then since singular homology is a homology theory, hence we get a triplet long exact sequence induced by inclusions as in Theorem 5.4.1.5. This is the third long exact sequence that one derives in singular homology, after l.e.s. of pairs and Mayer-Vietoris. This also holds for reduced homology.

We now showcase a result which we will meet again later, which relates fundamental group and first homology group.

Theorem 5.4.6.6 (Hurewicz for π_1)**.** *Let X be a path-connected space and $x_0 \in X$. The canonical map*

$$\begin{aligned} \varphi : \pi_1(X, x_0) &\longrightarrow H_1(X; \mathbb{Z}) \\ \langle \alpha \rangle &\longmapsto [\alpha] \end{aligned}$$

is surjective with $\operatorname{Ker}(\varphi) = [\pi_1(X, x_0) : \pi_1(X, x_0)]$.

Corollary 5.4.6.7. *Let (X, x_0) be a path-connected space and such that $\pi_1(X, x_0)$ is abelian. Then $\pi_1(X, x_0) \cong H_1(X; \mathbb{Z})$.* □

Remark 5.4.6.8 (Suspension isomorphism)**.** Let X be a space and SX be unreduces suspension. Then we have an isomorphism as in Theorem 5.4.4.2:

$$H_q(SX; \mathbb{Z}) \cong \tilde{H}_{q-1}(X; \mathbb{Z}).$$

One can also directly prove this by analyzing the Mayer-Vietoris for the $X_1 = SX - [x, 1]$ and $X_2 = SX - [x, 0]$.

5.4.7 Results & computations for singular homology

We now present many computations for singular homology theory, which showcases the strength of the tools available.

For this section, we may assume that our spaces are not compactly generated.

Remark 5.4.7.1. We begin with the list of topics that we cover here, for mental clarity and quick reference.

- Path components & relative homology.
- Map of long exact sequence of pairs.
- Immediate applications of Mayer-Vietoris.
- Degree of a map $f : S^n \to S^n$.
- Antipode preserving maps $f : S^n \to S^1$.
- Jordan-Brower separation theorem.

Path components & relative homology

Lemma 5.4.7.2. *Let $A \subseteq X$ be a non-empty subspace and X be path-connected. Then*

$$H_0(X, A; \mathbb{Z}) = 0.$$

Proof. Consider $\bar{d} : S_1(X, A) \to S_0(X, A)$. We claim that $\operatorname{Im}\left(\bar{d}\right) = S_0(X, A)$. Suffices to show that $\operatorname{Im}\left(\bar{d}\right)$ contains the class of generators $x : \Delta_0 \to X$. Pick any x as given. To show that there exists $\sigma + S_1(A) \in S_1(X, A)$ whose boundary is x. Indeed, as X is path-connected, so for any fixed point $x_0 \in A$, we may consider a path σ joinging x_0 to x. This defines an element $\sigma + S_1(A)$ whose boundary is $x - x_0 + S_0(A) = x + S_0(A)$, as needed. $\square$

Lemma 5.4.7.3. *Let $\{X_k\}$ be path-components of X and $A \subseteq X$ be non-empty. Then*

$$H_n(X, A; \mathbb{Z}) \cong \bigoplus_k H_n(X_k, A \cap X_k; \mathbb{Z}).$$

Proof. As $S_n(X, A) = \bigoplus_k S_n(X_k, A \cap X_k)$, the result then follows by quotienting. $\square$

Lemma 5.4.7.4. *Let $A \subseteq X$ be a non-empty subspace, then*

$$\operatorname{rank}(H_0(X, A; \mathbb{Z})) = \#\textit{ path components of } X \textit{ not intersecting } A.$$

Proof. By Theorem 5.4.6.2, 3 and Lemmas 5.4.7.2 and 5.4.7.3, the result is immediate. $\square$

Lemma 5.4.7.5. *Let X have r-path components. Then,*

$$H_0(X, \text{pt.}; \mathbb{Z}) \cong \mathbb{Z}^{\oplus r-1}$$

Proof. Use Lemma 5.4.7.4. $\square$

Example 5.4.7.6 (Homology of (D^n, S^{n-1}))**.** We claim that

$$\tilde{H}_i(D^n, S^{n-1}; \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{if } i = n \\ 0 & \text{else.} \end{cases}$$

Indeed, this follows immediately from les of a pair and Lemma 5.4.7.4.

The following is an important observation in geometry.

Proposition 5.4.7.7 (Künneth formula-1)**.** *Let X be a T_1-space and $x \in X$. If $U \subseteq X$ is an open set containing x, then we have*

$$H_i(X, X - x; \mathbb{Z}) \cong H_i(U, U - x; \mathbb{Z}).$$

Proof. For $A = U$ and $B = X - x$, we see that both of them are open (B is open as $\{x\}$ is closed). Then, (X, A, B) forms an excisive triple. Performing excision, we observe (as $A \cap B = U - x$) that

$$H_i(U, U - x; \mathbb{Z}) \cong H_i(X, X - x; \mathbb{Z}),$$

as required. □

Remark 5.4.7.8. It is really necessary for U in Künneth formula above to be open, for $(S^2 - x, I - x) \hookrightarrow (S^2, I)$ for some path $I \hookrightarrow X$ does not induces isomorphism in homology, as is readily visible a small computation in the associated les of pairs.

Map of long exact sequence of pairs

Proposition 5.4.7.9. *Let $f : (X, A) \to (Y, B)$ be a map of pairs. Then, we get a map in the long exact sequences of the corresponding pairs. That is, the following commutes*[12]

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & H_n(A) \arrow[r] \arrow[d, "f_*"'] & H_n(X) \arrow[r] \arrow[d, "f_*"'] & H_n(X, A) \arrow[r, "\Delta"] \arrow[d, "f_*"'] & H_{n-1}(A) \arrow[d, "f_*"'] \arrow[r, dash, dashed] & {} \\
\cdots \arrow[r, dashed] & H_n(B) \arrow[r] & H_n(Y) \arrow[r] & H_n(Y, B) \arrow[r, "\Delta"'] & H_{n-1}(B) \arrow[r, dash, dashed] & {}
\end{tikzcd}.$$

Proof. Since all maps in the long exact sequence of a pair except the connecting homomorphism are induced by inclusions, therefore we need only check the commutativity of the rightmost square. This follows from unravelling the definition of connecting homomorphism as constructed from the chain level. □

We also have a map in Mayer-Vietoris.

Proposition 5.4.7.10. *Let $f : (X, A, B) \to (Y, C, D)$ be a map of triples, where each is an excisive triple. Then we get a map in the Mayer-Vietoris sequences of the corresponding pairs. That is, the following commutes*

$$\begin{tikzcd}
\cdots \arrow[r, dashed] & H_n(A \cap B) \arrow[r] \arrow[d, "f_*"'] & H_n(A) \oplus H_n(B) \arrow[r] \arrow[d, "f_*"'] & H_n(X) \arrow[r, "\Delta"] \arrow[d, "f_*"'] & H_{n-1}(A \cap B) \arrow[d, "f_*"'] \arrow[r, dash, dashed] & {} \\
\cdots \arrow[r, dashed] & H_n(C \cap D) \arrow[r] & H_n(C) \oplus H_n(D) \arrow[r] & H_n(Y) \arrow[r, "\Delta"'] & H_{n-1}(C \cap D) \arrow[r, dash, dashed] & {}
\end{tikzcd}.$$

Proof. Follows directly from Proposition 5.4.7.9 and the proof of original Mayer-Vietoris (in which we show that Mayer-Vietoris is obtained by les of a pair and excision). □

[12] we drop the group $\mathbb{Z}$ in the following diagram.

Lemma 5.4.7.11. *If $f : (X, A) \to (Y, B)$ is a homotopy equivalence of pairs, that is, there exists $g : (Y, B) \to (X, A)$ such that $f : X \rightleftarrows Y : g$ and $f : A \rightleftarrows B : g$ are both homotopy equivalences, then*

$$f_* : H_n(X, A) \xrightarrow{\cong} H_n(Y, B)$$

is an isomorphism.

Proof. Use 5-lemma on the diagram in Proposition 5.4.7.9. □

Immediate applications of Mayer-Vietoris

Example 5.4.7.12 (Homology of spheres)**.** We wish to show that

$$\tilde{H}_i(S^n; \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{if } i = n \\ 0 & \text{else.} \end{cases}$$

Indeed, let $U = S^n - p$ and $V = S^n - q$ where p, q are north and south poles respectively. Note $U \cap V \simeq S^{n-1}$. Then (S^n, U, V) is an excisive triple and thus by Mayer-Vietoris (Remark 5.4.6.4), we deduce that the connecting homomorphism $H_q(S^n) \cong \tilde{H}_{q-1}(S^{n-1})$. We conclude by induction.

Example 5.4.7.13 (Homology of wedge of spheres)**.** We wish to show that for each $i \geq 0$, we have

$$\tilde{H}_i(S^m \vee S^n) \cong \tilde{H}_i(S^m) \oplus \tilde{H}_i(S^n)$$

Indeed this follows by considering U to be S^m with some open part of S^n and V to be S^n with some open part of S^m. We get that $U \cap V \simeq \text{pt.}$, $U \simeq S^m$, $V \simeq S^n$ and (X, U, V) an excisive triple. The result now follows by Mayer-Vietoris (Remark 5.4.6.4).

Using Example 5.4.7.12, we can prove the following seemingly obvious, but otherwise hard to prove statement.

Theorem 5.4.7.14. *Let $n, m \in \mathbb{N}$.*

1. *S^n is homeomorphic to S^m if and only if $n = m$.*
2. *$\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^m$ if and only if $n = m$.*

Proof. Item 1 is immediate application of computation in Example 5.4.7.12. Item 2 can be obtained from removing a point from the given homeomorphism $\varphi : \mathbb{R}^n \xrightarrow{\cong} \mathbb{R}^m$ to get a homotopy equivalence $S^{n-1} \to S^{m-1}$. Thus, they have same homology. Invoking Example 5.4.7.12, we win. □

Degree of a map $f : S^n \to S^n$

For a map $f : S^n \to S^n$, consider the map $f_* : \mathbb{Z} \to \mathbb{Z}$ obtained by $H_n(S^n) \to H_n(S^n)$. Thus, f_* takes a generator a to $k \cdot a$, $k \in \mathbb{Z}$. We define $\deg(f) = k$. We begin with some basics.

Lemma 5.4.7.15. *Let $f : S^n \to S^n$ be a map.*

1. *If $f : S^n \to S^n$ and $g : S^n \to S^n$,*

$$\deg(g \circ f) = \deg(g) \cdot \deg f.$$

2. *If $f, g : S^n \to S^n$ are homotopy equivalent, then* $\deg(f) = \deg(g)$.

Proof. Immediate. □

The main theorem is the following, which computes the degree of reflections.

Theorem 5.4.7.16 (Degree of reflections)**.** *Define the following map*

$$\begin{aligned} f : S^n &\longrightarrow S^n \\ (x_1, x_2 \dots, x_{n+1}) &\longmapsto (-x_1, x_2, \dots, x_{n+1}). \end{aligned}$$

Then,

$$\deg(f) = -1.$$

Proof. Use induction on n and observe that for $X_1 = S^n - p$ and $X_2 = S^n - q$, we get a map induced in Mayer-Vietoris (Proposition 5.4.7.10). This yields the following commutative square where connecting homomorphism is an isomorphism:

$$\begin{tikzcd} H_n(S^n) \arrow[r, "\Delta"] \arrow[d, "f_*"'] & \tilde{H}_{n-1}(S^{n-1}) \arrow[d, "f_*"] \\ H_n(S^n) \arrow[r, "\Delta"'] & \tilde{H}_{n-1}(S^{n-1}) \end{tikzcd}.$$

The result now follows by inductive hypothesis. □

Corollary 5.4.7.17. *Define the following map*

$$\begin{aligned} f : S^n &\longrightarrow S^n \\ (x_1, x_2 \dots, x_{n+1}) &\longmapsto (-x_1, -x_2, \dots, -x_{n+1}). \end{aligned}$$

Then,

$$\deg(f) = (-1)^{n+1}.$$

Proof. Immediate from Theorem 5.4.7.16. □

Remark 5.4.7.18 (Fixed points and degree)**.** It is an easy observation that if $f : S^n \to S^n$ has no fixed points, then f is homotopic to $a : S^n \to S^n$ which is the antipodal map. Thus the degree of a map $f : S^n \to S^n$ which has no fixed points is $(-1)^{n+1}$.

An easy corollary of this observation is that if $f : S^n \to S^n$ is null homotopic, then f has a fixed point. Indeed as $\deg f = 0$, therefore by contrapositive of above, we deduce that f must have a fixed point.

A simple use of above remark yields the following fact for maps $f : S^{2n} \to S^{2n}$.

Proposition 5.4.7.19. *Let $f : S^{2n} \to S^{2n}$ be a map. Then, there exists $x \in S^{2n}$ such that either $f(x) = x$ or $f(x) = -x$.*

Proof. Suppose f has no fixed points. Then by Remark 5.4.7.18, it follows that $f \simeq a$, where $a : S^{2n} \to S^{2n}$ is the antipodal map. Thus, $\deg f = \deg a = (-1)^{2n+1} = -1$. It follows that $\deg(-f) = 1$. Hence, $-f$ must have a fixed point by Remark 5.4.7.18. Consequently, there exists $x \in S^{2n}$ such that $-f(x) = x$, as required. □

We also have the following conclusion.

Proposition 5.4.7.20. *Let $f : S^n \to S^n$ be a degree 0 map. Then there exists $x, y \in S^n$ such that $f(x) = x$ and $f(y) = -y$.*

Proof. Indeed, by above we immediately conclude that both f and $-f$ has degree 0, thus have fixed points. □

A more non-trivial application of ideas surrounding degree is the following.

Lemma 5.4.7.21. *Any linear map $T : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ has an eigenvector.*

Proof. We may assume that T is a bijection. Thus, T takes one dimensional linear subspaces to one dimensional linear subspaces. We get in particular a map $g : S^{2n} \to S^{2n}$ given by $\frac{v}{\|v\|} \mapsto \frac{Tv}{\|Tv\|}$. Use Proposition 5.4.7.19 to conclude. □

Antipode preserving maps $f : S^n \to S^1$

Another interesting application of singular homology is to show that if $n > 1$, then there is no antipode preserving map $f : S^n \to S^1$, where a map $f : S^m \to S^n$ is antipode preserving if for all $x \in S^m$, we have $-f(x) = f(-x)$.

Theorem 5.4.7.22. *If $n > 1$, then there is no antipode preserving map $f : S^n \to S^1$.*

Remark 5.4.7.23. One can deduce Borsuk-Ulam theorem, that for any map $f : S^2 \to \mathbb{R}^2$ there exists $x \in S^2$ such that $f(x) = f(-x)$, from Theorem 5.4.7.22 as follows. By composing by linear shift, we may assume $\mathrm{Im}\,(f)$ does not contain origin. Composing with the map $y \mapsto \frac{y}{\|y\|}$, we obtain the map $g : S^2 \to S^1$ mapping as $x \mapsto \frac{f(x)}{\|f(x)\|}$. Applying the above theorem, Borsuk-Ulam follows.

Jordan-Brower separation theorem

We wish to show the following result.

Theorem 5.4.7.24 (JBST). *Suppose $C \subseteq S^n$ is a subspace of S^n homeomorphic to S^{n-1}. Then $S^n - C$ has two components and has boundary C.*

More important for us is the two homological results which will be used to prove the above theorem.

Definition 5.4.7.25 (Cells in a space). A k-cell in a space X is a subspace $A \subseteq X$ homeomorphic to D^k.

Theorem 5.4.7.26. *Let A be a k-cell in S^n. Then,*

$$\tilde{H}_i(S^n - A; \mathbb{Z}) = 0$$

for every $i \geq 0$.

Using above theorem, we have the following result.

Proposition 5.4.7.27. *Let $h : S^k \hookrightarrow S^n$ be an embedding where $n > k \geq 0$. Then*

$$\tilde{H}_i(S^n - h(S^k); \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{if } i = n - k - 1, \\ 0 & \text{else}. \end{cases}$$

Proof. This follows from Mayer-Vietoris and induction on k, where we define $X_1, X_2 \subseteq S^n - h(S^k) = X$ as follows. Let $E_k^+ = S^k - q$ and $E_k^- = S^k - p$, p, q are north, south poles, respectively. Then define $X_1 = S^n - h(E_k^+)$ and $X_2 = S^n - h(E_k^-)$. Then $X_1 \cap X_2 = S^n - h(S^k)$ and $X_1 \cup X_2 = S^n - h(S^{k-1})$. Using Theorem 5.4.7.26 will yield the isomorphism $\tilde{H}_q(S^n - h(S^{k-1})) \cong \tilde{H}_{q-1}(S^n - h(S^{k-1}))$. We conclude by inductive hypothesis. □

Remark 5.4.7.28. Note that Proposition 5.4.7.27 already shows the first statement of Theorem 5.4.7.24. Indeed, Using the result, we get for $k = n - 1$, that $\tilde{H}_0(S^n - h(S^{n-1}); \mathbb{Z}) = \mathbb{Z}$, that is, there are two path-components of $S^n - h(S^{n-1})$. As S^n is locally path-connected, so number of components and path-components are same.

An important application is the invariance of domain.

Theorem 5.4.7.29 (Invariance of domain)**.** *Let $U \subseteq \mathbb{R}^n$ be a n open set and consider a map $f : U \to \mathbb{R}^n$ which is a continuous bijection. Then,*

1. *$f(U)$ is open in $\mathbb{R}^n$,*
2. *$f : U \to f(U)$ is a homeomorphism.*

That is, f is an open embedding.

Proof. Pick any open ball $B \subseteq U$ such that $\bar{B} \subseteq U$. Observe $S^{n-1} \cong \bar{B} - B = \partial B$. Consider the composite $f : \partial B \to f(U) \hookrightarrow S^n$ where we consider $\mathbb{R}^n \hookrightarrow S^n$. By JBST, $f : S^{n-1} \to S^n$ separates S^n into two components, say $S^n - f(S^{n-1}) = W_1 \amalg W_2$. If $f(B) \subseteq W_1$, we claim that $f(B) = W_1$. Indeed, this follows from Theorem 5.4.7.26 which says that removing a k-cell still keeps S^n path-connected. □

5.4.8 Homology with local coefficients

5.5 Cohomology theories

5.6 Cohomology products and duality

5.7 CW-complexes & CW homotopy types

One of the important properties of compactly generated spaces is that any such space can be approximated upto homotopy by a class of spaces constructed in a rather simple manner. These are precisely the CW complexes. Once the above approximation theorems are set up, we can safely reduce a lot of computation in homology to such a CW-approximation. Moreover, the reductions run so deep that in-fact any homology theory E_q on general compactly generated spaces necessarily induces and comes from the restriction of E_q to CW-complexes. An application of Hurewicz theorem will then tell us that upto natural isomorphism, there is a unique homology theory over CW-complexes. Moreover, the fundamental result of Whitehead would allow us to interpret homotopy groups as a complete set of homotopical invariants for CW-complexes

5.7.1 Basic theory

5.7.2 Approximation theorems

5.7.3 CW homotopy types

We wish to prove some foundational results on homotopy equivalences of CW-complexes.

Whitehead's theorem

We wish to see the following important result.

Theorem 5.7.3.1 (Whitehead)**.** *Let X and Y be weakly equivalent CW-complexes. Then X and Y are homotopy equivalent.*

Applications of Whitehead's theorem

Lemma 5.7.3.2 (Weak uniqueness of universal covers)**.** *Let X be a CW-complex. If E is a CW-complex and $f : E \to X$ is such that*

$$f_* : \pi_k(E) \to \pi_k(X)$$

is an isomorphism for all $k \geq 2$ and $\pi_k(E) = 0$ for $k = 0, 1$, then E is homotopy equivalent to the universal cover $\tilde{X}$ of X.

Proof. As $\pi_0(E) = 0$, therefore E is connected. It follows by unique lifting (which is possible as $\pi_1(E) = 0$) that we have a commutative diagram of spaces:

$$\begin{tikzcd} & \tilde{X} \arrow[d, "p"] \\ E \arrow[ur, "\tilde{f}"] \arrow[r, "f"'] & X \end{tikzcd}.$$

Applying π_k for any $k \geq 2$, we deduce from our hypothesis that $\tilde{f}_* : \pi_k(E) \to \pi_k(\tilde{X})$ is an isomorphism. As $\pi_0(\tilde{X}) = \pi_1(\tilde{X}) = 0$, therefore $\tilde{f}$ is a weak equivalence. It follows by Whitehead's theorem (Theorem 5.7.3.1) that $\tilde{f}$ is a homotopy equivalence, as required. $\square$

5.8 Homotopy and homology

5.8.1 Hurewicz's theorem

Theorem 5.8.1.1 (Hurewicz-1). *Let X be an $(n-1)$-connected based space. Then the Hurewicz map*

$$h_n : \pi_n(X) \to H_n(X;\mathbb{Z})$$

is an isomorphism and

$$h_{n+1} : \pi_{n+1}(X) \to H_{n+1}(X;\mathbb{Z})$$

is a surjection.

It is also very beneficial to keep the following version of Hurewicz in mind as it is usually used to deduce conclusion about homology groups from some information about homotopy groups and vice-versa. The second item is often used after passing to universal covers.

Theorem 5.8.1.2 (Hurewicz-2). *Let X, Y be path-connected based spaces and $f : X \to Y$ be a based map. Let $n \in \mathbb{N}$.*

1. *If $f_* : \pi_k(X) \to \pi_k(Y)$ is an isomorphism for $k < n$ and a surjection for $k = n$, then $f_* : H_k(X;\mathbb{Z}) \to H_k(Y;\mathbb{Z})$ is an isomorphism for $k < n$ and a surjection for $k = n$.*
2. *If X, Y are simply connected and $f_* : H_k(X;\mathbb{Z}) \to H_k(Y;\mathbb{Z})$ is an isomorphism for $k < n$ and a surjection for $k = n$, then $f_* : \pi_k(X) \to \pi_k(Y)$ is an isomorphism for $k < n$ and a surjection for $k = n$.*

5.9 Homotopy & algebraic structures

5.9.1 H-spaces

Definition 5.9.1.1 (H-spaces & groups). Let (X, e) be a based space. Then X is said to be a an H-space if there exists a continuous map

$$\begin{aligned} m : X \times X &\longrightarrow X \\ (x, y) &\longmapsto x \cdot y \end{aligned}$$

such that

1. $e \cdot e = e$,
2. $m_e : X \to X, x \mapsto x \cdot e$ and $m^e : X \to X, x \mapsto e \cdot x$ are both homotopy equivalent to id_X rel $\{e\}$.

An H-space $(X, e, \cdot)$ is said to be an H-group if moreover the map m satisfies the following:

1. the two associativity maps $X \times X \times X \rightrightarrows X$ are homotopic to each other rel $\{(e,e,e)\}$,
2. there exists an inverse map $(-)^{-1} : X \to X$ such that $e^{-1} = e$ and that the two left/right multiplication by inverse maps $X \rightrightarrows X$, $x \mapsto x \cdot x^{-1}$ and $x \mapsto x^{-1} \cdot x$ is homotopic to constant map c_e rel $\{e\}$.

Example 5.9.1.2. Every topological group is a strict H-group.

Example 5.9.1.3. Every loop space ΩX is an H-group where the product of two loops is the concatenation and inverse is the inverse of the loop. The required conditions for ΩX to be an H-group is then immediate.

The following is one of the most important result for H-spaces. It says that the contravariant hom functor that they represent is group valued.

Theorem 5.9.1.4. *Let Y be an H-group. Then for any based space X, the based homotopy classes of maps $[X, Y]$ forms a group whose operation is*

$$(f \cdot g)(x) := f(x) \cdot g(x).$$

5.10 Model categories & abstract homotopy

5.11 Classifying spaces

5.11.1 Eilenberg-Maclane spaces

Remark 5.11.1.1 (The canonical map)**.** Let X be a connected space and $G = \pi_1(X)$. Then there is a natural map $i : X \to BG$ which identifies X as a subcomplex of BG. Moreover $i_* : \pi_1(X) \to \pi_1(BG) = G$ is an isomorphism.

5.12 Spectra

Spectra are objects which generalizes both the notion of cohomology theories and spaces, in that there are mappings from cohomology theories and spaces into the homotopy category of spectra. Thus, one needs to construct a good category of spectra, give a model structure on it and thus by Quillen's theory obtain this absolutely wonderful homotopy category of spectra, which unites the viewpoint of cohomology and spaces. However, we are getting ahead of ourselves, as finding the right homotopy category and giving a construction of category of spectra is easier said than done. We will meet this topic later in our discussion of ∞-categories (they will form a prototypical example of stable ∞-categories).

5.13 Lifting & extension problems

Chapter 6

Stable Homotopy Theory

In this chapter, we give an overview of stable homotopy theory.

Chapter 7

Classical Ordinary Differential Equations

Contents

We will prove some basic existence/uniqueness results about ODEs here, with a classical/analytic viewpoint in mind. Let us first begin by stating what is meant by an *initial value problem* and what is meant by *solving an initial value problem*. A main focus will be on doing analytical proofs, which is always extremely helpful. In particular, we will see how much weird and pathological behaviors can emerge after *passing to limit*, thus justifying why commuting with limits is a sought after property in all over analysis.

7.1 Initial value problems

Let us begin by understanding what is meant by a differential equation. Let $D \subseteq \mathbb{R}\times\mathbb{R}^n$ be an open set. Consider a continuous function $f : D \to \mathbb{R}^n$ mapping as $(t, x) \mapsto f(t, x)$ where $t \in \mathbb{R}, x \in \mathbb{R}^n$. A fundamental goal that one wishes to achieve is to find a "nice" function $x : I \subseteq \mathbb{R} \to D$ such that the function f can be known upto first derivatives, that is, we want to construct such a function $x : I \to \mathbb{R}^n$ such that it can tell us the following about f:

1. (*Correct domain*)$\forall t \in I$, we shall have $(t, x(t)) \in D$,
2. (*Differential equation*)$\frac{dx}{dt}(t_0) = f(t_0, x(t_0)), \forall t_0 \in I$. That is, the first derivative of x can give us exactly the values that f takes on D.

To find such a function x, the main difficulty is the condition 2 above, for this requires $x : I \to \mathbb{R}^n$ to be continuously differentiable (so of class C^1) and that we necessarily have to construct a function x by the knowledge only of it's first derivative (which is $f(t, x)$).

This problem of constructing a C^1 map $x : I \subseteq \mathbb{R} \to \mathbb{R}^n$ from only the data of it's continuous first derivative is called the process of solving a differential equation.

Clearly, many C^1 maps can have same first derivative (we need only add a scalar in front). So the uniqueness of the above problem is hopeless. However, one can add an extra data to the problem above that x shall satisfy and then we do get uniqueness at times. In particular, we demand the following from x:

3. (*Initial value*) for some fixed $s_0 \in I$ and $x_0 \in \mathbb{R}^n$, we require $x(s_0) = x_0$.

We then define an initial value problem (IVP) as follows:

Definition 7.1.0.1. (**IVP & solutions**) Let $f : D \to \mathbb{R}^n$ be a continuous map on an open set $D \subseteq \mathbb{R} \times \mathbb{R}^n$. An IVP is a construction problem where from the tuple of data $(f, (t_0, x_0))$ for some $(t_0, x_0) \in D$, we have to construct the following:

1. an interval $I \subseteq \mathbb{R}$ containing t_0,
2. a function $x : I \to \mathbb{R}^n$.

This function x should then satisfy the following:

1. $(t, x(t)) \in D \ \forall t \in I$,
2. $\frac{dx}{dt}(t) = f(t, x) \ \forall t \in I$,
3. $x(t_0) = x_0$.

We identify the above IVP with the tuple $(f, (t_0, x_0))$. If such a function $x : I \to \mathbb{R}^n$ exists, then it is called a *solution to the IVP* $(f, (t_0, x_0))$.

7.1.1 Existence: Peano's theorem

We have an elementary result which tells us that, if the solution exists, then what should be its form.

Lemma 7.1.1.1. *Let $f : D \to \mathbb{R}^n$ be a continuous map and $(f, (t_0, x_0))$ be an IVP. Then, a continuous map $x : I \to \mathbb{R}^n$ is a solution to the IVP $(f, (t_0, x_0))$ if and only if $\forall t \in I$, $x(t)$ is the following line integral*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

Proof. (L $\Rightarrow$ R) Since x is a solution, therefore $\frac{dx}{dt}(t) = f(t, x(t))$ $\forall t \in I$ and $t_0 \in I$. Then use fundamental theorem of calculus to calculate the line integral of the vector field ∇x along the line joining t_0 and t.

(R $\Rightarrow$ L) By continuity of x, we get that $t \mapsto f(t, x(t))$ is continuous. Since $x(t_0) = x_0$, therefore by continuity of $t \mapsto (t, x(t))$, there exists an open interval $I \ni t_0$ of $\mathbb{R}$ such that $(t, x(t)) \in D$ for all $t \in I$. It then follows by an application of fundamental theorem of calculus that $\frac{dx}{dt}(t) = f(t, x(t))$ for each $t \in I$. $\square$

We next do Peano's theorem, which tells us that indeed solutions exists. This when combined with above tells us that solutions to IVP $(f, (t_0, x_0))$ exists and is of same "form". However, it will require a classic result in analysis called Arzela-Ascoli theorem. Let us do that first.

Theorem 7.1.1.2. *(Arzela-Ascoli theorem) Let $x_n : [0, 1] \to \mathbb{R}^n$ be a sequence of continuous functions such that $\{x_n\}$ is a uniformly bounded and equicontinuous family of maps. Then there exists a subsequence of $\{x_n\}$ which is uniformly convergent.*

We can now approach the existence result.

Theorem 7.1.1.3. *(Peano's theorem) Let $f : D \to \mathbb{R}^n$ be a continuous map where $D \subseteq \mathbb{R} \times \mathbb{R}^n$ is open and let $(t_0, x_0) \in D$ so that the tuple $(f, (t_0, x_0))$ forms an IVP. Choose $r > 0$ and $c > 0$ such that $[t_0 - c, t_0 + c] \times \overline{B}_r(x_0) \subseteq D$[1]. Then, denoting $M := \max_{x \in [t_0 - c, t_0 + c] \times \overline{B}_r(x_0)} f(x)$ and $h := \min\{c, \frac{r}{M}\}$, there exists a solution to the IVP $(f, (t_0, x_0))$ given by*

$$x : [t_0 - h, t_0 + h] \longrightarrow \overline{B}_r(x_0).$$

Proof. We will construct the solution x in a limiting manner. First, we may replace t_0 by 0 as we can shift the solution to t_0 thus obtained. Second, we may define x on $[0, h]$ as we may translate and scale the solution as desired. Now, consider the sequence of functions defined as follows:

$$\begin{aligned} x_n(t) : [0, h] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \begin{cases} x_0 \text{ if } t \in [0, \frac{h}{n}], \\ x_0 + \int_0^{t - h/n} f(s, x_n(s)) ds & \text{if } t \in [\frac{h}{n}, h]. \end{cases} \end{aligned}$$

So we obtain a sequence of functions $\{x_n\}$ defined over $[0, h]$. Now, in the limiting case, we will have a function exactly of the form required by Lemma 7.1.1.1, so we reduce to showing that a subsequence of the above converges and converges to a continuous function. We will use the Arzela-Ascoli (Theorem 7.1.1.2) for showing this. We thus need only show that the sequence $\{x_n\}$ is uniformly bounded and equicontinuous. For uniform boundedness, we will simply show that

[1]That is, choose a basic closed set around (t_0, x_0) in D.

$x_n(t) \in \overline{B}_r(x_0)\ \forall t \in [0, h]$. This follows from the following:

$$\begin{aligned}|x_n(t) - x_0| &\leq \left|\int_0^{t-h/n} f(s, x_n(s))ds\right| \\ &\leq \int_0^{t-h/n} |f(s, x_n(s))|\, ds \\ &\leq M(t - \frac{h}{n}) \\ &\leq Mh \\ &\leq r.\end{aligned}$$

Finally, to see equicontinuity, we may simply observe that for any $\epsilon > 0$ and for any $n \in \mathbb{N}$,

$$\begin{aligned}|x_n(s) - x_n(t)| &\leq \left|\int_{t-h/n}^{s-h/n} f(u, x_n(u))\right| du \\ &\leq \int_{t-h/n}^{s-h/n} |f(u, x_n(u))|\, du \\ &\leq M(s - t).\end{aligned}$$

This shows equicontinuity. □

Remark 7.1.1.4. (*Comments on proof of Theorem 7.1.1.3*) The main idea of the proof was to *find the required function through a limiting procedure*, where to make sure that we do get the limit, we used Arzela-Ascoli. One of the foremost things we did as well was to *reduce to the nicest possible setting*, which will be very necessary to clear things around.

7.1.2 Uniqueness: Picard-Lindelöf theorem

We will now show that for an IVP $(f, (t_0, x_0))$, we may get unique solutions provided some hypotheses on f. In order to understand what this hypothesis on f is, we need to review Lipschitz and contractive functions.

Definition 7.1.2.1. (**(locally)Lipschitz functions**) A map $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz function if $\exists L > 0$ such that $\forall\ x, y \in R$, we have

$$\|f(x) - f(y)\| < L\|x - y\|.$$

The function f is called locally Lipschitz if $\forall x \in E$, there exists $r > 0$ such that $f|_{B_r(x)}$ is a Lipschitz map.

Example 7.1.2.2. The map $f : \mathbb{R} \to \mathbb{R}$ given by $x \mapsto x^{1/3}$ is not locally Lipschitz at $x = 0$. This is because if it is so, then $\exists \epsilon > 0$ such that on $B_\epsilon(0)$ the map f is Lipschitz. But for $x, y \in B_\epsilon(0)$ we get

$$\begin{aligned}|x - y| &= \left|(x^{1/3})^3 - (y^{1/3})^3\right| \\ &= \left|(x^{1/3} - y^{1/3})(x^{2/3} + y^{2/3} + (xy)^{1/3})\right| \\ &\leq 2\epsilon.\end{aligned}$$

Thus,

$$\left|x^{1/3} - y^{1/3}\right| \leq \frac{2\epsilon}{\left|x^{2/3} + y^{2/3} + (xy)^{1/3}\right|},$$

which shows that f can not be Lipschitz on $B_\epsilon(0)$.

We have that all continuously differentiable maps are locally Lipschitz.

Lemma 7.1.2.3. *Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a C^1-map on an open set E, then f is locally Lipschitz.*

Proof. Take $a \in E$. We reduce to showing that there exists a $\epsilon > 0$ such that $\overline{B}_\epsilon(0) \subset E$ so that the continuous map $Df : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ achieves maxima on the compact set. This follows from the fact that E is open. □

One definition that we will need is that of uniform Lipschitz.

Definition 7.1.2.4. (**Uniform Lipschitz**) Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map. Then f is called uniformly Lipschitz w.r.t. x if there exists $L > 0$ such that

$$\|f(t,x) - f(t,y)\| < L\|x - y\|$$

for all $(t,x), (t,y) \in D$.

A contraction is defined in an obvious manner.

Definition 7.1.2.5. (**Contractive mappings**) Let $f : X \to X$ be a continuous map of metric spaces. Then f is said to be contractive if there exists $0 < \lambda < 1$ such that

$$d(f(x), f(y)) < \lambda d(x,y)$$

for all $x, y \in X$.

Our goal is to find the conditions that one must impose on f for the IVP $(f, (t_0, x_0))$ to have a unique solution. This means we need to find a solution $x : I \to \mathbb{R}^n$ in such a manner that x is the unique solution possible on that interval I. Now a place uniqueness comes into the picture is Banach fixed point theorem. Indeed, we will use it to find such an interval I and map x so that it would be unique for the said IVP.

Theorem 7.1.2.6. *(Banach fixed point theorem) Let X be a complete metric space and $f : X \to X$ be a contractive mapping. Then, f has a unique fixed point.*

Proof. We will first show the existence of such a fixed point. There is an obvious process of doing so. Take any point $x_0 \in X$. We then form the sequence $\{x_n\}$ in X where $x_n = f^n(x_0)$. We claim that $\{x_n\}$ is Cauchy. Indeed, we have that for any $\epsilon > 0$ (we may take $n \geq m$):

$$\begin{aligned}
d(x_n, x_m) &= d(f^n(x_0), f^m(x_0)) \\
&< \lambda^m d(f^{n-m}(x_0), x_0) \\
&< \lambda^m \left(d(f^{n-m}(x_0), f(x_0)) + d(f(x_0), x_0)\right) \\
&< \lambda^m \left(\lambda d(f^{n-m-1}(x_0), x_0) + d(x_1, x_0)\right) \\
&= \lambda^{m+1} d(f^{n-m-1}(x_0), x_0) + \lambda^m d(x_1, x_0) \\
&< d(x_1, x_0) \left(\lambda^m + \cdots + \lambda^n\right) \\
&= \lambda^m \frac{1 - \lambda^{n-m}}{1 - \lambda} d(x_1, x_0) \\
&< \frac{\lambda^m}{1 - \lambda} d(x_1, x_0).
\end{aligned}$$

Next, by completeness of X, we have that there exists $x = \varinjlim_n x_n$ in X. Now, $f(x) = f(\varinjlim_n x_n) = \varinjlim_n f(x_n)$ by continuity and $\varinjlim_n f(x_n) = x$ by definition of x_n. The uniqueness is simple by contractive property of f. □

We now come to the main result, the uniqueness of solutions of IVP. Before stating it, let us state how we will be proving it, using the following bijection between solutions of $(f, (t_0, x_0))$ and fixed points of certain mapping.

Construction 7.1.2.7. Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping where D is open and let $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Now consider the following space for some $c > 0$

$$X := C^1\left[[t_0 - c, t_0 + c], \mathbb{R}^n\right]$$

and consider the following map

$$\begin{aligned}
T : X &\longrightarrow X \\
x(t) &\longmapsto T(x)(t) := x_0 + \int_{t_0}^{t} f(s, x(s))ds.
\end{aligned}$$

Then, by Lemma 7.1.1.1, we see that $x(t) \in X$ is a solution of $(f, (t_0, x_0))$ if and only if $T(x(t)) = x(t)$. Hence

$$\{\textit{Solutions of IVP } (f, (t_0, x_0))\} \cong \{\textit{Fixed points of } T : X \to X\}.$$

Theorem 7.1.2.8. *(Weak Picard-Lindelöf) Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where D is open and $(t_0, x_0) \in D$ such that $(f, (t_0, x_0))$ forms an IVP. Choose $c > 0$ and $r > 0$ such that $[t_0 - c, t_0 + c] \times \overline{B}_r(x_0) \subseteq D$. Denote $M := \max_{x \in [t_0 - c, t_0 + c] \times \overline{B}_r(x_0)} f(x)$. If the map*

$$f : [t_0 - c, t_0 + c] \times \overline{B}_r(x_0) \longrightarrow \mathbb{R}^n$$

is uniformly Lipschitz w.r.t. x and Lipschitz constant being L, then, denoting $h := \min\{c, \frac{r}{M}, \frac{1}{L}\}$, there exists a unique solution of IVP $(f, (t_0, x_0))$ given by

$$x : [t_0 - h, t_0 + h] \longrightarrow \overline{B}_r(x_0).$$

Proof. (Sketch) The main part of the proof will be the idea in Construction 7.1.2.7 and Banach fixed point theorem. Let X denote the following space

$$X := \left\{ y \in C^0\left[[t_0 - h, t_0 + h], \mathbb{R}^n\right] \mid y(t_0) = x_0 \;\&\; \sup_{x \in [t_0 - h, t_0 + h]} \|x_0 - y(t_0)\| \leq hM \right\}.$$

Consider the following function on X

$$\begin{aligned} T : X &\longrightarrow X \\ y &\longmapsto x_0 + \int_{t_0}^{t} f(s, y(s))ds. \end{aligned}$$

By Theorem 7.1.2.6, we reduce to showing that function X is complete and T is a contraction mapping. Let us first show completeness of X. One then shows that $X \hookrightarrow C\left[[t_0 - h, t_0 + h], \mathbb{R}^n\right]$ is a closed subspace and it will suffice since $C\left[[t_0 - h, t_0 + h], \mathbb{R}^n\right]$ is complete and closed subspaces of complete spaces are complete. □

We will now prove Picard-Lindelöf again but with a weakening of hypotheses as compared to Theorem 7.1.2.8. This is important because most of the time one doesn't has the information of Lipschitz constant L as is required in Theorem 7.1.2.8 while constructing the interval of the solution.

Lemma 7.1.2.9. *Something about Picard iterates: If f is Lipschitz with constant $L > 0$, then the Picard iterates $\{x_n(t)\}$ satisfies*

$$\|x_{n+1}(t) - x_n(t)\| \leq \frac{ML^n(t - t_0)^{n+1}}{(n+1)!}.$$

Theorem 7.1.2.10. *(Strong Picard-Lindelöf) Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map on an open set D and $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Choose $c > 0$ and $r > 0$ such that $[t_0 - c, t_0 + c] \times \overline{B}_r(x_0) \subseteq D$. Denote $M := \max_{x \in [t_0 - c, t_0 + c] \times \overline{B}_r(x_0)} f(x)$. If the map*

$$f : [t_0 - c, t_0 + c] \times \overline{B}_r(x_0) \longrightarrow \mathbb{R}^n$$

is uniformly Lipschitz w.r.t. x, then, for any $h < \min\{c, \frac{r}{M}\}$, there exists a unique solution of IVP $(f, (t_0, x_0))$ given by

$$x : [t_0 - h, t_0 + h] \longrightarrow \overline{B}_r(x_0).$$

The following corollary tells us an alternate sufficient condition on f for the existence of unique solution to an IVP on f.

Corollary 7.1.2.11. *Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a continuous map where D is open. If $\frac{\partial f_i}{\partial x_j} : D \to \mathbb{R}$ are continuous maps for all $1 \leq i, j \leq n$, then for each $(t_0, x_0) \in D$ there exists an open neighborhood around $(t_0, x_0) \in D$ in which there is a unique solution to IVP $(f, (t_0, x_0))$.*

Remark 7.1.2.12. In practice, to reduce to an open neighborhood where the solution is unique, the above corollary will be useful.

7.1.3 Continuation of solutions

Consider the map $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $(t, x) \mapsto x^2$ and $(0, 1) \in \mathbb{R} \times \mathbb{R}$. One sees that the IVP $(f, (0, 1))$ has a solution given by

$$\begin{aligned} x : (-1, 1) &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{1}{1-t}. \end{aligned}$$

However, this solution can be "extended"/"continued" to the following solution of the said IVP

$$\begin{aligned} y : (-\infty, 1) &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{1}{1-t}. \end{aligned}$$

These two are different solutions but the domain of one is inside the domain of the other. This concept of solutions extending from one domain to a larger domain will be investigated in this section.

The following definition is obvious.

Definition 7.1.3.1. (**Continuation of solutions**) Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where D is open and $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Let $x : I \to \mathbb{R}^n$ be a solution of $(f, (t_0, x_0))$. Then the solution x is said to be continuable if there exists a solution y of $(f, (t_0, x_0))$ given by $y : J \to \mathbb{R}^n$ where $J \supseteq I$ and $y|_I = x$.

The following theorem tells us a sufficient criterion on the solution which would make it continuable to some larger interval.

Theorem 7.1.3.2. *Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where D is open and $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Let $x : (a, b) \to \mathbb{R}^n$ be a solution of $(f, (t_0, x_0))$.*

1. *If $\varinjlim_{t \to b^-} x(t)$ exists and $(b, \varinjlim_{t \to b^-} x(t)) \in D$, then there exists $\epsilon > 0$ such that x can be continued to a solution $\tilde{x} : (a, b + \epsilon) \to \mathbb{R}^n$.*
2. *If $\varinjlim_{t \to a^+} x(t)$ exists and $(a, \varinjlim_{t \to a^+} x(t)) \in D$, then there exists $\epsilon > 0$ such that x can be continued to a solution $\tilde{x} : (a - \epsilon, b) \to \mathbb{R}^n$.*

The following lemma states that for mild conditions on f, the boundary limits might exist for a solution.

Lemma 7.1.3.3. *Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where D is open and $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. If f is bounded, then for any solution $x : (a, b) \to \mathbb{R}^n$, the limits*

$$\varinjlim_{t \to b^-} x(t) \;\&\; \varinjlim_{t \to a^+} x(t) \text{ exist}.$$

Proof. Use Lemma 7.1.1.1 to get that x is uniformly continuous over (a, b), so it has unique extension to its boundary. $\square$

7.1.4 Maximal interval of solutions

Let $(f,(t_0,x_0))$ be an IVP and let $x : I \to \mathbb{R}^n$ be a solution. A natural question is whether there is a "maximal continuation" of x in the sense of Definition 7.1.3.1. This is what we investigate here. The following definition is clear.

Definition 7.1.4.1. (**Maximal interval of solution**) Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map and $(t_0,x_0) \in D$ so that $(f,(t_0,x_0))$ forms an IVP. The maximal interval of solution x is an interval $J \subseteq \mathbb{R}$ such that there exists a continuation of x on J and there is no continuation of $z : L \to \mathbb{R}^n$ of y where $L \supsetneq J$.

Lemma 7.1.4.2. *Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map and $(t_0,x_0) \in D$ so that $(f,(t_0,x_0))$ forms an IVP. If $x : I \to \mathbb{R}^n$ is a solution of $(f,(t_0,x_0))$, then there exists a maximal interval of solution x.*

Proof. This is a simple application of Zorn's lemma on the poset

$$P = \{y : J \to \mathbb{R}^n \mid y \text{ is a continuation of } x\}$$

where $y \leq z$ iff z is a continuation of x. □

We wish to now find a characterization of maximal intervals of a solution. That is, we wish to know when can we say that a given solution is maximal.

Proposition 7.1.4.3. *Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map and $(t_0,x_0) \in D$ so that $(f,(t_0,x_0))$ forms an IVP. Let $x : (a,b) \to \mathbb{R}^n$ be a solution of $(f,(t_0,x_0))$. Then,*

1. *The interval $[t_0,b)$ is a right maximal interval of solution x if and only if for any compact subset $K \subseteq D$, there exists $t \in [t_0,b)$ such that $(t,x(t)) \notin K$.*
2. *The interval $(a,t_0]$ is a left maximal interval of solution x if and only if for any compact subset $K \subseteq D$, there exists $t \in (a,t_0]$ such that $(t,x(t)) \notin K$.*

Proof. (Sketch) By symmetry, we reduce to showing 1. The main idea is to use the maximality and the results of previous section. □

7.1.5 Solution on boundary

In this section, we investigate the limiting cases of solutions of ODEs on a maximal interval (see Lemma 7.1.4.2). We see that if the one-sided limit of a maximal solution exists, then it's graph has to lie on the boundary of the domain.

Theorem 7.1.5.1. *Let $f : D \to \mathbb{R}^n$ be a continuous map where $D \subseteq \mathbb{R} \times \mathbb{R}^n$ is open and let $(t_0,x_0) \in D$ so to make $(f,(t_0,x_0))$ an IVP. If $x : I \to \mathbb{R}^n$ is a solution to $(f,(t_0,x_0))$ and $I = (a,b)$ is a maximal interval of solution, then*

1. *If $\partial D \neq \emptyset$, $b < \infty$ and $\varinjlim_{t\to b^-} x(t)$ exists, then*

$$\left(b, \varinjlim_{t\to b^-} x(t)\right) \in \partial D.$$

2. *If $\partial D = \emptyset$, $b < \infty$ then*

$$\limsup_{t\to b^-} x(t) = \infty.$$

A similar statement holds for left sided limit towards a.

Proof. 1. Suppose not. Then $\left(b, \varinjlim_{t\to b^-} x(t)\right) \in D$ as D is open. It follows from Lemma 7.1.3.2 that $[t_0, b)$ is not maximal.

2. Suppose not. Then $\limsup_{t\to b^-} x(t) \neq \infty$. Hence, there exists $M > 0$ such that $\|x(t)\| < M$ for all $t \in [t_0, b)$. Now, construct $K = [t_0, b] \times C$ where C is a compact disc such that $\forall t \in [t_0, b)$, $x(t) \in C$, which can be chosen as an appropriate disc in $B_M(x_0)$. Since $K \subseteq D$, therefore by Proposition 7.1.4.3 we get a contradiction to maximality of $[t_0, b)$. □

That's all we have to say here, so far.

7.1.6 Global solutions

So far we have studied solutions $x(t)$ to IVP defined only on some small enough intervals I such that $(t, x(t)) \in D$. However, we defined $D \subseteq \mathbb{R} \times \mathbb{R}^n$ as an arbitrary open set. In this section we would restrict to certain type of domains D, namely of the form $D = I \times \mathbb{R}^n$ and will try to study whether we can obtain a solution $x(t) : I \to \mathbb{R}^n$ to an IVP $(f, (t_0, x_0))$. If they exists, we call such a solution to be a *global solution* of the IVP $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ with initial values $(t_0, x_0) \in I \times \mathbb{R}^n$.

Let $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where $I \subseteq \mathbb{R}$ is an open interval and choose $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Let $x : J \to \mathbb{R}^n$ be a solution of $(f, (t_0, x_0))$. The main result of this section says that every such solution $x(t)$ can be extended to a global solution on I given some regularity conditions of values of f.

Theorem 7.1.6.1. *Let $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map where $I \subseteq \mathbb{R}$ is an open interval and choose $(t_0, x_0) \in D$ so that $(f, (t_0, x_0))$ forms an IVP. Suppose*

$$\|f(t, x)\| \leq M(t) + \|x\| N(t)$$

where $M, N : I \to \mathbb{R}$ are non-negative continuous maps, $\forall (t, x) \in I \times \mathbb{R}^n$. Then any solution $x : J \to \mathbb{R}^n$ of $(f, (t_0, x_0))$ can be continued to a solution $\tilde{x} : I \to \mathbb{R}^n$.

We are more interested in the applications of the above theorem, which we now present.

Corollary 7.1.6.2. *Let $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ be uniformly Lipschitz w.r.t. x. Then, there exists a unique global solution $x : I \to \mathbb{R}^n$ of the IVP $(f, (t_0, x_0))$.*

Proof. We get that $\exists L > 0$ such that

$$\|f(t, x) - f(t, y)\| < L\|x - y\|$$

for all $t \in I$ and $x, y \in \mathbb{R}^n$. In particular for $y = 0$, we get

$$\begin{aligned} \|f(t, x)\| &\leq \|f(t, x) - f(t, 0)\| + \|f(t, 0)\| \\ &\leq L\|x\| + \|f(t, 0)\| \end{aligned}$$

where $N(t) = L$ and $M(t) = \|f(t, 0)\|$ in the notation of Theorem 7.1.6.1. Hence, by the same theorem, if there exists a solution of $(f, (t_0, x_0))$, say x on $J \subseteq I$, then it extends to a solution on I. Now by Strong Picard-Lindelöf (Theorem 7.1.2.10), we conclude that there is a unique solution on I; if there are two solutions on I, then by restriction on the interval obtained from Picard-Lindelöf, we would get a contradiction to it's uniqueness. □

For a system of equations linear in x, for $x \in \mathbb{R}^n$, we have the following result.

Corollary 7.1.6.3. *Let $f(t,x) = A(t)x + b(t)$ be a map from $I \times \mathbb{R}^n$ to $\mathbb{R}^n$ where $A(t) \in C(I, \mathbb{R}^{n\times n})$ and $b \in C(I, \mathbb{R}^n)$ for an open interval $I \subseteq \mathbb{R}$ and $x = (x_1, \ldots, x_n)$. For $(t_0, x_0) \in I \times \mathbb{R}^n$, consider the IVP $(f, (t_0, x_0))$. Then there exists a unique solution*

$$x : I \times \mathbb{R}^n \to \mathbb{R}^n.$$

Proof. Using triangle inequality, we obtain

$$\|f(t,x)\| \leq \|A(t)\|\|x\| + \|b\|.$$

The result follows by an application of Theorem 7.1.6.1 and Corollary 7.1.6.2. □

7.2 Linear systems

So far, we covered solutions of ODE of the form

$$\frac{dx}{dt} = f(t, x(t))$$

where $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ and $x : I \to \mathbb{R}^n$. In particular, $\frac{dx}{dt}$ is given as

$$\frac{dx}{dt}(t) = \begin{bmatrix} \frac{dx_1}{dt} & \frac{dx_2}{dt} & \cdots & \frac{dx_n}{dt} \end{bmatrix}$$

where each $x_i : I \to \mathbb{R}$. On the other hand, the right side consists of $f(t,x)$, which is a continuous function from a subset of $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$.

In this section, we would now study in detail a particular type of IVP in which the aforementioned function $f(t,x)$ is a linear map. In particular, the mapping f is given by

$$\begin{aligned} f : D \subseteq \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (t,x) &\longmapsto Ax \end{aligned}$$

for a real matrix A.

Remark 7.2.0.1. One should keep in mind that these are not new ODEs; a linear system is same as $\frac{dx}{dt} = f(t,x)$ where $f(t,x) = Ax$, so they are special cases of general ODEs and have special properties like uniqueness of solutions. In particular, all the results of the previous section on general ODEs will obviously hold in the linear case.

Remark 7.2.0.2. By Lemma 7.1.1.1, we know that a solution of $\frac{dx}{dt} = Ax$ is necessarily of the form

$$x(t) = x_0 + A\int_0^t x(s)ds.$$

7.2.1 Some properties of matrices

Let us begin by stating some of the properties of matrix algebra, especially of exponential of matrices as it will be used in Theorem 7.2.2.1. Since these are not fancy results so we omit the proof of all except the main observations required in each.

Theorem 7.2.1.1. *Let $A, B \in M_n(\mathbb{R})$. Then,*

1. $\|A+B\| \leq \|A\| + \|B\|$.
2. $\|AB\| \leq \|A\|\|B\|$.
3. *The series e^X defined by*

$$e^X := \sum_{n=0}^{\infty} \frac{X^n}{n!}$$

   converges for all $X \in M_n(\mathbb{R})$.
4. $e^0 = I$.
5. $(e^A)^T = e^{A^T}$.
6. *e^X is invertible and $(e^X)^{-1} = e^{-X}$ for all $X \in M_n(\mathbb{R})$.*
7. *If $AB = BA$, then $e^{A+B} = e^A e^B = e^B e^A$.*
8. *If $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, then $e^A = \mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.*
9. *If P is invertible, then $e^{PAP^{-1}} = Pe^A P^{-1}$.*

Proof. We omit the proof of all but the 3. To show that the series converges, by M-test, we reduce to showing that $\sum_n \frac{\|X\|^n}{n!}$ converges as

$$\|\frac{X^n}{n!}\| \leq \frac{\|X\|^n}{n!}.$$

Indeed, it converges to $e^{\|X\|}$. □

Out of the above, perhaps the most important is the last one, as it tells us that if we have a diagonalizable matrix $A = PDP^{-1}$, then knowing its eigenvalues (that is, knowing D) and the matrix P is enough for us to calculate the e^A. Indeed, one should note that the exponent of a matrix is not easy to compute all the time!

We now give the lemma which will be quite useful for our goals, that the derivative of exponential of matrices is the obvious one.

Lemma 7.2.1.2. *Let $X \in M_n(\mathbb{R})$. Then,*

$$\frac{d}{dt} e^{At} = Ae^{At}.$$

Proof. One would need to interchange two limits at one point, which could only be done if the convergences are uniform. This could be shown by M-test. □

7.2.2 Fundamental theorem of linear systems

The most important theorem for linear systems of the form $\frac{dx}{dt} = Ax$ is that that they have a unique solution.

Theorem 7.2.2.1. *Let $A \in M_n(\mathbb{R})$. Then for any $x_0 \in \mathbb{R}^n$, the IVP*

$$\frac{dx}{dt} = Ax(t)$$

with $x(0) = x_0$ has a unique solution given by

$$x(t) = e^{At}x_0.$$

Proof. Suppose $y(t)$ is another solution. Then, define $z(t) = e^{-At}y(t)$. Differentiating this, we get

$$\frac{d}{dt}z(t) = -Ae^{-At}y(t) + e^{-At}\frac{dy}{dt}(t).$$

Since $\frac{dy}{dt} = Ay$, thus the above equation gives $\frac{d}{dt}z(t) = -Ae^{-At}y + e^{-At}Ay = 0$. Hence $z(t) = c$ is constant, therefore $y(t) = ce^{At}$. Since $y(0) = x_0 = c$, therefore $y = x$. □

Non-homogeneous linear systems

A non-homogeneous linear system is a linear IVP with an *offset*; they are of the form:

$$\frac{dx}{dt} = Ax(t) + b(t)$$

with $x(0) = x_0$. Their solution have a peculiar form.

Lemma 7.2.2.2. *Let $\frac{dx}{dt} = Ax(t) + b(t)$ with $x(0) = x_0$ be a non-homogeneous IVP for $A \in M_n(\mathbb{R})$. Then x is a solution if and only if*

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}b(s)ds.$$

Proof. We can multiply the IVP by e^{-At} to obtain

$$\begin{aligned}
e^{-At}\frac{dx}{dt} &= Ae^{-At}x + e^{-At}b(t) \\
e^{-At}\frac{dx}{dt} - Ae^{-At}x &= e^{-At}b(t) \\
\frac{d}{dt}[e^{-At}x] &= e^{-At}b(t) \\
x(t) &= e^{At}x_0 + e^{At}\int_0^t e^{-As}b(s)ds.
\end{aligned}$$

One can easily check that the given form satisfies the IVP, by an application of fundamental theorem of calculus. □

7.3 Stability of linear systems in $\mathbb{R}^2$

Consider the linear IVP given by

$$\frac{dx}{dt} = Ax(t)$$

with $x(0) = x_0$ where $x(t) = (x_1(t), x_2(t)) \in \mathbb{R}^2$ and $A \in M_2(\mathbb{R})$. From the fundamental theorem, we know that the solution is of the form $x(t) = e^{At}x_0$. By Jordan form, we know that there exists base change matrix $P \in GL_2(\mathbb{R})$ such that $A = P^{-1}BP$ where B is in Jordan form and hence it is of either of the three forms:

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}, \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

By Theorem 7.2.1.1, 9, we get

$$x(t) = e^{At}x_0 = e^{P^{-1}BPt}x_0 = P^{-1}e^{Bt}Px_0$$

so we reduce to understanding the plots of $e^{Bt}x_0$ for the aforementioned three cases, in order to understand the plot of $e^{At}x_0$ as both are related by coordinate transformation by P.

A phase portrait of a linear system

$$\frac{dx}{dt} = Ax(t)$$

is a plot of $x_1(t)$ vs $x_2(t)$ for various choices of initial points. Indeed, the choice of initial points is paramount if one ought to find the behavior of solutions. On the basis of the analysis of the three cases for B, we make the following definitions.

Definition 7.3.0.1. Let $\frac{dx}{dt} = Ax$ be a linear system where $\det A \neq 0$ and $A \in M_2(\mathbb{R})$. Then, the system is said to have

1. **saddle at origin** if $A \sim \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ where $\lambda < 0 < \mu$,
2. **node at origin** if
   (a) $A \sim \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ where λ, μ have same sign,
   (b) $A \sim \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$,
3. **focus at origin** if $A \sim \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$,
4. **center at origin** if $A \sim \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$.

7.4 Autonomous systems

An IVP is said to be *autonomous* if the governing equation

$$\frac{dx}{dt} = f(x(t))$$

is such that the continuous map $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is independent of time parameter t and we further assume that $f \in C^1$. In such a case we write $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$, and for a fixed initial datum, the maximal interval of existence is unique as well (see Corollary 7.1.2.11)

One calls a point $x_0 \in D$ to be an *equilibrium point* of the $\frac{dx}{dt} = f(x(t))$ if $f(x_0) = 0$.

7.4.1 Flows and Liapunov stability theorem

In our attempt at a better understanding of the autonomous system's dependence on initial point, we develop a basic machinery to handle it. The phase plots were a tool only available for linear systems, but we are not dealing with then in this section. Note however that a linear system is also autonomous.

The first tool we want to make is the notion of flows.

Definition 7.4.1.1. (**Flows**) Consider the following autonomous ODE

$$\frac{dx}{dt} = f(x(t))$$

where $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a continuous map. Denote by $\varphi(-, y) : I_y \to E$ to be a solution of the IVP $(f, (0, y))$ defined on **the** maximal interval of existence I_y for $\varphi(-, y)$ (Lemma 7.1.4.2). The map

$$\begin{aligned} \varphi : I \times E &\longrightarrow E \\ (t, y) &\longmapsto \varphi(t, y) \end{aligned}$$

is called the flow of the system and the map $\varphi(t, -) : E \to E$ is called the *flow of the system at time t*. As we argued in the beginning, there is only one maximal interval of existence for each initial datum.

Remark 7.4.1.2. For a pair $(t, y) \in I \times E$, the value of the flow $\varphi(t, y) \in E$ tells us where the solution $\varphi(-, y)$ takes the initial point y at time t.

We have some obvious observations.

Lemma 7.4.1.3. *Consider the following autonomous ODE*

$$\frac{dx}{dt} = f(x(t)).$$

Let $\varphi : I \times E \to E$ be the flow of the system. Then,

1. $\varphi(0, y) = y$.
2. $\varphi(s, \varphi(t, y)) = \varphi(s + t, y)$.
3. $\varphi(-t, \varphi(t, y)) = y$.

Proof. Trivial. □

We now define the important notions surrounding stability.

Definition 7.4.1.4. (**Stability**) Consider the following autonomous ODE

$$\frac{dx}{dt} = f(x(t)).$$

Let $\varphi : I \times E \to E$ be the flow of the system.

1. An equilibrium point $x_0 \in E$ is said to be *(Liapunov)stable* if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $x \in B_\delta(x_0) \implies \varphi(t,x) \in B_\epsilon(x_0)\ \forall t \geq 0$.
2. An equilibrium point $x_0 \in E$ is said to be *unstable* if it is not stable.
3. An equilibrium point $x_0 \in E$ is said to be *asymptotically stable* if it is stable and $\exists r > 0$ such that

$$x \in B_r(x_0) \implies \lim_{\overrightarrow{t \to \infty}} \varphi(t,x) = x_0.$$

We are now ready to state one of the most important results in stability theory, the Liapunov stability theorem. This result gives a sufficient condition for stability of a given point in the domain of $f : E \to \mathbb{R}^n$ of an autonomous system.

Theorem 7.4.1.5. *(Liapunov stability theorem) Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map and*

$$\frac{dx}{dt} = f(x(t))$$

be a given autonomous system with $x_0 \in E$ being an equilibrium point. If there exists a map of class C^1

$$V : E \to \mathbb{R}$$

such that $V(x_0) = 0$ and $V(x) > 0$ for all $x \in E \setminus \{x_0\}$, then

1. *if $V'(x) \leq 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is stable,*
2. *if $V'(x) < 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is asymptotically stable,*
3. *if $V'(x) > 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is unstable.*

Remark 7.4.1.6. It is important to note that for most of the autonomous systems in nature, the function V as above which will do the job will be the energy functional of the physical system, that is, sum of kinetic and potential energy.

7.5 Linearization and flow analysis

Consider the following *system*:

$$x' = f(x)$$

where $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a continuous map and E is an open set. In the terminology of what we have covered so far, we have an autonomous system. In general, the above system may not be linear, as we studied previously. However, we can *linearize* the system at an equilibrium point x_0, as we shall show below. Indeed, this allows us to analyze the general autonomous system around each point as if it were linear.

Construction 7.5.0.1. (*Linearization of system at a point*) Let $E \subseteq \mathbb{R}^n$ be an open set and $f : E \to \mathbb{R}^n$ be a C^1 map. Let $x_0 \in E$ be an equilibrium point. For any $x \in E$, by Taylor's theorem, we get

$$\begin{aligned} f(x) &= f(x_0) + Df(x_0)(x - x_0) + \text{higher order terms} \\ &= Df(x_0)(x - x_0) + \text{higher order terms} \\ &= A(x - x_0) + \text{higher order terms.} \end{aligned}$$

We thus call the $x' = Df(x_0)x$ to be the *linearization of the system f at point x_0.*

Few definitions are in order.

Definition 7.5.0.2. (**Hyperbolic, sink, source & saddle points**) Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1-map. An equilibrium point $x_0 \in E$ is said to be:

1. *hyperbolic* if all eigenvalues of $Df(x_0)$ has non-zero real part,
2. *sink* if all eigenvalues of $Df(x_0)$ has negative real part,
3. *source* if all eigenvalues of $Df(x_0)$ has positive real part,
4. *saddle* if there exists eigenvalues λ, μ of $Df(x_0)$ such that real part of λ is > 0 and real part of μ is < 0.

7.5.1 Stable manifold theorem

"*For a non-linear system, there are stable and unstable submanifolds, so that once you are in either of them, the flow will constrain you to remain there.*"

We will do an important theorem in the theory of linearization of autonomous systems. We shall avoid the proof this theorem. A reference is pp 107, [cite Perko]. Let us first define three important subspaces corresponding to a linear system.

Definition 7.5.1.1. (**Stable, unstable & center subspaces**) Let

$$x' = Ax$$

be a linear system where $A \in M_n(\mathbb{R})$. Let $\lambda_j = a_j + ib_j$ be eigenvalues of A and $w_j = u_j + iv_j$ be a generalized eigenvector of λ_j. Then,

1. the stable subspace E^s is defined to be the span of all u_j, v_j in $\mathbb{R}^n$ for those $j = 1, \dots, n$ such that $a_j < 0$,
2. the unstable subspace E^u is defined to be the span of all u_j, v_j in $\mathbb{R}^n$ for those $j = 1, \dots, n$ such that $a_j > 0$,
3. the center subspace E^c is defined to be the span of all u_j, v_j in $\mathbb{R}^n$ for those $j = 1, \dots, n$ such that $a_j = 0$.

Lemma 7.5.1.2. *Let $x' = Ax$ be a linear system for $A \in M_n(\mathbb{R})$. Then,*

1. $\mathbb{R}^n = E^s \oplus E^u \oplus E^c$,
2. *E^s, E^u and E^c are invariant under the flow $\varphi(t, x)$ of the linear system, which as we know is given by $e^{At}x$.*

Proof. 1. This is easy, as generalized eigenvectors always span the whole space.
2. We need only show that for a generalized eigenvector w_j corresponding to $\lambda_j = a_j + ib_j$ with $a_j < 0$, the vector $A^k w_j$ is again a genralized eigenvector. Indeed, this follows from definition of a generalized eigenvector as $(A - \lambda_j I)w_j$ is again a generalized eigenvector. □

We now come to the real deal.

Theorem 7.5.1.3. *(Stable manifold theorem) Let $E \subseteq \mathbb{R}^n$ be an open subset with $0 \in E$, consider $f : E \to \mathbb{R}^n$ to be a C^1-map and consider the system that it defines. Denote E^s and E^u to be the stable and unstable subspaces of the system $x' = Df(0)x$. If,*

- *$f(0) = 0$,*
- *$Df(0) : \mathbb{R}^n \to \mathbb{R}^n$ has k eigenvalues with negative real part and $n-k$ eigenvalues with positive real part,*

then:

1. *There exists a k-dimensional differentiable manifold S inside E such that*
   (a) *$T_0 S = E^s$,*
   (b) *for all $t \geq 0$ and for all $x \in S$, we have*
   $$\varphi(t, x) \in S,$$
   (c) *for all $x \in S$, we have*
   $$\lim_{\overrightarrow{t \to \infty}} \varphi(t, x) = 0.$$
2. *There exists an $n-k$-dimensional differentiable manifold inside E such that*
   (a) *$T_0 U = E^u$,*
   (b) *for all $t \leq 0$ and for all $x \in U$, we have*
   $$\varphi(t, x) \in S,$$
   (c) *for all $x \in U$, we have*
   $$\lim_{\overrightarrow{t \to -\infty}} \varphi(t, x) = 0.$$

Let us explain via an example

Example 7.5.1.4. Consider the system

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} -x_1 \\ -x_2 + x_1^2 \\ x_3 + x_1^2 \end{bmatrix}.$$

This is not a linear system as for $f((x_1, x_2, x_3)) = (-x_1, -x_2 + x_1^2, x_3 + x_1^2)$, the above system is given by

$$x' = f(x) \tag{7.1}$$

and $f(x)$ is clearly not linear in x. However, note that $f(0) = 0$. Thus, linearizing the system (7.1) at 0, we obtain the linear system

$$x' = Ax \tag{7.2}$$

where

$$A := Df(0) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So A has two eigenvalues with negative real part, namely -1 and -1 and one eigenvalue with positive real part, namely 1. In particular A is diagonalizable, hence E^s and E^u are just span of the eigenvectors (as all generalized eigenvectors in this case are just your regular eigenvectors). Hence we see

$$E^s = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$
$$E^u = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Hence $E^s = x$-y plane and E^u is the z-axis of $\mathbb{R}^3$.

By an application of stable manifold theorem on this system, the stable manifold S is of dimension 2 and unstable manifold U is of dimension 1. Now, by elementary calculations, we can actually solve the linear system (7.2) and we thus obtain the following solution

$$\begin{aligned} x_1(t) &= c_1 e^{-t} \\ x_2(t) &= c_2 e^{-t} + c_1^2(e^{-t} - e^{-2t}) \\ x_3(t) &= c_3 e^t + \frac{c_1^2}{3}(e^t - e^{-2t}). \end{aligned}$$

Hence, the flow of the system is given by

$$\begin{aligned} \varphi : \mathbb{R} \times \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (t, (c_1, c_2, c_3)) &\longmapsto \begin{pmatrix} cc_1 e^{-t} \\ c_2 e^{-t} + c_1^2(e^{-t} - e^{-2t}) \\ c_3 e^t + \frac{c_1^2}{3}(e^t - e^{-2t}) \end{pmatrix}. \end{aligned}$$

Now, notice the following for any $c = (c_1, c_2, c_3) \in \mathbb{R}^3$

$$\begin{aligned} \varinjlim_{t \to \infty} \varphi(t, c) = 0 &\iff c_3 + \frac{c_1^2}{3} = 0 \\ \varinjlim_{t \to -\infty} \varphi(t, c) = 0 &\iff c_1 = c_2 = 0. \end{aligned}$$

Notice the fact that the above equivalence is very particular to this example. But this leads us to the following conclusions

$$\begin{aligned} S &= \{(c_1, c_2, c_3) \in \mathbb{R}^3 \mid c_3 + c_1^2/3\} \\ U &= \{(c_1, c_2, c_3) \in \mathbb{R}^3 \mid c_1 = c_2 = 0\} \cong z\text{-axis}. \end{aligned}$$

Note that it is indeed true that for all $c \in S$ and any $t \geq 0$, $\varphi(t, c) \in S$. Similarly for U. Finally, one can check that $T_0 S = E^s$ and $T_0 U = E^u$, where the latter is immediate.

7.5.2 Poincaré-Bendixon theorem

So far, for a system we have defined its flow. Flow or integral curves of the system holds important information about the system at hand. However, we have not done any serious analysis with them. We shall begin the analysis of flows of a system now and prove the aforementioned theorem. It's use is predominantly to find closed trajectories of a system, which most of the times appears as a boundary of two differing phenomenon of the system, hence the importance of closed trajectories and of the theorem.

We first set up the terminology to be used in order to define basic objects of analysis of flow of a system.

Definition 7.5.2.1. (ω & α -limit set) Let $E \subseteq \mathbb{R}^n$ be an open set and $f : E \to \mathbb{R}^n$ be a C^1 map. Let $\varphi : \mathbb{R} \times E \to \mathbb{R}^n$ be the flow of the system. Then,

1. a point $y \in E$ is said to be a ω-limit point of $x \in E$ if there exists a sequence $t_1 < t_2 < \cdots < t_n < \dots$ in $\mathbb{R}$ such that $\varinjlim_{n\to\infty} t_n = \infty$ and $\varinjlim_{n\to\infty} \varphi(t_n, x) = y$.
2. a point $y \in E$ is said to be an α-limit point of $x \in E$ if there exists a sequence $t_1 > t_2 > \cdots > t_n > \dots$ in $\mathbb{R}$ such that $\varinjlim_{n\to\infty} t_n = -\infty$ and $\varinjlim_{n\to\infty} \varphi(t_n, x) = y$.

Let $x \in E$, the set of all ω and α limit points of x are denoted $L_\omega(x)$ and $L_\alpha(x)$ respectively.

The following are some simple observations from the definition

Lemma 7.5.2.2. *Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map on an open set E and consider the system given by it.*

1. *If $y \in L_\omega(x)$ and $z \in L_\omega(y)$ then $z \in L_\omega(x)$.*
2. *If $y \in L_\omega(x)$ and $z \in L_\alpha(y)$ then $z \in L_\omega(x)$.*
3. *For any $x \in E$, the limit sets $L_\omega(x)$ and $L_\alpha(x)$ are closed in E.*

Using the concept of limit points, we can define certain nice subspaces of E conducive to them.

Definition 7.5.2.3. (**Positively invariant set**) Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map on an open set E and consider the system defined by it. A region $D \subseteq E$ is said to be positively invariant if for all $x \in D$, $\varphi(t, x) \in D$ for all $t \geq 0$ where $\varphi : \mathbb{R} \times E \to E$ is the flow.

We then have the following simple result.

Lemma 7.5.2.4. *Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map on an open set E and consider the system defined by it.*

1. *If x, z are on same flow line/trajectory, then $L_\omega(x) = L_\omega(z)$.*
2. *For any $x \in E$, the limit set $L_\omega(x)$ is positively invariant.*
3. *If $D \subseteq E$ is a closed positively invariant set, then for all $x \in D$, $L_\omega(x) \subseteq D$.*

We now define another set of tools helpful in doing flow analysis. First is a notion which will come in handy while trying to discuss both the topology of underlying space and the flow together. A hyperplane in $\mathbb{R}^n$ is a codimension 1 linear subspace.

Definition 7.5.2.5. (**Local sections**) Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map on an open set E and consider the system defined by it. Let $0 \in E$. A local section S of f is an open connected subset of a linear hyperplane $H \subseteq \mathbb{R}^n$ such that $0 \in S$ and H is transverse to f, that is, $f(x) \notin H$ for all $x \in S$.

The next tool helps to "straighten" out flow around a local section.

Definition 7.5.2.6. (**Flow box**) Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map on an open set E and consider the system defined by it. Let $0 \in E$ and S be a local section of f. A flow box around S is a diffeomorphism Φ between $(-\epsilon, \epsilon) \times S \subseteq \mathbb{R} \times E$ and $V_\epsilon \subseteq E$ given by $V_\epsilon := \{\varphi(t, x) \mid t \in (-\epsilon, \epsilon), x \in S\}$:

$$\begin{aligned} \Phi : (-\epsilon, \epsilon) \times S &\longrightarrow V_\epsilon \\ (t, x) &\longmapsto \varphi(t, x). \end{aligned}$$

We identify $(-\epsilon, \epsilon) \times S$ as the flow box around S.

For a flow box, the diffeomorphism is important as it tells us that we can assume WLOG in a flow box that flow line are identical to the orthogonal coordinate system of $(-\epsilon, \epsilon) \times S \subseteq \mathbb{R}^{n+1}$.

We would now like to do flow analysis for the special case of planar systems. Indeed, the main theorem of this section is about the behaviour of certain limit sets of planar systems.

Let us first observe that for a planar system, any local section intersects a flow line at only discretly many points.

Lemma 7.5.2.7. *Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. Let $x \in E$ and consider a local section S around x. Let*

$$\Sigma := \{\varphi(t, x) \in E \mid t \in [-l, l]\}.$$

Then $\Sigma \cap S$ is discrete.

Next, we see that if a sequence of points in a local section S of a planar system is monotonous in S and those same points appear in a trajectory, then it is monotonous in that trajectory as well. Indeed, a sequence of points $\{\varphi(t_n, x)\}$ along a trajectory is said to be *monotonous* if $\varinjlim_{n \to \infty} t_n = \infty$. Note that for a planar system, a codimension 1 linear subspace is a line, hence it has an inherent order and thus we can talk about monotonous sequences in a local section.

Proposition 7.5.2.8. *Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. Let S be a local section of the system. If $x_n = \varphi(t_n, x)$ is a sequence of points monotonous along the trajectory and $x_n \in S$, then $\{x_n\}$ are monotonous in S as well.*

One can use the above proposition to deduce some "eventual" properties of points in a local section by observing their intersection points with a flow line (which are discrete). Further it can be used for replacing a sequence along a trajectory to a sequence along a local section, which might be easier to analyze (as it's behaviour will just be that of monotonous sequences in $\mathbb{R}$).

Next we see an important observation, that trajectories of some special points cannot intersect a local section at more than one point(!)

Lemma 7.5.2.9. *Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. For some $x \in E$, let $y \in L_\omega(x) \cup L_\alpha(x)$. Then the trajectory of y intersects any local section at not more than single point.*

The next result is interesting, for it says that if the trajectory of a point intersects a local section, then there is a whole neighborhood worth of point around it, each of whose trajectories will intersect the local section(!) In some sense, this corresponds to the continuity of flow.

Proposition 7.5.2.10. *Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. Let $\varphi : \mathbb{R} \times E \to \mathbb{R}^2$ denote the flow of the system. Let S be a local section around $y \in E$. If there exists $z_0 \in E$ such that for some $t_0 > 0$ we have $\varphi(t_0, z_0) = y$, then*

1. *there exists an open set $U \ni z_0$,*
2. *there exists a unique C^1-map $\tau : U \to \mathbb{R}$,*

where τ has the property that $\tau(z_0) = t_0$ and

$$\varphi(\tau(z), z) \in S \; \forall z \in U.$$

With this, we define the main object of study, a closed orbit.

Definition 7.5.2.11. (**Closed orbits**) Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. A closed orbit is a periodic trajectory which doesn't contain an equilibrium point.

Note that if a trajectory contains an equilibrium point, then it will terminate after some finite time, hence the above requirement.

We now come to the main theorem of this section, which tells us a sufficient condition to find a closed orbits of a planar system.

Theorem 7.5.2.12. *(Poincaré-Bendixon theorem) Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. Let $x \in E$ be such that $L_\omega(x)$ ($L_\alpha(x)$) is a non-empty compact limit set which doesn't contain an equilibrium point. Then $L_\omega(x)$ ($L_\alpha(x)$) is a closed orbit.*

Let us now give some applications of the above theorem. First, we can classify limit sets $L_\omega(x)$ completely.

Theorem 7.5.2.13. *(Classification of limit sets) Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. Let $x \in E$ be such that $L_\omega(x)$*

- *is connected,*
- *is compact,*
- *has finitely many equilibrium points.*

Then one of the following holds

1. *$L_\omega(x)$ is a singleton.*
2. *$L_\omega(x)$ is periodic trajectory with no equilibrium points.*
3. *$L_\omega(x)$ consists of equilibrium points $\{x_j\}$ and a set of non-periodic trajectories $\{\gamma_i\}$ such that for all i, the trajectory γ_i tends to some x_j as $t \to \pm\infty$.*

The main use of Poincaré-Bendixon is to find limit cycles.

Definition 7.5.2.14. (**Limit cycles**) Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. A limit cycle is a periodic trajectory γ such that there exists $x \in E$ for which $\gamma \subseteq L_\omega(x)$ or $\gamma \subseteq L_\alpha(x)$.

We now state the corollary of Poincaré-Bendixon which allows us to find the existence of limit cycles.

Corollary 7.5.2.15. *Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map on an open set E and consider the planar system defined by it. If there exists a subseteq $D \subseteq E$ such that D*

1. *is compact,*
2. *is positively invariant,*
3. *has no equilibrium points,*

then there exists a limit cycle in D.

Proof. By Poincaré-Bendixon, we need only find $x \in D$ such that $L_\omega(x)$ is compact, as then $L_\omega(x)$ itself will be the limit cycle. This is straightforward, as D is positively invariant and compact, so $L_\omega(x)$ is inside D and is closed (hence compact). $\square$

7.6 Second order ODE

We now discuss some basic theory of second order ordinary differential equations.

Definition 7.6.0.1. (**Second order system and solutions**) Let $I \subseteq \mathbb{R}$ be an interval of $\mathbb{R}$ and consider $a_0, a_1, a_2, g \in C(I)$ to be four continuous maps $I \to \mathbb{R}$ such that $a_0(x) > 0\ \forall x \in I$. Then, a second order system with parameters a_0, a_1, a_2, g is given by

$$a_0(x)y'' + a_1(x)y' + a_2(x)y = g(x). \tag{7.3}$$

Note that $y' := \frac{dy}{dx}$. A *solution* of a second order system (q, r, f) is a $C^2(I)$ map $y(x)$ such that it satisfies (7.3).

Remark 7.6.0.2. A second order ODE can be written in the form

$$y'' + q(x)y' + r(x)y = f(x)$$

where $q, r, f \in C(I)$. This form is the one that we shall use and will identify a second order system by the tuple (q, r, f).

Remark 7.6.0.3. On the $\mathbb{R}$-vector space $C^2(I)$ of twice continuously differentiable functions, every 2nd order system (q, r, f) defines a linear transformation

$$\begin{aligned} L : C^2(I) &\longrightarrow C(I) \\ y(x) &\longmapsto (D^2 + q(x)D + r(x))y \end{aligned}$$

where $D : C^2(I) \to C(I)$ is the derivative transformation $y \mapsto y'$, which is evidently linear. In this notation, we can write a second order system (q, r, f) as

$$Ly = f$$

where $L = D^2 + qD + r$. We call this linear transformation L the *transform associated to* (q, r, f).

Definition 7.6.0.4. (**Solution space**) Let (q, r, f) be a 2nd order system and $L : C^2(I) \to C(I)$ be the associated transform. The solution space of (q, r, f) is defined as the $\mathrm{Ker}\,(L) \subseteq C^2(I)$. Note that the set of all solutions of (q, r, f) in $C^2(I)$ is given by $L^{-1}(f) \subseteq C^2(I)$.

Lemma 7.6.0.5. *Let (q, r, f) be a 2nd order system and L be the associated transform. Then* $\dim_{\mathbb{R}}(\mathrm{Ker}\,(L)) = 2$.

We now observe that one can obtain all solutions of the 2nd order system $S := (q, r, f)$ by obtaining a basis of the solution space of S and one solution of S.

Lemma 7.6.0.6. *Let $S = (q, r, f)$ be a 2nd order system and L be the associated transform. Then, for any $y_p \in L^{-1}(f)$*

$$L^{-1}(f) = y_p + \mathrm{Ker}\,(L).$$

Proof. Observe that $y - y_p \in \mathrm{Ker}\,(L)$ and a linear transformation has all fibers of same size. □

We define a tool which helps in distinguishing independent or dependent solutions of a homogeneous system.

Definition 7.6.0.7. (**Wronskian**) Let $f, g \in C^1(I)$. The Wronskian of f and g is given by

$$W(f, g) : I \to \mathbb{R}$$

where for any $x \in I$, we have

$$\begin{aligned} W(f, g)(x) :&= \det \begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix} \\ &= f(x)g'(x) - g(x)f'(x). \end{aligned}$$

Lemma 7.6.0.8. *Let $(q, r, 0)$ be a homogeneous system and let $y_1, y_2 \in C^2(I)$ be two solutions. Then,*

1. *$W(y_1, y_2)$ is either constant 0 for all $x \in I$ or $W(y_1, y_2)(x) \neq 0$ for all $x \in I$.*
2. *y_1, y_2 are linearly independent if and only if $W(y_1, y_2) \neq 0\ \forall x \in I$.*

7.6.1 Zero set of homogeneous systems

Let (q, r, f) be a 2nd order system and let y be a solution. There are some peculiar properties of the zero set $Z(y) := \{x \in I \mid y(x) = 0\} \subseteq \mathbb{R}$. We first show that the set $Z(y)$ is discrete if the system is homogeneous.

Lemma 7.6.1.1. *Let $(q, r, 0)$ be a 2nd order homogeneous system and let y be a solution. The zeroes of $y(x)$ are isolated, that is, $Z(y)$ is discrete.*

Strum separation and comparison theorems

These theorems are at the heart of the analysis of zeros of homogeneous systems.

Theorem 7.6.1.2. *(Strum separation theorem) Let $(q, r, 0)$ be a 2nd order homogeneous system. Let y_1, y_2 be two distinct linearly independent solutions of the system. Then,*

1. *$Z(y_1)$ and $Z(y_2)$ are disjoint.*
2. *$Z(y_1)$ and $Z(y_2)$ are braided, that is, for any two x_1^1 and x_2^1 in $Z(y_1)$, there exists $x_1^2 \in Z(y_2)$ between them, and vice versa.*

Theorem 7.6.1.3. *(Strum comparison test) Consider two homogeneous 2nd order systems* $(0, r_1, 0)$ *and* $(0, r_2, 0)$. *Let* y *be a solution of* $(0, r_1, 0)$ *and* u *be a solution of* $(0, r_2, 0)$, *both non-trivial. Let* $x_1, x_2 \in Z(u)$ *such that*

1. $r_1(x) \geq r_2(x)$ *for all* $x \in (x_1, x_2)$,
2. $\exists\, x_k \in (x_1, x_2)$ *such that* $r_1(x_k) > r_2(x_k)$.

Then, there exists $z \in Z(y)$ *such that* $z \in (x_1, x_2)$.

7.6.2 Boundary value problems

A *boundary value problem* (BVP) is a second order system on an interval $I = [a, b]$ given by

$$y'' + qy' + ry = f$$

for $q, r, f \in C(I)$ such that its solutions has to satisfy certain conditions on the boundary given by

$$\begin{aligned} B_a(y) &:= \alpha_1 y(a) + \beta_1 y'(a) = 0 \\ B_b(y) &:= \alpha_2 y(b) + \beta_2 y'(b) = 0 \end{aligned}$$

where $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2$. This is clearly a different problem than that of IVP. However, with some construction, we can convert this problem into a pair of 2nd order IVPs. It will turn out that the solution of this pair has important consequences for the original IVP at hand.

Reduction to a pair of 2nd order IVPs and criterion for uniqueness of BVP solution

Theorem 7.6.2.1. *Let* $I = [a, b]$ *and* $q, r, f \in C(I)$. *Consider the 2nd order system* (q, r, f) *and denote the associated transform as* $L : C^2(I) \to C^2(I)$. *From the system* (q, r, f) *consider the BVP given explicitly by*

$$\begin{aligned} Ly &:= y'' + qy' + ry = f \\ B_a(y) &:= \alpha_1 y(a) + \beta_1 y'(a) = 0 \\ B_b(y) &:= \alpha_2 y(b) + \beta_2 y'(b) = 0 \end{aligned} \tag{7.4}$$

where $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2$. *Construct the following two 2nd order IVPs*

$$\begin{aligned} Ly &:= y'' + qy' + ry = 0 \\ y(a) &= \beta_1 \\ y'(a) &= -\alpha_1 \end{aligned} \tag{7.5}$$

and

$$\begin{aligned} Ly &:= y'' + qy' + ry = 0 \\ y(b) &= \beta_2 \\ y'(b) &= -\alpha_2. \end{aligned} \tag{7.6}$$

Then the following are equivalent

1. *Let* y_1 *be a solution of* (7.5) *and* y_2 *be a solution of* (7.6). *Then* y_1 *and* y_2 *are linearly independent in the solution space* $\mathrm{Ker}\,(L)$.

2. *The homogeneous BVP*

$$\begin{aligned} Ly := y'' + qy' + ry = 0 \\ B_a(y) = 0 \\ B_b(y) = 0 \end{aligned} \tag{7.7}$$

has only 0 as solution.
3. *The BVP (7.4) has a unique solution.*

Variation of parameters

Variation of parameters can give us a general form of a particular solution of $Ly = f$, in terms of the solutions of IVPs (7.5) and (7.6). Indeed, we have the following theorem.

Theorem 7.6.2.2. *Let y_1 be a solution of (7.5) and y_2 be a solution of (7.6). Let*

$$\begin{aligned} c_1(x) = \int_a^x \frac{-f(s)y_2(s)}{W(y_1,y_2)(s)}ds \\ c_2(x) = \int_a^x \frac{f(s)y_1(s)}{W(y_1,y_2)(s)}ds. \end{aligned} \tag{7.8}$$

Then,

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x) \tag{7.9}$$

is a particular solution of $Ly = f$ with $y_p(a) = 0$.

Further, we obtain a general form of solution of BVP (7.4).

Theorem 7.6.2.3. *Consider the notations of Theorems 7.6.2.1 and 7.6.2.2.*
1. *Any solution y of BVP (7.4) is*

$$y = y_p - c_1(b)y_1$$

where y_1 is a solution of (7.5) and $c_1(x)$ is defined in (7.8).
2. *Any solution of the BVP (7.4) is given by the integral*

$$y(x) = \int_a^b G(x,s)f(s)ds \tag{7.10}$$

for all $x \in I$, where

$$G(x,s) = \begin{cases} \frac{y_1(x)y_2(s)}{W(y_1,y_2)(s)} & \text{if } x \le s \le b \\ \frac{y_1(s)y_2(x)}{W(y_1,y_2)(s)} & \text{if } a \le s \le x. \end{cases} \tag{7.11}$$

This map G is called the Green's function for the transformation $L : C^2(I) \to C(I)$.

Strum-Liouville system

Let $p, q \in C^2(I)$ and $f \in C(I)$ with $p > 0$. Define the 2nd order system

$$py'' + p'y' + qy = f.$$

We can write it in neater terms as follows

$$(py')' + qy = f. \tag{7.12}$$

We will call this the *Strum-Liouville system*, denoted by (p, q, f), and the associated transform as $L : C^2(I) \to C(I)$ mapping $y \mapsto (py')' + qy$. Consequently, (7.12) can be written as

$$Ly := (py')' + qy = f.$$

We have some basic results about the associated transform L.

Lemma 7.6.2.4. *Let (p, q, f) be a Strum-Liouville system and L be the associated transform.*

1. *(Lagrange's identity) If $y_1, y_2 \in C^2(I)$, then*

$$y_1 L y_2 - y_2 L y_1 = (pW(y_1, y_2))'.$$

2. *(Abel's formula) If y_1, y_2 are solutions of $Ly = 0$, that is, they are solutions of the Strum-Liouville system defined by $(p, q, 0)$, then*

$$W(y_1, y_2) = c/p$$

for some constant $c \in \mathbb{R}$.

Strum-Liouville Boundary Value Problems (SL-BVPs)

Consider a homogeneous Strum-Liouville system $(p, q, 0)$ and let L be the associated transform. Consider $r \in C(I)$ and $\lambda \in \mathbb{C}$. Then, a *Strum-Liouville boundary value problem* is a following type of 2nd order BVP

$$Ly + \lambda r y = 0 \tag{7.13}$$

with

$$B_a(y) = 0$$
$$B_b(y) = 0.$$

Strum-Liouville EigenValue Problems (SL-EVPs)

An SL-EVP consists of an SL-BVP (7.13) and the following question: find $\lambda \in \mathbb{C}$ such that the SL-BVP (7.13) admits a non-zero solution $y_\lambda \in C^2(I)$. In such a case λ is called the *eigenvalue* and y_λ the *eigenfunction* of the corresponding SL-EVP. We then call the tuple (p, q, r) as the SL-EVP.

Types of SL-EVPs

We further classify an SL-EVP (p, q, r) based on the properties of the underlying functions.

1. **regular** if $p > 0$ and $r > 0$ on $[a, b]$,
2. **singular** if $p > 0$ on (a, b), $p(a) = 0 = p(b)$ and $r \geq 0$ on $[a, b]$,
3. **periodic** if $p > 0$ on $[a, b]$, $p(a) = p(b)$ and $r > 0$ on $[a, b]$.

We next see that any eigenvalue of SL-EVP is always real.

Lemma 7.6.2.5. *Let (p, q, r) be a regular SL-EVP. Then all eigenvalues of (p, q, r) are real.*

Chapter 8

K-Theory of Vector Bundles

Chapter 9

Jet Bundles

It is through the concepts explained in this chapter that we shall begin doing some geometry over our base manifolds. A classical use of differential equations is elaborated in Chapter 7, whereas here we shall be more conceptual in order to elucidate the underlying structure of the notion of "differential equations".

Part IV

The Analytic Viewpoint

Chapter 10

Analysis on Complex Plane

Contents

We will here review some of the classical results of complex function theory of one variable, namely the following four topics:

- Analytic functions; Cauchy-Riemann equations, harmonic functions.
- Complex integration; Zeroes of analytic functions, winding numbers, Cauchy's formula and theorem, Liouville's theorem, Morera's theorem, open-mapping theorem, Schwarz's lemma.
- Singularities; Classification, Laurent series, Casorati-Weierstrass theorem, residues and applications, meromorphic maps, Rouché's theorem.
- Conformal maps; Möbius transformations, normality and compactness, Riemann mapping theorem.

All this is important as it will build one's intuition of geometry in complex case, which is something we don't learn as early in our studies as, say, real geometry. Of-course this would be of immense use in complex algebraic geometry, some if which we shall cover in later chapters. Moreover, a complex manifold by definition locally looks like $\mathbb{C}^n$, hence it is imperative that we understand the geometry and analysis of complex plane and make it as second nature as the usual geometry over $\mathbb{R}^2$ is to us.

10.1 Holomorphic functions

Let $\Omega \subseteq \mathbb{C}$ denote an open subset of the complex plane $\mathbb{C}$ for the rest of this chapter. Consider a function $f : \Omega \to \mathbb{C}$. Motivated by the classical case of real differentiability in one variable, we can define a notion of differentiation for f at $a \in \Omega$.

Definition 10.1.0.1. (**$\mathbb{C}$-differentiable/holomorphic functions**) A function $f : \Omega \to \mathbb{C}$ is $\mathbb{C}$-differentiable or holomorphic at $a \in \Omega$ if the following limit exists:

$$\varinjlim_{z \to 0} \frac{f(a+z) - f(a)}{z}$$

in which case it's value is said to be the derivative of f at a and is denoted by $\frac{df}{dz}(a) = f'(a) \in \mathbb{C}$.

Remark 10.1.0.2. As we shall soon see, this seemingly innocuous definition for some surprising reason gives the following fantastic results:

1. Theorems 10.1.1.2 and **??** tells us:

$$\begin{gathered}\{\textit{All } \mathbb{C}\textit{-differentiable maps } f : \Omega \to \mathbb{C}\} \\ \cong \\ \{\textit{All pairs of differentiable maps } u, v : \Omega \to \mathbb{R}, \textit{ related by CR-equations}\}\end{gathered}.$$

2. Corollary 10.1.2.2 and Theorem **??** tells us:

$\mathbb{C}$-differentiable maps are conformal.

3. Theorem **??** tells us:

$\mathbb{C}$-differentiable functions are harmonic.

Moreover, Theorem **??** tells us that if Ω is simply connected, then

$$\begin{gathered}\{\textit{Harmonic functions } \Omega \subseteq \mathbb{R}^2 \cong \mathbb{C}\} \\ \cong \\ \{\mathbb{C}\textit{-differentiable functions on } \Omega \subseteq \mathbb{C}\}\end{gathered}.$$

4. Theorem **??** tells us:

Contour integral of a $\mathbb{C}$-differentiable map around a loop is 0.

5. Theorem **??** tells us:
   A $\mathbb{C}$-differentiable function inside a disc is determined by its values on the disc's boundary.
6. Corollary 10.2.3.5 tells us:

$$\begin{array}{c}\{\mathbb{C}\text{-}\textit{differentiable maps } f : \Omega \to \mathbb{C}\} \\ \cong \\ \{\textit{Analytic maps } f : \Omega \to \mathbb{C}\}\end{array}.$$

This shows the sheer importance of the notion of $\mathbb{C}$-differentiability, which we will explore later in a more *local* setting. Our goal in the rest of this chapter is to provide rather quick proofs to these results while portraying the main ideas employed in them.

Let us start by analyzing some elementary properties of holomorphic maps.

10.1.1 Cauchy-Riemann equations

Let $f : \Omega \to \mathbb{C}$ be a holomorphic map on an open subset $\Omega \subseteq \mathbb{C}$. Now, there is a homeomorphism $\varphi : \mathbb{R}^2 \to \mathbb{C}$ given by $(x, y) \mapsto x + iy$. Composing f with this map, we get that f can equivalently be stated as the data of two real valued maps $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : \mathbb{R}^2 \to \mathbb{R}$ given by $u(x, y) = \Re f(\varphi(x, y))$ and $v(x, y) = \Im f(\varphi(x, y))$.

Like in the case of $\mathbb{R}$-differentiability, in our case we can also define partial differential operators of f w.r.t. x, y and z.

Definition 10.1.1.1. (**Partial differential operators on f**) Let $f : \Omega \to \mathbb{C}$ be a holomorphic map on an open subset Ω of $\mathbb{C}$. Then, we define the following quantities in an obvious manner:
1. $\frac{\partial f}{\partial x} := \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$.
2. $\frac{\partial f}{\partial y} := \frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}$.
3. $\frac{\partial f}{\partial z} := \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right)$.
4. $\frac{\partial f}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)$.

Then the fact that f is holomorphic can be equivalently stated in terms of real differentiability of the maps u and v as the following theorem states.

Theorem 10.1.1.2. *Suppose $f : \Omega \to \mathbb{C}$ is any $\mathbb{C}$-valued function on an open set Ω of $\mathbb{C}$. Then write $f(x + iy) = u(x, y) + iv(x, y)$ where $u, v : \mathbb{R}^2 \rightrightarrows \mathbb{R}$.*
1. *$f : \Omega \to \mathbb{C}$ is holomorphic at $z_0 \in \Omega$ if and only if u, v are real differentiable and satisfy any of the following equivalent PDEs at z_0:*
   (a) $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ & $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$.
   (b) $\frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y}$.
   (c) $\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}$.
   (d) $\frac{\partial f}{\partial \bar{z}} = 0$.
2. *If $u, v : \Omega \to \mathbb{R}$ is a pair of C^1-maps satisfying the CR-equations, then $f := u + iv$ is a holomorphic map.*

Proof. Equivalence of the four PDEs is straightforward. Now let $f : \Omega \to \mathbb{C}$ be a holomorphic map. This means that for any $a \in \Omega$, we have

$$\frac{\partial f}{\partial z}(a) = \varinjlim_{z \to 0} \frac{f(a+z) - f(a)}{z}.$$

The required PDEs for u and v follows by letting z approach 0 first from real axis and then from imaginary axis and deeming them equal.

Next, we may write $R(z) = f(a+z) - f(a) - cz$ for some $c = c_1 + ic_2$ and then $R(z) = R_u(z) + iR_v(z)$ where $R_u(z) = u(a+z) - u(a) - c_1 x + c_2 y$ and $R_v(z) = v(a+z) - v(a) - c_2 x - c_1 y$. Then, f is holomorphic at a with $\frac{df}{dz}(a) = c$ if and only if $\varinjlim_{z\to 0} \frac{R(z)}{z} = 0$. But the latter happens if and only if $\varinjlim_{z\to 0} \frac{R_u(z)}{z} = 0 = \varinjlim_{z\to 0} \frac{R_v(z)}{z}$. Now $\frac{R_u(z)}{z} = 0$ if and only if $c_1 = \frac{\partial u}{\partial x}(a)$ and $c_2 = -\frac{\partial u}{\partial y}(a)$. Similarly, $\varinjlim_{z\to 0} \frac{R_v(z)}{z} = 0$ if and only if $c_2 = \frac{\partial v}{\partial x}(a)$ and $c_1 = \frac{\partial v}{\partial y}(a)$. □

10.1.2 Conformal maps

We will now show that holomorphic maps "preserves angles". The meaning of angle is not well-defined a-priori on the complex plane, so we will have to develop that first.

A *curve* in $\mathbb{C}$ is a continuous map $\gamma : I \to \mathbb{C}$. It is said to be *differentiable* if $\Re\gamma : I \to \mathbb{R}$ and $\Im\gamma : I \to \mathbb{R}$ are differentiable $\mathbb{R}$-valued functions. It is said to be regular at $t \in I$ if $\gamma'(t) \neq 0 \in \mathbb{C}$. Now, let $\gamma_1, \gamma_2 : I \to \mathbb{C}$ be two curves which intersect at $\gamma_1(t_1) = \gamma_2(t_2)$ for $t_1, t_2 \in I$ such that γ_i is regular at t_i, $i = 1, 2$. Such an intersection is said to be *regular*. Then, the angle of intersection of γ_1 and γ_2 at a regular point is defined to be:

$$\angle\gamma_1(t_1), \gamma_2(t_2) := \arg\gamma_2'(t_2) - \arg\gamma_1'(t_1).$$

A function $f : \Omega \to \mathbb{C}$ is *conformal* at $z_0 \in \Omega$ if f preserves angles of all regular intersections of two curves at z_0.

It is now an easy observation that holomorphic maps will be conformal.

Lemma 10.1.2.1. *Let $f : \Omega \to \mathbb{C}$ be a holomorphic map on an open set Ω of $\mathbb{C}$. If $z_0 \in \Omega$ such that $f'(z_0) \neq 0$, then for any two curves γ_1, γ_2 such that $\gamma_1(t_1) = z_0 = \gamma_2(t_2)$ and γ_1, γ_2 are regular at t_1, t_2 respectively, then*

$$\angle\gamma_1(t_1), \gamma_2(t_2) = \angle f \circ \gamma_1(t_1), f \circ \gamma_2(t_2).$$

Proof. The result follows from chain rule and the fact that $\arg wz = \arg w + \arg z$. □

A map $f : \Omega \to \mathbb{C}$ is called *conformal* if it preserves angles of all regularly intersecting curves. Thus,

Corollary 10.1.2.2. *All holomorphic functions are conformal except at those points at which derivative is zero.* □

We will now show that even an arbitrary conformal map $f : \Omega \to \mathbb{C}$ is also holomorphic.

Theorem 10.1.2.3. *Let $f : \Omega \to \mathbb{C}$ be a conformal map such that $\Re f$ and $\Im f$ are of class C^1. Then,*

1. *f is holomorphic.*
2. *$f'(z) \neq 0$ for all $z \in \Omega$.*

Proof. Simple thus omitted. □

10.1.3 Harmonic maps

A function $f : \Omega \to \mathbb{C}$ is said to be harmonic if $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$. Below are some straightforward implications of Cauchy-Riemann equations.

Lemma 10.1.3.1. *Let $f = u + iv : \Omega \to \mathbb{C}$ be a function where $u, v : \Omega \rightrightarrows \mathbb{R}$. Then, f is harmonic if and only if u and v are harmonic (in $\mathbb{R}$-sense).*

Lemma 10.1.3.2. *All holomorphic maps are harmonic.*

Lemma 10.1.3.3. *All conformal maps are harmonic.*

10.1.4 Linear fractional transformations

A linear fractional transformation is a map

$$\begin{aligned} \varphi : \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\longmapsto \frac{az+b}{cz+d} \end{aligned}$$

where $a, b, c, d \in \mathbb{C}$. These are important as they provide a class of workable examples of rational functions, which are pretty much the bread and butter of algebraic geometry. Moreover, these maps arrange themselves in a group and it then follows that it contains as a subgroup the biholomorphic automorphism group of lots of geometric objects of interest (see Lemmas 10.1.4.2, 10.1.4.3). However, these maps makes the most sense on the complex projective line, $\mathbb{C}P^1$, the quotient of $\mathbb{C}^2$ by all lines passing through origin, which is the usual Riemann sphere $\bar{\mathbb{C}}$.

Let us work out this connection in detail. We have the following maps:

$$\begin{tikzcd} \alpha : \mathbb{C}^2 \arrow[r, two heads] & \mathbb{C}P^1 \arrow[r, "\cong"] & \bar{\mathbb{C}} \\ (w,z) \arrow[r, mapsto] & {[w,z]} \arrow[r, mapsto] & \frac{w}{z} \end{tikzcd}$$

Notice that $L_f(\bar{\mathbb{C}})$, the collection of all linear fractional transforms on $\bar{\mathbb{C}}$ forms a group where the identity is given when $a = 0 = c$. The multiplication of two fractional transforms is again a fractional transform, as can be checked easily. Hence, it follows that $L_f(\bar{\mathbb{C}})$ is a subgroup of all biholomorphic maps of $\bar{\mathbb{C}}$, the Aut $\left(\bar{C}\right)$. Hence we have a hold on one type of global biholomorphic maps of the Riemann sphere(!)

We then have the following result.

Lemma 10.1.4.1. *Let $\bar{\mathbb{C}}$ denote the Riemann sphere. Then,*

$$L_f(\bar{\mathbb{C}}) \cong GL_2(\mathbb{C})/\mathbb{C}^\times I_2$$

Proof. There's a natural map

$$\begin{aligned} \kappa : GL_2(\mathbb{C}) &\longrightarrow L_f(\bar{\mathbb{C}}) \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} &\longmapsto \frac{az+b}{cz+d}. \end{aligned}$$

This is a group homomorphism, as can be checked easily. The kernel of this homomorphism consists of matrices

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

such that $\frac{az+b}{cz+d} = z$. Unravelling, we get $c = 0 = b$ and $a = d \neq 0$. □

This group is also known by *projective general linear group*, $PGL_2(\mathbb{C}) := L_f(\bar{\mathbb{C}})$. The group $L_f(\bar{\mathbb{C}})$ also has some special subgroups. For example, it consists of all biholomorphic maps of $D^\circ := \{z \in \mathbb{C} \mid |z| < 1\}$.

Lemma 10.1.4.2. *For the open unit ball D°, we have*

$$\mathrm{Aut}\,(D^\circ) \cong \left\{ \frac{t(z-a)}{1-\bar{a}z} \mid |t| = 1 \;\&\; a \in D^\circ \right\}.$$

Similarly, it also contains an isomorphic copy of all biholomorphic maps of upper half plane $\mathbb{H}$.

Lemma 10.1.4.3. *For the upper half plane $\mathbb{H} \subset \mathbb{C}$, we have*

$$\mathrm{Aut}\,(\mathbb{H}) \cong SL_2(\mathbb{R}) \subset GL_2(\mathbb{C}).$$

Properties

Let us now state some basic properties of fractional transforms.

Lemma 10.1.4.4. *If $\varphi : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ is a non-identity fractional transform, then either it has one or two fixed points, but not zero.*

Proof. A non-identity fractional transform $\varphi(z) = \frac{az+b}{cz+d}$ follows that either $b, d \neq 0$ or $a \neq c$. Suppose the former is not the case. Now if $\varphi(z) = z$, then it follows that $cz^2 + (d-a)z - b = 0$ where $b = d = 0$. Thus we obtain $z(cz - a) = 0$, which gives atleast one and atmost two solutions. Similarly, if $a = c$, then $b, d \neq 0$. It then follows that the above quadratic has either one or two solutions. □

Another property of fractional transforms is that they are uniquely determined by how they map on three points.

Lemma 10.1.4.5. *If z_1, z_2, z_3 and w_1, w_2, w_3 are two pair of distinct points in $\bar{\mathbb{C}}$, then there exists a unique fractional transform $\varphi \in L_f(\bar{\mathbb{C}})$ such that*

$$f(z_i) = w_i \; \forall i = 1, 2, 3.$$

Proof. Uniqueness follows from the fact that if $\varphi, \varpi : \bar{\mathbb{C}} \rightrightarrows \bar{\mathbb{C}}$ are two fractional transforms taking $z_i \mapsto w_i$, then the fractional transform $\varphi \circ \varpi^{-1}$ has 3 fixed points. It follows from Lemma 10.1.4.4 that $\varphi \circ \varpi^{-1} = \mathrm{id}$.

To show existence, take any arbitrary triple $v_1, v_2, v_3 \in \bar{\mathbb{C}}$. We will show that one can construct a fractional transform depending on v_i mapping as $z_i \mapsto v_i$. Denote then the map φ, $z_i \mapsto v_i$ and ϖ, $w_i \mapsto v_i$. Then $\varpi^{-1} \circ \varphi$ would be the required map. Since v_i can be arbitrary, therefore we choose it as per our convenience. It is perhaps easier to write it for $\infty, 0, 1$. □

One last basic property that may be observed for fractional transforms is that they are conformal.

Lemma 10.1.4.6. *All fractional transforms $\varphi : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ are conformal.*

Proof. Since fractional transforms are holomorphic, therefore by Corollary 10.1.2.2, we reduce to showing that $\varphi'(z) \neq 0$ for all $z \in \bar{\mathbb{C}}$. Indeed, we have

$$\phi'(z) = \frac{ad - bc}{(cz + d)^2},$$

where since $ad - bc \neq 0$ by definition, therefore $\phi'(z) \neq 0$. □

Example : The Cayley transform

We will discuss here the properties of the following fractional transform, known by Cayley's name:

$$\begin{aligned} \varphi : \bar{\mathbb{C}} &\longrightarrow \bar{\mathbb{C}} \\ z &\longmapsto \frac{z+i}{z-i}. \end{aligned}$$

10.2 La théorie des cartes holomorphes

The theory of holomorphic maps. We now begin another part of complex function theory which is at the heart of a lot of sources of interest in it. We first consider the line integrals.

10.2.1 Line integrals

Let $\gamma : [a, b] \to \mathbb{C}$ be a continuous function. Suppose $G \subseteq \mathbb{C}$ is an open subset containing γ and it's interior and let $f \in \mathcal{C}^{\mathrm{hol}}(G)$ be a holomorphic map $f : G \to \mathbb{C}$. Then, the *line integral* of f along γ is defined as

$$\int_\gamma f(z)dz := \int_a^b f(\gamma(t))\gamma'(t)dt$$

where definite integral of a complex valued function $g : [a, b] \to \mathbb{C}$ where $g = u + iv$ is given simply as the Riemann integral on each of the real and imaginary parts:

$$\int_a^b g(t)dt = \int_a^b u(t)dt + i\int_a^b v(t)dt.$$

A continuous map $\gamma : [a, b] \to \mathbb{C}$ is called piecewise C^1 if γ is C^1 at all but finitely many points and where it isn't differentiable, one sided derivative exists.

Few properties of line integrals are in order.

Theorem 10.2.1.1. *Let $\gamma : [a, b] \to \mathbb{C}$ be a curve in $\mathbb{C}$ and let $G \subseteq \mathbb{C}$ be an open subset containing γ. Let $f \in \mathcal{C}^{\text{hol}}(G)$ be a holomorphic map over G. Then,*

1. *(FTOC) If γ is piecewise C^1, then*
$$\int_a^b \gamma'(t)dt = \gamma(b) - \gamma(a).$$
2. *If $f \in \mathcal{C}^{\text{hol}}(G)$ where G contains γ, then*
$$\int_\gamma f'(z)dz = f(\gamma(b)) - f(\gamma(a)).$$
So if γ is a closed loop, then integral of f' along it is 0.
3. *If $f \in \mathcal{C}^{\text{hol}}(G)$ and $\tilde{\gamma}$ is a reparametrization of γ, then $\int_\gamma f(z)dz = \int_{\tilde{\gamma}} f(z)dz$.*
4. *(Estimate) If $f \in \mathcal{C}^{\text{hol}}(G)$ and $M = \sup_{t\in[a,b]} |f(\gamma(t))|$, then*
$$\left| \int_\gamma f(z)dz \right| \leq ML(\gamma)$$
where $L(\gamma) = \int_a^b |\gamma'(t)|\, dt$ is the arc-length.

Proof. Assuming 1 by FTOC on each piece, all results follows from basic analysis. ☐

10.2.2 Cauchy's theorem - I

We will now state the Cauchy's theorems on holomorphic maps and integrals. This will be a special case of the general version, which we shall do later, for we will find almost all of the traditional applications without needing that generality. We will begin with it's infantile version, which is quite simple to state now with line integrals in our pouch.

Theorem 10.2.2.1. *(Cauchy's theorem) Let $\gamma : [a, b] \to \mathbb{C}$ be a closed piecewise C^1 loop in $\mathbb{C}$ and let $G \subseteq \mathbb{C}$ be a convex open set containing γ and it's interior $\text{Int}(\gamma)$. If $f \in \mathcal{C}^{\text{hol}}(G)$, then*

$$\int_\gamma f(z)dz = 0.$$

Then there is the Cauchy integral formula.

Theorem 10.2.2.2. *(Cauchy's integral formula) Let C be a circle oriented in the counterclockwise manner and let $G \subseteq \mathbb{C}$ be an open set containing C and its interior* $\mathrm{Int}(C)$*. Then,*

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(w)}{w-z}dw$$

for all $z \in \mathrm{Int}(C)$.

Remark 10.2.2.3. Let $f \in \mathcal{C}^{\mathrm{hol}}(G)$ be a holomorphic map on open $G \subseteq \mathbb{C}$. The integral formula tells us that the value of f at $z \in G$ can be given in terms of line integral of f around a small enough circle C in the CCW orientation centered at z so that $C \subseteq G$. Hence the integral formula tells us that holomorphic maps are pretty much completely determined by taking their line integrals around circles.

We will provide some results which can be derived from them. In particular, using these, we would be able to show that a holomorphic map is analytic (Corollary 10.2.3.5).

Proof of Cauchy's theorem : Holomorphic maps have primitives

A *primitive* of a holomorphic map f is a holomorphic map g such that $g' = f$. We first state the following theorem without proof using which we will prove the Cauchy's theorem.

Theorem 10.2.2.4. *(Cauchy's triangle theorem) Let T be a triangle in $\mathbb{C}$ and $G \subseteq \mathbb{C}$ be an open set containing T and* $\mathrm{Int}(T)$*. If $f \in \mathcal{C}^{\mathrm{hol}}(G)$, then*

$$\int_T f(z)dz = 0.$$

Proof. [**??**] [Sarason]. □

Now, we will prove the following lemma using the above triangle theorem.

Lemma 10.2.2.5. *Let $G \subseteq \mathbb{C}$ be a convex open set and $f \in \mathcal{C}^{\mathrm{hol}}(G)$. Then there exists a map $g \in \mathcal{C}^{\mathrm{hol}}(G)$ such that $g' = f$.*

Proof. For a fixed $z_0 \in G$, define

$$\begin{aligned} g : G &\longrightarrow \mathbb{C} \\ z &\longmapsto \int_{[z_0,z]} f(z)dz \end{aligned}$$

where $[z_0, z]$ denotes the path formed by line joining z_0 and z in G. We claim that for all $z \in G$, $g'(z) = f(z)$. Indeed, pick any $z_1 \in G$ to form triangle $T = (z_0, z_1, z)$ inside G (G is convex). Then, by Theorem 10.2.2.4, we get the following

$$\begin{aligned} 0 &= \int_T f(w)dw \\ &= \int_{[z_0,z_1]} f(w)dw + \int_{[z_1,z]} f(w)dw + \int_{[z,z_0]} f(w)dw \\ g(z) - g(z_1) &= \int_{[z_1,z]} f(w)dw. \end{aligned}$$

We wish to estimate

$$\begin{aligned}\left|\frac{g(z)-g(z_1)}{z-z_1}-f(z_1)\right| &= \left|\frac{1}{z-z_1}\int_{[z_1,z]} f(w)dw - f(z_1)\right| \\ &= \left|\frac{1}{z-z_1}\int_{[z_1,z]} (f(w)-f(z_1))dw\right| \\ &\leq \frac{1}{|z-z_1|}\int_{[z_1,z]} |f(w)-f(z_1)|\, dw.\end{aligned}$$

Since f is continuous, therefore for any $\epsilon > 0$, there is a $\delta > 0$ such that $|w - z_1| < \delta$ implies $|f(w) - f(z_1)| < \epsilon$. Hence, for $|w - z_1| < \delta$, we get

$$\begin{aligned}&\leq \frac{1}{|z-z_1|}\int_{[z_1,z]} \epsilon dw \\ &= \epsilon.\end{aligned}$$

Hence as $z \to z_1$, the above difference $\to 0$. □

Proof of Theorem 10.2.2.1. Since $f \in \mathcal{C}^{\text{hol}}(G)$, therefore by Lemma 10.2.2.5, there exists $g \in \mathcal{C}^{\text{hol}}(G)$ such that $g' = f$. Hence the result follows by Theorem 10.2.1.1, 2. □

Proof of Cauchy's integral formula : Cauchy integrals

We would like to present the proof of Cauchy integral formula as it portrays how to use the fact that integral of holomorphic maps around closed loops are zero (Theorem 10.2.2.1).

Proof of Theorem 10.2.2.2. Pick any $z_0 \in \text{Int}(C)$. We shall show the result for this chosen z_0. We shall use the Cauchy's theorem 10.2.2.1 in an essential manner. Indeed, consider the following figure on the complex plane inside G: Integrating the holomorphic map $\frac{f(w)}{w-z_0}$ over the each

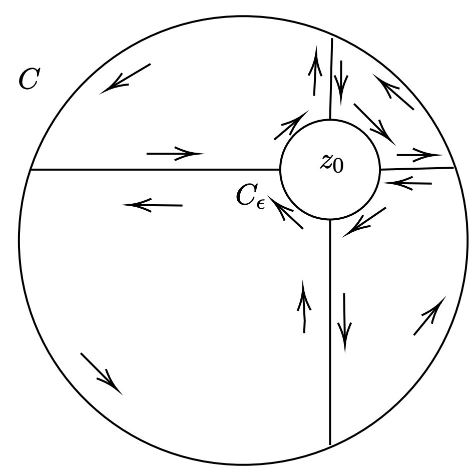


Figure 10.1: Contour over which to integrate $\frac{f(w)}{w-z_0}$.

of the four regions will give zero by Theorem 10.2.2.1. However, summing them up, one can see that we get the difference $\int_C \frac{f(w)}{w-z_0}dw - \int_{C_\epsilon} \frac{f(w)}{w-z_0}dw$, which should thus be zero, yielding us $\int_C \frac{f(w)}{w-z_0}dw = \int_{C_\epsilon} \frac{f(w)}{w-z_0}dw$. Note this is true for all $\epsilon < d(z_0, C)$.

Now recall that $\int_C \frac{1}{z}dz = 2\pi i$. Hence, we get the following estimate for any chosen $\epsilon < d(z_0, C)$

$$\begin{aligned} \left|\frac{1}{2\pi i}\int_C \frac{f(w)}{w-z_0}dw - f(z_0)\right| &= \left|\frac{1}{2\pi i}\int_{C_\epsilon} \frac{f(w)}{w-z_0}dw - f(z_0)\right| \\ &= \left|\frac{1}{2\pi i}\int_{C_\epsilon} \frac{f(w)-f(z_0)}{w-z_0}dw\right| \end{aligned}$$

Now, by Theorem 10.2.1.1, 4, let $M_\epsilon = \sup_{w\in C_\epsilon}\left|\frac{f(w)-f(z_0)}{w-z_0}\right|$ to obtain the following inequality

$$\begin{aligned} &\leq \frac{M_\epsilon}{2\pi}L(C_\epsilon) \\ &= \frac{M_\epsilon}{2\pi}2\pi\epsilon \\ &= \epsilon M_\epsilon. \end{aligned}$$

Since f is holomorphic, therefore $\varinjlim_{\epsilon\to 0} M_\epsilon = |f'(z_0)|$. Hence, $\varinjlim_{\epsilon\to 0} \epsilon M_\epsilon = 0$, which gives the desired result. □

10.2.3 Theory of holomorphic maps

We now present applications of the two highly useful results of Cauchy (Theorems 10.2.2.1, 10.2.2.2). The results covered here are as follows:

- Mean value property.
- Power series representation of Cauchy integrals.
- Morera's theorem.
- Derivatives.
- Liouville's theorem.
- Identity theorem.
- Maximum modulus theorem.
- Schwarz's lemma.
- Classification of bijective holomorphic maps of open unit ball.
- Open mapping theorem.
- Fundamental theorem of algebra.
- Inverse function theorem.
- Local m^{th} power property.
- Harmonic conjugates.

These results lie at the heart of complex analysis.

Let us begin by understanding the behavior of a holomorphic map around a circle centered at a point.

Mean value property of holomorphic maps

Proposition 10.2.3.1. *Let $G \subseteq \mathbb{C}$ be an open set and $f \in \mathcal{C}^{\text{hol}}(G)$. Then, for all $z_0 \in G$ and C_r a circle of radius r centered at z_0 contained inside G together with its interior* $\text{Int}(C)$*, we have*

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{it})dt.$$

Proof. Using the integral formula (Theorem 10.2.2.2) and using $\gamma(t) = z_0 + re^{it}$ as a parameterization of C_r, the result follows. □

Power series representation of Cauchy integrals

We will in this section show that functions defined by Cauchy integrals are analytic. Since holomorphic maps are given by Cauchy integrals, thus we would be able to show that holomorphic maps are analytic.

Definition 10.2.3.2. (**Maps given by Cauchy integral**) Let $\gamma : [a, b] \to \mathbb{C}$ be a piecewise C^1 curve in $\mathbb{C}$ and $f \in \mathcal{C}^{\text{hol}}(G)$ be a holomorphic map on an open subset $G \subseteq \mathbb{C}$ where G contains $\text{Im}\,(\gamma)$. Define the following map

$$\begin{aligned} \tilde{f} : \mathbb{C} \setminus \text{Im}\,(\gamma) &\longrightarrow \mathbb{C} \\ z &\longmapsto \int_\gamma \frac{f(w)}{w - z}dw. \end{aligned}$$

Then $\tilde{f}$ is called the Cauchy integral associated to $f \in \mathcal{C}^{\text{hol}}(G)$ and $\gamma : [a, b] \to G$.

We first show that holomorphic maps are given by Cauchy integrals.

Lemma 10.2.3.3. *Let $f \in \mathcal{C}^{\text{hol}}(G)$ be a holomorphic map on an open set $G \subseteq \mathbb{C}$. Then f is locally given by a Cauchy integral.*

Proof. Indeed, by Theorem 10.2.2.2, we see that for all $z \in G$, choosing a small circle C_z around z and such that C_z and $\text{Int}(C_z)$ are inside G, we can write

$$f(z) = \frac{1}{2\pi i}\int_{C_z} \frac{f(w)}{w - z}dw.$$

Hence locally f looks like a Cauchy integral. □

We now show that Cauchy integrals are analytic, making holomorphic maps analytic by above lemma.

Proposition 10.2.3.4. *Maps defined by Cauchy integrals are analytic.*

Proof. Let $f \in \mathcal{C}^{\text{hol}}(G)$ where G is open and let $\gamma : [a, b] \to G$ be a piecewise C^1 curve. We wish to show that $\tilde{f}$ defined on $\mathbb{C} \setminus \text{Im}\,(\gamma)$ is given locally by power series. Indeed, pick any $z \in \mathbb{C} \setminus \text{Im}\,(\gamma)$. Since $\text{Im}\,(\gamma)$ is closed, therefore there exists a ball of radius r, B_r, such that $B_r \subseteq \mathbb{C} \setminus \text{Im}\,(\gamma)$. In order to expand $\tilde{f}(z) = \int_\gamma \frac{f(w)}{w-z}dw$ as a power series, we first focus on $\frac{1}{w-z}$, where $w \in \text{Im}\,(\gamma)$ and

z is as above. Indeed, for any $z_0 \in B_r$, we have $|z - z_0| < r$ and $|w - z_0| > r$, thus yielding that $\left|\frac{z-z_0}{w-z_0}\right| < 1$ and hence we can write

$$\begin{aligned}\frac{1}{w-z} = \frac{1}{(w-z_0)-(z-z_0)} &= \frac{1}{(w-z_0)(1-\frac{z-z_0}{w-z_0})} \\ &= \frac{1}{w-z_0}\sum_{n=0}^{\infty}\left(\frac{z-z_0}{w-z_0}\right)^n \\ &= \sum_{n=0}^{\infty}\frac{(z-z_0)^n}{(w-z_0)^{n+1}}.\end{aligned}$$

Moreover the convergence is uniform as we are within the radius of convergence. Now, $f(w) \leq M$ for all $w \in \operatorname{Im}(\gamma)$ as $\operatorname{Im}(\gamma)$ is compact and f is continuous over it. Hence we get that that following holds for all $w \in \operatorname{Im}(\gamma)$

$$\frac{f(w)}{w-z} = \sum_{n=0}^{\infty}\frac{f(w)(z-z_0)^n}{(w-z_0)^{n+1}}.$$

Taking integral both sides, it thus follows from uniform convergence of above series that

$$\begin{aligned}\tilde{f}(z) = \int_\gamma \frac{f(w)}{w-z}dw &= \int_\gamma \sum_{n=0}^{\infty}\frac{f(w)(z-z_0)^n}{(w-z_0)^{n+1}}dw \\ &= \sum_{n=0}^{\infty}\int_\gamma \frac{f(w)(z-z_0)^n}{(w-z_0)^{n+1}}dw \\ &= \sum_{n=0}^{\infty}\left(\int_\gamma \frac{f(w)}{(w-z_0)^{n+1}}dw\right)(z-z_0)^n.\end{aligned}$$

Hence locally $\tilde{f}$ looks like a power series, i.e. it is analytic. □

Corollary 10.2.3.5. *Holomorphic maps are analytic.*

Proof. By Lemma 10.2.3.3, holomorphic maps are given by Cauchy integrals. By Proposition 10.2.3.4, maps given by Cauchy integrals are analytic. □

Morera's theorem : Converse of Cauchy's triangle theorem

Proposition 10.2.3.6. *If $f : G \to \mathbb{C}$ is a continuous map on an open set $G \subseteq \mathbb{C}$ such that for all triangles $T \subseteq G$ where $\operatorname{Int}(T) \subseteq G$ as well we get*

$$\int_T f(z)dz = 0,$$

then f is holomorphic.

Proof. □

Derivatives of a holomorphic map

Proposition 10.2.3.7. *Let $f \in \mathcal{C}^{\mathrm{hol}}(G)$ be a holomorphic map on an open set $G \subseteq \mathbb{C}$. Then, f is differentiable to all orders and*

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_{C_r} \frac{f(w)}{(w-z)^{n+1}}dw$$

where C_r is a circle in CCW orientation of radius r such that $C_r \subseteq G$ and $\mathrm{Int}(C_r) \subseteq G$. Moreover, for all $z \in \mathrm{Int}(C_r)$ with z_0 as center, we have that

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)(z-z_0)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi i}\left(\int_{C_r} \frac{f(w)}{(w-z)^{n+1}}dw\right)(z-z_0)^n \end{aligned}$$

Liouville's theorem

A holomorphic map f on the entire complex plane, that is $f \in \mathcal{C}^{\mathrm{hol}}(\mathbb{C})$, is said to be *entire*.

Proposition 10.2.3.8. *Any entire bounded function $f : \mathbb{C} \to \mathbb{C}$ is constant.*

Zeroes of holomorphic maps

Proposition 10.2.3.9. *Let $G \subseteq \mathbb{C}$ be an open connected subset of $\mathbb{C}$. If $f \in \mathcal{C}^{\mathrm{hol}}(G)$ is a holomorphic map on G, then the zero set $V(f) = \{z \in G \mid f(z) = 0\}$ has no limit point in G i.e. either $V(f) = G$ or $V(f)$ is discrete.*

Identity theorem

Proposition 10.2.3.10. *Let $f, g \in \mathcal{C}^{\mathrm{hol}}(G)$ be two holomorphic maps defined on an open connected set $G \subseteq \mathbb{C}$. Then $f = g$ on G if and only if there exists a set $A \subseteq G$ which has a limit point in G such that $f|_A = g|_A$.*

Corollary 10.2.3.11. *Let f, g be two holomorphic maps on open connected subset $G \subseteq \mathbb{C}$ such that there exists an open set $U \subsetneq G$ contained inside of G such that $\partial U \neq \emptyset$ and $\bar{U} \subseteq G$ and $f|_U = g|_U$. Then $f = g$ on G.*

Proof. Indeed, since any element in ∂U is a limit point of U in G and $f|_U = g|_U$, then the result follows by Proposition 10.2.3.10. □

Corollary 10.2.3.12. *Let f, g be two holomorphic maps on open connected subset $G \subseteq \mathbb{C}$ such that there exists a closed ball $B \subset G$ on which $f|_B = g|_B$, then $f = g$ on G.*

Proof. A closed ball has non-empty interior. The result follows by Corollary 10.2.3.11. □

Maximum modulus principle

Proposition 10.2.3.13. *Let $G \subseteq \mathbb{C}$ be an open connected set and $f \in \mathcal{C}^{\mathrm{hol}}(G)$ be a holomorphic map on G. Then $|f|$ doesn't achieves local maxima in G.*

Schwarz's lemma

Lemma 10.2.3.14. *Let $D = \{z \in \mathbb{C} \mid |z| < 1\}$ be the open unit disc. If $f \in \mathcal{C}^{\text{hol}}(D)$ is a holomorphic map $f : D \to D$ such that $f(0) = 0$, then*

1. *$|f(z)| \leq |z|$ for all $z \in D$.*
2. *$|f'(0)| \leq 1$.*
3. *If f is not of the form λz for $\lambda \in S^1$, then the inequality in 1. & 2. is strict at all points $z \in D \setminus \{0\}$. In particular, if there exists $z_0 \in D \setminus \{0\}$ such that $|f(z_0)| = |z_0|$, then $f(z) = \lambda z$ for $|\lambda| = 1$ and $\lambda = f'(0)$.*

Proof. Consider the map defined by

$$\begin{aligned} g : D &\longrightarrow \mathbb{C} \\ z &\longmapsto \begin{cases} \frac{f(z)}{z} & \text{if } z \in D \setminus \{0\} \\ f'(0) & \text{if } z = 0. \end{cases} \end{aligned}$$

Clearly g is holomorphic. Now, for any $r \in (0,1)$, for $C_r \subset D$, by maximum modulus, Proposition 10.2.3.13, we have

$$|g(z)| < \frac{1}{r}$$

for all $z \in \text{Int}(C_r)$. Taking limit as $r \to 1$, we obtain $|g(z)| \leq 1$ for all $z \in D$. Now, if $\exists w \in D$ such that $|f(w)| = |w|$, then $|g(w)| = 1$. Since $|g(z)| < 1$ for all $z \in D$ as shown above, therefore by another use of maximum modulus, Proposition 10.2.3.13, it follows that $g(z) = \lambda$ is a constant where $|\lambda| = 1$. Thus $f(z) = \lambda z$. □

Corollary 10.2.3.15. *(Pick's lemma) Let $f : D \to D$ be a holomorphic map where $D = \{z \in \mathbb{C} \mid |z| < 1\}$. Then, for any two points $z, w \in D$*

$$\left| \frac{f(z) - f(w)}{1 - f(z)\overline{f(w)}} \right| \leq \left| \frac{z - w}{1 - z\bar{w}} \right|$$

except if f is a linear fractional transform mapping disc onto itself.

Proof. Define for each $w \in D$ the following fractional transform

$$\begin{aligned} g_w : D &\longrightarrow D \\ z &\longmapsto \frac{z - w}{1 - z\bar{w}}. \end{aligned}$$

Then apply Schwarz's lemma (Lemma 10.2.3.14) on $g_{f(w)} \circ f \circ g_w^{-1} : D \to D$ as fractional transforms are biholomorphic. □

Classification of bijective holomorphic maps of open unit ball

We shall classify all bijective holomorphic maps $f : D \to D$ for $D := \{z \in \mathbb{C} \mid |z| < 1\}$ and see that in the process that they are biholomorphic as well. For this, we first define the following important map which we encountered in Pick's lemma (Corollary 10.2.3.15). Define the following map for each $\alpha \in D$:

$$\begin{aligned} \varphi_\alpha : \bar{D} &\longrightarrow \bar{D} \\ z &\longmapsto \frac{z-\alpha}{1-\bar{\alpha}z}. \end{aligned}$$

This is indeed a holomorphic map over $\bar{D}$. We now see that this is biholomorphic.

Theorem 10.2.3.16. *For any $\alpha \in D$, the map $\varphi_\alpha : \bar{D} \to \bar{D}$ is such that*

1. *φ_α takes D to D,*
2. *φ_α takes ∂D to ∂D,*
3. *φ_α is injective,*
4. *φ_α is surjective,*
5. *φ_α has a holomorphic inverse given by $\varphi_{-\alpha}$.*

Proof. Fix an $\alpha \in D$. We first show 2. For any $z \in \partial D$, we can write $z = e^{it}$ for $t \in \mathbb{R}$. Thus we have

$$\begin{aligned} \left|\varphi_\alpha(e^{it})\right| &= \left|\frac{e^{it}-\alpha}{1-\bar{\alpha}e^{it}}\right| \\ &= \left|\frac{e^{it}-\alpha}{1-\bar{\bar{\alpha}}\bar{e}^{it}}\right| \\ &= \left|\frac{e^{it}-\alpha}{1-\alpha e^{-it}}\right| \\ &= \left|\frac{e^{it}-\alpha}{e^{it}-\alpha}\right| \\ &= 1. \end{aligned}$$

Thus, $\varphi_\alpha(e^{it}) \in \partial D$. This shows 2. Now we show 1. For this, by maximum modulus (Proposition 10.2.3.13), we have that $|\varphi_\alpha|$ achieves maxima on ∂D, and by 1., that maxima is 1, hence at every point of ∂D does $|\varphi_\alpha|$ achieves maxima. Hence $\varphi_\alpha(D) \subseteq D$. This shows 1. Next, it is a matter of simple calculation to see that $\varphi_\alpha \circ \varphi_{-\alpha} = \mathrm{id}_{\bar{D}}$ and thus by symmetry $\mathrm{id}_{\bar{D}} = \varphi_{-\alpha} \circ \varphi_\alpha$. Hence, φ_α is a biholomorphic map taking D onto D and ∂D onto ∂D. □

We would now like to see that all biholomorphic maps of open unit ball are given by some unit modulus scalar multiples of φ_α. However, we need an idea to do so, which is provided by the following result.

Proposition 10.2.3.17. *(Extremality) For fixed $\alpha, \beta \in D$, denote $\mathcal{C}_{\alpha,\beta}$ to be the class of holomorphic maps into the unit disc $f : D \to D$ such that $f(\alpha) = \beta$. Then,*

1. *we have*

$$\sup_{f \in \mathcal{C}_{\alpha,\beta}} \left|f'(\alpha)\right| = \frac{1-|\beta|^2}{1-|\alpha|^2}.$$

2. *The map $f \in \mathcal{C}_{\alpha,\beta}$ achieving the suprema is given by the following rational map*

$$f = \varphi_{-\beta} \circ \lambda \circ \varphi_\alpha$$

where $\lambda \in \partial D$ is a scalar.

Proof. 1. We need only show that for each $f \in \mathcal{C}_{\alpha,\beta}$, we get

$$|f'(\alpha)| \leq \frac{1 - |\beta|^2}{1 - |\alpha|^2}.$$

Indeed, this simply follows from a similar idea as used in the proof Pick's lemma (Corollary 10.2.3.15) above; consider the map $g = \varphi_\beta \circ f \circ \varphi_{-\alpha}$ and use Schwarz's lemma (Lemma 10.2.3.14) on it to get the bound $|g'(0)| \leq 1$. Now use chain rule while keeping in mind that $\varphi'(0) = 1 - |\alpha|^2$ and $\varphi'_\alpha(\alpha) = \frac{1}{1-|\alpha|^2}$.

2. From proof of 1, it follows that the equality is achieved if and only if $|g'(0)| = 1$. By Schwarz's lemma (Lemma 10.2.3.14) this happens only if $g(z) = \lambda z$ for $\lambda \in \partial D$. Rest follows by composing with inverses of φ_β and $\varphi_{-\alpha}$ which we know from Theorem 10.2.3.16, 5. □

We now come to the real deal. The following shows that all bijective holomorphic maps $D \to D$ are biholomorphic and are given by unit modulus scalar multiples of φ_α for some $\alpha \in D$. However we shall need a topic which we will cover in the next few sections, namely the inverse function theorem for one complex variable (see Section **??**, Theorem **??**). Moreover we shall also need another result which we do only in a further section called Rouché's theorem (Section **??**, Theorem **??**).

Theorem 10.2.3.18. *(Bijective holomorphic maps $D \to D$) Let $f : D \to D$ be a bijective holomorphic map. Denote $\alpha \in D$ to be the unique element such that $f(\alpha) = 0$. Then, there exists $\lambda \in \partial D$ such that*

$$f = \lambda \varphi_\alpha.$$

Proof. Consider the set-theoretic inverse of f, denoted $g : D \to D$. By Rouché's theorem (Theorem **??**) and by inverse function theorem (Theorem **??**), we obtain that $g \in \mathcal{C}^{\mathrm{hol}}(D)$. Now by chain rule we obtain $g'(f(\alpha))f'(\alpha) = 1$, that is, $g'(0) = 1/f'(\alpha)$. Now by Proposition 10.2.3.17, we obtain the following inequality for f and g where $f(\alpha) = 0$ and $g(0) = \alpha$:

$$|f(z)| \leq \frac{1}{1 - |\alpha|^2}$$
$$|g(z)| \leq 1 - |\alpha|^2.$$

In particular, we obtain that $1 - |\alpha|^2 \geq g'(0) = 1/f'(\alpha) \geq 1 - |\alpha|^2$, thus $g'(0) = 1 - |\alpha|^2$. Similarly, $|f'(\alpha)| = \frac{1}{1-|\alpha|^2}$. Hence f achieves the suprema of Proposition 10.2.3.17, 1. By Proposition 10.2.3.17, the result follows. □

Corollary 10.2.3.19. *There is a bijection*

$$\left\{\text{Bijective holomorphic maps } f : D \to D\right\}$$

$$\cong$$

$$\left\{\text{Rational functions of the form } \lambda\tfrac{z-\alpha}{1-\bar{\alpha}z},\ \alpha \in D,\ \lambda \in \partial D\right\}.$$

□

Using this and Schwarz's lemma, we can show that a holomorphic map $f : D \to D$ can have atmost one fixed point.

Corollary 10.2.3.20. *Let $f : D \to D$ be a holomorphic map. Then f has atmost one fixed point.*

Proof. The idea is quite simple and we have used it already in the proof of Pick's lemma (Corollary 10.2.3.15). Indeed, we will construct $\varphi_\alpha : D \to D$ in such a manner that Schwarz's lemma can be applied to $\varphi \circ f \circ \varphi^{-1}$ and will use the results about the function φ_α (Theorem 10.2.3.16).

Suppose $z_1 \neq z_2 \in D$ are two fixed points of f. Consider the map $\varphi_{z_1}(z) := \frac{z-z_1}{1-\bar{z}_1 z}$. This is a biholomorphic mapping $\varphi_{-z_1} : D \to D$. Consider

$$h = \varphi_{z_1} \circ f \circ \varphi_{-z_1}.$$

Then $h : D \to D$ is a holomorphic map and $h(0) = 0$. Applying Schwarz's lemma (Lemma 10.2.3.14), we obtain that $|h(z)| \leq |z|$. But notice that $h(\varphi_{z_1}(z_2)) = \varphi_{z_1}(z_2)$. Thus $\varphi_{z_1}(z_2)$ is a fixed point of h. Moreover, $\varphi_{z_1}(z_2) \neq 0$ as other wise $z_2 = z_1$, a contradiction. Thus, there exists $w \in D$ such that $|h(w)| = |w|$ (in particular, for $w = \varphi_{z_1}(z_2)$). Hence by contrapositive of Lemma 10.2.3.14, 3, we obtain that $h(z) = \lambda z$. Since $h(w) = w = \lambda w$, we obtain that $\lambda = 1$. Hence $h = \mathrm{id}$, thus $f = \mathrm{id}$. □

Open mapping theorem

This theorem is quite an important result in the theory of holomorphic maps. It says a very simple thing, *all holomorphic maps on open connected sets are open maps*(!)

Theorem 10.2.3.21. *Let $G \subseteq \mathbb{C}$ be an open connected subset and let $f \in \mathcal{C}^{\text{hol}}(G)$ be a non-constant holomorphic map $f : G \to \mathbb{C}$. Then f is an open map.*

Fundamental theorem of algebra

Proposition 10.2.3.22. *Every non-constant polynomial $f(x) \in \mathbb{C}[x]$ can be factored into linear factors.*

Proof. Suppose $f(x) \in \mathbb{C}[x]$ is a non-constant polynomial given by

$$f(x) = a_n x^n + \cdots + a_1 x + a_0.$$

Suppose to the contrary that f has no zeros in $\mathbb{C}$. Then $g(x) = \frac{1}{f(x)} : \mathbb{C} \to \mathbb{C}$ is an entire map. We wish to use Liouville's theorem (Proposition 10.2.3.8) on $g(x)$ in order to obtain a contradiction. Indeed, to get an upper bound for $|g(x)|$, we need a lower bound for $|f(x)|$. To this end we have

$$\begin{aligned} |f(x)| &\geq |a_n x^n + \cdots + a_1 x + a_0| \\ &\geq |a_n x^n| \left| \left(1 + \frac{a_{n-1}}{a_n x} + \cdots + \frac{a_1}{a_n x^{n-1}} + \frac{a_0}{a_n x^n}\right) \right| \\ &\geq |a_n x^n| \left(1 - \left|\frac{a_{n-1}}{a_n x}\right| - \cdots - \left|\frac{a_1}{a_n x^{n-1}}\right| - \left|\frac{a_0}{a_n x^n}\right|\right) \end{aligned}$$

where the last inequality comes from triangle inequality. Now write $h(x) = 1 - \left|\frac{a_{n-1}}{a_n x}\right| - \cdots - \left|\frac{a_1}{a_n x^{n-1}}\right| - \left|\frac{a_0}{a_n x^n}\right|$. In order to get a further lower bound for $|f(x)|$, we need to get an upper bound for $h(x)$. Since $h(x) \to 0$ as $x \to \infty$, therefore for some $R > 0$, we shall have $h(x) \leq \frac{1}{3}$ for $|x| > R$. Thus, we get

$$|f(x)| \geq |a_n R^n| \frac{2}{3}$$

for $|x| > R$. Now on $|x| \leq R$, by continuity of $|f|$ on a compact domain, we get that it achieves a minima, and hence $|f|$ is a lower bounded map and hence $g(x)$ is an upper bounded map. $\square$

Inverse function theorem

Remember that for a differentiable map $f : \mathbb{R}^n \to \mathbb{R}^n$, if $x_0 \in \mathbb{R}^n$ is a point such that Df_{x_0} is invertible, the inverse function theorem tells us that f is a diffeomorphism in some neighborhood around x_0. A similar statement is true for holomorphic maps $f : G \subseteq \mathbb{C} \to \mathbb{C}$.

Theorem 10.2.3.23. *(Inverse function theorem) Let $G \subseteq \mathbb{C}$ be an open connected set and $\varphi \in \mathcal{C}^{\text{hol}}(G)$ be a holomorphic map on G. If for $z_0 \in G$ we have that $f'(z_0) \neq 0$, then there exists a neighborhood $z_0 \in V \subseteq G$ such that*

1. *$\varphi|_V : V \to \varphi(V)$ is bijective,*
2. *$\varphi(V) \subseteq G$ is open,*
3. *the map $\psi : \varphi(V) \to V$ given by $\varphi(z) \mapsto z$ is in $\mathcal{C}^{\text{hol}}(\varphi(V))$,*
4. *the map $\varphi|_V : V \to \varphi(V)$ is biholomorphic.*

We will now prove it. Let us begin with the following simple lemma.

Lemma 10.2.3.24. *Let $G \subseteq \mathbb{C}$ be an open connected set. If $f : G \to \mathbb{C}$ is a holomorphic map, then the map defined by*

$$\begin{aligned} g : G \times G &\longrightarrow \mathbb{C} \\ (z, w) &\longmapsto \begin{cases} \frac{f(z)-f(w)}{z-w} & \text{if } z \neq w, \\ f'(z) & \text{if } z = w \end{cases} \end{aligned}$$

is continuous.

Proof. Clearly g is continuous for all (z,w) with $z \neq w$. Pick any $a \in G$. We will show that g is continuous at (a,a). For that, we wish to estimate $|g(z,w) - g(a,a)|$. For this, note that we can write $g(z,w)$ as follows where γ is the straight path $\gamma(t) = (1-t)z + tw$:

$$\begin{aligned} g(z,w) &= \frac{f(z) - f(w)}{z - w} \\ &= \frac{f(\gamma(0)) - f(\gamma(1))}{\gamma(0) - \gamma(1)} \\ &= \frac{1}{w - z}\int_\gamma f'(z)dz \\ &= \int_0^1 f'(\gamma(t))dt \end{aligned}$$

where the third equality follows from Theorem 10.2.1.1, 2. Thus we can write

$$\begin{aligned} |g(z,w) - g(a,a)| &= \left|\int_0^1 f'(\gamma(t))dt - f'(a)\right| \\ &= \left|\int_0^1 f'(\gamma(t)) - f'(a)dt\right| \\ &\leq \int_0^1 \left|f'(\gamma(t)) - f'(a)\right| dt. \end{aligned}$$

Now by continuity of f', the estimate follows. □

We can now prove the inverse function theorem.

Proof of Theorem 10.2.3.23. 1. The surjectivity is clear. For injectivity, we will show that for two $z_1 \neq z_2 \in V$, $|\varphi(z_1) - \varphi(z_2)| \geq M$ for some $M > 0$ using the lemma just proved. Indeed, using Lemma 10.2.3.24 and triangle inequality, we obtain for $\epsilon = \frac{1}{2}|\varphi'(z_0)|$ an open set $\tilde{V}$ containing (z_0, z_0) such that for all $(z_1, z_2) \in \tilde{V}$ with $z_1 \neq z_2$ we get the following

$$\left|\left|\frac{\varphi(z_1) - \varphi(z_2)}{z_1 - z_2}\right| - \left|\varphi'(z_0)\right|\right| \leq \left|\frac{\varphi(z_1) - \varphi(z_2)}{z_1 - z_2} - \varphi'(z_0)\right| < \epsilon = \frac{1}{2}\left|\varphi'(z_0)\right|.$$

Using this, we obtain that

$$|\varphi(z_1) - \varphi(z_2)| \geq \frac{1}{2}\left|\varphi'(z_0)\right| |z_1 - z_2|.$$

Thus for $z_1, z_2 \in V \subseteq G$ where V is obtained by projecting a small open set inside $\tilde{V}$ back to G, we see that on that V φ is injective.

2. This is just open mapping theorem, Theorem 10.2.3.21.

3. Shrink V enough so that $\varphi'(z) \neq 0$ for all $z \in V$. Then everything is straightforward using

$$|\varphi(z_1) - \varphi(z_2)| \geq \frac{1}{2}\left|\varphi'(z_0)\right| |z_1 - z_2|$$

which we obtained in 1. □

Local m^{th} power property

Any holomorphic map around a point can be represented by the m^{th} power of some other special holomorphic map. Indeed, this is what the following theorem tells us.

Theorem 10.2.3.25. *Let $G \subseteq \mathbb{C}$ be an open-connected subset of $\mathbb{C}$ and let $f \in \mathcal{C}^{\text{hol}}(G)$ be a holomorphic map on G. Let $z_0 \in G$ and denote $w_0 = f(z_0)$. Let m be the order of zero that $f - w_0$ has at z_0. Then, there exists an open set $z_0 \in V \subseteq G$ and a holomorphic map*

$$\varphi : V \to \mathbb{C}$$

in $\mathcal{C}^{\text{hol}}(G)$ such that
1. *$f(z) = w_0 + (\varphi(z))^m$ for all $z \in V$,*
2. *φ' is nowhere vanishing in V, i.e. has no zero in V,*
3. *there exists $r > 0$ such that φ is biholomorphic onto $D_r(0)$, the open disc of radius r around 0. Thus, $\varphi : V \to D_r(0)$ is bijective.*

Proof. The main point of the proof is to try to represent the desired φ as $\exp \frac{??}{m}$. We just need to fill ?? correctly. Since $f - w_0$ has zero of order m at z_0, therefore there exists $g \in \mathcal{C}^{\text{hol}}(G)$ such that

$$f(z) - w_0 = (z - z_0)^m g(z).$$

Now, by appropriately shrinking G away from zeros of g, we may assume $g \neq 0 \forall z \in G \setminus \{z_0\}$[1]. Thus we have that $\frac{g'}{g}$ is holomorphic on G (this is our V). By Lemma 10.2.2.5, we get $h \in \mathcal{C}^{\text{hol}}(G)$ such that $h' = \frac{g'}{g}$. We now claim that $g = \exp h$. Indeed, it is a simple matter to see that the derivative of $g \exp -h$ is zero. Thus, by using surjectivity of exp, we can absorb the additive constant into h to obtain the above claim. One then sees that

$$\varphi(z) = (z - z_0) \exp \frac{h(z)}{m}$$

does the job for 1. The rest is straightforward. □

Harmonic conjugates

We will now show that any real valued harmonic map $u : G \subseteq \mathbb{R}^2 \to \mathbb{R}$ defines a unique (upto some constant) holomorphic map $g : G \to \mathbb{C}$ whose real part is u.

Theorem 10.2.3.26. *Let $G \subseteq \mathbb{C}$ be a convex open connected set. Let $u : G \to \mathbb{R}$ be a harmonic real valued function. Then, there exists holomorphic map $g : G \to \mathbb{C}$ unique upto an additive constant such that*

$$\Re g = u.$$

Proof. The main idea is to construct a holomorphic map f on G via the data of partial derivatives of u, and then use the Lemma 10.2.2.5, to get a primitive g, which will do the job. Indeed, we can make f via the following observation: u is harmonic real valued function if and only if $\frac{\partial^2}{\partial \bar{z} \partial z} = 0$. Using this, just define $f = u_x - i u_y$ and to show that f is holomorphic, observe that $\frac{\partial f}{\partial \bar{z}} = 0$. □

[1]We are implicitly using the isolated zeros theorem (Theorem **??**) which we shall do later.

In combination with Lemma 10.1.3.2, we get that

Corollary 10.2.3.27. *Let $G \subseteq \mathbb{C}$ be open connected. Then,*

$$\{g : G \to \mathbb{C} \text{ is holomorphic}\} \cong \{u : G \to \mathbb{R} \text{ is harmonic}\}$$

where we identify functions upto additive constant. □

10.3 Singularities

Consider the map $f(z) = 1/z$ on $\mathbb{C}^\times$. It is holomorphic. However, at $z = 0$, it is not holomorphic. Such points are called singularities of f, as we shall define more clearly later. Our goal is to study this phenomenon more carefully in this section. For this, we first need to develop a tool for local analysis of such "bad" points (some may also call it "the" points).

10.3.1 Laurent series

Definition 10.3.1.1. (**Laurent series**) A Laurent series centered at $z_0 \in \mathbb{C}$, denoted by $\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$, is a series of functions defined on some annulus $A_{z_0}(R_1, R_2) := \{z \in \mathbb{C} \mid R_1 < |z - z_0| < R_2\}$ centered at z_0 for $0 \leq R_1 < R_2$ such that the series converges at all points $z \in A_{z_0}(R_1, R_2)$. That is, the sequence of holomorphic maps $\{\sum_{n=-N}^{n=N} a_n(z-z_0)^n\}_N$ on $A_{z_0}(R_1, R_2)$ converges uniformly and absolutely to a holomorphic function $f : A_{z_0}(R_1, R_2) \to \mathbb{C}$ (by Weierstrass theorem).

For a Laurent series, we can find the coefficents in terms of Cauchy integral of the function it represents.

Lemma 10.3.1.2. *Let $f(z) = \sum_{n=-\infty}^{n=\infty} a_n(z-z_0)^n$ be a Laurent series around $z_0 \in \mathbb{C}$ in an annulus. Then for all $n \in \mathbb{Z}$*

$$a_n = \frac{1}{2\pi i} \int_{C_r} \frac{f(w)}{(w - z_0)^{n+1}} dw$$

where $R_1 < r < R_2$.

Proof. Use the uniform convergence of the Laurent series on $\frac{f(z)}{(z-z_0)^{n+1}}$ (so to limits out of integrals) and the fact that $\int_{C_r} (z - z_0)^n dz = 2\pi i$. □

By Cauchy-Hadamard theorem for calculation of radius of convergence we also get the parameters for the maximum annulus on which a Laurent series can exist.

Lemma 10.3.1.3. *For a Laurent series $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$, the smallest value of R_1 and largest value of R_2 such that $f(z)$ converges on $A_{z_0}(R_1, R_2)$ is given by*

1. $R_1 = \limsup_{n\to\infty} |a_{-n}|^{\frac{1}{n}}$
2. $R_2 = \frac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}}$

Proof. Straightforward use of Cauchy-Hadamard. □

The following is the main theorem here.

Theorem 10.3.1.4. *Consider any $0 < R_1 < R_2$ and any $z_0 \in \mathbb{C}$. If $f : A_{z_0}(R_1, R_2) \to \mathbb{C}$ is holomorphic, then it is represented by a Laurent series.*

10.3.2 Isolated singularities : Removable, poles and essential

We now come to the main matter of the present study, the notion of singularities. A holomorphic function $f : G \to \mathbb{C}$ is said to have an *isolated singularity* at $z_0 \notin G$ if there exists a punctured disc $A_{z_0}(0, r) \hookrightarrow G$. Consequently, by Theorem 10.3.1.4, we obtain a Laurent series expansion of f in $A_{z_0}(0, r)$. Let us denote it by

$$f(z) = \sum_{n=-\infty}^{n=\infty} a_n z^n.$$

We can then classify the isolated singularity z_0 into three types:
1. z_0 is a *removable singularity* if $a_n = 0$ for all $n < 0$,
2. z_0 is a *pole of order m* if $\min\{n < 0 \mid a_n \neq 0\} = m$,
3. z_0 is an *essential singularity* if $\min\{n < 0 \mid a_n \neq 0\} = -\infty$ or unbounded.

There are three characterizing theorems of each of the three kinds of singularities.

Theorem 10.3.2.1. *(Riemann's extension theorem) Let $f : G \to \mathbb{C}$ be a holomorphic map. Then the following are equivalent.*
1. *The point $z_0 \in \mathbb{C} \setminus G$ is a removable singularity of f.*
2. *There exists a punctured disc $A_{z_0}(0, r) \hookrightarrow G$ such that f is bounded on it.*

Theorem 10.3.2.2. *(Criterion for a pole) Let $f : G \to \mathbb{C}$ be a holomorphic map. Then the following are equivalent.*
1. *The point $z_0 \in \mathbb{C} \setminus G$ is a pole of f of some order.*
2. *We have*

$$\varinjlim_{z \to z_0} |f(z)| = \infty.$$

Theorem 10.3.2.3. *(Casorati-Weierstrauss theorem) Let $f : G \to \mathbb{C}$ be a holomorphic map. If the point $z_0 \in \mathbb{C} \setminus G$ is an essential singularity of f, then there exists a punctured disc $A_{z_0}(0, r) \hookrightarrow G$ such that $f(A_{z_0}(0, r))$ is dense in $\mathbb{C}$.*

The last theorem in particular shows the chaotic behaviour of essential singularities.

10.4 Cauchy's theorem - II

Let $\gamma : I \to \mathbb{C}$ be a piecewise closed C^1 curve in $\mathbb{C}$ and let $\Omega = \mathbb{C}^\times \setminus \mathrm{Im}\,(\gamma)$. We define the *index of* γ to be the following map over Ω:

$$\begin{aligned} \mathrm{Ind}_\gamma(z) : \Omega &\longrightarrow \mathbb{C} \\ z &\longmapsto \frac{1}{2\pi i} \int_\gamma \frac{1}{w - z} dw. \end{aligned}$$

Lemma 10.4.0.1. *Let $\gamma : I \to \mathbb{C}$ be a piecewise closed C^1 curve in $\mathbb{C}$ and let $\Omega = \mathbb{C}^\times \setminus \mathrm{Im}\,(\gamma)$. Then $\mathrm{Ind}_\gamma(z)$ is a holomorphic map on Ω.*

Proof. This follows from Proposition 10.2.3.4, as $\mathrm{Ind}_\gamma(z)$ is the Cauchy integral of the constant function 1. □

The following is the main theorem that we shall use.

Theorem 10.4.0.2. *Let $\gamma : I \to \mathbb{C}$ be a piecewise closed C^1 curve in $\mathbb{C}$ and let $\Omega = \mathbb{C}^\times \setminus \mathrm{Im}\,(\gamma)$. Then,*

1. $\mathrm{Ind}_\gamma(z)$ *is an integer valued map,*
2. $\mathrm{Ind}_\gamma(z)$ *is constant on each connected component of* Ω*,*
3. $\mathrm{Ind}_\gamma(z)$ *is 0 on unbounded component of* Ω*.*

We now introduce the main Cauchy's theorem.

10.4.1 General Cauchy's theorem

To state the Cauchy's theorem in full generality, we first need to build the small language of chains, which is just a slight generalization of curves. Let $\{\gamma_i : I_i \to \mathbb{C}\}_{i=1}^n$ be a finite collection of piecewise C^1 curves over $\mathbb{C}$. A *chain generated by* $\{\gamma_i\}$ is a formal sum of the form

$$\Gamma = \gamma_1 + \cdots + \gamma_n.$$

One can be more precise here by treating Γ as an element of the free abelian group of all singular 1-chains, but we don't need that technology right now. We denote

$$\mathrm{Im}\,(\Gamma) := \bigcup_{i=1}^n \mathrm{Im}\,(\gamma_i).$$

Moreover, for a continuous map $f : \mathrm{Im}\,(\Gamma) \to \mathbb{C}$, we further denote

$$\int_\Gamma f(z)dz := \sum_{i=1}^n \int_{\gamma_i} f(z)dz.$$

We can further define the *index of a chain* Γ as simply the sum of indices of individual curves:

$$\mathrm{Ind}_\Gamma(z) := \sum_{i=1}^n \mathrm{Ind}_{\gamma_i}(z)$$

for all $z \in \Omega$, where $\Omega = \mathbb{C}^\times \setminus \mathrm{Im}\,(\Gamma)$. Note that the set Ω here will have multiple components if each element of the cycle is a distinct loop. Indeed, if $\Gamma = \gamma_1 + \cdots + \gamma_n$ is a cycle where each γ_i is a closed loop, then we call Γ a *cycle*. The general Cauchy's theorem is then a statement about integral over cycles.

Theorem 10.4.1.1. *(Cauchy's theorem) Let $\Omega \subseteq \mathbb{C}$ be an open set and $\Gamma \hookrightarrow \Omega$ be a cycle such that*

$$\mathrm{Ind}_\Gamma(z) = 0 \;\; \forall z \in \mathbb{C}^\times \setminus \Omega.$$

Let $f : \Omega \to \mathbb{C}$ be a holomorphic map. Then,

1. *(Integral formula)*

$$\mathrm{Ind}_\Gamma(z) f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f(w)}{w - z} dw$$

for all $z \in \Omega \setminus \mathrm{Im}\,(\Gamma)$.

2. *(Integral theorem)*

$$\int_{\Gamma} f(z)dz = 0,$$

3. *if* $\Gamma_0, \Gamma_1 \hookrightarrow \Omega$ *are two cycles such that* $\mathrm{Ind}_{\Gamma_0}(z) = \mathrm{Ind}_{\Gamma_1}(z)$ *for all* $z \notin \Omega$*, then*

$$\int_{\Gamma_0} f(z)dz = \int_{\Gamma_1} f(z)dz.$$

The most important of the above triad of conclusions is the first one, which clearly generalizes the known integral formula.

10.4.2 Homotopy & Cauchy's theorem

Theorem 10.4.2.1. *Let* $\Omega \subseteq \mathbb{C}$ *be an open-connected set. If* $\gamma_0, \gamma_1 \hookrightarrow \Omega$ *are two piecewise* C^1 *closed loops in* Ω *such that they are homotopic in* Ω*, then*

$$\mathrm{Ind}_{\gamma_0}(z) = \mathrm{Ind}_{\gamma_1}(z) \ \ \forall z \notin \Omega.$$

This has some major corollaries in combination with Theorem 10.4.1.1.

Corollary 10.4.2.2. *Let* $\Omega \subseteq \mathbb{C}$ *be an open-connected set and let* $f : \Omega \to \mathbb{C}$ *be a holomorphic map. If* $\gamma_0, \gamma_1 \hookrightarrow \Omega$ *are two piecewise* C^1 *closed loops in* Ω *such that they are homotopic in* Ω*, then*

$$\int_{\gamma_0} f(z)dz = \int_{\gamma_1} f(z)dz.$$

Corollary 10.4.2.3. *Let* $\Omega \subseteq \mathbb{C}$ *be an open-connected set and let* $f : \Omega \to \mathbb{C}$ *be a holomorphic map. If* $\gamma \hookrightarrow \Omega$ *is a piecewise* C^1 *closed loop in* Ω *and* Ω *is simply connected, then*

$$\int_{\gamma} f(z)dz = 0.$$

10.5 Residues and meromorphic maps

Let $\Omega \subseteq \mathbb{C}$ be an open-connected set and $f : \Omega \to \mathbb{C}$ be holomorphic. Let $z_0 \notin \Omega$ be a point of isolated singularity of f. The *residue of* f *at* z_0 is then defined to be the coefficient a_{-1} of the Laurent series

$$\sum_{n=-\infty}^{\infty} a_n z^n$$

of the map f around z_0. We denote residue of f at z_0 by $\mathrm{res}_{z_0}(f) := a_{-1}$. For example, consider the following integral where C_r is a circle of radius r centered at z_0

$$\int_{C_r} \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n.$$

Since all terms $a_n(z-z_0)^n$ are $n \neq -1$ contributes zero integral as the positive parts of holomorphic in the interior of the loop and the negative parts are derivatives of constant 1, which is zero, therefore the only non-zero term is contributed by $n = -1$. Consequently, we have

$$\begin{aligned}\int_{C_r} \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n &= 2\pi i a_{-1} \\ &= 2\pi i \mathrm{res}_{z_0}(f)\end{aligned}$$

where $f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$.

We now define a class of holomorphic maps which one encounters often in complex analysis.

Definition 10.5.0.1. (**Meromorphic maps**) Let $\Omega \subseteq \mathbb{C}$ be an open-connected set and $f : \Omega \to \mathbb{C}$ be any function. We say that f is meromorphic if

1. there exists a set $A \subset \Omega$ which has no limit points in Ω,
2. $f : \Omega \setminus A \to \mathbb{C}$ is holomorphic,
3. every point of A is a pole of f.

One often calls the set A as the *set of poles of* f.

There are some observations to be made.

Lemma 10.5.0.2. *Let $f : \Omega \to \mathbb{C}$ be a meromorphic map on an open-connected set Ω. Then, the set of poles of f is atmost countable.*

Proof. Let $A \subset \Omega$ be the set of poles of f. Covering Ω by countably many compact sets $\{K_i\}$, we observe that intersection of each of $K_i \cap A$ has to be atmost finite, otherwise there exists a sequence in $K_i \cap A$, which consequently admits a convergent subsequence, that is, a limit point in Ω. Consequently, A is a countable union of finite sets. □

Remark 10.5.0.3. For the purposes of residue of f at $a \in A$, one can replace analysis of f with analysis of f by the analysis of $Q = \sum_{n=-m}^{-1} a_n(z-a)^n$, called the principal part of f at a where m is the order of pole of f at $a \in A$. Clearly, $\mathrm{res}_a Q = \mathrm{res}_a f$. Moreover, one sees that

$$\mathrm{res}_a(f)\mathrm{Ind}_\gamma(a) = \frac{1}{2\pi i}\int_\gamma Q(z)dz$$

where γ is a piecewise C^1-loop centered at a, in $\Omega \setminus A$. This is again a consequence of the fact that all terms inside the integral are zero except the one corresponding to a_{-1}. Indeed, this hints at a general phenomenon, which is clarified by the following theorem.

Theorem 10.5.0.4. *(The residue theorem) Let $\Omega \subseteq \mathbb{C}$ be an open-connected set. If $f : \Omega \to \mathbb{C}$ is a meromorphic map with $A \subseteq \Omega$ its set of poles and Γ a cycle in $\Omega \setminus A$ such that*

$$\mathrm{Ind}_\Gamma(z) = 0 \ \ \forall z \notin \Omega,$$

then

$$\frac{1}{2\pi i}\int_\Gamma f(z)dz = \sum_{a\in A} \mathrm{res}_a(f)\mathrm{Ind}_\Gamma(a).$$

We now an important result, which gives us information of zeroes of holomorphic maps on certain subsets.

Theorem 10.5.0.5. *Let $\gamma : I \to \mathbb{C}$ be a piecewise closed C^1-loop in an open-connected set $\Omega \subseteq \mathbb{C}$ such that*

1. $\mathrm{Ind}_\gamma(z) = 0$ *for all* $z \notin \Omega$,
2. $\mathrm{Ind}_\gamma(z) = 0$ *or 1 for all* $z \in \Omega \setminus \mathrm{Im}\,(\gamma)$.

Then we have that for any holomorphic maps $f, g : \Omega \to \mathbb{C}$, denoting $\Omega_1 := \{z \in \Omega \backslash \mathrm{Im}\,(\gamma) \mid \mathrm{Ind}_\gamma(z) = 1\}$ and $N_f = \#Z(f) \cap \Omega_1$, we get that

1. *if f has no zeros on $\mathrm{Im}\,(\gamma) \subseteq \Omega$, then*
$$N_f = \frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = \mathrm{Ind}_{f \circ \gamma}(0).$$
2. *(Rouché's theorem) if*
$$|f(z) - g(z)| < |g(z)| \quad \forall z \in \mathrm{Im}\,(\gamma),$$
then $N_g = N_f$.

10.5.1 Riemann mapping theorem

The following is a very strong rigidity result for holomorphic maps.

Theorem 10.5.1.1. *Let $\Omega \subsetneq \mathbb{C}$ be a proper simply connected domain. Then Ω is biholomorphic to the open unit disc.*

This is a starting point for the uniformization.

Chapter 11

Riemann Surfaces

Contents

We discuss the theory of one dimensional complex manifolds, aka Riemann surfaces. The tools which we will use here will yield a good motivation to study them in the scheme theoretic setting, as done in Chapter 1. Our main goal in these notes is to reach Riemann-Roch theorem and discuss some applications.

Let us begin by defining a Riemann surface and then see some examples.

11.1 Introduction

These are some notes on Riemann surfaces. We wish to prove three main results here: monodromy, Riemann-Hurwitz formula and the infamous Riemann-Roch theorem. We also wish to portray example uses of them. A philosophical goal in our mind is also to see how analytic world behaves in comparison to algebraic world. We do this in part so that we can get more insights into the latter, before going into an involved study of it.

We make references to our notes "Facets of Geometry" by writing Theorem FoG.23.2.1.4.

11.1.1 Definitions and basic properties

After defining Riemann surfaces and giving basic examples, we will cover some basic lemmas some of which generalizes results which we have seen in complex analysis of one variable.

Definition 11.1.1.1 (Riemann surface). A conformal atlas $\mathcal{A} = (\{U_i\}_i, \{z_i\}_i)$ on a second countable Hausdorff space X is the data of an open cover $\{U_i\}_{i\in I}$ of X together with open embeddings $z_i : U_i \to \mathbb{C}$ such that if $U_i \cap U_j \neq \emptyset$ then the composite

$$z_j \circ z_i^{-1} : z_i(U_i \cap U_j) \to z_j(U_i \cap U_j)$$

is a holomorphic map between two open subsets of $\mathbb{C}$. Two conformal atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are equivalent if for any U, z in $\mathcal{A}_1$ and any V, w in $\mathcal{A}_2$, the transition map

$$w \circ z^{-1} : z(U \cap V) \to w(U \cap V)$$

is a conformal map. A Riemann surface X is a connected Hausdorff space with an equivalence class of conformal atlas. We usually fix one atlas in a class which is maximal in that it is the union of all atlases in that class.

Example 11.1.1.2. Here are few examples of Riemann surfaces.

1. Any open subset $U \subseteq \mathbb{C}$ is a Riemann surface. Indeed, consider $\mathrm{id} : U \to U$, this defines a conformal atlas on U. Thus $\mathbb{C}$ and the open unit disc $\mathbb{D}$ are Riemann surfaces.
2. The Riemann sphere $\bar{\mathbb{C}}$ or usually called complex projective line $\mathbb{P}^1_{\mathbb{C}}$ (see Proposition 11.1.3.2) is a Riemann surface. Topologically, $\mathbb{P}^1_{\mathbb{C}}$ is S^2. We give a conformal structure on S^2 as follows. Consider the open sets $U_+ = S^2 - p$ and $U_- = S^2 - q$ where p and q are north and south poles respectively. Consider

$$\begin{aligned} z_+ : U_+ &\longrightarrow \mathbb{C} \\ (x_1, x_2, x_3) &\longmapsto \frac{x_1 + ix_2}{1 - x_3} \\ z_- : U_- &\longrightarrow \mathbb{C} \\ (x_1, x_2, x_3) &\longmapsto \frac{x_1 - ix_2}{1 + x_3}. \end{aligned}$$

These are obtained by usual stereographic projection from north pole p. One can observe that

$$\begin{aligned} z_+(U_+) &= \mathbb{C} \\ z_-(U_-) &= \mathbb{C} - \{0\} \end{aligned}$$

and are thus homeomorphisms. Furthermore $U_+ \cap U_- = S^2 - \{p, q\}$. It follows that $z_+(U_+ \cap U_-) = \mathbb{C}^\times = z_-(U_+ \cap U_-)$, the punctured complex plane. The transition map can be checked to be

$$\begin{aligned} z_+ \circ z_-^{-1} : z_-(U_+ \cap U_-) &\longrightarrow z_+(U_+ \cap U_-) \\ w &\longmapsto \frac{1}{w} \end{aligned}$$

which as a map $\mathbb{C}^\times \to \mathbb{C}^\times$ is conformal.

Here is how we can define maps of Riemann surfaces.

Definition 11.1.1.3 (**Holomorphic maps of Riemann surfaces**). Let X and Y be two Riemann surfaces with atlases (U_i, z_i) and (V_i, w_i) on X and Y respectively and $f : X \to Y$ be a continuous map. Then, f is said to be holomorphic if for each $x \in X$ and charts $U_i \ni x$ and $V_j \ni f(x)$, the composite

$$w_j \circ f \circ z_i^{-1} : z_i(U_i) \to w_j(V_j)$$

is a holomorphic map between two open sets of $\mathbb{C}$. We denote by $\mathcal{O}(X) = \{f : X \to \mathbb{C} \mid f \text{ is holomorphic}\}$. This is a $\mathbb{C}$-algebra under pointwise addition and multiplication.

Lemma 11.1.1.4. *Let $f : X \to Y$ and $g : Y \to Z$ be a holomorphic map of Riemann surfaces. Then $g \circ f : X \to Z$ is a holomorphic map.*

Proof. Denote $h = g \circ f : X \to Z$. Pick any $x \in X$ and pick any coordinate charts $(U_x, \varphi_x) \ni x$ and $(W_{h(x)}, \varphi_{h(x)}) \ni h(x)$. We wish to show that $\varphi_{h(x)} \circ h \circ \varphi_x^{-1} : \varphi_x(U_x) \to \varphi_{h(x)}(W_{h(x)})$ is holomorphic. Pick any chart $V_{f(x)} \ni f(x)$. Then we have

$$\varphi_{h(x)} \circ h \circ \varphi_x^{-1} = \varphi_{h(x)} \circ g \circ \varphi_{f(x)}^{-1} \circ \varphi_{f(x)} \circ f \circ \varphi_x^{-1}$$

where $\varphi_{h(x)} \circ g \circ \varphi_{f(x)}^{-1}$ and $\varphi_{f(x)} \circ f \circ \varphi_x^{-1}$ are holomorphic as f and g are holomorphic. This completes the proof. $\square$

Remark 11.1.1.5. We get a category of Riemann surfaces, denoted **RS**.

Definition 11.1.1.6 (**Subsurface**). Let X be a Riemann surface and $U \subseteq X$ be an open set. Then U is also a Riemann surface with the charts obtained by restrictions of that of X.

There is an identity principle for Riemann surfaces, which would be used quite often.

Lemma 11.1.1.7 (Identity principle). *Let X, Y be Riemann surface and X be connected. If $f, g : X \to Y$ are holomorphic and there exists $A \subseteq X$ which has a limit point in X such that $f|_A = g|_A$, then $f = g$.*

Proof. Let $a \in X$ be a limit point of A and let (U, z) be a chart of a. Then $f|_U = g|_U$ by usual identity principle of $\mathbb{C}$. Now pick any point $a \neq b \in X$. As X is locally path-connected and connected, therefore it is path-connected. Let $\gamma : a \to b$ be a path joining a and b in X. We claim that f is constant along this path. Indeed, cover $\operatorname{Im}(\gamma)$ by finitely many charts of X denoted U_i such that $U_i \cap U_{i+1} \neq \emptyset$ with $U_1 = U$. As f and g agree on an open subset of U_2, therefore by identity principle of $\mathbb{C}$, it follows that $f|_{U_2} = g|_{U_2}$. Continuing this, we conclude that $f = g$ on $\operatorname{Im}(\gamma)$ and thus $f(b) = g(b)$, as required. $\square$

Corollary 11.1.1.8. *Let $f : X \to \mathbb{C}$ be a non-zero holomorphic where X is a Riemann surface. Then $D(f) := \{x \in X \mid f(x) = 0\}$ is a discrete set in X.*

Proof. If $D(f)$ is not discrete, then it has a limit point and thus by Lemma 11.1.1.7, $f = 0$, a contradiction. $\square$

We can define meromorphic maps between Riemann surfaces as well.

Definition 11.1.1.9 (**Meromorphic maps**)**.** Let X be a Riemann surface. A meromorphic map on X is a holomorphic map $f : X \to \mathbb{P}^1_{\mathbb{C}}$ such that $f \neq c_\infty$, c_∞ being the constant infinity map. By identity principle (Lemma 11.1.1.7), thus, $f^{-1}(\infty)$ has to be a discrete set. We denote the set of all meromorphic functions on X as $\mathcal{M}(X)$. Clearly, $\mathcal{M}(X)$ is a $\mathbb{C}$-algebra.

Meromorphic maps form a field!

Lemma 11.1.1.10. *Let X be a connected Riemann surface. Then $\mathcal{M}(X)$ is a field.*

Proof. Let $f : X \to \mathbb{P}^1_{\mathbb{C}}$ be a non-zero meromorphic map. Then consider $g := 1/f$ on $D(f) = \{x \in X \mid f(x) \neq 0\}$ and $g := \infty$ on $X \setminus D(f)$. Clearly, g is holomorphic on $D(f)$, which is open. Since $D(f)$ is discrete by Corollary 11.1.1.8. Thus, g is indeed meromorphic. Observe that $f \cdot g = 1$ on X as it is one on $D(f)$ and then we may apply identity principle (Lemma 11.1.1.7). This completes the proof. □

Remark 11.1.1.11. As there is a natural inclusion $\mathcal{O}(X) \hookrightarrow \mathcal{M}(X)$, thus it follows $\mathcal{O}(X)$ is a domain. By universal property of fraction fields, $Q(\mathcal{O}(X)) \subseteq \mathcal{M}(X)$.

We now see when $\mathcal{O}(X)$ itself is a field.

Lemma 11.1.1.12 (General Liouville)**.** *Let X be a compact connected Riemann surface*[1]*, then $\mathcal{O}(X)$ is isomorphic to $\mathbb{C}$ as the only elements in $\mathcal{O}(X)$ are constants.*

Proof. Pick any $f \in \mathcal{O}(X)$. We wish to show that f is a constant. Consider the composite $X \to \mathbb{C} \to \mathbb{R}$ given by $|f|$. As X is compact, thus $|f|$ achieves maxima, say at $x_0 \in X$ and $a = |f(x_0)|$. For a chart $(U, z) \ni x_0$, we have by maximum-modulus for $\mathbb{C}$ that $|f|$ is constant and thus f is constant c_a on U. By identity principle (Lemma 11.1.1.7), it follows that f is constant c_a on the entire X. □

Open mapping theorem is also true for maps of Riemann surfaces.

Lemma 11.1.1.13 (Open mapping theorem)**.** *Let X be a connected Riemann surface and $f : X \to Y$ be a holomorphic map. Then f is an open map.*

Proof. Pick any open set $U \subseteq X$ and consider $f(U) \subseteq Y$. We wish to show that $f(U)$ is open. Pick any point $f(x) \in f(U)$ where $x \in U$. Pick any chart $(V, z) \ni x$ and $(W, w) \ni f(x)$ such that $V \subseteq U$. Thus the map $w \circ f \circ z^{-1} : z(V) \to w(W)$ is a holomorphic map. By open mapping theorem for $\mathbb{C}$, it follows that $w \circ f \circ z^{-1}$ is an open map. Thus, let $x \in V' \subseteq V$ be an open set. Then $w \circ f \circ z^{-1}(z(V')) = w(f(V')) \subseteq w(W)$ is open and thus $f(x) \in f(V') \subseteq W$ is open, as required. □

There is an intimate connection between covering spaces and Riemann surfaces, whose first piece we explain as follows. We first need a small lemma.

Lemma 11.1.1.14. *Let $p_i : X_i \to Y$ for $i = 1, 2$ be a holomorphic map where X_i, Y are Riemann surfaces and X_i are connected. If there exists a continuous $f : X_1 \to X_2$ such that*

$$\begin{tikzcd} & X_2 \arrow[d, "p_2"] \\ X_1 \arrow[ur, "f"] \arrow[r, "p_1"'] & Y \end{tikzcd}$$

[1] say, for example, $\mathbb{P}^1_{\mathbb{C}}$!

commutes, then f is holomorphic.

Proof. Pick any point $x \in X_1$, charts $(U_1, z_1) \ni x$ and $(U_2, z_2) \ni f(x)$. We wish to show that $z_2 \circ f \circ z_1^{-1} : z_1(U_1) \to z_2(U_2)$ is holomorphic. Indeed, we may first assume by continuity of f and p_i that U_i is a sheet of an evenly covered neighborhood $V \subseteq Y$ under p_i for $i = 1, 2$. Now, the restricted maps $p_i : U_i \to z_i(U_i)$ are biholomorphic maps. Now, $z_i \circ p_i^{-1} : V \to z_i(U_i)$ are two charts in Y. As transition maps has to be holomorphic, therefore we get

$$(z_2 \circ p_2^{-1}) \circ (z_1 \circ p_1^{-1})^{-1} = z_2 \circ f \circ z_1^{-1}$$

is holomorphic, as required. □

Proposition 11.1.1.15. *Let $p : X \to Y$ be a covering map where Y is a Riemann surface. Then there exists a unique conformal structure on X such that p is holomorphic.*

Proof. We first do uniqueness, as it is easy by Lemma 11.1.1.14. Indeed, if there are two non-equivalent conformal structures on X, then we get two holomorphic covering maps $p_i : X_i \to Y$. As underlying space and maps of each (X_i, p) is same, therefore by lifting criterion for covering maps, we deduce that there is a continuous map $f : X_1 \to X_2$ which is furthermore holomorphic by above lemma and $p \circ f = p$. Now, we may similarly get $g : X_2 \to X_1$ holomorphic such that $p \circ g = p$. As these lifts are based lifts, we get that $f \circ g$ is unique with respect to the fact that it fixes a point, thus it is id, similarly for the other side. Hence $X_1 \cong X_2$, that is, they are biholomorphic and thus have equivalent conformal structure.

We thus need only construct a conformal structure on X via p. Indeed, we may first assume that Y has an atlas (V_i, z_i) fine enough that each $V_i \subseteq Y$ is an evenly covered neighborhood. Hence for each V_i, the map $p : W_{i,j} \to V_i$ is a homeomorphism where $p^{-1}(V_i) = \coprod_j W_{i,j}$. Define an open cover of X by $(W_{i,j}, z_i \circ p)$. We claim that this is an atlas. Indeed, $z_i \circ p : W_{i,j} \to z_i(V_i)$ is a homeomorphism and for any $(i, j), (k, l)$, we have $(z_i \circ p) \circ (z_j \circ p)^{-1} = z_i \circ z_j^{-1}$, which is a holomorphic map. This completes the proof. □

11.1.2 Structure sheaf and modules

We wish to show that the structure sheaf of a Riemann surface $\mathcal{O}_X$ is such that the meromorphic sheaf $\mathcal{M}$ is an $\mathcal{O}_X$-module. So we first define the structure sheaf.

Remark 11.1.2.1 (Riemann surface as a locally ringed space)**.** Let X be a Riemann surface with an atlas (U_i, z_i). As discussed in Chapter 8, §8.1.2 on "Sheaves and atlases" in FoG, by Theorem FoG.8.1.2.4, it follows that we get an atlas sheaf (Definition FoG.8.1.2.1) $\mathcal{O}_X$ on X w.r.t which $(X, \mathcal{O}_X)$ is a locally ringed space which is a complex manifold (Definition FoG.8.1.1.3) of dimension 1. Recall that in particular for an open subset $U \subseteq X$, $\mathcal{O}_X(U)$ is defined by

$$\mathcal{O}_X(U) = \{f : U \to \mathbb{C} \mid f \circ x_i^{-1} : x_i(U \cap U_i) \to \mathbb{C} \text{ is holomorphic}\},$$

that is, $\mathcal{O}_X$ is the sheaf of holomorphic maps on X. The $\mathcal{O}_X$ is also called the structure sheaf of X. Thus, giving a conformal structure on X is equivalent to giving an atlas sheaf.

We will be using this sheaf very frequently, as it will be of fundamental importance to us to translate over working working knowledge of algebraic geometry to this analytic language[2].

[2]Note that explicit charts are rarely used in schemes, whereas in geometry, one uses it quite frequently.

Remark 11.1.2.2. There might be apparent addition of complexity to think of a Riemann surface as a locally ringed space with a sheaf of holomorphic maps without any reference to a chart. But we wish to portray that one can prove results similar to that in previous section from this point of view as well, as this allows us to reduce to *local affine patch* (i.e. local chart) quite immediately.

For example, general Liouville (Lemma 11.1.1.12) can also be seen by the following argument. Considering that $|f| : X \to \mathbb{R}$ achieves maximum at $x_0 \in X$, for an affine open set containing x_0 say U, the restriction $f|_U : U \to \mathbb{R}$ can be thought of as a map on an open subset of $\mathbb{C}$ which achieves maximum on interior, so $f|_U$ is constant. Thus by identity priniciple, we are done.

An important and crucial observation from complex analysis of one variable is the following:

Proposition 11.1.2.3. *Let $(X, \mathcal{O}_X)$ be a Riemann surface. Then for any $x \in X$, the stalk is isomorphic to power series ring over $\mathbb{C}$:*

$$\mathcal{O}_{X,x} \cong \mathbb{C}[[z]].$$

Proof. Let $U \ni x$ be an affine open subset of x. Then, $\mathcal{O}_{X,x} = \mathcal{O}_{U,x}$. Let $\varphi : U \to \mathbb{C}$ be a chart. As it is an open embedding, therefore, $\mathcal{O}_{U,x} \cong \mathcal{O}_{\mathbb{C},\varphi(x)}$, where $\mathcal{O}_{\mathbb{C}}$ is the sheaf of holomorphic maps on $\mathbb{C}$. As any homolomorphic map has a power series representation at each point, thus, power series forms a cofinal system in the representation of a holomorphic map in the stalk. The result now follows. $\square$

Remark 11.1.2.4. This proposition immediately tells us what type of information is stored in the stalk. That is, it tells you how a function locally around a point looks like.

We next see that meromorphic maps form a sheaf as well.

Definition 11.1.2.5. Let $(X, \mathcal{O}_X)$ be a Riemann surface. The assignment for each open $U \subseteq X$

$$\mathcal{M}_X(U) = \{f : U \to \mathbb{P}^1_{\mathbb{C}} \mid f \neq c_\infty \text{ holomorphic}\}$$

forms a presheaf under restrictions. This is called the sheaf of meromorphic maps on X.

We first see that $\mathcal{M}_X$ is a constant sheaf!

Proposition 11.1.2.6. *Let X be a Riemann surface and let $K = \mathcal{M}_X(X)$ the field of global meromorphic maps. Then*

$$\mathcal{M}_X \cong \underline{K},$$

where the latter is the constant sheaf on field K.

We'll see its proof later. An important property is that the stalks of $\mathcal{M}_X$ are again quite simple.

Proposition 11.1.2.7. *Let $(X, \mathcal{O}_X)$ be a Riemann surface with meromorphic sheaf $\mathcal{M}_X$. Then for any $x \in X$,*

$$\mathcal{M}_{X,x} \cong \mathbb{C}((z)).$$

Proof. Same as Proposition 11.1.2.3 except that in the end we use the fact that any meromorphic function locally has a Laurent series expansion at each point. $\square$

We now study some important class of Riemann surfaces, those coming from non-singular projective plane curves.

11.1.3 Smooth algebraic plane curves

We wish to study a class of examples of Riemann surfaces coming from algebra. This will give us a tight intuition about algebraic curves which will guide further development.

We begin by giving an alternate construction of Riemann surface.

Example 11.1.3.1 (Complex projective line $\mathbb{P}^1_{\mathbb{C}}$)**.** Topologically, we first define $\mathbb{P}^1_{\mathbb{C}} = \mathbb{C}^2/\sim$ where $(z_0, z_1) \sim (\lambda z_0, \lambda z_1)$ for all $\lambda \in \mathbb{C}$. Denote any point in $\mathbb{P}^1_{\mathbb{C}}$ by $[z_0 : z_1]$ where $z_i \in \mathbb{C}$. We now give a conformal structure on $\mathbb{P}^1_{\mathbb{C}}$. Consider $U_0 = \{[1 : z] \mid z \in \mathbb{C}\}$ and $U_1 = \{[z : 1] \mid z \in \mathbb{C}\}$. These are open subspaces of $\mathbb{P}^1_{\mathbb{C}}$ since under the quotient map $\pi : \mathbb{C}^2 \twoheadrightarrow \mathbb{P}^1_{\mathbb{C}}$, $\pi^{-1}(U_0) = \{(z_0, z_1) \in \mathbb{C}^2 \mid z_0 \neq 0\} = D(z_0)$, the plane minus the z_1-axis, which is open. Similarly for U_1.

Now consider the maps which we will show makes (U_i, φ_i) into an affine chart

$$\begin{aligned} \varphi_0 : U_0 &\longrightarrow \mathbb{C} \\ [z_0 : z_1] &\longmapsto \frac{z_1}{z_0} \\ \varphi_1 : U_1 &\longrightarrow \mathbb{C} \\ [z_0 : z_1] &\longmapsto \frac{z_0}{z_1}. \end{aligned}$$

Note that these are homeomorphisms as the image the whole complex plane which is open and φ_i are homeomorphisms onto it. Indeed, φ_i can be seen to be bijective to $\mathbb{C}$ quite easily and an inverse of φ_0, say, can be constructed by defining $\psi_0 : \mathbb{C} \to U_0$ given by $z \mapsto [1 : z]$. This is continuous and an inverse of φ_0.

Now observe that $U_0 \cap U_1 = \{[z_0 : z_1] \mid z_0, z_1 \neq 0\} = U$. Observe that $\varphi_i(U) = \mathbb{C}^\times$. The transition maps then are

$$\begin{aligned} \varphi_1 \circ \varphi_0^{-1} : \varphi_0(U) &\longrightarrow \varphi_1(U) \\ z &\longmapsto \frac{1}{z}, \end{aligned}$$

which is a holomorphic map $\mathbb{C}^\times \to \mathbb{C}^\times$. Thus, we have obtained a Riemann surface $\mathbb{P}^1_{\mathbb{C}}$ with structure sheaf $\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}$ whose sections on an open subset $U \subseteq \mathbb{P}^1_{\mathbb{C}}$ are those functions $f : U \to \mathbb{C}$ which are holomorphic with respect to the chart $(U_i, \varphi_i)_{i=1,2}$. The Riemann surface $(\mathbb{P}^1_{\mathbb{C}}, \mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}})$ is called the *projective line* over $\mathbb{C}$.

Proposition 11.1.3.2. *Let $\overline{\mathbb{C}}$ be the Riemann sphere. Then $\overline{\mathbb{C}}$ is biholomorphic to $\mathbb{P}^1_{\mathbb{C}}$*

Proof. Indeed, consider the map

$$\begin{aligned} f : \overline{\mathbb{C}} &\longrightarrow \mathbb{P}^1_{\mathbb{C}} \\ z &\longmapsto \begin{cases} [1 : z] & \text{if } z \neq \infty \\ [0 : 1] & \text{if } z = \infty. \end{cases} \end{aligned}$$

Indeed, this is continuous since on any neighborhood of 0, this is the inverse of the chart map φ_0 and on any neighborhood of ∞ it is the inverse of the chart φ_1. As $\bar{\mathbb{C}}$ is compact and $\mathbb{P}^1_{\mathbb{C}}$ Hausdorff, it follows that f is a homeomorphism.

Using charts of Example 11.1.1.2, it is immediate to see that this is holomorphic. The inverse of this map is $[z_0 : z_1] \mapsto \frac{z_1}{z_0}$. Again this is continuous and holomorphic by same reasons. □

We now introduce a space where most of our geometry will take place.

Construction 11.1.3.3 ($\mathbb{CP}^2$, the projective plane[3]). Topologically, $\mathbb{CP}^2$ is $\mathbb{C}^3/\sim$ where $(z_0, z_1, z_2) \sim (\lambda z_0, \lambda z_1, \lambda z_2)$. This can be given the structure of a complex 2-manifold by giving an atlas consisting of three charts $(U_i, \varphi_i)_{i=0,1,2}$ where $U_i = \{[z_0 : z_1 : z_2] \mid z_i \neq 0\}$. The maps are given by

$$\begin{aligned}
\varphi_0 : U_0 &\longrightarrow \mathbb{C}^2 \\
[z_0 : z_1 : z_2] &\longmapsto \left(\frac{z_1}{z_0}, \frac{z_2}{z_0}\right) \\
\varphi_1 : U_1 &\longrightarrow \mathbb{C}^2 \\
[z_0 : z_1 : z_2] &\longmapsto \left(\frac{z_0}{z_1}, \frac{z_2}{z_1}\right) \\
\varphi_2 : U_2 &\longrightarrow \mathbb{C}^2 \\
[z_0 : z_1 : z_2] &\longmapsto \left(\frac{z_0}{z_2}, \frac{z_1}{z_2}\right).
\end{aligned}$$

One can check that this makes $\mathbb{CP}^2$ a complex 2-manifold by showing all transitions are holomorphic maps from open subsets of $\mathbb{C}^2$ to $\mathbb{C}^2$ (which would require a knowledge of several complex variables, but we skip over that as really don't require that here).

We would like to know a class of closed (thus compact) subsets of $\mathbb{CP}^2$ formed by polynomials in two variables. These will motivate algebraic counterparts of the analytic geometry that we are consider currently.

Definition 11.1.3.4 (Projective algebraic plane curves). Let $\bar{p}(z_1, z_2) \in \mathbb{C}[z_1, z_2]$ be a polynomial and let $p(z_1, z_2, z_3) \in \mathbb{C}[z_1, z_2, z_3]$ be its homogenization so that p is homogeneous of degree $d \geq 1$. Consider the set

$$V(p) = \{[z_0 : z_1 : z_2] \in \mathbb{CP}^2 \mid p(z_1, z_2, z_3) = 0\}.$$

This defines a closed subset of $\mathbb{CP}^2$ since it is $\mathbb{CP}^2 \setminus V(p)$ is the image of $\mathbb{C}^3 \setminus V(\bar{p})$ under the quotient map $\pi : \mathbb{C}^3 \to \mathbb{CP}^2$. We call $V(p) \subseteq \mathbb{CP}^2$ a projective algebraic plane curve.

We see that a projective algebraic plane curve Z is formed by three pieces of affine algebraic plane curves.

Lemma 11.1.3.5. *Let $\bar{p} \in \mathbb{C}[z_0, z_1]$, $p \in \mathbb{C}[z_0, z_1, z_2]$ be its homogenization and $Z = V(p)$ be the projective algebraic curve. Let $(U_i, \varphi_i)_{i=0,1,2}$ be the standard chart of $\mathbb{P}^2_{\mathbb{C}}$ (see Construction 11.1.3.3). Then the image of $Z \cap U_i$ under φ_i in $\mathbb{C}$ is $V(\bar{p}_i)$ where $\bar{p}_0 = p(1, z_1, z_2)$, $\bar{p}_1 = p(z_0, 1, z_2)$ and $\bar{p}_2 = p(z_0, z_1, 1)$.*

Proof. Indeed, since, say $Z \cap U_0 = \{[1 : z_1 : z_2] \mid p(1, z_1, z_2) = 0\}$, therefore

$$\varphi_0(Z \cap U_0) = \{(z_1, z_2) \mid p(1, z_1, z_2) = 0\} = V(\bar{p}_0).$$

The other cases are same. $\square$

[3]We would freely interchange between $\mathbb{CP}^2$ and $\mathbb{P}^2_{\mathbb{C}}$, depending on the temperature outside.

We now show that a certain type of algebraic plane curves define Riemann surfaces.

Definition 11.1.3.6 (**Smooth algebraic plane curves**). Let $f \in \mathbb{C}[z_1, z_2, z_3]$ be a homogeneous polynomial. Then, the polynomial f is called non-singular or smooth if for all points $p \in V(f) \subseteq \mathbb{CP}^2$, we have that $\frac{\partial f}{\partial z_i}\Big|_p \neq 0$ for atleast one i from $0, 1, 2$. In this case, the projective plane curve $V(f)$ that it defines is called the smooth projective algebraic plane curve. A similar definition gives smooth affine algebraic plane curves in $\mathbb{C}^2$.

We now show that every smooth projective plane curve defined by an irreducible smooth homogeneous polynomial in three variables gives a Riemann surface. For that we need following two preliminary results.

Theorem 11.1.3.7. *Let $p \in \mathbb{C}[z_0, z_1, z_2]$ be a homogeneous polynomial.*
1. *If p is non-singular, then $V(p) \subseteq \mathbb{P}^2_{\mathbb{C}}$ is irreducible.*
2. *If p is irreducible, then $V(p) \subseteq \mathbb{P}^2_{\mathbb{C}}$ is connected.*

We now state the main theorem. Its proof can be seen by implicit function theorem for $\mathbb{C}$, but we omit all such checks.

Theorem 11.1.3.8. *Let $p \in \mathbb{C}[z_0, z_1, z_2]$ be a non-singular homogeneous polynomial. Then, $V(p) \subseteq \mathbb{P}^2_{\mathbb{C}}$ is a compact connected Riemann surface.*

11.2 Ramified coverings & Riemann-Hurwitz formula

11.3 Monodromy & analytic continuation

11.4 Holomorphic & meromorphic forms

Having differentials on a given geometric object gives us a sense of direction of each point. Exploiting this, one can define very many types of forms (differentiable, holomorphic, meromorphic...) and their interrelations which allows us to study the object in question more deeply.

11.4.1 Differentials

We will We first construct the sheaf of differentiable maps on a Riemann surface.

Definition 11.4.1.1 (**Sheaf of differentiable maps**). Let X be a Riemann surface. Consider the assignment for each open $U \subseteq X$

$$\mathcal{E}_X(U) := \{f : U \to \mathbb{C} \mid \forall \text{ charts } (U_i, z_i),\ f \circ z_i^{-1} : z_i(U \cap U_i) \to \mathbb{C} \text{ is differentiable.}\}$$

This assignments with restrictions naturally forms a sheaf, called the sheaf of differentiable maps on X. This is a sheaf of $\mathbb{C}$-algebras. Moreover, this is an $\mathcal{O}_X$-algebra as well since pointwise product of holomorphic and differentiable map is again differentiable.

We will use the sheaf $\mathcal{E}_X$ to build many other sheaves which will be of prime importance to us. Let us first introduce few operators on the seaf $\mathcal{E}_X$.

Construction 11.4.1.2 (Operators on $\mathcal{E}_X$). Define $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ as two operators on $\mathcal{E}_X$ as follows. For any open $U \subseteq X$, define

$$\begin{aligned} \frac{\partial}{\partial x} : \mathcal{E}_X(U) &\longrightarrow \mathcal{E}_X(U) \\ f : U \to \mathbb{C} &\longmapsto \frac{\partial f}{\partial x} : U \to \mathbb{C} \end{aligned}$$

where $\frac{\partial f}{\partial x} : U \to \mathbb{C}$ is defined as follows. Let (U_i, z_i) be a chart. As f is differentiable, therefore $f \circ z_i^{-1} : z_i(U \cap U_i) \to \mathbb{C}$ is differentiable. Define

$$\frac{\partial f}{\partial x} \circ z_i^{-1} = \frac{\partial}{\partial x}\left(f \circ z_i^{-1}\right)$$

for each chart (U_i, z_i). Similarly, one defines $\frac{\partial}{\partial y}$. Note that these maps commutes with restrictions. Hence we get sheaf maps $\frac{\partial}{\partial x}, \frac{\partial}{\partial y} : \mathcal{E}_X \to \mathcal{E}_X$. Note that both of these are $\mathbb{C}$-linear.

Consider the two operators

$$\begin{aligned} \frac{\partial}{\partial z} &:= \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \\ \frac{\partial}{\partial \bar{z}} &:= \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right). \end{aligned}$$

These two also define $\mathbb{C}$-linear operators on $\mathcal{E}_X$.

We observe some of the immediate consequences.

Lemma 11.4.1.3. *Let $(X, \mathcal{O}_X)$ be a Riemann surface. Then,*

$$0 \longrightarrow \mathcal{O}_X \longrightarrow \mathcal{E}_X \xrightarrow{\frac{\partial}{\partial \bar{z}}} \mathcal{E}_X$$

is exact.

Proof. Note that the map $\mathcal{O}_X \to \mathcal{E}$ is obtained by thinking of a holomorphic map as a real differentiable map. By Cauchy-Riemann, $f : V \subseteq \mathbb{C} \to \mathbb{C}$ is holomorphic if and only if $\frac{\partial f}{\partial \bar{z}} = 0$. It follows that on an open $U \subseteq X$, we have $\operatorname{Ker}\left(\frac{\partial}{\partial \bar{z}}\right) = \{f \in \mathcal{O}_X(U) \mid f \text{ is holomorphic}\} = \mathcal{O}_X(U)$, as required. □

Remark 11.4.1.4 ($\mathcal{E}_X$ is locally ringed)**.** Consider the sheaf of differentiable maps $\mathcal{E}_X$ on a Riemann surface X. We observe that for any point $x \in X$, the stalk $\mathcal{E}_{X,x}$ is a local ring where the maximal ideal $\mathfrak{m}_x$ consists of those germs which vanishes at point x. Thus, $\mathcal{E}_X$ is a locally ringed $\mathcal{O}_X$-algebra.

Definition 11.4.1.5 (Cotangent space at a point). Let $(X, \mathcal{O}_X)$ be a Riemann surface and $\mathcal{O}_X$ be the $\mathcal{O}_X$-algebra of differentiable maps. Then, the cotangent space at point $x \in X$ is the $\mathbb{C}$-vector space given by

$$T_x^{(1)} = \frac{\mathfrak{m}_x}{\mathfrak{m}_x^2}$$

where $(\mathcal{E}_{X,x}, \mathfrak{m}_x)$ is the local ring at point x of sheaf $\mathcal{E}_X$. A point in $T_x^{(1)}$ is referred to as a cotangent vector at $x \in X$.

Remark 11.4.1.6 (Cotangent vectors and "direction")**.** For a point $a \in X$, pick a covector $(U, f)_a \in T_x^{(1)}$. As $(U, f)_a \in T_x^{(1)}$ and (U_i, z_i) is a chart containing $a \in X$, therefore $f \circ z_i^{-1} : z_i(U \cap U_i) \to \mathbb{C}$ is differentiable. We may write the Taylor expansion of $f \circ z_i^{-1}$ at the point $z_i(a) = (a_1, a_2)$ to get that

$$f \circ z_i^{-1}(x, y) = f(a) + \frac{\partial f}{\partial x}(a)(x - a_1) + \frac{\partial f}{\partial y}(a)(y - a_2) + \text{terms of degree} \geq 2.$$

As $\mathfrak{m}_x^2$ consists of products of those germs vanishing at a and "terms of degree ≥ 2" vanishes at a, therefore, we get that

$$(U, f)_a = \frac{\partial f}{\partial x}(a)(x - a_1) + \frac{\partial f}{\partial y}(a)(y - a_2) + \mathfrak{m}_a^2.$$

The above motivates the following definition.

Definition 11.4.1.7 (Differential of a map). Let X be a Riemann surface and $f \in \mathcal{O}_X(U)$ be a differentiable map on open $U \subseteq X$. Let $a \in U$. Define the following $\mathbb{C}$-linear transformation

$$\begin{aligned} d_a : \mathcal{E}_{X,a} &\longrightarrow T_a^{(1)} \\ (U, f)_a &\longmapsto (f - f(a)) + \mathfrak{m}_a^2. \end{aligned}$$

As is evident from Remark 11.4.1.6, the differentials of maps $x, y : U \to \mathbb{C}$ which on a chart (U_i, z_i) is defined by $x \circ z_i^{-1} : z_i(U \cap U)i \to \mathbb{C}$ mapping $(x, y) \mapsto x$ and similarly for y, holds special position amongst all differentials.

Proposition 11.4.1.8. *Let X be a Riemann surface and $a \in X$ contained in open U. Then,*

1. *$T_a^{(1)}$ has $\{d_a x, d_a y\}$ as basis.*
2. *$T_a^{(1)}$ has $\{d_a z, d_a \bar{z}\}$ as a basis.*
3. *For any $f \in \mathcal{E}_X(U)$,*

$$\begin{aligned} d_a f &= \frac{\partial f}{\partial x}(a) d_a x + \frac{\partial f}{\partial y}(a) d_a y \\ &= \frac{\partial f}{\partial z}(a) d_a z + \frac{\partial f}{\partial \bar{z}}(a) d_a \bar{z}. \end{aligned}$$

Proof. A simple exercise in reduction to affine charts and using properties of it (in this case, Taylor series). □

Notation 11.4.1.9 (Decomposition of cotangent space)**.** By Proposition 11.4.1.8, it follows that we can write

$$\begin{aligned} T_a^{(1)} &= \mathbb{C} d_a z \oplus \mathbb{C} d_a \bar{z} \\ &=: T_a^{1,0} \oplus T_a^{0,1}. \end{aligned}$$

Elements of $T_a^{1,0}$ are called covectors of type (1,0), same for the other case. For any $f \in \mathcal{E}_X(U)$, we further denote

$$d_a f = d_a' f + d_a'' f$$

for unique $d_a' f \in T_a^{1,0}$ and $d_a'' f \in T_a^{0,1}$, where

$$\begin{aligned} d_a' f &= \frac{\partial f}{\partial z}(a) d_a z, \\ d_a'' f &= \frac{\partial f}{\partial \bar{z}}(a) d_a \bar{z}. \end{aligned}$$

Taking exterior powers of $T_a^{(1)}$ gives us other vector spaces which we will use to define differential k-forms.

Definition 11.4.1.10 (Differential k-forms). Let X be a Riemann surface and $U \subseteq X$ be an open subset. Let $T_a^{(k)} = \wedge^k T_a^{(1)}$ be the k^{th}-exterior power of $T_a^{(1)}$. Note that $\dim_{\mathbb{C}} \wedge^k T_a^{(1)} =$. A differential k-form is a section

$$\omega : U \to \coprod_{a \in U} T_a^{(k)}$$

where $\omega(a) \in T_a^{(k)}$ (that is, a differential k-form is a section of k^{th}-exterior power of the cotangent bundle). A differential 1-form ω over U is of type (1,0) if for all $a \in U$, $\omega(a) \in T_a^{1,0}$. Similarly for differential 1-form of type $(0, 1)$.

Using differential forms, we can define differentiable, holomorphic and meromorphic 1-forms. Before that we quickly define Laurent expansion and residuce of holomorphic maps.

Remark 11.4.1.11 (Laurent expansion & residue)**.** Let X be a Riemann surface and $a, b \in X$. Let (U, z) be a chart containing a, where we may assume $z(a) = 0$. Let $f \in \mathcal{O}_X(U \setminus \{a\})$. Then $f \circ z^{-1} : z(U) \setminus \{0\} \to \mathbb{C}$ is holomorphic. Thus, around point $0 \in z(U)$, there is a Laurent series representation of $f \circ z^{-1}$:

$$(f \circ z^{-1})(x) = \sum_{n=-\infty}^{\infty} c_n x^n,$$

which we may then write in terms of coordinates z as

$$f(z) = \sum_{n=-\infty}^{\infty} c_n z^n.$$

Thus, f has a removable singularity or pole of order k at a if and only if so does $f \circ z^{-1}$ at $z(a) = 0$.

Let $\omega = f dz \in \Omega^1_X(U \setminus \{a\})$ be a holomorphic 1-form. Then $f = \sum_n a_n z^n$, we define residue of f at a as $\mathrm{res}_a f = c_{-1}$.

Definition 11.4.1.12 (**Differentiable, holomorphic and meromorphic 1-forms**)**.** Let X be a Riemann surface and and $U \subseteq X$ open. Let (U_i, z_i) be any chart. A differential 1-form ω is said to be:

1. differentiable if on $U \cap U_i$ we have
$$\omega = f dz + g d\bar{z}$$
where $f, g \in \mathcal{E}_X(U \cap U_i)$, denote $\omega \in \mathcal{E}^{(1)}_X(U)$. If $\omega = f dz$, then we say that ω is a differentiable 1-form of type (1,0), denoted $\omega \in \mathcal{E}^{1,0}_X(U)$. Similarly, if $\omega = g d\bar{z}$, then ω is a differentiable 1-form of type (0,1), denoted $\omega \in \mathcal{E}^{0,1}_X(U)$;
2. holomorphic if on $U \cap U_i$ we have
$$\omega = f dz$$
where $f \in \mathcal{O}_X(U \cap U_i)$, denote $\omega \in \Omega^{(1)}_X(U)$,
3. meromorphic if there exists an open subseteq $V \subseteq U$ such that ω on V is a holomorphic 1-form, $U \setminus V$ contains isolated points and ω has a pole at each point in $U \setminus V$. Denote $\omega \in \mathcal{M}^{(1)}_X(U)$.

One can also define a differential 2-form ω to be differentiable if $\omega = f dz \wedge d\bar{z}$ where $f \in \mathcal{E}(U \cap U_i)$. Differentiable 2-forms on $U \subseteq X$ are denoted $\mathcal{E}^{(2)}_X(U)$. Note in all of the above, say in differentiable 2-forms, when we wrote $\omega = f dz \wedge d\bar{z}$, we meant that for any $a \in U$, we have

$$\omega(a) = f(a) d_a z \wedge d_a \bar{z} \in T^{(2)}_a = T^{(1)}_a \wedge T^{(1)}_a.$$

Finally, all $\mathcal{E}^{(1)}_X$, $\mathcal{E}^{(2)}_X$, $\Omega^{(1)}_X$ and $\mathcal{M}^{(1)}_X$ are sheaves of $\mathbb{C}$-vector spaces. One also calls $\mathcal{M}^{(1)}_X$ sheaf of abelian differentials.

Construction 11.4.1.13 (Exterior derivative)**.** We now construct the following two maps:

$$\mathcal{E}_X \xrightarrow{d} \mathcal{E}_X^{(1)} \xrightarrow{d} \mathcal{E}_X^{(2)}.$$

Indeed, on an open set $U \subseteq X$, define

$$\begin{aligned} d : \mathcal{E}_X(U) &\longrightarrow \mathcal{E}_X^{(1)}(U) \\ f &\longmapsto df \end{aligned}$$

where $df : U \to \coprod_{a\in U} T_a^{(1)}$ is given by $a \mapsto d_a f$. Next, define for

$$\begin{aligned} d : \mathcal{E}_X^{(1)}(U) &\longrightarrow \mathcal{E}_X^{(2)}(U) \\ \omega &\longmapsto d\omega \end{aligned}$$

where if $\omega = \sum_k f_k dg_k$ for $f_k, g_k \in \mathcal{E}_X(U \cap U_i)$ for some chart (U_i, z_i), then $d\omega$ is defined as

$$d\omega = \sum_k df_k \wedge dg_k.$$

Definining for any $f \in \mathcal{E}_X(U)$ elements $d'f \in \mathcal{E}_X^{1,0}(U)$ given by $a \mapsto d'_a f$ and $d''f \in \mathcal{E}_X^{0,1}(U)$ given by $a \mapsto d''_a f$ and similarly the maps $d', d'' : \mathcal{E}_X^{(1)} \to \mathcal{E}_X^{(2)}$, we thus have the following two chains as well:

$$\mathcal{E}_X \xrightarrow{d'} \mathcal{E}_X^{(1)} \xrightarrow{d'} \mathcal{E}_X^{(2)}$$

and

$$\mathcal{E}_X \xrightarrow{d''} \mathcal{E}_X^{(1)} \xrightarrow{d''} \mathcal{E}_X^{(2)}.$$

11.4.2 Dolbeault's lemma

Theorem 11.4.2.1 (Dolbeault's lemma for $\mathbb{C}$)**.** *Let $X = \{z \in \mathbb{C} \mid |z| < R\}$ for $0 < R \leq \infty$. If $f : X \to \mathbb{C}$ is differentiable, then there exists $g : X \to \mathbb{C}$ differentiable such that*

$$\frac{\partial g}{\partial \bar{z}} = f.$$

Theorem 11.4.2.2 (Dolbeault's lemma for Riemann surfaces)**.** *Let X be a Riemann surface and $U \subseteq X$ be an open set. Then for any $f \in \mathcal{E}_X(U)$, there exists $g \in \mathcal{E}_X(U)$ such that*

$$\frac{\partial g}{\partial \bar{z}} = f.$$

Proof. Let f be as above. Pick any chart (U_i, z_i) of X. Then $f_i := f \circ z_i^{-1} : z_i(U \cap U_i) \to \mathbb{C}$ is a differentiable map where we may assume $z_i(U \cap U_i)$ to be an open disc by considering finer charts. By Dolbeault's lemma for $\mathbb{C}$ (Theorem 11.4.2.1), we get that there exists differentiable $g_i : z_i(U \cap U_i) \to \mathbb{C}$ such that $\frac{\partial g_i}{\partial \bar{z}} = f_i$. Thus, we get a differentiable $g : U \to \mathbb{C}$ which on chart (U_i, z_i) is given by $g \circ z_i^{-1} = g_i$ so that $\frac{\partial g}{\partial \bar{z}} \circ z_i^{-1} = \frac{\partial}{\partial \bar{z}}(g \circ z_i^{-1}) = \frac{\partial g_i}{\partial \bar{z}} = f_i$. Thus, $\frac{\partial g}{\partial \bar{z}}$ agrees with f on each chart, hence $\frac{\partial g}{\partial \bar{z}} = f$. $\square$

11.5 Riemann-Roch theorem

Our goal is to prove and showcase the uses of the following theorem.

Theorem 11.5.0.1 (Riemann-Roch theorem)**.** *Let X be a compact Riemann surface of genus g and let D be a divisor on X. Then:*

1. *The cohomology groups $H^0(X, \mathcal{O}(D))$ and $H^1(X, \mathcal{O}(D))$ are finite-dimensional $\mathbb{C}$-vector spaces.*
2. *The dimensions of the 0^{th} and 1^{st} cohomology groups satisfy*

$$\dim_{\mathbb{C}} H^0(X, \mathcal{O}(D)) - \dim_{\mathbb{C}} H^1(X, \mathcal{O}(D)) = 1 - g + \deg D.$$

Remark 11.5.0.2. Note that $H^0(X, \mathcal{O}(D)) = \Gamma(X, \mathcal{O}(D))$, so one can interpret $\dim_{\mathbb{C}} H^1(X, \mathcal{O}(D))$ as the correction term to the inequality $\dim_{\mathbb{C}} \Gamma(X, \mathcal{O}(D)) \geq 1 - g + \deg D$ so that it becomes an equality.

In the process of proving the above statement, we have to understand the following notions on a Riemann surface : cohomology of sheaves, divisors and genus. We undertake the last two, as we have covered cohomology of sheaves in detail in Chapter FoG.27. However, as we need some results on cohomology of sheaves of differentials and some important long exact sequences, we spend a section setting up the results which we will use.

11.5.1 Cohomology

Recall from §FoG.27.7 that for a given space X and a sheaf $\mathcal{F}$ on X, we can define its Čech-cohomology groups $H^i(X, \mathcal{F})$. For us the most important is the first cohomology, as is evident in Theorem 11.5.0.1.

Remark 11.5.1.1. We first recollect the sheaves that we have so far constructed on any Riemann surface X.

1. $\mathcal{O}_X$ of holomorphic maps on X.
2. $\mathcal{M}_X$ of meromorphic maps on X.
3. $\mathcal{E}_X$ of differentiable maps on X.
4. $\mathcal{E}_X^{(k)}$ of differentiable k-forms on X, $k = 1, 2$.
5. $\mathcal{E}_X^{1,0}$ and $\mathcal{E}_X^{0,1}$ of differentiable 1-forms of type (1,0) and (0,1), repsectively.
6. $\Omega_X^{(1)}$ of holomorphic 1-forms on X.
7. $\mathcal{M}_X^{(1)}$ of meromorphic 1-forms on X.

Every sheaf from 2-7 is an $\mathcal{O}_X$-module. Using these seven sheaves, we can extract quite a bit of geometric information about Riemann surfaces.

We first explore the many maps that one has amongst the above seven sheaves.

Example 11.5.1.2. Let $(X, \mathcal{O}_X)$ be a Riemann surface. Here are some maps between above sheaves.

1. [Exterior derivative] We have maps

$$\mathcal{E}_X \xrightarrow{d} \mathcal{E}_X^{(1)} \xrightarrow{d} \mathcal{E}_X^{(2)},$$

$$\mathcal{E}_X \xrightarrow{d'} \mathcal{E}_X^{(1)} \xrightarrow{d'} \mathcal{E}_X^{(2)}$$

and

$$\mathcal{E}_X \xrightarrow{d''} \mathcal{E}_X^{(1)} \xrightarrow{d''} \mathcal{E}_X^{(2)}.$$

as constructed in Construction 11.4.1.13.

2. [Natural inclusions] We have following inclusions:

$$\mathcal{O}_X \hookrightarrow \mathcal{E}_X$$
$$\Omega_X^{(1)} \hookrightarrow \mathcal{E}_X^{1,0}$$

3. [Exponential map] Let $\mathcal{O}_X^\times$ be a sheaf of abelian groups obtained as follows. For each open $U \subseteq X$, define

$$\mathcal{O}_X^\times(U) := \{f \in \mathcal{O}_X(U) \mid f : U \to \mathbb{C}^\times\}.$$

That is, $\mathcal{O}_X^\times$ is the multiplicative abelian group of units of sheaf of $\mathbb{C}$-algebra $\mathcal{O}_X$. We then define the following map

$$\exp : \mathcal{O}_X \longrightarrow \mathcal{O}_X^\times$$

which on an open $U \subseteq X$ is

$$\begin{aligned}\exp_U : \mathcal{O}_X(U) &\longrightarrow \mathcal{O}_X^\times(U)\\ f &\longmapsto e^{2\pi i f}.\end{aligned}$$

This is clearly a map of sheaves. This is called the exponential map and it plays an important role in geometry.

We have an example of a situation where the image presheaf is not a sheaf (hence justifies why we need to sheafify to get the image preseheaf).

Example 11.5.1.3 (Image presheaf may not be a sheaf)**.** For $X = \mathbb{C}$, consider the open cover $U = \mathbb{C} \setminus (-\infty, 0]$ and $V = \mathbb{C} \setminus [0, \infty)$. Consider the image presheaf of the exponential map $\exp : \mathcal{O}_X \to \mathcal{O}_X^\times$, denoted F. Let $\mathrm{id} \in \mathcal{O}_X^\times(U)$ and $\mathrm{id} \in \mathcal{O}_X^\times(V)$. Observe that they agree on intersection $U \cap V$. Observe further that U and V are simply connected, therefore they have an analytic branch of log, that is, $\mathrm{id} \in \mathrm{Im}\,(\exp_U), \mathrm{Im}\,(\exp_V)$. We claim that there is no section in $F(U \cup V)$ whose restriction to U and V are id. Indeed, since $U \cup V = \mathbb{C}^\times$, therefore if the above two sections glue, then we will have an analytic branch of log on $\mathbb{C}^\times$, not possible.

The cohomology long exact sequence

We now state the main theorem, after proving two lemmas which are nice exercises in general sheaf theory.

Recall that a sheaf map is injective, surjective, if it is so at the level of stalks.

Lemma 11.5.1.4. [4] *Let $\varphi : \mathcal{F} \to \mathcal{G}$ be an injective map of sheaves of abelian groups. Then $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective.*

[4] To remove before addition to FoG and add relevant reference.

Proof. Indeed, if $\varphi_U(f) = 0$ for some $f \in \mathcal{F}(U)$, then at stalks at each point $x \in U$, we get $\varphi_x((U, f)_x) = (U, \varphi_U(f))_x = 0$. Thus, by injectivity, $(U, f)_x = 0$ in $\mathcal{F}_x$ for all $x \in U$. Consequently, f is locally zero at each point of U, thus it is zero at each point of U. □

Lemma 11.5.1.5. [5] *Let X be a space and $0 \to \mathcal{F} \xrightarrow{\varphi} \mathcal{G} \xrightarrow{\psi} \mathcal{H}$ be an exact sequence of sheaves of abelian groups on X. If $U \subseteq X$ is any open, then*

$$0 \to \mathcal{F}(U) \xrightarrow{\varphi_U} \mathcal{G}(U) \xrightarrow{\psi_U} \mathcal{H}(U)$$

is exact.

Proof. By Lemma 11.5.1.4, φ_U is injective. We first show that $\mathrm{Ker}\,(\psi_U) \supseteq \mathrm{Im}\,(\varphi_U)$. Pick any $f \in \mathcal{F}(U)$. Need only show that $\psi_U(\varphi_U(f)) = 0$. Indeed, it suffices to show that $(U, \psi_U(\varphi_U(f)))_x = 0$ in $\mathcal{H}_x$ for all $x \in U$. Pick any $x \in U$. Observe that $(U, \psi_U(\varphi_U(f)))_x = \psi_x \circ \phi_x((U, f)_x)$. The latter is zero by exactness, as needed.

Next, we wish to show that $\mathrm{Ker}\,(\psi_U) \subseteq \mathrm{Im}\,(\varphi_U)$. Indeed, pick $g \in \mathrm{Ker}\,(\psi_U)$ and consider the germ $(U, g)_x \in \mathcal{G}_x$ for any $x \in U$. Observe that $(U, g)_x \in \mathrm{Ker}\,(\psi_x) = \mathrm{Im}\,(\varphi_x)$. Thus, there exists $(V_x, f_x)_x \in \mathcal{F}_x$ such that

$$\varphi_x\left((V_x, f_x)_x\right) = (V_x, \varphi_{V_x}(f_x))_x = (U, g)_x.$$

By definition of germs, we may assume that $\varphi_{V_x}(f_x) = g$ on $V_x \subseteq U$ for all $x \in U$. Hence we have an open covering $\{V_x\}_{x \in U}$ of U and $\varphi_{V_x}(f_x) = g$ on $\mathcal{G}(V_x)$. We claim that $\{f_x\}_{x \in U}$ can be glued. To this end, we wish to show that on $V_x \cap V_y$, we have an equality $f_x = f_y$ in $\mathcal{F}(V_x \cap V_y)$. By Lemma 11.5.1.4, it suffices to show that $\varphi_{V_x \cap V_y}(f_x) = \varphi_{V_x \cap V_y}(f_y)$ in $\mathcal{G}(V_x \cap V_y)$. Observe that the element $\varphi_{V_x \cap V_y}(f_x) = g|_{V_x \cap V_y} = \varphi_{V_x \cap V_y}(f_y)$ in $\mathcal{G}(V_x \cap V_y)$. Thus, we have the required equality.

It follows that $\{f_x\}_{x \in U}$ can be glued to $f \in \mathcal{F}(U)$ such that $\varphi_U(f)$ in $\mathcal{G}(U)$ is such that its restriction to each V_x is g, thus by sheaf axioms, $\varphi_U(f) = g$, that is, $g \in \mathrm{Im}\,(\varphi_U)$, as needed. □

Remark 11.5.1.6 (Surjective maps of sheaves)**.** Recall that if $\varphi : \mathcal{F} \to \mathcal{G}$ is surjective on sections, then it is a surjective map, but the converse is not true. Indeed for $X = \mathbb{C}$, the map of sheaves $\exp : \mathcal{O}_X \to \mathcal{O}_X^\times$ is surjective as any germ in the latter locally has a logarithm, but $\exp_{\mathbb{C}^\times}$ is not surjective on sections as the constant map $\mathrm{id} \in \mathcal{O}_X^\times(\mathbb{C}^\times)$ does not have a logarithm.

However, we do have the following "local surjectivity": φ is surjective if and only if for any open $U \subseteq X$ and any $s \in \mathcal{G}(U)$, there exists an open cover $\{U_i\}_{i \in I}$ of U and $t_i \in \mathcal{F}(U_i)$ such that $\varphi_{U_i}(t_i) = s|_{U_i}$.

Moreover, some of the above sheaves are obtained by kernels and gives us several short exact sequences, which will be used later.

Example 11.5.1.7. [Important short exact sequences] Let $(X, \mathcal{O}_X)$ be a Riemann surface. Some of the sheaves in Remark 11.5.1.1 are kernels of some other map of sheaves and they give rise to some important short exact sequences.

1. The sheaf of holomorphic maps $\mathcal{O}_X$ is obtained as the kernel

$$\mathcal{O}_X = \mathrm{Ker}\left(d'' : \mathcal{E}_X \to \mathcal{E}_X^{0,1}\right).$$

[5] To remove before addition to FoG and add relevant reference.

Thus, we have a s.e.s.

$$0 \longrightarrow \mathcal{O}_X \longrightarrow \mathcal{E}_X \xrightarrow{d''} \mathcal{E}_X^{0,1} \longrightarrow 0$$

where d'' is surjective by Dolbeault's lemma (Theorem 17.3.2.3).

2. The sheaf of holomorphic 1-forms $\Omega_X^{(1)}$ is obtained from the kernel

$$\Omega_X^{(1)} = \mathrm{Ker}\left(d : \mathcal{E}_X^{1,0} \to \mathcal{E}_X^{(2)}\right).$$

Thus we have a s.e.s.

$$0 \longrightarrow \Omega_X^{(1)} \longrightarrow \mathcal{E}_X^{1,0} \xrightarrow{d} \mathcal{E}_X^{(2)} \longrightarrow 0$$

where d is surjective as follows. For any $\omega = f dz \wedge d\bar{z}$ in $\mathcal{E}_X^{(2)}(U \cap U_i)$ for (U_i, z_i) a chart and $f \in \mathcal{E}_X(U \cap U_i)$, we get by Dolbeault's lemma (Theorem 17.3.2.3) that there exists $g \in \mathcal{E}_X(U \cap U_i)$ such that $\frac{\partial g}{\partial \bar{z}} = f$. Thus, $d(-g dz) = -\left(\frac{\partial g}{\partial z} dz + \frac{\partial g}{\partial \bar{z}} d\bar{z}\right) \wedge dz = -\frac{\partial g}{\partial \bar{z}} d\bar{z} \wedge dz = f dz \wedge d\bar{z}$. This shows that d is surjective on sections, as required.

3. Let $\mathcal{L}_X = \mathrm{Ker}\left(d : \mathcal{E}_X^{(1)} \to \mathcal{E}_X^{(2)}\right)$ be the sheaf of closed 1-forms. Then we claim that the following is a s.e.s.

$$0 \longrightarrow \underline{\mathbb{C}} \longrightarrow \mathcal{E}_X \xrightarrow{d} \mathcal{L}_X \longrightarrow 0.$$

Indeed, $\mathrm{Ker}\,(d : \mathcal{E}_X \to \mathcal{L}_X)$ is given on an open-connected $U \subseteq X$ by those differentiable maps $f : U \to \mathbb{C}$ such that $df = d' f dz + d'' f d\bar{z} = 0$, that is, $\partial f / \partial z = 0$ and $\partial f / \partial \bar{z} = 0$ on U. It follows that f is holomorphic with zero derivative, that is, f is constant (U is connected). Hence, we get the inclusion $\underline{\mathbb{C}} \hookrightarrow \mathcal{E}_X$ whose image is $\mathrm{Ker}\,(d)$.
Now, d is surjective as locally any closed form is exact by local existence of primitives from one variable complex analysis.

4. For the exponential map, observe that we have a map $\underline{\mathbb{Z}} \to \mathcal{O}_X$ which on an open-connected $U \subseteq X$ is given by

$$\begin{aligned} \mathbb{Z} = \underline{\mathbb{Z}}(U) &\longrightarrow \mathcal{O}_X(U) \\ c_n &\longmapsto c_n. \end{aligned}$$

For some arbitrary open set $U \subseteq X$, $\underline{\mathbb{Z}}(U)$ is given by functions which are constant on each open connected component (any Riemann surface is locally connected), so they are in particular also holomorphic. We thus get a s.e.s.

$$0 \longrightarrow \underline{\mathbb{Z}} \longrightarrow \mathcal{O}_X \xrightarrow{\exp} \mathcal{O}_X^\times \to 0$$

where exp is surjective as locally any non-zero holomorphic map has an analytic branch of log.

We will now discuss the map in cohomology induced by a map of sheaves and how the connecting homomorphism works.

Construction 11.5.1.8 (Map in cohomology). [6] Any map of abelian sheaves over X yields a map in the cohomology as well. Indeed, let $\varphi : \mathcal{F} \to \mathcal{G}$ be a map of sheaves. Then we get a map

$$\begin{aligned} \varphi^q : C^q(\mathcal{U}, \mathcal{F}) &\longrightarrow C^q(\mathcal{U}, \mathcal{G}) \\ s = (s(\alpha_0, \dots, \alpha_q)) &\longmapsto \varphi^q(s) = \left(\varphi_{\alpha_0 \dots \alpha_q}(s(\alpha_0, \dots, \alpha_q))\right) \end{aligned}$$

where $\varphi_{\alpha_0 \dots \alpha_q} = \varphi_{U_{\alpha_0} \cap \dots \cap U_{\alpha_q}}$.

It then follows quite immediately from the fact that each $\varphi_{\alpha_0 \dots \alpha_q}$ is a group homomorphism that $d\varphi^q = \varphi^{q+1} d$. It follows that we get a map of chain complexes

$$\varphi^\bullet : C^\bullet(\mathcal{U}, \mathcal{F}) \longrightarrow C^\bullet(\mathcal{U}, \mathcal{G}).$$

Hence, we get a map in cohomology

$$\varphi^q : H^q(\mathcal{U}, \mathcal{F}) \longrightarrow H^q(\mathcal{U}, \mathcal{G}).$$

Finally, this gives by universal property of direct limits a unique map

$$\varphi^q : \check{H}^q(X, \mathcal{F}) \longrightarrow \check{H}^q(X, \mathcal{G})$$

such that for every open cover $\mathcal{U}$, the following diagram commutes:

$$\begin{tikzcd} \check{H}^q(X, \mathcal{F}) \arrow[r, dashed, "\varphi^q"] & \check{H}^q(X, \mathcal{G}) \\ H^q(\mathcal{U}, \mathcal{F}) \arrow[u] \arrow[r, "\varphi^q"'] & H^q(\mathcal{U}, \mathcal{G}) \arrow[u] \end{tikzcd}$$

where vertical maps are the maps into direct limits.

Construction 11.5.1.9 (Connecting homomorphism). [7] Let X be a topological space and

$$0 \longrightarrow \mathcal{F} \xrightarrow{\varphi} \mathcal{G} \xrightarrow{\psi} \mathcal{H} \longrightarrow 0$$

be an exact sequence of sheaves on X. We define the connecting homomorphism

$$H^0(X, \mathcal{H}) \xrightarrow{\delta} H^1(X, \mathcal{F})$$

as follows. First, pick any $h \in H^0(X, \mathcal{H}) = \Gamma(X, \mathcal{H})$. As ψ is surjective therefore there exists an open covering $\mathcal{U} = \{U_i\}_{i \in I}$ of X and $g_i \in \mathcal{G}(U_i)$ such that $\psi_{U_i}(g_i) = h|_{U_i}$. Using (g_i) and (U_i) we construct a 1-cocycle for $\mathcal{F}$ as follows. Observe that for each $i, j \in I$, we have $\psi_{U_i \cap U_j}(g_i - g_j) = 0$ in $\mathcal{H}(U_i \cap U_j)$. Thus, $g_i - g_j \in \operatorname{Ker}\left(\psi_{U_i \cap U_j}\right)$. By exactness guaranteed by Lemma 11.5.1.5, it follows that there exists $f_{\alpha_0 \alpha_1} \in \mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1})$ such that $\varphi_{U_{\alpha_0} \cap U_{\alpha_0}}(f_{\alpha_0 \alpha_1}) = g_{\alpha_0} - g_{\alpha_1}$, for each $\alpha_0, \alpha_1 \in I$. We claim that the element

$$f := (f_{\alpha_0 \alpha_1})_{\alpha_0, \alpha_1} \in \prod_{(\alpha_0, \alpha_1) \in I^2} \mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1}) = C^1(\mathcal{U}, \mathcal{F})$$

[6] To remove before addition to FoG and add relevant reference.

[7] To remove before addition to FoG and add relevant reference.

is a 1-cocycle. Indeed, we need only check that $df = 0$ in $C^2(\mathcal{U}, \mathcal{F})$. Pick any $(\alpha_0, \alpha_1\alpha_2) \in I^3$. We wish to show that $df(\alpha_0, \alpha_1\alpha_2) = 0$. Indeed,

$$\begin{aligned} df(\alpha_0, \alpha_1\alpha_2) &= \sum_{j=0}^{2} (-1)^j \rho_j \left(f_{\alpha_0 \hat{\alpha}_j \alpha_2}\right) \\ &= f_{\alpha_1\alpha_2} - f_{\alpha_0\alpha_2} + f_{\alpha_0\alpha_1} \end{aligned}$$

in $\mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1} \cap U_{\alpha_2})$. We claim the above is zero. Indeed, By Lemma 11.5.1.5 on $V := U_{\alpha_0} \cap U_{\alpha_1} \cap U_{\alpha_2}$ we get that φ_V is injective. But since

$$\begin{aligned} \varphi_V(f_{\alpha_1\alpha_2} - f_{\alpha_0\alpha_2} + f_{\alpha_0\alpha_1}) &= \varphi_V(f_{\alpha_1\alpha_2}) - \varphi_V(f_{\alpha_0\alpha_2}) + \varphi_V(f_{\alpha_0\alpha_1}) \\ &= g_{\alpha_1} - g_{\alpha_2} - (g_{\alpha_0} - g_{\alpha_2}) + g_{\alpha_0} - g_{\alpha_1} \\ &= 0, \end{aligned}$$

hence it follows that $df(\alpha_0, \alpha_1\alpha_2) = 0$, as required. Hence $f \in C^1(\mathcal{U}, \mathcal{F})$ is a 1-cocycle. Thus we get an element $[f] \in H^1(\mathcal{U}, \mathcal{F})$. This defines a group homomorphism $H^0(X, \mathcal{H}) \to H^1(\mathcal{U}, \mathcal{F})$. Further by passing to direct limit, we get an element $[f] \in H^1(X, \mathcal{F})$. We thus define

$$\delta(f) := [f] \in H^1(X, \mathcal{F}).$$

This defines the required group homomorphism δ.

Theorem 11.5.1.10 (Long exact cohomology sequence)**.** [8] *Let X be a topological space and*

$$0 \longrightarrow \mathcal{F} \xrightarrow{\varphi} \mathcal{G} \xrightarrow{\psi} \mathcal{H} \longrightarrow 0$$

be an exact sequence of sheaves on X. Then there exists a long exact sequence

$$\begin{tikzcd} 0 \arrow[r] & H^0(X,\mathcal{F}) \arrow[r, "\varphi^0"] & H^0(X,\mathcal{G}) \arrow[r, "\psi^0"] & H^0(X,\mathcal{H}) \arrow[dll, "\delta"'] \\ & H^1(X,\mathcal{F}) \arrow[r, "\varphi^1"'] & H^1(X,\mathcal{G}) \arrow[r, "\psi^1"'] & H^1(X,\mathcal{H}) \end{tikzcd}$$

where δ is as in Construction 11.5.1.9.

Applications

We now state and prove three big results, which follows from cohomology l.e.s. quite naturally.

The first result is an immediate corollary of the cohomology l.e.s. good to get the muscles moving, which states what happens when the middle sheaf has no first cohomology.

Proposition 11.5.1.11. *Let X be a topological space and*

$$0 \longrightarrow \mathcal{F} \xrightarrow{\varphi} \mathcal{G} \xrightarrow{\psi} \mathcal{H} \longrightarrow 0$$

be an exact sequence of sheaves on X. If $H^1(X, \mathcal{G}) = 0$, then

$$H^1(X, \mathcal{F}) \cong \frac{\Gamma(X, \mathcal{H})}{\psi_X(\Gamma(X, \mathcal{G}))}$$

[8] To remove before addition to FoG and add relevant reference.

Proof. Write the cohomology l.e.s. (Theorem 11.5.1.10) and use first isomorphism theorem. □

We next give Dolbeault's theorem which for a Riemann surface calculates first cohomology of structure sheaf and holomorphic 1-forms purely in terms of differentiable functions and differentiable 1 and 2-forms. This is essentially already clear from the first two s.e.s. in Example 11.5.1.7 and above result.

Theorem 11.5.1.12 (Dolbeault's theorem)**.** *Let X be a Riemann surface.*

1. *All sheaves $\mathcal{E}_X, \mathcal{E}_X^{1,0}, \mathcal{E}_X^{0,1}, \mathcal{E}_X^{(1)}$ and $\mathcal{E}_X^{(2)}$ have first cohomology group 0.*
2. *We have isomorphisms*

$$H^1(X, \mathcal{O}_X) \cong \frac{\Gamma(X, \mathcal{E}_X^{0,1})}{d''_X(\Gamma(X, \mathcal{E}_X))},$$
$$H^1(X, \Omega_X^{(1)}) \cong \frac{\Gamma(X, \mathcal{E}_X^{(2)})}{d_X(\Gamma(X, \mathcal{E}_X^{1,0}))}.$$

Proof. We omit the proof of item 1, for it can be found in Forster, Theorem 12.6 cite[Forster].

By Proposition 11.5.1.11 and the first two s.e.s. in Example 11.5.1.7, we need only show that $H^1(X, \mathcal{E}_X) = 0 = H^1(X, \mathcal{E}_X^{1,0})$, which we know to be true from item 1. □

Corollary 11.5.1.13. *Let $X = B_R(0) \subseteq \mathbb{C}$ be an open ball considered as a Riemann surface. Then $H^1(X, \mathcal{O}_X) = 0$.*

Proof. Need only show that $d''_X : \Gamma(X, \mathcal{E}_X) \to \Gamma(X, \mathcal{E}_X^{0,1})$ is surjective. That is, for any differentiable 1-form of type $(0,1)$, i.e. $\omega = f d\bar{z}$ on X, we wish to find a differentiable map g on X such that $d''g = \frac{\partial g}{\partial \bar{z}} d\bar{z} = f d\bar{z}$. Indeed, by Dolbeault's lemma for $\mathbb{C}$ (Theorem 11.4.2.1), we get such a g. □

Remark 11.5.1.14 (deRham cohomology)**.** Let X be a Riemann surface and denote $H^1_{\mathrm{dR}}(X)$ the deRham cohomology of X, that is,

$$H^1_{\mathrm{dR}}(X) = \frac{\operatorname{Ker}\left(d_X : \mathcal{E}_X^{(1)}(X) \to \mathcal{E}_X^{(2)}(X)\right)}{\operatorname{Im}\left(d_X : \mathcal{E}_X(X) \to \mathcal{E}_X^{(1)}(X)\right)}.$$

We can now easily see by cohomology l.e.s. that $H^1_{\mathrm{dR}}(X)$ is same as $H^1(X, \underline{\mathbb{C}})$.

Theorem 11.5.1.15 (deRham isomorphism for Riemann surfaces)**.** *Let X be a Riemann surface and $\underline{\mathbb{C}}$ be the constant sheaf associated to field $\mathbb{C}$. Then we have an isomorphism*

$$H^1_{\mathrm{dR}}(X) \cong H^1(X, \underline{\mathbb{C}}).$$

Proof. By Example 11.5.1.7, 3, we have a s.e.s.

$$0 \longrightarrow \underline{\mathbb{C}} \longrightarrow \mathcal{E}_X \xrightarrow{d} \mathcal{L}_X \longrightarrow 0.$$

where $H^1(X, \mathcal{E}_X) = 0$ by Dolbeault's theorem (Theorem 11.5.1.12). By Proposition 11.5.1.11, it follows that

$$H^1(X, \underline{\mathbb{C}}) \cong \frac{\Gamma(X, \mathcal{L}_X)}{d_X \Gamma(X, \mathcal{E}_X)}$$

where $\Gamma(X, \mathcal{L}_X)$ is the set of all global closed differentiable 1-forms and $d_X\Gamma(X, \mathcal{E}_X)$ is the image of all differentiable functions, that is $\frac{\Gamma(X, \mathcal{L}_X)}{d_X\Gamma(X, \mathcal{E}_X)} =: H^1_{\mathrm{dR}}(X)$, as required. □

11.5.2 Divisors

11.5.3 Proof of Riemann-Roch theorem

11.5.4 Applications

Chapter 12

Foundational Analytic Geometry

Part V

The Categorical Viewpoint

Chapter 13

Classical Topoi

Contents

13.1 Towards the axioms of a Topos

Consider the categories **Sets**, **Sets**^{n}, **G-Sets**, **Sets**$^{\mathbf{N}}$ for **N** being the totally ordered $\mathbb{N}$ regarded as a category and the slice category **Sets**$/S$ for some set S . These all, as we will see later, are important examples of topoi. But the more interesting fact arises from the realization that all of the above examples except the last one, can be shown to be a special case of another important example of a topos, the presheaf category:

$$\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}.$$

For example, we can see that if we fix **C** as the n object discrete category, then we get back the category **Sets**^{n}. The arrows in **Sets**^{n} is just the arrow (natural transf.) in the functor category **Sets**$^{\mathbf{C}^{\mathrm{op}}}$. To see this, consider two presheaves $F, G : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$. A natural transformation $\eta : F \longrightarrow G$ would take some object $*_i$ in **C** to the component $\eta *_i : F*_i \to G*_i$ for $0 \leq i \leq n$. Hence we have n arrows between two n-tuple of sets, $(F*_1, \ldots, F*_n)$ and $(G*_1, \ldots, G*_n)$. This is exactly the arrow in **Sets**^{n}, so that we have established an isomorphism of categories. Similar constructions for the other examples realize the importance of the presheaf category in this aspect. However, the slice category **Sets**$/S$ is different in the sense that it is *almost* a presheaf category. *Almost* because one can only construct the presheaf category **Sets**^{S} of S indexed family of sets which will not be isomorphic to the slice, but equivalent only upto a natural isomorphism.

Next thing we notice about the above examples is that they all have finite limits. Recall that any category have finite limits if and only if it has terminal object and all pullbacks. Now one may see the importance of having pullbacks by the following trivial example. Suppose we are working in the **Sets** and we have $f : X \to B$ and $g : Y \hookrightarrow B$ for $Y \subseteq B$, the latter of which is clearly a monic. Then the pullback of g along f would be the set P characterized as:

$$P = \{(x, y) \in X \times Y \ : \ f(x) = g(y) = y\}.$$

Clearly, we have that $P \cong (f)^{-1}(Y)$. That is, we have a set P which is same as the inverse image of f for some subset Y of it's co-domain. We can hence quite easily turn this idea into any category in order to find a generalized notion of *inverse* of an arrow.

The most immediate use, however, of existence of terminal object and pullbacks is in the generalization of the fact that in **Sets**, any subset $S \subset X$ can be written either as a monic $m : S \rightarrowtail X$

or as a characteristic function $\chi_S : X \to \mathbf{2}$ where $\mathbf{2} = \{0, 1\}$ is a 2 object set which maps as following:

$$\chi_S(x) = \begin{cases} 0 & \text{if } x \in S \\ 1 & \text{if } x \notin S \end{cases}$$

where we regard $\mathbf{2} = \{0, 1\}$ as the **Truth** object, which contains the possible truth values of the category of sets, which clearly exists and has two elements. This latter convention for a subset of a set X can be generalized to an arbitrary category by the following definition:

Definition 13.1.0.1. (**Subobject Classifier**) *Suppose* $\mathbf{C}$ *has all finite limits. A subobject classifier is a monic global element*

$$\mathsf{true} : \mathbf{1} \rightarrowtail \Omega$$

where Ω *is the truth object, such that for any subobject* $m : S \rightarrowtail X$*, there exists a unique arrow* $\phi : X \to \Omega$ *such that the following forms a Pullback Square:*

$$\begin{tikzcd} S \arrow[r] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, tail, "\mathsf{true}"] \\ X \arrow[r, dashed, "\phi"'] & \Omega \end{tikzcd}$$

We can in-fact turn the whole construction of a Subobject to a Presheaf itself!

Definition 13.1.0.2. (**Subobject Functor**) *In a well-powered*[1] *category* $\mathbf{C}$*, we have the following presheaf called the subobject functor*

$$\mathrm{Sub}_{\mathbf{C}}(-) : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$$

which takes:

1. ***Object*** X *to the small set* $\mathrm{Sub}_{\mathbf{C}}(X)$,
2. ***Arrow*** $f : Y \to X$ *to the arrow:*

$$\mathrm{Sub}_{\mathbf{C}}(f) : \mathrm{Sub}_{\mathbf{C}}(X) \to \mathrm{Sub}_{\mathbf{C}}(Y)$$

which takes any subobject $m : S \rightarrowtail X$ *in* $\mathrm{Sub}_{\mathbf{C}}(X)$ *to the pullback of* s *along* f*, i.e.:*

$$\begin{tikzcd} S \arrow[d, tail, "m"'] & \longmapsto & P \arrow[r] \arrow[d, tail, "m'"'] \arrow[dr, phantom, "\lrcorner", very near start] & S \arrow[d, tail, "m"] \\ X & & Y \arrow[r, "f"'] & X \end{tikzcd}$$

In-fact, the functor $\mathrm{Sub}_{\mathbf{C}}(-)$ is a Representable Functor represented exactly by the Truth object Ω!

Proposition 13.1.0.3. *(**Representability of*** $Sub_{\mathbf{C}}(-)$*) Suppose* $\mathbf{C}$ *is a small category with finite limits. Then,* $\mathbf{C}$ *has a subobject classifier if and only if* $\exists$ *an object* Ω *such that* $\exists\ \theta$

$$\theta : Sub_{\mathbf{C}}(-) \Longrightarrow \mathrm{Hom}_{\mathbf{C}}(-, \Omega)$$

which is a natural isomorphism.

[1] A category which has small set of all subobjects for any object.

Proof. (L $\implies$ R) Suppose $\mathbf{C}$ has a subobject classifier. Take any $f : Y \to X$ in $\mathbf{C}$. The $\mathrm{Sub}_{\mathbf{C}}(f)$ identifies an arrow $\theta_X : \mathrm{Sub}_{\mathbf{C}}(X) \longrightarrow \mathrm{Hom}_{\mathbf{C}}(X, \Omega)$, given by $\theta_X(m) =$ characteristic arrow of subobject m. This θ_X can be seen to be natural in X by direct verification.
(R $\implies$ L) Suppose $\theta : \mathrm{Sub}_{\mathbf{C}}(-) \Rightarrow \mathrm{Hom}_{\mathbf{C}}(-, \Omega)$ is a natural isomorphism. Take any monic $m : A \rightarrowtail E$ in $\mathbf{C}$. Consider the unique arrow $\theta_E(m) : A \to \Omega$. We also have by the isomorphism a unique subobject $t : \mathbf{1} \rightarrowtail \Omega$ obtained by such t with $\theta_\Omega(t) = 1_\Omega$. The result then follows by observing the naturality condition on $\theta_X(f)$ for a monic $f : A \rightarrowtail E$. $\square$

13.1.1 Subobject Classifier in $\mathbf{Sets}^{\mathbf{C}^{op}}$

Since we saw earlier that most of the examples of topoi are particular instances of presheaf categories, therefore the construction of a subobject classifier for $\mathbf{Set}^{\mathbf{C}^{op}}$ constructs a subobject classifier for them too.
To begin with, clearly, every monic in $\mathbf{Set}^{\mathbf{C}^{op}}$ is just a natural transformation whose each component is injective. That is, for a monic arrow $\theta : Q \Longrightarrow P$ in the presheaf category, for any object C of $\mathbf{C}$, the component $\theta C : QC \rightarrowtail PC$ is monic, so that $QC \subset PC$. Moreover, by naturality, for any arrow $f : D \to C$ in $\mathbf{C}$, we must have that $Qf = Pf|_{QC \subset PC}$. This exactly is the definition of Q being the subfunctor of P. Therefore, subobjects in presheaf category are subfunctors. With this, we now construct the classifier object Ω.

Suppose Ω in $\mathbf{Sets}^{\mathbf{C}^{op}}$ is the classifier object. Since it classifies all subobjects, therefore it must also classify the contravariant functor $\mathrm{Hom}_{\mathbf{C}}(-, C)$ for any object C of $\mathbf{C}$. But, by Proposition 13.1.0.3, we have that:

$$\begin{aligned} \mathrm{Sub}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C)) &\cong \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C), \Omega) \\ &\cong \Omega C \qquad\qquad \text{By Yoneda Lemma} \end{aligned}$$

where $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{op}}$. Hence, Ω is that object/presheaf in $\widehat{\mathbf{C}}$, which takes object C of $\mathbf{C}$ to the set of all subobjects of the representable functor $\mathrm{Hom}_{\mathbf{C}}(-, C)$ in $\widehat{\mathbf{C}}$. But this set would contain just the subfunctors of $\mathrm{Hom}_{\mathbf{C}}(-, C)$. Now, if we define the following:

Definition 13.1.1.1. (**Sieve over an Object**) *In a category* $\mathbf{C}$, *a sieve on an object* C *is the following set :*

$$S_C = \{f : A \to C \mid \text{ if } h : B \to A, \text{ then } f \circ h \in S_C \text{ for any such } A\}.$$

Then we notice that a sieve on C is actually the same thing as a subfunctor of $\mathrm{Hom}_{\mathbf{C}}(-, C)$. Hence, ΩC is just the set of all sieves over the object C. But what about the action of Ω on the arrows of $\mathbf{C}$?
To answer this, we must remind ourselves first that pullback preserves the subobjects. This means that if $g : B \to C$ in $\mathbf{C}$, then pullback of the subobject identified by the monic $Q \Longrightarrow \mathrm{Hom}_{\mathbf{C}}(-, C)$ along the arrow $\mathbf{Yon}(g) : \mathrm{Hom}_{\mathbf{C}}(-, B) \Longrightarrow \mathrm{Hom}_{\mathbf{C}}(-, C)$ is a subobject of $\mathrm{Hom}_{\mathbf{C}}(-, B)$. This translates to the fact that pullback of a sieve over C, S_C is just the sieve over B, S_B defined by:

$$S_B = S_C \cdot g = \{h \mid g \circ h \in S_C\}.$$

Hence, Ωg for $g : B \to C$ is the set function:

$$\begin{aligned} \Omega g : \Omega C &\longrightarrow \Omega B \\ S_C &\longmapsto S_B = S_C \cdot g. \end{aligned}$$

Hence the monic subobject classifier in $\widehat{\mathbf{C}}$ is simply:

$$\begin{aligned} \mathsf{true} : \mathbf{1} &\Longrightarrow \Omega \\ \mathsf{true}_C : 1C = \{\star\} &\rightarrowtail \Omega C \\ \star &\mapsto \text{Maximal Sieve over } C. \end{aligned}$$

The final piece left to settle is the unique characteristic function $\phi : P \Longrightarrow \Omega$ for a given subobject $Q \Longrightarrow P$. It can be seen that the following choice of ϕ does make the corresponding classifier a pullback diagram:

$$\begin{aligned} \phi : P &\Longrightarrow \Omega \\ \phi_C : PC &\longrightarrow \Omega C \\ p &\mapsto \{f : E \to C \mid (Pf)(p) \in Q(E) \text{ for any such } f : E \to C\} \end{aligned}$$

Note that component ϕ_C maps each element to that sieve which contains those arrows of $\mathbf{C}$ whose image under P takes that element to the subset mapped to by the Q under the domain of that arrow. Clearly, when $p \in QC \subset PC$, then for any $f : E \to C$, $Pf(p) \in QE$ because $Qf = Pf|_{QC \subset PC}$, hence $\phi_C(p)$ is the maximal sieve over C.

13.1.2 Colimits in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$

It turns out that the presheaf category has a peculiar property that each object/presheaf in it is the colimit of some particular diagram of representable presheaves, and that too in the most obvious way. Before stating it precisely, let us look at a very general result, whose corollary gives us the above result.

Definition 13.1.2.1. (**Category of Elements for a Presheaf**) *Suppose $P : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$. Then we can construct a category called the category of elements, $\int_{\mathbf{C}} P$, which has:*

1. ***Objects** as the pairs (C, p) where $C \in \mathrm{Ob}(\mathbf{C})$ and $p \in PC$,*
2. ***Arrows** as $u : (C', p') \to (C, p)$ where u is just $u : C' \to C$ in $\mathbf{C}$ but with property that*

$$Pu(p) = p'.$$

where composition is defined as in $\mathbf{C}$. Also, there clearly exists a functor:

$$\begin{aligned} \pi_P : \int_{\mathbf{C}} P &\longrightarrow \mathbf{C} \\ (C, p) &\longmapsto C \\ \big(u : (C', p') \to (C, p)\big) &\longmapsto \big(u : C' \to C\big) \end{aligned}$$

Theorem 13.1.2.2. *Suppose $\mathbf{C}$ is a small category and $A : \mathbf{C} \longrightarrow \mathbf{E}$ is a functor to a co-complete category $\mathbf{E}$. Then, the following functor:*

$$\begin{aligned} R : \mathbf{E} &\longrightarrow \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \\ E &\longmapsto \mathrm{Hom}_{\mathbf{E}}\left(A(-), E\right) \end{aligned}$$

has a Left Adjoint given by:

$$L : \mathbf{Sets}^{\mathbf{C}^{op}} \longrightarrow \mathbf{E}$$
$$P \longmapsto \varinjlim \left(\int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{A} \mathbf{E} \right)$$

Proof. Suppose E is some object in $\mathbf{E}$. Take any $\lambda : P \Longrightarrow RE$ from $\mathrm{Hom}_{\widehat{\mathbf{C}}}(P, RE)$. For any object C of $\mathbf{C}$, $\lambda C : PC \to \mathrm{Hom}_{\mathbf{E}}(AC, E)$ is therefore just the collection of (C, p) for $p \in PC$. Hence, λC is a subset of $\mathrm{Ob}(\int_{\mathbf{C}} P)$. The projection π_P of (C, p) for any p is simply the object C ov $\mathbf{C}$. Further application of functor A on $\pi_P(C, p)$ would be AC. Similarly, for $u : (C', p') \longrightarrow (C, p)$, the arrow $Au : A\pi_P(C', p') \to A\pi_P(C, p)$ would be such that the following commutes, because λ is a natural transformation:

$$\begin{tikzcd}
A\pi_P(C,p) \arrow[dr, "\lambda C(p)"] & \\
& E \\
A\pi_P(C',p') \arrow[uu, "u"] \arrow[ur, "\lambda C'(p)"'] &
\end{tikzcd}$$

Clearly, this means that E forms a cocone over the diagram $\int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{A} \mathbf{E}$. Therefore, there exists a unique arrow $E \dashrightarrow LP$. This assignment establishes the corresponding natural isomorphism between $\mathrm{Hom}_{\widehat{\mathbf{C}}}(P, RE)$ and $\mathrm{Hom}_{\mathbf{E}}(LP, E)$. □

This Theorem now naturally leads to the result discussed in the beginning:

Proposition 13.1.2.3. *In the presheaf category $\widehat{\mathbf{C}}$, any presheaf P is the colimit of a particular diagram of representable functors, in a canonical way.*

Proof. In the above Theorem 13.1.2.2, if we set $\mathbf{E} = \widehat{\mathbf{C}}$ and $A = \mathbf{Yon}(-)$, we get the desired result, where the diagram is the following:

$$\int_{\mathbf{C}} P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{\mathbf{Yon}(-)} \widehat{\mathbf{C}}$$

of index as category of elements of P. □

13.1.3 Exponentials in $\mathbf{Sets}^{\mathbf{C}^{op}}$

The next important general result about the presheaf category is that every object/presheaf in it is exponentiable. Now since $\widehat{\mathbf{C}}$ is already complete, therefore existence of exponentials would imply that $\widehat{\mathbf{C}}$ is Cartesian Closed! As we will note much later in a bit more detail, this is a general property of any topoi.

Proposition 13.1.3.1. *Suppose $\mathbf{C}$ is a small category, then the presheaf category $\widehat{\mathbf{C}}$ is Cartesian Closed.*

Proof. It can be verified that for any object P, Q in $\widehat{\mathbf{C}}$, the another object Q^P defined by:

$$\begin{aligned} Q^P : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ C &\longmapsto \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C) \times P, Q) = \mathrm{Nat}(\mathrm{Hom}_{\mathbf{C}}(-, C) \times P, Q) \end{aligned}$$

makes Q^P the exponential of P and Q. $\square$

Remark 13.1.3.2. With all the discussion above, we now see that presheaves is an important category and the examples seen in the beginning of topoi follows the following three rules: (1) it has finite (co)limits, (2) each object in it is exponentiable & (3) it has a subobject classifier. In fact, these three exactly constitute the definition of a topos, as we will see later!

13.2 Grothendieck Topologies & Sheaves

Suppose we wish to generalize sheaves on to arbitrary categories. The first problem that one might face to this goal would be that defining a sheaf requires a notion of *cover* of an open subset of the given topological space. Therefore sheaves on arbitrary categories would require a notion of *cover* of each object in that category. This is precisely the problem handled (among many others) by Grothendieck topologies and the notion of a site. Reconciling such topologies on any categories would lead us to a vast generalization of the notion of a *cover* of a given space, but now arbitrary objects, which has very deep connotations within algebraic geometry, which we would discuss after defining them.

13.2.1 Grothendieck Topologies

Definition 13.2.1.1. (**Grothendieck Topology & Site**) *Suppose* $\mathbf{C}$ *is a small category. A* ***Grothendieck topology*** *on* $\mathbf{C}$ *is a functor:*

$$\begin{aligned} J(-) : \mathrm{Ob}(\mathbf{C}) &\longrightarrow \mathbf{Sets} \\ C &\longmapsto JC := \{S \mid S \text{ is a sieve over } C\} \end{aligned}$$

where $\mathrm{Ob}(\mathbf{C})$ *is the discrete category of objects of* $\mathbf{C}$, *such that for any object* C, *the collection* JC *of sieves over* C *in category* $\mathbf{C}$ *must satisfy:*

GT.1 *(**Maximal Cover**) The maximal sieve over* C, S_C^{max}, *must be in* JC.

GT.2 ***(Stability of Covers)*** *For any* $S \in JC$ *and* $f : D \to C$ *in* $\mathbf{C}$, *we must have that:*

$$f^*(S) := \underbrace{\{g \mid f \circ g \in S\}}_{\text{a sieve over } D} \in JD$$

GT.3 *(**Transitivity of Covers**) If* $S \in JC$ *and* R *is any sieve over* C *such that*

$$\forall\, f \in S\,, \quad f^*(R) \in J(\mathrm{dom}(f))$$

then we must have that

$$R \in JC.$$

A ***site*** $(\mathbf{C}, J)$ *is just a small category* $\mathbf{C}$ *coupled with a Grothendieck topology* J *on* $\mathbf{C}$.

Lemma 13.2.1.2. *Suppose* $(\mathbf{C}, J)$ *is a site. Then for any object* C*,*

$$R, S \in JC \implies R \cap S \in JC$$

Proof. Take any $f : D \to C$ and note that $f^*(R \cap S) = f^*(S)$. Then use GT.3 on $R \cap S$, in which GT.2 would be used. ☐

Example 13.2.1.3. Suppose X is a topological space. Then a functor $J : \mathrm{Ob}(\mathbf{O(X)}) \longrightarrow \mathbf{Sets}$ where $\mathrm{Ob}(\mathbf{O(X)})$ is the discrete category of open sets of X defined by:

$$S \in JU \iff U \subseteq \bigcup_{V \in S} V$$

forms a Grothendieck topology over **O(X)**.

Proof. GT.1 : The set of all arrows over U in **O(X)**, when we take all their union, would equal U. So maximal sieve over U is in JU.
GT.2 : Let $S \in JU$ and $V \subset U$. Then $V^*(S) = \{W \in S \mid W \subset V \subset U\}$ is such that $\bigcup_{W \in V^*(S)} W \supseteq V$.
GT.3 : Let $S \in JU$ and R be any sieve over U such that for any $V \subset U$ with $V \in S$, the $V^*(R) \in JV$. Since $\bigcup_{W \in S} W \supseteq U$ and $V^*(R) = \{T \mid T \subset V \subset U\ ,\ \ T \in R\} \in JV \implies \bigcup_{T \in V^*(R)} T \supseteq V$ and $V^*(R) \subset R\ \forall\ V \in S$ as sieve, therefore $\bigcup_{Q \in R} Q \supseteq \bigcup_{W \in S} W \supseteq U$. Hence $R \in JU$. ☐

13.2.2 Basis for a Grothendieck Topology

Note that in **O(X)**, the usual notion of an open cover of U as $U = \bigcup_{i \in I} U_i$ does not makes $\{U_i\}_{i \in I}$ a sieve over U. But one could generate a sieve from this open cover, by collection all $V \subseteq U$ with $V \subseteq U_i$ for some $i \in I$. Therefore the collection $\{U_i\}_{i \in I}$ forms a *base cover* for U from which we can generate a usual cover over U. We now extend this to any category with pullbacks[2].

Definition 13.2.2.1. (**Basis for a Grothendieck Topology**) *Suppose* $\mathbf{C}$ *is a small category. A basis for Grothendieck topology is a functor*

$$K : \mathrm{Ob}(\mathbf{C}) \longrightarrow \mathbf{Sets}$$

where $\mathrm{Ob}(\mathbf{C})$ *is the discrete category of objects of* $\mathbf{C}$*, such that for any object* C *of* $\mathbf{C}$*, the collection* KC *of sets of arrows over* C *must satisfy:*

BGT.1 *(**Isomorphic Objects are Bases**) If* $f : C' \to C$ *is an isomorphism, then*

$$\{f\} \in KC$$

BGT.2 *(**Pullback Stability of Bases**) If* $\{f_i : C_i \to C \mid i \in I\} \in KC$ *and* $g : D \to C$ *is any arrow, then*

$$\{\pi_i : C_i \times_C D \to D \mid i \in I\} \in KD$$

BGT.3 *(**Transitivity of Bases**) If* $\{f_i : C_i \to C \mid i \in I\} \in KC$ *and* $\{f_i^j : C_i^j \to C_i \mid j \in I_i\} \in KC_i$*, then*

$$\{f_i \circ f_i^j : C_i^j \to C \mid j \in I_i\ ,\ \ i \in I\} \in KC$$

[2]because the pullback in **O(X)** of $U \hookrightarrow X$ and $V \hookrightarrow X$ is just the intersection of U and V.

One uses the same terminology of sites, even when we just have a base. That is, we will call $(\mathbf{C}, K)$ a site, even when K is a base, not a Grothendieck topology. One also generates a Grothendieck topology J from a base K as follows:

Lemma 13.2.2.2. *Suppose* $\mathbf{C}$ *is small and* K *is a base defined over it. Consider the topology* J *generated by the base* K *as follows:*

$$S \in JC \iff \exists\, R \in KC \text{ such that } R \subseteq S.$$

Then J *is indeed a Grothendieck Topology.*

Proof. GT.1 : Any $R \in KC$ is subset of the maximal sieve over C. So maximal sieve is in JC.
GT.2 : Suppose $S = \{f_i : C_i \to C \mid i \in I\} \in JC$ and $g : D \to C$. We want to show $g^*(S) \in JD$. By BGT.2, we have $\{\pi_i : C_i \times_C D \to D \mid i \in I\} \in KD$. Remember that $g \circ \pi_i = f_i \circ \sigma_i$ where $\sigma_i : C_i \times_C D \to C_i$ because of pullback square. But $f_i \in S$, therefore $f_i \circ \sigma_i = g \circ \pi_i \in S$. Hence, $\pi_i \in g^*(S)$, which shows that $\{\pi_i : C_i \times_C D \to D \mid i \in I\} \subseteq S$ so that $g^*(S) \in JD$.
GT.3 : The canonical idea works. □

Remark 13.2.2.3. Conversely to Lemma 13.2.2.2, given a site $(\mathbf{C}, J)$, we can construct a basis K of J as follows:

$$R \in KC \iff (R) \in JC$$

where (R) is the sieve generated by the basic cover R. More precisely:

$$(R) = \{f \circ g \mid f \in R\,,\ \ g \text{ is any composable arrow}\}.$$

Similar to Lemma 13.2.1.2, we have the following for basic covers:

Lemma 13.2.2.4. *Suppose we have a site* $(\mathbf{C}, K)$*. For any two basic covers* $R, P \in KC$ *of* C*, there exists a common* refinement *of* R *and* P *in* KC*.*

Proof. A collection $\{f_i : C_i \to C \mid i \in I\}$ is said to be a *refinement* of $\{g_j : D_j \to C \mid j \in J\}$ if every f_i factors through some g_j. Suppose J is the Grothendieck topology generated from K. Therefore $(R), (P) \in JC$. By Lemma 13.2.1.2, $(R) \cap (P) \in JC$, so that $\exists\, Q \in KC$ such that $Q \subseteq (R) \cap (P)$. Moreover, any arrow $f \in Q$ is such that $f \in (R)$ and $f \in (P)$. This means that f factors through some arrow in R and P, so that Q is a common refinement of R and P. □

13.2.3 Sheaves on a Site

With the notion of Grothendieck topologies over a small category in place, we have now a notion of what it means to *cover* an object in a small category. Therefore, the next step would be now to generalize the notion of sheaves over sites.

Matching Families and Amalgamations

In ordinary definition of sheaves over topological spaces (Definition **??**), the gluing condition requires the elements of the members of the cover of some open set which match over all restrictions of their corresponding members to their intersections, to be collatable to form an element of the whole open set under the sheaf functor. The analogue of the elements of members of the cover which satisfy the above property in sheaves over a site is known as a *matching family*:

Definition 13.2.3.1. (**Matching Family of a Cover for a Presheaf**) *Suppose* $(\mathbf{C}, J)$ *is a site. Let* $S \in JC$ *be an arbitrary cover of object* C *and* P *be a presheaf over* $\mathbf{C}$*. Then a matching family of* S *for the presheaf* P *is a set* M_S^P *given by:*

$$M_S^P = \{x_f \in P(dom\,(f)) \mid f \in S \textit{ and } \forall\ g \textit{ post-composable with } f\ ,\quad P(g)(x_f) = x_{f\circ g}\}$$

Note that M_S^P *may not be the only matching family for the cover* S*.*

The notion of a global collation is then captured by the following:

Definition 13.2.3.2. (**Amalgamation of a Matching Family**) *Suppose* $(\mathbf{C}, J)$ *is a site,* $S \in JC$ *is a cover of* C*,* P *is a presheaf over* $\mathbf{C}$ *and* M_S^P *is a matching family of* S*. An amalgamation of* M_S^P *is then the following element of* PC*:*

$$x \in PC \textit{ such that } Pf(x) = x_f\ \forall\ f \in S\ ,\quad \textit{where } x_f \in M_S^P.$$

Finally, a sheaf over a site is then defined as the following:

Definition 13.2.3.3. (**Sheaves over a Site - Ordinary Defn.**) *Suppose* $(\mathbf{C}, J)$ *is a site. A presheaf* $P : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ *is a sheaf if and only if:*

any matching family M_S^P *for any cover* $S \in JC$ *for any object* $C \in \mathrm{Ob}(\mathbf{C})$ *has a* ***unique amalgamation****.*

In-fact, the above definition can be written purely in categorical language as follows:

Definition 13.2.3.4. (**Sheaves over a Site - Categorical Defn.**) *Suppose* $(\mathbf{C}, J)$ *is a site. A presheaf* $P : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ *is a sheaf if and only if* $\forall C \in \mathrm{Ob}(\mathbf{C})$ *and* $\forall S \in JC$*, the following is an equalizer diagram*

$$\begin{tikzcd}
PC \arrow[dr, "e"] & & \\
& \prod_{f\in S} P(\mathrm{dom}\,(f)) \arrow[r, shift left, "p"] \arrow[r, shift right, "a"'] & \prod_{f\in S, g,\ \mathrm{dom}(f)=\mathrm{cod}(g)} P(\mathrm{dom}\,(g))
\end{tikzcd}$$

where e*,* p *and* a *are:*

$$\begin{aligned} e : PC &\longrightarrow \prod_{f\in S} P(\mathrm{dom}\,(f)) \\ x &\longmapsto \{Pf(x)\}_{f\in S} \end{aligned}$$

$$\begin{aligned} a : \prod_{f\in S} P(\mathrm{dom}\,(f)) &\longrightarrow \prod_{f\in S, g,\ \mathrm{dom}(f)=\mathrm{cod}(g)} P(\mathrm{dom}\,(g)) \\ \{x_f\}_{f\in S} &\longmapsto \{Pg(x_f)\}_{\mathrm{dom}(f)=\mathrm{cod}(g)} \end{aligned}$$

$$\begin{aligned} p : \prod_{f\in S} P(\mathrm{dom}\,(f)) &\longrightarrow \prod_{f\in S, g,\ \mathrm{dom}(f)=\mathrm{cod}(g)} P(\mathrm{dom}\,(g)) \\ \{x_f\}_{f\in S} &\longmapsto \{x_{f\circ g}\}_{\mathrm{dom}(f)=\mathrm{cod}(g)} \end{aligned}$$

Note that a *is just mapping the* x_f *to the corresponding member* $x_{f\circ g} \in \{x_f\}_{f\in S}$ *because* $f \circ g \in S$*.*

Sheaves in terms of Basis

A sheaf over a site $(\mathbf{C}, K)$ where K is a basis (Definition 13.2.2.1) can also be realized, albeit we require different notion of matching families for a basic cover:

Definition 13.2.3.5. (**Matching Family of a Basic Cover for a Presheaf**) *Suppose $(\mathbf{C}, K)$ is a site where K is a basis, $R \in KC$ is a basic cover of C and P is a presheaf over $\mathbf{C}$. A matching family M_R^P of the basic cover R for the presheaf P is then defined as the following set:*

$$M_R^P = \left\{x_{f_i} \in PC_i \mid f_i : C_i \to C \in R \text{ such that } P(\pi_{ij}^1)(x_{f_i}) = P(\pi_{ij}^2)(x_{f_j})\right\}$$

where π_{ij}^1 and π_{ij}^2 are as follows:

$$\begin{tikzcd}
C_i \times_C C_j \arrow[r, "\pi_{ij}^2"] \arrow[d, "\pi_{ij}^1"'] \arrow[dr, phantom, "\lrcorner", very near start] & C_j \arrow[d, "f_j"] \\
C_i \arrow[r, "f_i"'] & C
\end{tikzcd}$$

Remark 13.2.3.6. An **amalgamation** of a matching family of a basic cover for a presheaf is then defined as done previously (Definition 13.2.3.2), i.e. $x \in PC$ is an amalgamation of M_R^P if $Pf(x) = x_f \in M_R^P \; \forall \; f \in R$.

A sheaf is over a basic site is then obtained as follows (note the similarity of the result with that of Definition 13.2.3.3) :

Theorem 13.2.3.7. *Suppose $(\mathbf{C}, K)$ is a site and K is a basis and P is a presheaf over $\mathbf{C}$. Then, P is a sheaf if and only if any basic cover $\{f_i : C_i \to C \mid i \in I\} \in KC$ for any object C has a unique amalgamation.*

Proof. Proof is a bit long so we only provide a very brief sketch.
(L $\implies$ R) Take the topology J generated from K. P is a sheaf so $(R) \in JC$ has an unique amalgamation. Canonical observations lead to the realization that this unique amalgamation of (R) is an amalgamation for R too.
(R $\implies$ L) Again take the J as above. Take any matching family of a cover S from JC. Since $\exists \; R \subseteq S$ where $R \in KC$ and premise of the question says that this R has a unique amalgamation, therefore argue that this unique amalgamation of R is a unique amalgamation for S too. □

An example of sheaves on a site $(\mathbf{Top}, K)$ where K is the open cover topology is the usual contravariant hom-functor:

$$\mathbf{Yon}\,(Y) = \mathrm{Hom}_{\mathbf{C}}\,(-, Y)$$

This is quite trivial to see by a simple un-ravelling of definitions.

13.2.4 The Grothendieck Topos

We are now at a good footing to understand one of the central themes of this text, the Grothendieck topos. Denote the category of sheaves over a site $(\mathbf{C}, J)$ and natural transformations between them as the following:

$$\mathrm{Sh}\,(\mathbf{C}, J)$$

Note that this category of sheaves is a full subcategory of the presheaf category $\widehat{\mathbf{C}}$. Hence we have the following inclusion functor:

$$\mathrm{Sh}\left(\mathbf{C}, J\right) \rightarrowtail \widehat{\mathbf{C}}.$$

Hence, we have the following definition:

Definition 13.2.4.1. (**Grothendieck Topos**) *A Grothendieck Topos is a category* $\mathbf{T}$ *which is functorially equivalent to a category* $\mathrm{Sh}\left(\mathbf{C}, J\right)$ *of sheaves over some site* $(\mathbf{C}, J)$.

Remark 13.2.4.2. Clearly, the category $\mathrm{Sh}\left(\mathbf{C}, J\right)$ is a trivial example of a Grothendieck Topos.

We will later see some of the basic properties of the category of sheaves.

13.2.5 The Sheafification Functor

Suppose $(\mathbf{C}, J)$ is a site and we have the sheaf category $\mathrm{Sh}\left(\mathbf{C}, J\right)$. As mentioned in Section **??** for $\widehat{\mathbf{O(X)}}$, since the inclusion of $\mathrm{Sh}\left(\mathbf{C}, J\right)$ into $\widehat{\mathbf{C}}$ is the *simplest* way to get a presheaf from a sheaf, then it's adjoint has to be the *simplest* way to get a sheaf from a presheaf. Of-course, the aspect of the construction which transforms a presheaf to a sheaf is interesting, but as we will see (as we had already seen in **??**) this construction is a bit non-trivial.
The construction which we would now study is known as the $(-)^{+}$**-Construction**:

The $(-)^{+}$-Construction

Suppose $(\mathbf{C}, J)$ is a site and $P : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ be a presheaf. We define a new presheaf P^{+} as follows:

- Define

$$\mathrm{Match}\left(R, P\right)_{C} := \text{Set of all matching families of } R \in JC$$

- Define the following functor:

$$\begin{aligned} \mathrm{Match}\left(-, P\right)_{C} : \mathbf{(JC)}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ R &\longmapsto \mathrm{Match}\left(R, P\right)_{C} \\ S \subset R &\longmapsto \mathrm{Match}\left(R, P\right)_{C} \to \mathrm{Match}\left(S, P\right)_{C} \end{aligned}$$

  where $\mathrm{Match}\left(R, P\right)_{C} \to \mathrm{Match}\left(S, P\right)_{C}$ takes a matching family of R to that of S by restricting the elements of the family to that of $S \subset R$.
- Define presheaf P^{+} as the following:

$$\begin{aligned} P^{+} : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ C &\longmapsto \varinjlim \mathrm{Match}\left(-, P\right)_{C} \end{aligned}$$

A more illuminating equivalent definition of P^{+} by reminding ourselves the definition of colimits in **Sets**, however, would be the following:

Definition 13.2.5.1. (($(-)^{+}$**-Construction of a Presheaf**) *Suppose* $(\mathbf{C}, J)$ *is a site and* P *is a presheaf over* $\mathbf{C}$. *The presheaf* P^{+} *is given as follows:*

- *Define an equivalence relation on the set* $\bigcup_{S \in JC} \bigcup_{M_S^P \in Match(S,P)_C} M_S^P$ *where two matching families are related as follows:*

$$\begin{aligned} M_S^P \sim M_R^P \iff & \exists \text{ a refinement } T \subseteq S \cap R \,,\; T \in JC \text{ such that} \\ & x_f = y_f \; \forall\, f \in T \,,\; x_f \in M_S^P \;\&\; y_f \in M_R^P \end{aligned}$$

- *Define action of* P^+ *on objects as:*

$$\begin{aligned} P^+ : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ C &\longmapsto \text{Set of equivalence classes } \left(\bigcup_{S \in JC} \bigcup_{M_S^P \in \mathrm{Match}(S,P)_C} M_S^P \right) / \sim \end{aligned}$$

- *Define action of* P^+ *on arrows as:*

$$\begin{aligned} P^+ : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ f : D \to C &\longmapsto P^+ f : P^+ C \to P^+ D \end{aligned}$$

  where $P^+ f$ *is as follows:*

$$\begin{aligned} P^+ f : P^+ C &\longrightarrow P^+ D \\ [M_S^P] &\longmapsto [f \circ M_S^P] \end{aligned}$$

  and $f \circ M_S^P$ *is defined as (let* $M_S^P = \{x_g \in P(\mathrm{dom}\,(g)) \mid g \in S\}$*):*

$$f \circ M_S^P := \{x_{f \circ h} \in M_S^P \mid h \in f^*(S)\}.$$

Remark 13.2.5.2. (P^+ **is well defined**) One may wonder whether the $P^+ f$ as mentioned above is well-defined or not. That is, does $M_S^P \sim M_R^P \implies f \circ M_S^P \sim f \circ M_R^P$? Well it can be seen quite easily that this is true because if we take the refinement $T \subset R \cap S$ and it's pullback $f^*(T)$, one can see that it would be a refinement too of $f^*(R) \cap f^*(S)$. Clearly, elements of $f \circ M_S^P$ and $f \circ M_R^P$ are same when restricted to $f^*(T)$. So P^+ is indeed well-defined.

There is a canonical natural transformation which takes any presheaf to it's $(-)^+$ presheaf. This would be important in the following constructions.

Lemma 13.2.5.3. *Suppose* $(\mathbf{C}, J)$ *is a site and* P *is a presheaf over* $\mathbf{C}$*. Then the following is an important natural transformation:*

$$\eta : P \Longrightarrow P^+$$

defined by

$$\begin{aligned} \eta_C : PC &\longrightarrow P^+ C \\ x &\longmapsto \left[M^P_{S_C^{max}} \right] \end{aligned}$$

where

$$M^P_{S_C^{max}} := \{Pf(x) \in P(dom(f)) \mid f \in S_C^{max}\}$$

A Presheaf to the Sheaf via $(-)^{++}$

We will now see how we would transform a presheaf to a sheaf via the above construction. More specifically, we would take a presheaf P and then do the following cascade of transformations:

$$P \xmapsto{\eta^P} P^+ \xmapsto{\eta^{P^+}} (P^+)^+$$

where η^P is the natural transformation $\eta^P : P \Longrightarrow P^+$, as given in Lemma 13.2.5.3. As we will see now, the $(P^+)^+$ is guaranteed to be a sheaf over $(\mathbf{C}, J)$ for any presheaf P.
We now quickly state some canonical lemmas without proof (in the following, we assume $(\mathbf{C}, J)$ is a given site):

Lemma 13.2.5.4. *A presheaf P is separated $\iff$ $\eta : P \Longrightarrow P^+$ is a monic.*
Similarly, a presheaf P is a sheaf $\iff$ $\eta : P \Longrightarrow P^+$ is an isomorphism.

Lemma 13.2.5.5. *(**Universality of P^+**) If F is a sheaf and P is a presheaf over $\mathbf{C}$ and we are given a map $\theta : P \Longrightarrow F$ in $\widehat{\mathbf{C}}$, then $\exists$ a unique map $P^+ \Longrightarrow F$ such that the following commutes:*

$$\begin{tikzcd}
P \arrow[d, Rightarrow, "\eta"'] \arrow[dr, Rightarrow, "\theta"] & \\
P^+ \arrow[r, Rightarrow, dashed] & F
\end{tikzcd}$$

Lemma 13.2.5.6. *For any presheaf P, the presheaf P^+ is separated.*

Lemma 13.2.5.7. *For any separated presheaf, the presheaf P^+ is a sheaf.*

With the above lemmas we can state the following theorem, whose proof is now quite easy with their help:

Theorem 13.2.5.8. *Suppose $(\mathbf{C}, J)$ is a site. Then sheaves over this site $Sh(\mathbf{C}, J)$ forms a reflective subcategory of $\widehat{\mathbf{C}}$, where the corresponding adjunction is given by the following:*

$$\begin{tikzcd}
\widehat{\mathbf{C}} \arrow[rr, bend left, "a"] \arrow[rr, phantom, "\perp"] & & Sh(\mathbf{C}, J) \arrow[ll, bend left, hook, "i"]
\end{tikzcd}$$

where

$$\begin{aligned}
a : \widehat{\mathbf{C}} &\longrightarrow Sh(\mathbf{C}, J) \\
P &\longmapsto a(P) := \eta^{P^+} \circ \eta^P = (P^+)^+.
\end{aligned}$$

Moreover, we have:

$$a \circ i \cong_{Nat} \mathrm{id}_{Sh(\mathbf{C},J)}.$$

Proof. Lemma 13.2.5.6 means P^+ is separated and Lemma 13.2.5.7 means that $(P^+)^+$ is a sheaf. That a as defined above is the left adjoint of inclusion can be seen via the Lemma 13.2.5.5. $\square$

Remark 13.2.5.9. The **sheafification functor** $a(-)$ **preserves finite limits**. This is due to the fact that the functor $\mathrm{Match}(R, -)$ preserves finite limits as it is isomorphic to $\mathrm{Hom}_{\widehat{\mathbf{C}}}(R, -)$ (Proposition 20.6.2.4 generalized to sites), which preserves limits, for a fixed cover R. But the P^+ is defined to be a filtered colimit. But filtered colimit commutes with finite limits. Therefore $P \mapsto P^+$ preserves limits and hence $a(-)$ preserves limits. This shows **that $i \vdash a$ is a geometric morphism between topoi**.

13.2.6 Properties of $\mathrm{Sh}(\mathbf{C}, J)$

We now look at some of the basic properties of categories of sheaves over a site. In particular, we will see that $\mathrm{Sh}(\mathbf{C}, J)$ satisfies all properties of an elementary topos. In the following, we assume that a site $(\mathbf{C}, J)$ is given to us.

Proposition 13.2.6.1. *$Sh(\mathbf{C}, J)$ has all small limits.*

Proof. Take any diagram in $\mathrm{Sh}(\mathbf{C}, J)$. Compute their limit in the $\widehat{\mathbf{C}}$. But limits in $\widehat{\mathbf{C}}$ are computed point-wise. Then use the Definition 13.2.3.4 on each member of the diagram in $\widehat{\mathbf{C}}$ evaluated at some object C to get an equalizer diagram. Take the limit of each member of the equalizer diagram. This new diagram would also be an equalizer. Hence limit of the original diagram also follows the equalizer definition of sheaves over a site and hence it is also in $\mathrm{Sh}(\mathbf{C}, J)$. $\square$

Remember the following functor from our previous discussions:

$$\begin{aligned} \Delta : \mathbf{Sets} &\longrightarrow \widehat{\mathbf{C}} \\ S &\longmapsto \Delta S := \text{Constant presheaf to } S \end{aligned}$$

The following is an useful adjunction, called the *global sections adjunction*:

Definition 13.2.6.2. (**Global Sections Adjunction**) *Suppose $(\mathbf{C}, J)$ is a site and left adjoint of inclusion $a : \widehat{\mathbf{C}} \longrightarrow \mathrm{Sh}(\mathbf{C}, J)$ as given in Theorem 13.2.5.8. We then have the following adjunction:*

$$\begin{tikzcd} \mathbf{Sets} \arrow[rr, bend left, "a\circ\Delta"] \arrow[rr, phantom, "\perp"] & & \mathrm{Sh}(\mathbf{C}, J) \arrow[ll, bend left, "\Gamma"] \end{tikzcd}$$

*where $\Gamma : \mathrm{Sh}(\mathbf{C}, J) \longrightarrow$ **Sets** takes a sheaf F to $\mathrm{Nat}(\mathbf{1}, F)$, i.e. Γ is the global sections functor.*

We now see that sheaf category has small colimits.

Proposition 13.2.6.3. *$Sh(\mathbf{C}, J)$ has all small colimits.*

Proof. From Theorem 13.2.5.8, we have that a preserves colimits as it is the left adjoint. Therefore, to find colimit of a diagram in $\mathrm{Sh}(\mathbf{C}, J)$, first take it's limit in $\widehat{\mathbf{C}}$ and then apply a to it, since it would preserves colimits, we then have colimit of the original diagram in $\mathrm{Sh}(\mathbf{C}, J)$. $\square$

Finally, we show without proof that the exponentials in $\mathrm{Sh}(\mathbf{C}, J)$ exists and how they are given by:

Proposition 13.2.6.4. *$Sh(\mathbf{C}, J)$ has all exponentials and for two sheaves F & G, the exponential F^G is given by:*

$$\begin{aligned} F^G : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ C &\longmapsto F^G(C) := \mathrm{Nat}\left(\mathrm{Hom}_{\mathbf{C}}(-, C) \times G, F\right) \end{aligned}$$

Remark 13.2.6.5. Since $\mathrm{Sh}(\mathbf{C}, J)$ is a full subcategory of $\widehat{\mathbf{C}}$, therefore it doesn't matter whether we take all natural transformations in $\widehat{\mathbf{C}}$ or $\mathrm{Sh}(\mathbf{C}, J)$ in the above. Also note that the above construction of exponential was already proved in Proposition 13.1.3.1.

Hence we are just one step away from proving that $\mathrm{Sh}(\mathbf{C}, J)$ is an elementary topos; we just need to show that $\mathrm{Sh}(\mathbf{C}, J)$ has a subobject classifier (Definition 13.1.0.1) which would classify one subobject from the other. Constructing that requires a bit more insight, which we gain now.

Subobject Classifier in Sh $(\mathbf{C}, J)$

The truth object and the subobject classifier in $\mathrm{Sh}(\mathbf{C}, J)$ are to be constructed now. We first define a closed sieve:

Definition 13.2.6.6. (**Closed Sieve over an object**) *Let $(\mathbf{C}, J)$ be a category. A closed sieve S over an object C in the given site is such a sieve which follows:*

If for any $f : D \to C$, $f^*(S) \in J(C)$, then $f \in S$.

Remark 13.2.6.7. It's quite trivial to see that closed sieves are stable under pullback under any compatible arrow. Also, all maximal sieves are closed sieves.

Then, the truth object is given by the following :

Proposition 13.2.6.8. *Suppose $(\mathbf{C}, J)$ is a site and $Sh(\mathbf{C}, J)$ is the sheaf category over it. Then, the truth object $\Omega : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ is given by:*

$$\begin{aligned} \Omega : \mathbf{C}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ C &\longmapsto \Omega C := \textit{Set of all } J\textit{-closed sieves over } C \\ f : B \to C &\longmapsto \Omega f : \Omega C \to \Omega B \\ &\qquad\qquad S \mapsto f^*(S) \end{aligned}$$

and this truth object Ω is indeed a sheaf.

Proof. (*Sketch*) First prove that Ω is a separated presheaf. For this, take any matching family R of any object C and take two amalgamations S^1 and S^2 in ΩC. Note now that $S^1 \cap R = S^2 \cap R$. With this take $g \in S^1$ and show that $g \in S^2$. Similarly the converse to assert that $S^1 = S^2$.
After this, now show that Ω has amalgamations for any matching family. For this, take any matching family of $R \in JC$ and form a different sieve over C by collecting all closed sieves' pre-composed with the corresponding member of R to get arrows over C. Argue that the closure[3] of this sieve is the amalgamation of that matching family. □

[3]The closure of a sieve S over C in a site $(\mathbf{C}, J)$ is defined as $\bar{S} := \{g \mid \mathrm{cod}(g) = C\ ,\ \ g^*(S) \in J(\mathrm{dom}(g))\}$.

Now, the **"truth" monic** and the unique **characteristic map** is then given as (the proof requires a small result but is fairly trivial afterwards, so is not presented below):

Proposition 13.2.6.9. *Suppose* $(\mathbf{C}, J)$ *is a site and* $Sh(\mathbf{C}, J)$ *is the sheaf category. Then the monic*

$$\mathit{true} : \mathbf{1} \Longrightarrow \Omega$$

in $Sh(\mathbf{C}, J)$*, given as*

$$\begin{aligned} \mathit{true}_C : \{\star\} &\longrightarrow \Omega C \\ \star &\longmapsto S_C^{max} \end{aligned}$$

is the subobject classifier for $Sh(\mathbf{C}, J)$*. The unique characteristic map* $\chi : G \Longrightarrow \Omega$ *corresponding to the monic* $m : F \Longrightarrow G$ *in* $Sh(\mathbf{C}, J)$ *is given as:*

$$\begin{aligned} \chi : G &\Longrightarrow \Omega \\ \chi_C : GC &\longrightarrow \Omega_C \\ x &\longmapsto \{f \mid cod(f) = C\ ,\ \ Gf(x) \in F(dom(f))\}. \end{aligned}$$

Hence by Propositions 13.2.6.1,13.2.6.3,13.2.6.4 & 13.2.6.9, Sh $(\mathbf{C}, J)$ is an elementary topos, as was required to be shown.

13.3 Basic Properties and Results in Topoi

We now study the first properties observed from the definition of a topos. We will see that there are a lot of striking similarity between **Sets** and a topos **E**. For example, each subobject in a topos will have a clearly defined way of identifying "elements" as whether they are indeed in the given subobject or not, similarly, whether two "elements" are same or not, image of a subobject along other arrow and so on, purely in categorical terms, and all of which is generalized from their usual notions in **Sets**.
Let's begin with the definition of a topos, with the underlying help of sets:

Definition 13.3.0.1. (**Elementary Topos - I**) *A category* **E** *is a topos if:*

ETI.1 **E** *has all finite limits.*

ETI.2 *The subobject functor (Definition 13.1.0.2)* $\mathrm{Sub}_{\mathbf{E}}(-) : \mathbf{E}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ *is representable and the object which represents it, denoted* Ω*, is called the subobject classifier. That is,*

$$\mathrm{Sub}_{\mathbf{E}}(A) \cong_{\mathrm{Nat}} \mathrm{Hom}_{\mathbf{E}}(A, \Omega).$$

ETI.3 *The functor* $\mathrm{Hom}_{\mathbf{E}}(B \times -, \Omega)$ *is representable for all objects* B *and the representing object is denoted as* PB*, called the power object of object* B*. That is,*

$$\mathrm{Hom}_{\mathbf{E}}(B \times A, \Omega) \cong_{\mathrm{Nat}} \mathrm{Hom}_{\mathbf{E}}(A, PB).$$

Remark 13.3.0.2. One can combine ETI.2 and ETI.3 to get

$$\mathrm{Sub}_{\mathbf{E}}(B \times A) \cong_{\mathrm{Nat}} \mathrm{Hom}_{\mathbf{E}}(A, PB)$$

As pointed earlier, we can actually state a definition of topos in complete categorical language without any use of underlying sets. Hence, the following is the first order theory of an elementary topos:

Definition 13.3.0.3. (**Elementary Topos -II**) *A category* **E** *is a topos if:*

ETII.1 *All pullbacks exists.*

ETII.2 *A terminal object* **1** *exists.*

ETII.3 *There exists an object Ω and a monomorphism* **true** : $\mathbf{1} \rightarrowtail \Omega$ *in* **E** *such that for any monomorphism $m : E \rightarrowtail B$, there exists a unique map $\phi : B \to \Omega$ such that the following is a pullback square:*

$$\begin{tikzcd} E \arrow[r] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, tail, "\text{true}"] \\ B \arrow[r, dashed, "\phi"'] & \Omega \end{tikzcd}$$

ETII.4 *For any object B, there exists an object PB in* **E** *and an arrow:*

$$\epsilon_B : B \times PB \longrightarrow \Omega$$

such that for any object A and any arrow $f : B \times A \to \Omega$, we have a unique arrow $\hat{f} : A \longrightarrow PB$ such that the following commutes:

$$\begin{tikzcd} A \arrow[d, dashed, "\hat{f}"'] & & B \times A \arrow[d, "1\times\hat{f}"'] \arrow[dr, "f"] & \\ PB & & B \times PB \arrow[r, "\epsilon_B"'] & \Omega \end{tikzcd} .$$

The $\hat{f}$ is usually called the P-transpose of f.

The construction of power object can be defined as the following functor:

Definition 13.3.0.4. (**Power Object Functor**) : *Suppose* **E** *is a topos. The power object for each object B of* **E** *is a functor*

$$\begin{aligned} P : \mathbf{E}^{\mathrm{op}} &\longrightarrow \mathbf{E} \\ B &\longmapsto PB \quad \text{(the object defined in Definition 13.3.0.3, ETII.4)} \\ f : B \to C &\longmapsto Pf : PC \to PB \end{aligned}$$

where the Pf is the unique map for which the corresponding diamond commutes:

$$\begin{tikzcd} & B \times PB \arrow[dr, "\epsilon_B"] & \\ B \times PC \arrow[ur, "1\times Pf"] \arrow[dr, "f\times 1"'] & & \Omega \\ & C \times PC \arrow[ur, "\epsilon_C"'] & \end{tikzcd} .$$

That is,

$$\epsilon_B \circ (1 \times Pf) = \epsilon_C \circ (f \times 1) .$$

Remark 13.3.0.5. This means that the arrow $\epsilon_B : B \times PB \longrightarrow \Omega$ is *dinatural* in B.

13.3.1 Extension, Characteristic & Name of a subobject

Remember that $A \cong A \times 1$ in any category. Therefore by the Definition 13.3.0.1, we have the following isomorphisms in a topos $\mathbf{E}$:

$$\mathrm{Sub}_{\mathbf{E}}(A) \cong \mathrm{Hom}_{\mathbf{E}}(A, \Omega) \cong \mathrm{Hom}_{\mathbf{E}}(1, PA)$$

Hence there are three equivalent ways to talk about a subobject. These are denoted as follows:

Definition 13.3.1.1. (**Extension, Characteristic & Name**) *Suppose* $\mathbf{E}$ *is a topos. Let* $m : S \rightarrowtail A$ *be a subobject of* A*. Then we denote the total three corresponding arrows (from above isomorphisms) as follows*:

$$\begin{array}{ccc} S \overset{m}{\rightarrowtail} A & A \xrightarrow{\phi} \Omega & 1 \xrightarrow{s} PA \\ & & \\ \text{Extension of } \phi & \text{Characteristic of } S & \text{Name of } \phi \\ S{:=}\{m|\phi\} & \phi{:=}\text{char } S & s{:=}\ulcorner\phi\urcorner \end{array}.$$

This convention becomes useful when we realize that an arrow $b : X \to A$ can be treated as a "generalized element" of A. Hence, we can talk whether an "element" b of A is in a subobject of A. The role of characteristic map $\phi = \text{char } S$ is illuminating. It tells us about the property/predicate that is followed by the subobject that it characterizes. For example, the statement of the subobject classifier (Definition 13.3.0.3, ETII.3) can be reformulated as the condition that an "element" of A, $a : X \to A$, is in the subobject $m : S \rightarrowtail A$ if and only if $(\text{char } S) \circ a = \mathsf{true} \circ (X \to \mathbf{1})$. Let us denote $\mathsf{true}_X := \mathsf{true} \circ (X \to \mathbf{1})$. Note that $\mathsf{true} \circ (X \to \mathbf{1})$ is simply the arrow $X \to \Omega$ corresponding to all truth. Therefore, the fact that $(\text{char } S) \circ a = \mathsf{true}_X$ means that the element X forms a cone over the pullback square and hence we have a unique arrow $X \to S$, meaning that X is also a "generalized element" of S.

We now develop more interesting predicates, like the one which tells us whether an element is in a given subobject named by a name $\ulcorner\phi\urcorner$.

Membership Predicate

Suppose we are given an element $b : X \to B$ of B and a subobject named $s := \ulcorner\phi\urcorner$ of B. We then have the following commutative diagram:

$$\begin{tikzcd} X \arrow[r, "b"] \arrow[d, equal] & B \arrow[r, "\phi"] \arrow[d, equal] & \Omega \\ X \times 1 \arrow[r, "b\times 1"'] & B \times 1 \arrow[r, "1\times s"'] & B \times PB \arrow[u, "\epsilon_B"'] \end{tikzcd} \tag{13.1}$$

This means that:

$$\boxed{\epsilon_B \circ (b \times s) = \mathsf{true}_{X\times 1} \textbf{ if and only if } \phi \circ b = \mathsf{true}_X.}$$

That is, $\epsilon_B \circ (b \times s)$ is true exactly when $\phi \circ b$ is true. But the latter means that b factors through the subobject (is in) named by s because of universality of pullback. Hence $\epsilon_B \circ (b \times s)$ is true only when b is in the subobject named by s.

This is exactly the reason why ϵ_B is called the **membership predicate** for object B.

Equality Predicate

Suppose now that we have two elements $b, b' : X \to B$ of B. How can we know that these are same elements, that is, what is the condition for $b = b'$? Clearly, this is a property of the elements of B, hence there must be a predicate for B to answer this question. This predicate for an object which tells us when two elements of it are same is exactly what we construct now.

Remember first the diagonal arrow of B, $\Delta_B : B \longrightarrow B \times B$, which is such that $p_1 \circ \Delta_B = p_2 \circ \Delta_B = 1$, where p_1, p_2 are projections of $B \times B$. It can be seen without much effort that Δ_B is actually a monic. We can hence talk about the subobject $\Delta_B : B \rightarrowtail B \times B$ of $B \times B$:

$$\begin{tikzcd}
B \arrow[r] \arrow[d, tail, "\Delta_B"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, tail, "\mathsf{true}"] \\
B \times B \arrow[r, "\mathrm{char}\ \Delta_B"'] & \Omega
\end{tikzcd} \tag{13.2}$$

where $\mathrm{char}\ \Delta_B$ is the unique characteristic map of subobject Δ_B. But now, by power object adjunction (Definition 13.3.0.3, ETII.4), we have the following unique arrow:

$$\begin{tikzcd}
B \arrow[d, dashed, "\mathrm{char}\hat{\ }\Delta_B"'] \\
PB
\end{tikzcd}
\qquad
\begin{tikzcd}
B \times B \arrow[r, "\mathrm{char}\ \Delta_B"] \arrow[d, "1 \times \mathrm{char}\hat{\ }\Delta_B"'] & \Omega \\
B \times PB \arrow[ur, "\epsilon_B"'] &
\end{tikzcd}$$

where $\mathrm{char}\hat{\ }\Delta_B$ is the P-transpose of $\mathrm{char}\ \Delta_B$. Let us denote this transpose as

$$\{\cdot\}_B := \mathrm{char}\hat{\ }\Delta_B.$$

Now, take any two elements $b, b' : X \to B$ of B. We then have a unique arrow $\langle b, b' \rangle : X \to B \times B$ due to the universality of the product. Hence, by universality of the pullback in (13.2), we get the following condition:

$$\boxed{(\mathrm{char}\ \Delta_B) \circ \langle b, b' \rangle = \mathsf{true}_X \textbf{ if and only if } b = b'.}$$

Due to this exact reason, the characteristic map of subobject Δ_B, $\mathrm{char}\ \Delta_B$ is called the **equality predicate** for object B[4].

The arrow $\{\cdot\}_B : B \longrightarrow PB$ is always a monic:

[4] At this point one should notice how the above two predicates have been constructed. We are using the subobject classifier to *distinguish* between elements based on the object's *properties*. Hence it is the subobject classifier (Definition 13.3.0.3, ETII.3) which provides us with the opportunity to talk about *properties* of an object in a topos, which can not be done otherwise in any arbitrary category.

Proposition 13.3.1.2. *Suppose* **E** *is a topos. For any object* B *of* **E**, *the arrow*

$$\{\cdot\}_B := \hat{char}\,\Delta_B : B \longrightarrow PB$$

is always a monic.

Proof. Let $b, b' : X \to B$ be two arrows such that

$$\begin{aligned}
\{\cdot\}_B \circ b &= \{\cdot\}_B \circ b' \\
(1 \times \{\cdot\}_B) \circ (1 \times b) &= (1 \times \{\cdot\}_B) \circ (1 \times b') \\
\epsilon_B \circ (1 \times \{\cdot\}_B) \circ (1 \times b) &= \epsilon_B \circ (1 \times \{\cdot\}_B) \circ (1 \times b') \\
(\mathrm{char}\ \Delta_B) \circ (1 \times b) &= (\mathrm{char}\ \Delta_B)(1 \times b')
\end{aligned}$$

Now note the following diagram:

$$\begin{tikzcd}
X \arrow[r, shift left, "b"] \arrow[r, shift right, "b'"'] \arrow[d, shift right, "\langle b,1\rangle"'] \arrow[d, shift left, "\langle b',1\rangle"] & B \arrow[r] \arrow[d, "\Delta_B"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, "\mathsf{true}"] \\
B \times X \arrow[r, shift left, "1\times b'"] \arrow[r, shift right, "1\times b"'] & B \times B \arrow[r, "\mathrm{char}\ \Delta_B"'] & \Omega
\end{tikzcd}$$

The right square is a pullback, whereas the two left squares are also pullback by an easy observation. Hence the whole two bigger rectangles in the above diagram are pullbacks. But this means that the $\mathrm{char}\ \langle b, 1\rangle = \mathrm{char}\ \Delta_B \circ (1 \times b) = \mathrm{char}\ \Delta_B \circ (1 \times b') = \mathrm{char}\ \langle b', 1\rangle$. Hence $\langle b, 1\rangle$ and $\langle b', 1\rangle$ are same pullbacks. Therefore $\exists f : X \to X$ isomorphism such that $\langle b, 1\rangle = \langle b', 1\rangle \circ f$. But this means that $b = b' \circ f$ and $1 = f$ because $\langle a, b\rangle \circ f = \langle a \circ f, b \circ f\rangle$. Hence $b = b' \circ 1 = b'$. $\square$

Logical Morphism

The fact that a topos has it's own defining properties like a subobject classifier and power objects means that any functor may or may not preserve these properties. To distinguish such a functor, we define what we call a logical morphism. Before formally defining them, let's look at one of the more defining properties of a topos; that a topos is *balanced*:

Proposition 13.3.1.3. *Suppose* **E** *is a topos. Any monomorphism* $f : A \rightarrowtail B$ *is an equalizer of some parallel pair.*

Proof. The arrow $f : A \rightarrowtail B$ equalizes the following:

$$\begin{tikzcd}
B \arrow[r, shift left, "\mathsf{true}\circ !_B"] \arrow[r, shift right, "\mathrm{char}\ f"'] & \Omega
\end{tikzcd}$$

where $!_B : B \longrightarrow \mathbf{1}$, because of the following:

$$\begin{aligned}
\mathrm{char}\ f \circ f &= \mathsf{true}\circ !_A \\
&= \mathsf{true}\circ !_B \circ f
\end{aligned}$$

as $!_A = !_B \circ f$. Now because A is universal with this property (note we are just stating the subobject classifier's definition) so we have that $f : A \rightarrowtail B$ equalizes the above pair. $\square$

Corollary 13.3.1.4. *In a topos* **E**, *every arrow* $f : A \to B$ *which is both monic and epic is an isomorphism.*

Proof. If $f : A \to B$ is monic, so it equalizes two parallel pairs $x, y : B \to C$. But then $x \circ f = y \circ f \implies x = y$ because f is epic. Therefore f is an equalizer of the same pair of arrows, which is always the identity at the domain. Hence f is an isomorphism. $\square$

Remark 13.3.1.5. A category in which every monic + epic arrow is an isomorphism is called a **balanced** category. Hence every topos is balanced.

Definition 13.3.1.6. (**Logical Morphism**) *Suppose* **E** *and* **E**$'$ *are two topoi. A functor*

$$T : \mathbf{E} \longrightarrow \mathbf{E}'$$

is called a logical morphism if it preserves all finite limits, subobject classifier and all exponentials, upto isomorphism.

Direct Image

Suppose in **Sets**, we have a monic $m : S' \longrightarrow B'$ and some other function $f : B' \longrightarrow B$. With this, we have another function $PB' \longrightarrow PB$, which takes a subset S' in PB' to the set $f(S')$. We can generalize it to an arbitrary topos.

Definition 13.3.1.7. (**Direct Image Arrow**) *Suppose* **E** *is a topos and let* $m : S' \rightarrowtail B'$ *be a monic and* $k : B' \to B$ *be any arrow. Then there is an arrow*

$$\exists_k : PB' \longrightarrow PB$$

called the direct image arrow of k *is defined by:*

$$\begin{tikzcd}
U \arrow[r] \arrow[d, tail, dashed, "u_{B'}"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[r, equal] \arrow[d, tail, "\text{true}"] & \mathbf{1} \arrow[dd, tail, dashed, "\text{true}"] \\
B' \times PB' \arrow[r, "\epsilon_B"] \arrow[d, "k \times 1"'] & \Omega & \\
B \times PB' \arrow[rr, dashed, "e_k := \text{char } (k \times 1) \circ u_{B'}"] & {} \arrow[d, squiggly] & \Omega \\
PB' \arrow[rr, dashed, "\exists_k"'] & {} & PB
\end{tikzcd}$$

Direct image arrow preserves names of the subobject, as was expected from the beginning discussion of the same arrow in **Sets**:

Proposition 13.3.1.8. *Suppose* **E** *is a topos. Then, for monics*

$$S \overset{m}{\rightarrowtail} B' \overset{k}{\rightarrowtail} B$$

we have that the following commutes:

$$\begin{tikzcd}
\mathbf{1} \arrow[d, "\ulcorner char\ m \urcorner"'] \arrow[dr, "\ulcorner char\ k \circ m \urcorner"] & \\
PB' \arrow[r, "\exists_k"'] & PB
\end{tikzcd}$$

13.3.2 Factorization in a Topos

A distinguishing feature of a topos is that each arrow in it can be factored into a product of a monomorphism and an epimorphism:

Proposition 13.3.2.1. *Suppose* **E** *is a topos and let f be any arrow in it. Then there is a monic m and an epic e composable such that*

$$f = m \circ e.$$

Proof. (*Sketch*) The m is constructed by taking the equalizer of the cokernel pair[5] of f and e as the universal arrow from A to the object representing this equalizer. In more concrete setting, for the given f, denote the x and y as the following pushout components:

$$\begin{tikzcd} C & B \arrow[l, "y"'] \\ B \arrow[u, "x"] & A \arrow[l, "f"] \arrow[u, "f"'] \end{tikzcd}$$

Then the m and e are the following arrows (with their defining properties mentioned above):

$$\begin{tikzcd} A \arrow[rr, bend left, "f"] \arrow[r, dashed, "e"'] & M \arrow[r, dashed, "m"'] & B \arrow[r, shift left, "x"] \arrow[r, shift right, "y"'] & C \end{tikzcd} .$$

It then follows that m is monic and e is epic. □

We now show that the collection of all subobjects for an object in a topos forms a lattice. Later we will show that it actually forms a *Heyting Algebra*.

Definition 13.3.2.2. (**Lattice**) *Suppose* **C** *is a posetal category.* **C** *is a lattice when it has all binary products and coproducts. The product is alternatively called "meet" and coproduct the "join" and denoted $\wedge$ & $\vee$ respectively.*

Now, we get the following important theorem:

Theorem 13.3.2.3. *Suppose* **E** *is a topos. Let $k : A \to B$ be any arrow in* **E**. *Then,*

1. *$Sub_{\mathbf{E}}(D)$ is a lattice for any object D in* **E**.
2. *The functor $\exists_k : Sub_{\mathbf{E}}(A) \longrightarrow Sub_{\mathbf{E}}(B)$[6] which takes each subobject to the object which mono-epi factorizes[7] it, and the pullback arrow forms the following adjunct pair between lattices (regarded as*

[5] The pushout of $f : A \to B$ with itself.

[6] Note that $\exists_k$ here is the "external" direct image functor, in contrast to Definition 13.3.1.7, which was internal.

[7] That is, for $u : S \rightarrowtail A$, $\exists_k(u)$ would be the monic $m_{k \circ u} : \exists_k S \rightarrowtail B$ in the figure given below:

$$\begin{tikzcd} S \arrow[d, two heads, "e_{k \circ u}"'] \\ \exists_k S \arrow[d, tail, "m_{k \circ u}"] \\ B \end{tikzcd}$$

composes to $k \circ u$.

categories):

$$\begin{tikzcd} Sub_{\mathbf{E}}(A) \arrow[r, bend left, "\exists_k"] \arrow[r, phantom, "\perp"] & Sub_{\mathbf{E}}(B) \arrow[l, bend left, "(k)^{-1}"] \end{tikzcd}$$

where

$$\begin{aligned} (k)^{-1} : Sub_{\mathbf{E}}(B) &\longrightarrow Sub_{\mathbf{E}}(A) \\ m : S \rightarrowtail B &\longmapsto \pi_1 : A \times_B S \rightarrowtail A. \end{aligned}$$

Proof. To show that $\mathrm{Sub}_{\mathbf{E}}(D)$ is a lattice, we first note that it is partially ordered by inclusion, as if $S, T, U \in \mathrm{Sub}_{\mathbf{E}}(D)$ then $S \subseteq S$, $S \subseteq T$ & $T \subseteq S$ implies $S = T$ and finally if $S \subseteq T$ and $T \subseteq U$ then $S \subseteq U$. Moreover, for any two subobjects S, T, we can form two more subobjects as follows:

$$\begin{tikzcd} S \wedge T := S \times_D T \arrow[r, tail] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & T \arrow[d, tail] \\ S \arrow[r, tail] & D \end{tikzcd}$$

and

$$\begin{tikzcd} & D & \\ & M \arrow[u, tail, "m_f"] & \\ S \arrow[uur, tail] \arrow[r, "i_1"'] & S \amalg T \arrow[u, two heads, "e_f"] \arrow[uu, dashed, bend right, "f"'] & T \arrow[uul, tail] \arrow[l, "i_2"] \end{tikzcd}$$

where $S \vee T := M$ and f is the unique universal arrow from the coproduct of S and T to D and m_f, e_f are it's mono-epic factors (Proposition 13.3.2.1). Therefore $(\mathrm{Sub}_{\mathbf{E}}(D), \vee, \wedge)$ is a lattice.

Now, to show that $\exists_k$ and $(k)^{-1}$ are adjoints as required in the theorem, we first take any subobject $m_A : S_A \rightarrowtail A$ of A and $m_B : S_B \rightarrowtail B$ of B. We want to establish the following natural isomorphism:

$$\mathrm{Hom}_{\mathrm{Sub}_{\mathbf{E}}(B)}(\exists_k m_A, m_B) \cong \mathrm{Hom}_{\mathrm{Sub}_{\mathbf{E}}(A)}\left(m_A, (k)^{-1} m_B\right).$$

Hence, take any arrow $g : \exists_k S_A \longrightarrow S_B$, so that g is just an arrow $\exists_k m_A \longrightarrow m_B$. Then consider the following diagram:

$$\begin{tikzcd} S_A \arrow[r, dashed, "f"] \arrow[d, tail, "m_A"'] \arrow[rrr, bend left, two heads, "e_{k \circ m_A}"] & k^{-1}S_B \arrow[r] \arrow[d, tail, "k^{-1}m_B"] \arrow[dr, phantom, "\lrcorner", very near start] & S_B \arrow[d, tail, "m_B"'] & \exists_k S_A \arrow[l, "g"'] \arrow[d, tail, "m_{k \circ m_A}"] \\ A \arrow[r, equal] & A \arrow[r, "k"'] & B \arrow[r, equal] & B \end{tikzcd}$$

where, because the pair m_A and $g \circ e_{k \circ m_A}$ forms a cone over the middle pullback as $m_B \circ g \circ e_{k \circ m_A} = m_{k \circ m_A} \circ e_{k \circ m_A} = k \circ m_A$, $\exists\, !\ f : S_A \longrightarrow (k)^{-1} S_B$ which is the required arrow to establish the adjunction. $\square$

Remark 13.3.2.4. In continuation of "external"/"internal" debate mentioned in footnote 15, we note that $\mathrm{Sub}_{\mathbf{E}}(A)$ is the external subobject lattice. We will later see that the internal analogue of $\mathrm{Sub}_{\mathbf{E}}(A)$ is the power object PA and it too forms an appropriate notion of an *internal lattice*. In-fact, both $\mathrm{Sub}_{\mathbf{E}}(A)$ and PA forms a *Heyting algebra* and an *internal Heyting algebra*, respectively.

$\mathbf{Sub}_{\mathbf{E}}(1) \equiv \mathbf{Open}\ (\mathbf{E})$

There's more to the story than just the fact that all subobjects of an object forms a lattice. As we will see now, the subobject lattice of terminal object is a bit special in the sense that it is equivalent to a particular category of all objects which have a monic to the terminal. Let's retrospect this in the category of **Sets**, in which the terminal object is the singleton $\{\star\}$. Clearly, any subset of the singleton is either the singleton $\{\star\}$ itself or null-set ϕ. But note that $\{\star\}$ and ϕ are also the only sets with a monic arrow to $\{\star\}$. Therefore the result mentioned in the beginning clearly holds in the prototypical elementary topos **Sets**. Let's now prove this in any arbitrary elementary topos **E**:

Definition 13.3.2.5. (**Open Objects**) *Suppose* **E** *is a topos. Then an object B is an open object if the unique arrow to the terminal object is a monic, $B \rightarrowtail \mathbf{1}$.*

We then have the following:

Lemma 13.3.2.6. *Suppose* **E** *is a topos. An object U is open if and only if* $\mathrm{Hom}_{\mathbf{E}}(X, U) = \{\star\}\ \forall\ X \in \mathrm{Ob}(\mathbf{E})$.

Proof. U is open $\iff$ $m : U \rightarrowtail \mathbf{1}$ $\iff$ any parallel pair $x, y : X \to U$ would be such that $m \circ x = m \circ y$, but because m is monic, we get $x = y$. ☐

Proposition 13.3.2.7. *Suppose* **E** *is a topos. The lattice $Sub_{\mathbf{E}}(\mathbf{1})$ is equivalent to the full subcategory $Open\ (\mathbf{E})$ of open objects of* **E**, *i.e.*

$$Sub_{\mathbf{E}}(\mathbf{1}) \equiv Open\ (\mathbf{E}).$$

Proof. Since $\mathrm{Hom}_{\mathrm{Sub}(\mathbf{1})}(S, T) = \{\star\} = \mathrm{Hom}_{\mathrm{Open}\ (\mathbf{E})}(X, Y)$ where last equation follows from Lemma 13.3.2.6, therefore $\mathrm{Sub}_{\mathbf{E}}(\mathbf{1}) \equiv \mathrm{Open}\ (\mathbf{E})$. ☐

$\mathbf{Sub}_{\mathbf{E}}(B)$ is reflective into $\mathbf{E}/B$

We now show that the subobject category (lattice) is a reflective subcategory of the slice over that object.

Proposition 13.3.2.8. *Suppose* **E** *is a topos. For any object B in* **E**, *the inclusion $Sub_{\mathbf{E}}(B) \hookrightarrow \mathbf{E}/B$ has a left adjoint σ, i.e.*

$$\begin{tikzcd} \mathbf{E}/B \arrow[rr, bend left, "\sigma"] \arrow[rr, phantom, "\perp"] & & Sub_{\mathbf{E}}(B) \arrow[ll, bend left, "i"] \end{tikzcd}$$

where

$$\sigma : \mathbf{E}/B \longrightarrow Sub_{\mathbf{E}}(B)$$
$$(f : A \to B) \longmapsto m_f : S \rightarrowtail B, \textit{ where}$$

$$\begin{tikzcd} & S \arrow[dr, tail, "m_f"] & \\ A \arrow[ur, two heads, "e_f"] \arrow[rr, "f"'] & & B \end{tikzcd}.$$

Proof. Take any arrow $\sigma f \subset g$ in $\mathrm{Sub}_{\mathbf{E}}(B)$ where $f : A \to B$ is an object in $\mathbf{E}/B$ and $g : S \rightarrowtail B$ is an object in $\mathrm{Sub}_{\mathbf{E}}(B)$. Now note the following diagram:

$$\begin{tikzcd} S \arrow[d, tail, "g"'] & A \arrow[l, "h \circ e_f"] \arrow[r, two heads, "e_f"] \arrow[d, "f"] & T \arrow[ll, bend right, "h"'] \arrow[dl, tail, "\sigma f := m_f"] \\ B \arrow[r, equal] & B & \end{tikzcd}$$

where $h : T \to S$ is the arrow corresponding to the subobject inclusion $\sigma f \subset g$. We claim that the arrow $h \circ e_f : A \to S$ is the unique arrow $f \to ig$ in $\mathbf{E}/B$, because, firstly $(g \circ h) \circ e_f = (m_f) \circ e_f = f$ and, secondly, for any other arrow $k : A \to S$ with $g \circ k = f$, since $g \circ h \circ e_f = f$ too, therefore by monic nature of g, we would have $k = h \circ e_f$. □

13.3.3 Internal Structures

Any mathematical structure which admits definition in set theory can be translated into *internal object* in a suitable category with enough structure. One prime example of such internalization is the group object in a category:

Definition 13.3.3.1. (**Internal Group Object**) *Suppose* $\mathbf{C}$ *is a category with binary products and a terminal object. An object* G *in* $\mathbf{C}$ *is said to be a group object if there are:*

1. *An arrow* $m : G \times G \longrightarrow G$,
2. *An arrow* $i : G \longrightarrow G$,
3. *An arrow* $0 : \mathbf{1} \longrightarrow G$

which satisfy the following three commutative diagrams:

$$\begin{tikzcd} G \times G \times G \arrow[r, "m \times 1"] \arrow[d, "1 \times m"'] & G \times G \arrow[d, "m"] \\ G \times G \arrow[r, "m"'] & G \end{tikzcd}$$

Associativity

$(g+h)+k=g+(h+k)$

$$\begin{tikzcd} G \times G \arrow[r, "m"] & G \\ G \arrow[u, "\langle 1,i \rangle"] \arrow[ur, "0 \circ !_G"] \arrow[r, "\langle i,1 \rangle"'] & G \times G \arrow[u, "m"'] \end{tikzcd}$$

Unital

$g+(-g)=0=-g+g$

$$\begin{tikzcd} & G \times G \arrow[r, "m"] & G \arrow[r, dash, "\cong"] & G \times \mathbf{1} \\ G \times \mathbf{1} \arrow[r, dash, "\cong"'] & G \arrow[u, "\langle 1,0 \rangle"] \arrow[ur, "1"] \arrow[r, "\langle 0,1 \rangle"'] & G \times G \arrow[u, "m"'] & \end{tikzcd}$$

Identity

$g+0=g=0+g$

where $!_G : G \longrightarrow \mathbf{1}$ *is the unique terminal arrow.*

In a similar fashion, we define an internal meet semilattice in the truth object Ω in any topos, which would prove to be very beneficial in the following discussion on generalization of sheaves on an arbitrary topos.

However, the definition of this internal meet semilattice in Ω is not as direct as Definition 13.3.3.1.

Internal Meet in Ω

Definition 13.3.3.2. (**Internal Meet Semilattice**[8] **in** Ω) Suppose $\mathbf{E}$ is a topos and Ω is it's truth object. There is an arrow $\bigwedge : \Omega \times \Omega \longrightarrow \Omega$ which is given by the following construction:

1. By Theorem 13.3.2.3, 2, for any $k : A \to B$ in $\mathbf{E}$, $(k)^{-1}$ preserves finite limits, hence we get:
$$(k)^{-1}(S \cap T) \cong (k)^{-1}(S) \cap (k)^{-1}(T)$$
[9]for any two subobjects S, T of B.
2. This determines the following functor:
$$- \cap - : \mathrm{Sub}_{\mathbf{E}}(B) \times \mathrm{Sub}_{\mathbf{E}}(B) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(B)$$
which is natural in B, due to the 1.
3. We then have the following diagram:
$$\begin{tikzcd}
\mathrm{Sub}_{\mathbf{E}}(B) \times \mathrm{Sub}_{\mathbf{E}}(B) \arrow[r, "- \cap -"] \arrow[d, "\cong"] & \mathrm{Sub}_{\mathbf{E}}(B) \arrow[d, "\cong"] \\
\mathrm{Hom}_{\mathbf{E}}(B, \Omega) \times \mathrm{Hom}_{\mathbf{E}}(B, \Omega) \arrow[d, "\cong"] & \mathrm{Hom}_{\mathbf{E}}(B, \Omega) \arrow[d, equal] \\
\mathrm{Hom}_{\mathbf{E}}(B, \Omega \times \Omega) \arrow[r, "\bigwedge_B"'] & \mathrm{Hom}_{\mathbf{E}}(B, \Omega)
\end{tikzcd}$$
which gives rise to the above $\bigwedge_B : \mathrm{Hom}_{\mathbf{E}}(B, \Omega \times \Omega) \longrightarrow \mathrm{Hom}_{\mathbf{E}}(B, \Omega)$, which is also natural in B because $- \cap -$ was too.
4. Now, because $\bigwedge_B$ is natural in B, this would translate that $\bigwedge_{(-)}$ is the following natural transformation:
$$\bigwedge_{(-)} : \mathrm{Hom}_{\mathbf{E}}(-, \Omega \times \Omega) \Longrightarrow \mathrm{Hom}_{\mathbf{E}}(-, \Omega)$$
But by Yoneda Lemma, we would have that
$$\mathrm{Nat}(\mathrm{Hom}_{\mathbf{E}}(-, \Omega \times \Omega), \mathrm{Hom}_{\mathbf{E}}(-, \Omega)) \cong \mathrm{Hom}_{\mathbf{E}}(\Omega \times \Omega, \Omega)$$
Therefore by the above, the natural transformation $\bigwedge_{(-)}$ determines a unique arrow $\bigwedge$ as follows:
$$\bigwedge_{\Omega \times \Omega}(1_{\Omega \times \Omega}) = \bigwedge : \Omega \times \Omega \longrightarrow \Omega$$

[8]a poset which has a meet for any nonempty finite subset.
[9]The "$\cap$" is the meet of two subobjects in the lattice (Theorem 13.3.2.3,1).

The collection $(\Omega, \bigwedge, \mathsf{true} : \mathbf{1} \longrightarrow \Omega)$ forms what we call an internal meet semilattice in Ω.

Remark 13.3.3.3. (**$\bigwedge$ gives the characteristic of intersection**) The internal meet operation $\bigwedge$ in Ω gives the characteristic map of the intersection of two subobjects S, T of some arbitrary object B. That is, if $s, t : B \longrightarrow \Omega$ are characteristic maps for S and T respectively then, the characteristic map for $S \cap T$ would be given by:

$$B \xrightarrow{\langle s,t \rangle} \Omega \times \Omega \xrightarrow{\bigwedge} \Omega \ .$$

This follows from the construction in Definition 13.3.3.2 and the trivial natural isomorphism $\mathrm{Sub}_B(\mathbf{E}) \cong \mathrm{Hom}_{\mathbf{E}}(B, \Omega)$ (Definition 13.3.0.1, ETI.2).

Remark 13.3.3.4. (**Internal meet in PA**) Moreover, one can consider an internal meet not just in Ω but in any power object PA in a topos $\mathbf{E}$. The construction is roughly similar and is done in Proposition 13.3.5.9.

13.3.4 Slice Topos

One important result in the theory of topoi is that a slice of a topos is itself a topos. Moreover, the change of base functor between two slices provides a way of giving *sets-like* properties of topoi, which we will discuss later.

Theorem 13.3.4.1. *Suppose* $\mathbf{E}$ *is a topos. The slice category* $\mathbf{E}/B$ *is a topos for any object* B *in* $\mathbf{E}$.

Proof. Let us not derive power object here as it's construction is a bit involved and would defeat the purpose of the notes. It is done in detail in Theorem 1, pp 190 of [**MacMoer**]. Let's show that $\mathbf{E}/B$ has finite limits. For this, the terminal object in $\mathbf{E}/B$ is the identity $1 : B \to B$. The equalizer of two parallel arrows $a, c : f \to g$ where $f : A \to B, g : C \to B$ are two objects in $\mathbf{E}/B$ is just the equalizer of $a, c : A \to C$ in $\mathbf{E}$ itself. Binary product of $a : A \to B$ and $c : C \to B$ in $\mathbf{E}/B$ is given by their pullback in $\mathbf{E}$. Hence, $\mathbf{E}/B$ has all finite limits.
Next, the subobject classifier of $\mathbf{E}/B$ is given as follows: for a subobject $m : s \rightarrowtail a$ in $\mathbf{E}/B$ where $s : S \to B$ and $a : A \to B$ is given by the following diagram:

$$\begin{tikzcd}
S \arrow[rr, "!_S"] \arrow[dd, tail, "m"'] \arrow[dr, "s"] & & B \arrow[r, dash, "\cong"] \arrow[dl, "1"'] \arrow[dd, tail, "\mathsf{true}_{\mathbf{E}/B}"] & \mathbf{1} \times B \arrow[dd, tail, "\langle \mathsf{true}, 1\rangle"] \\
& B & & \\
A \arrow[rr, "\mathrm{char}\ m"'] \arrow[ur, "a"] & & \Omega \times B \arrow[r, equal] \arrow[ul, "\pi_2"'] & \Omega \times B
\end{tikzcd}$$

Hence, the subobject classifier in $\mathbf{E}/B$ is $\langle \mathsf{true}, 1\rangle : \mathbf{1} \times B \rightarrowtail \Omega \times B$. ☐

Change of base functor

The following theorem would be used to derive *sets-like* properties of a topos:

Theorem 13.3.4.2. *Suppose* $\mathbf{E}$ *is a topos and* $k : B \longrightarrow A$ *is any arrow in it. The change of base functor defined as:*

$$k^* : \mathbf{E}/A \longrightarrow \mathbf{E}/B$$

which takes an object of $\mathbf{E}/A$ *to it's pullback along* k*, that is:*

$$\begin{tikzcd}
X \arrow[d, "x"'] & & k^*(X) \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "x"] \\
A \arrow[r, phantom, ""] & {} \arrow[u, phantom, ""] & B \arrow[r, "k"'] & A
\end{tikzcd} \qquad X \overset{x}{\to} A \;\longmapsto\; k^*(X) \overset{\pi_1}{\to} B$$

has both left and right adjoints where left adjoint is given by pre-composition with k*:*

$$\begin{tikzcd}
\mathbf{E}/B \arrow[rr, bend left, "\sum_k := k\circ -"] \arrow[rr, phantom, "\perp"] & & \mathbf{E}/A \arrow[ll, bend left, "k^*"]
\end{tikzcd} \qquad \begin{tikzcd}
\mathbf{E}/A \arrow[rr, bend left, "k^*"] \arrow[rr, phantom, "\perp"] & & \mathbf{E}/B \arrow[ll, bend left, "\prod_k"]
\end{tikzcd} .$$

Moreover, k^* *is a logical morphism of topoi (Definition 13.3.1.6).*

Proof. Since a slice category in a topos is also a topos (Theorem 13.3.4.1), therefore each slice is cartesian closed, so we have both the adjoints by Theorem 4, Chapter I, pp 59 of [**MacMoer**]. It remains to be seen that k^* is a logical morphism.
Since k^* is a right adjoint, therefore it preserves all finite limits. To see that k^* preserves the subobject classifier, first take the pullback of each member of the subobject classifier diagram in $\mathbf{E}/A$ so that we obtain a diagram in $\mathbf{E}/B$. Since pullbacks preserves monic, so we just need to show that $B \times_A (\Omega \times A) \cong \Omega \times B$, where Ω is the truth object of $\mathbf{E}$. Now since it can be verified quite easily by universality of pullbacks that $B \times_A (\Omega \times A) \cong \Omega \times (B \times_A A) \cong \Omega \times B$, we hence have that k^* preserves the subobject classifier. Lastly, we need to show that k^* preserves exponentials, that is, there is a following natural isomorphism:

$$k^*(-) \circ (-)^x \cong (-)^{k^*x} \circ k^*(-)$$

where $y : Y \to A$ and $x : X \to A$ are objects in slice $\mathbf{E}/A$. If we could show that the corresponding left adjoints of the above equation is also isomorphic, then the result would follow. Therefore the left transpose equivalent to be shown is:

$$(X \times_A -) \circ \Sigma_k(-) \cong \Sigma_k(-) \circ (k^*(X) \times_B -).$$

For some object $z : Z \to B$ in $\mathbf{E}/B$, we have by definition of Σ_k:

$$(X \times_A -)(\Sigma_k z) = (X \times_A -)(k \circ z) = \overline{k \circ z}_x : X \times_A Z \to X$$

where $\overline{k \circ z}_x$ is the pullback of $k \circ z : Z \to A$ along $x : X \to A$. Similarly, we have:

$$\Sigma_k(-) \circ (k^*(X) \times_B z) = \Sigma_k\left(k^*(X) \times_B z\right) = k \circ \left(\overline{z}_{\overline{x}_k}\right)$$

where $\overline{z}_{\overline{x}_k}$ is the pullback of $z : Z \to B$ along the pullback of x along k. Clearly, both the objects are isomorphic. Hence the left adjoint commutes, then so does the right adjoint. ☐

"Sets-like" properties of topoi

From Theorem 13.3.4.2, we can derive various *sets-like* properties of topoi, the proofs of all which depends on the translation of the problem from the topos $\mathbf{E}$ to some appropriate slice topos and using the fact that pullback would preserve both finite limits and colimits and would be a logical morphism. We derive some such results below.

The first result draws motivation from the fact that in **Sets**, for a surjective function $e : A \to B$ and any function $f : C \to B$, the set $C \times_B A = \{(c, a) \in C \times A \mid e(a) = f(c)\}$ has an obvious surjection to C given by $(c, a) \mapsto c$ as for all $c \in C$, $\exists a \in A$ with $f(c) = e(a)$ because $\mathrm{im}\,(f) \subseteq B = \mathrm{im}\,(e)$. It's generalization in an arbitrary topos is the following:

Proposition 13.3.4.3. *In a topos* $\mathbf{E}$*, the pullback of an epimorphism is an epimorphism.*

Proof. Suppose $e : A \to B$ is an epimorphism in a topos $\mathbf{E}$. The fact that e is an epimorphism can be equivalently stated as the pushout condition on the left below:

$$\begin{tikzcd} B & B \arrow[l, "1"'] \\ B \arrow[u, "1"] \arrow[ur, phantom, "\ulcorner"] & A \arrow[l, two heads, "e"] \arrow[u, two heads, "e"'] \end{tikzcd} \xrightarrow{\text{In Slice } \mathbf{E}/B} \begin{tikzcd} 1_B & 1_B \arrow[l, "1"'] \\ 1_B \arrow[u, "1"] \arrow[ur, phantom, "\ulcorner"] & e \arrow[l, two heads, "e"] \arrow[u, two heads, "e"'] \end{tikzcd}$$

Now treat e as an object in the slice $\mathbf{E}/B$, where this pushout condition would translate to the pushout condition on the right. For any arrow $f : C \to B$, the corresponding change of base $f^* : \mathbf{E}/B \longrightarrow \mathbf{E}/C$ preserves colimits (Theorem 13.3.4.2), therefore we would have the same pushout condition in $\mathbf{E}/C$ by the pullback along f. We hence have that epics are pullback stable in a topos. □

Next, consider the initial object in **Sets**, which is the null set $\varnothing$. By definition, any map $f : A \to \varnothing$ is an isomorphism by default. In general:

Proposition 13.3.4.4. *In a topos* $\mathbf{E}$*, any arrow* $k : A \to 0$ *is an isomorphism.*

Proof. Denote $!_A : 0 \to A$ as the unique initial arrow. Since $k \circ !_A = 1_0$ by uniqueness, we hence need to show $!_A \circ k = 1_A$. Focus on the objects k and 1_0 in the topos $\mathbf{E}/0$. Clearly, 1_0 is the initial object in $\mathbf{E}/0$. But it is also terminal in this slice. Now, since the pullback $!_A : 0 \to A$ of $1_0 : 0 \to 0$ along $k : A \to 0$ would also be initial and final in $\mathbf{E}/A$, then, because $1_A : A \to A$ and $k : A \to 0$ forms a cone over this pullback, therefore we get $!_A \circ k = 1_A$. □

Another consequence in **Sets** of null-set $\varnothing$ is that the unique arrow $f : \varnothing \to A$ is always injective. In general:

Corollary 13.3.4.5. *In a topos* $\mathbf{E}$*, the unique arrow* $!_A : 0 \longrightarrow A$ *for any object* A *is a monomorphism.*

Proof. Let $x, y : B \to 0$ be two arrows such that $!_A \circ x = !_A \circ y$. But since x and y are isomorphisms by Proposition 13.3.4.4, therefore, $(x)^{-1}, (y)^{-1} : 0 \to B$ are two arrows from initial 0, hence $(x)^{-1} = (y)^{-1} \implies x = y$, so $!_A$ is a monic. □

In **Sets**, the product of two surjective functions is also surjective. This holds in general:

Proposition 13.3.4.6. *In a topos* $\mathbf{E}$*, if* $f : X \to Y$ *and* $g : W \to Z$ *are epimorphisms, then* $f \times g : X \times W \to Y \times Z$ *is also an epimorphism.*

Proof. Note $f \times g = (f \times 1_Z) \circ (1_X \times g)$. So if we could show that $f \times 1_Z$ and $1_X \times g$ are epics then we would be done. To this end, note that $(X \times Z) \times_Z W \cong X \times W$ follows from universality of pullbacks and products, and then we have the following diagram:

$$\begin{tikzcd}
X \times W \arrow[r, "\cong", no head] \arrow[d, "1_X \times g"', two heads] & (X \times Z) \times_Z W \arrow[r, "\pi_2"] \arrow[d, "\pi_1"', two heads] \arrow[dr, phantom, "\lrcorner", very near start] & W \arrow[d, "g", two heads] \\
X \times Z \arrow[r, equal] & X \times Z \arrow[r, "p_2"'] & Z
\end{tikzcd}$$

where $\pi_1 \cong 1_X \times g$. But by Proposition 13.3.4.3, π_1 is an epimorphism, then so is $1_X \times g$. Similarly, $f \times 1_Z$ is also an epimorphism, so we have our result. □

We finally have the following important result:

Theorem 13.3.4.7. *In a topos* $\mathbf{E}$*, every epimorphism is the coequalizer of it's kernel pair.*

Proof. Suppose $f : C \to B$ is an epimorphism. Let $\pi_1, \pi_2 : C \times_B C \to C$ be the kernel pair of f. The coequalizer of this kernel pair would be denoted:

$$C \times_B C \underset{\pi_2}{\overset{\pi_1}{\rightrightarrows}} C \xrightarrow{c} Q \ .$$

Let's translate this coequalizer to a diagram in slice topos $\mathbf{E}/B$. For this, the object would be the epic $f : C \to B$. Also notice that f being epi means that $!_f : f \to 1$ is an epic in $\mathbf{E}/B$. Take the product $f \times f$ in $\mathbf{E}/B$. By Theorem 13.3.4.1, the product is exactly the pullback of f along itself. The projections $f \times f \to f$ is therefore exactly the π_1 and π_2 as above. Hence, to find the coequalizer in $\mathbf{E}$ of pullback projections $\pi_1, \pi_2 : C \times_B C \to C$ is same as finding coequalizer of product projections $\pi_1, \pi_2 : f \times f \to f$ in $\mathbf{E}/B$.
Now, for any topos $\mathbf{F}$, let X be an object and $p_1, p_2 : X \times X \to X$ be the projection of the product. Let the coequalizer of p_1, p_2 be denoted as L as shown:

$$X \times X \underset{p_2}{\overset{p_1}{\rightrightarrows}} X \xrightarrow{v} L$$

We wish to show that when $!_X : X \to 1$ is an epic, then $L \cong 1$. Our motivation for this stems from the fact that $!_f : f \to 1$ is an epic in $\mathbf{E}/B$ and $L \cong 1$ would mean directly that $f : C \to B$ would be the coequalizer of $\pi_1, \pi_2 : C \times_B C \to C$. To this end, we use the fact that in a topos, any arrow which is both monic and epic is an isomorphism. $!_L : L \to 1$ is monic, because if we let two arrows x and y in $\mathbf{F}$ with domain some arbitrary K be such that $!_L \circ x = !_L \circ y$, which just means that $!_K = !_K$, then we have the following:

$$\begin{tikzcd}
 & K \arrow[d, "\langle x,y \rangle"] & \\
X \times X \arrow[r, two heads, "v \times v"] \arrow[d, "p_i"'] & L \times L \arrow[d, "q_i"] & i = 1, 2 \\
X \arrow[r, two heads, "v"'] & L &
\end{tikzcd}$$

Remember the coequalizer (here, q) is always an epimorphism and so $q \times q$ is an epimorphism by Proposition 13.3.4.6. Since both the squares commute, $q_1 \circ (v \times v) = v \circ p_1 = v \circ p_2 = q_2 \circ (v \times v) \implies q_1 = q_2$. Therefore $q_1 \circ (\langle x, y \rangle) = q_2 \circ (\langle x, y \rangle) \implies x = y$, so that $!_L : L \to 1$ is a monic. Now to show $!_L$ is an epic, if $a \circ !_L = b \circ !_L$, then $a \circ !_L \circ v = b \circ !_L \circ v \implies a \circ !_X = b \circ !_X \implies a = b$ as $!_X : X \to 1$ is a given epic. Hence $L \cong 1$. □

13.3.5 Internal Lattices, Heyting Algebras and Subobject Lattice

We saw in Section 13.3.3 that one can define internal group objects and internal meet "in" some object. In particular, we saw internal meet in the truth object Ω, which determines the characteristic of intersection or meet of two subobjects in a given subobject lattice $\mathrm{Sub}_B(\mathbf{E})$. This would be important to define the closure of a subobject when generalizing sheaves over arbitrary topoi.
We now continue this line of *internalization* of algebraic structures and, in the same vein as internal group object, introduce lattice and Heyting algebra objects and then study the "external" subobject lattice $\mathrm{Sub}_{\mathbf{E}}(A)$ in more detail.

Internal Lattices

Definition 13.3.5.1. (**Internal Lattice**) *Suppose* **C** *is a category with finite limits. An internal lattice or a lattice object* $(L, \wedge, \vee)$ *is an object* L *in* **C** *with two arrows*

$$\bigwedge : L \times L \longrightarrow L \quad \& \quad \bigvee : L \times L \longrightarrow L$$

which satisfy the following commutative diagrams:

Associativity

$x\wedge(y\wedge z)=(x\wedge y)\wedge z$

$$\begin{tikzcd}
L \times L \times L \arrow[r, "1\times\wedge"] \arrow[d, "\wedge\times 1"'] & L \times L \arrow[d, "\wedge"] \\
L \times L \arrow[r, "\wedge"'] & L
\end{tikzcd}$$

$x\vee(y\vee z)=(x\vee y)\vee z$

$$\begin{tikzcd}
L \times L \times L \arrow[r, "1\times\vee"] \arrow[d, "\vee\times 1"'] & L \times L \arrow[d, "\vee"] \\
L \times L \arrow[r, "\vee"'] & L
\end{tikzcd}$$

Commutativity

$x\wedge y=y\wedge x$

$$\begin{tikzcd}
L \times L \arrow[r, "\wedge"] \arrow[d, "\tau"'] & L \\
L \times L \arrow[ur, "\wedge"'] &
\end{tikzcd}$$

$x\vee y=y\vee x$

$$\begin{tikzcd}
L \times L \arrow[r, "\vee"] \arrow[d, "\tau"'] & L \\
L \times L \arrow[ur, "\vee"'] &
\end{tikzcd}$$

Idempotency

$x\wedge x=x$

$$\begin{tikzcd}
L \times L \arrow[r, "\wedge"] \arrow[d, "1\times 1"'] & L \\
L \times L \arrow[ur, "\wedge"'] &
\end{tikzcd}$$

$x\vee x=x$

$$\begin{tikzcd}
L \times L \arrow[r, "\vee"] \arrow[d, "1\times 1"'] & L \\
L \times L \arrow[ur, "\vee"'] &
\end{tikzcd}$$

Absorption

$x\wedge(y\vee x)=x=(x\wedge y)\vee x$

$$\begin{tikzcd}
L & & L \times L \arrow[ll, "\vee"'] \\
L \times L \arrow[u, "p_1"] \arrow[d, "p_1"'] \arrow[r, "\Delta_L\times 1"'] & L \times L \times L \arrow[r, "1\times\tau"'] & L \times L \times L \arrow[u, "1\times\wedge"'] \arrow[d, "\vee\times 1"] \\
L & & L \times L \arrow[ll, "\wedge"]
\end{tikzcd}$$

where $\Delta_L := \langle 1, 1\rangle : L \longrightarrow L \times L$ *is the diagonal map and* $\tau : L \times L \longrightarrow L \times L$ *is the twist arrow, given as* $\tau := \langle p_2, p_1\rangle$ *where* p_1, p_2 *are the projection arrows of* $L \times L$.

An internal lattice can also have *top* and *bottom* elements:

Definition 13.3.5.2. (**Internal Lattice with $\top$ and $\bot$**) *Suppose category* $\mathbf{C}$ *has finite limits and* L *is an internal lattice in* $\mathbf{C}$. L *is said to be an internal lattice with top* ($\top$) *and bottom* ($\bot$) *elements if there are additionally two arrows* $\top : 1 \longrightarrow L$ *and* $\bot : 1 \longrightarrow L$ *which satisfy the following commutative diagrams:*

$$
\begin{tikzcd}
 & L \times L \arrow[r, "\wedge"] & L \\
L \arrow[r, dash, "\cong"'] & 1 \times L \arrow[u, "\top\times 1"] \arrow[ur, "1"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & L \times L \arrow[r, "\vee"] & L \\
L \arrow[r, dash, "\cong"'] & 1 \times L \arrow[u, "\bot\times 1"] \arrow[ur, "1"'] &
\end{tikzcd}
.
$$

$\top \wedge x = x$ $\qquad$ $\bot \vee x = x$

We now define an internal Heyting algebra object:

Internal Heyting Algebras

Definition 13.3.5.3. (**Internal Heyting Algebra - I**) *Suppose* $\mathbf{C}$ *is a category with finite limits and* H *is an internal lattice with* $\top$ *and* $\bot$ *in* $\mathbf{C}$. H *is then said to be an internal Heyting algebra if there is an additional arrow* $(\Rightarrow) : H \times H \longrightarrow H$ *such that the following commutes*

$$
\begin{tikzcd}
H \times H \arrow[r, "(\Rightarrow)"] & H \\
H \arrow[u, "\Delta_H"] \arrow[ur, "\top\circ !_H"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
H \times H \times H \arrow[r, "1\times(\Rightarrow)"] & H \times H \arrow[d, "\wedge"] \\
H \times H \arrow[u, "\Delta_H\times 1"] \arrow[r, "\wedge"'] & H
\end{tikzcd}
$$

$(x \Rightarrow x) = \top$ $\qquad$ $x \wedge (x \Rightarrow y) = x \wedge y$

$$
\begin{tikzcd}
H \times H \times H \arrow[r, "1\times\tau"] & H \times H \times H \arrow[r, "1\times(\Rightarrow)"] & H \times H \arrow[d, "\wedge"] \\
H \times H \arrow[u, "\Delta_H\times 1"] \arrow[rr, "p_1"'] & & H
\end{tikzcd}
\qquad
\begin{tikzcd}
H \times H \times H \times H \arrow[r, "(\Rightarrow)\times1\times1"] & H \times H \times H \arrow[r, "1\times(\Rightarrow)"] & H \times H \arrow[dd, "p_2"] \\
H \times H \times H \times H \arrow[u, "1\times\tau\times1"] & & \\
H \times H \times H \arrow[u, "\Delta_H\times1\times1"] \arrow[r, "1\times\wedge"'] & H \times H \arrow[r, "(\Rightarrow)"'] & H
\end{tikzcd}
$$

$x \wedge (y \Rightarrow x) = x$ $\qquad$ $x \Rightarrow (y \wedge z) = (x \Rightarrow y) \wedge (x \Rightarrow z)$

where $\Delta_H : H \longrightarrow H \times H$ *is the diagonal map and* $\tau : H \times H \longrightarrow H \times H$ *is the twist map.*

For any lattice $(S, \wedge, \vee)$, we have an partial order induced by the meet, given by $x \leq y \iff x = x \wedge y$. We can do the same in an internal lattice:

Definition 13.3.5.4. (**Internal Partial Order in an Internal Lattice**) *Suppose* $(L, \wedge, \vee)$ *is an internal lattice in a category* $\mathbf{C}$ *with finite limits. Then* $(L, \leq_L)$ *is called an internal partial order in* L *or an internal* *poset where the internal order in L,$\leq_L$, is given by the following equalizer diagram*[10]*:*

$$\leq_L \overset{e}{\rightarrowtail} L \times L \underset{p_1}{\overset{\wedge}{\rightrightarrows}} L \ .$$

Remark 13.3.5.5. Note that internal partial order $\leq_L$ is hence a subobject of $L \times L$.

One may remember from the usual set-theoretic definition of a Heyting algebra $(H, \wedge, \vee, \top, \bot, \Rightarrow)$ that it was just a lattice with top and bottom where each object additionally was exponentiable, meaning that for the partial order induced from the meet of the lattice, $\forall x, y \in H$, $\exists (x \Rightarrow y) \in H$ such that for any $z \in H$:

$$z \leq (x \Rightarrow y) \text{ if and only if } z \wedge x \leq y.$$

Since by Definition 13.3.5.4, we now have a way to induce the internal partial order in an internal lattice, we can hence redefine internal Heyting algebra as follows:

Definition 13.3.5.6. (**Internal Heyting Algebra - II**) *Suppose* **C** *is a category with finite limits. Let* $(H, \wedge, \vee, \top, \bot)$ *be an internal lattice with* $\top$ *and* $\bot$ *in* **C**. H *is said to be an internal Heyting algebra if there exists an additional arrow* $(\Rightarrow) : H \times H \longrightarrow H$ *such that the subobjects* P *and* Q *are equivalent in the following:*

$$\begin{tikzcd}
P \arrow[r] \arrow[d, tail, dashed] \arrow[dr, phantom, "\lrcorner", very near start] & \leq_H \arrow[d, tail, "e"] & Q \arrow[l] \arrow[d, tail, dashed] \arrow[dl, phantom, "\llcorner", very near start] \\
H \times H \times H \arrow[r, "\wedge \times 1"'] & H \times H & H \times H \times H \arrow[l, "1 \times (\Rightarrow)"]
\end{tikzcd} .$$

$$x \wedge y \leq z \qquad\qquad x \leq y \Rightarrow z$$

A lattice homomorphism $f : (A, \wedge, \vee, \top, \bot) \longrightarrow (B, \wedge, \vee, \top, \bot)$ is defined to be the one which respects all the structure ($\wedge$ and $\vee$) and preserves $\top$ and $\bot$. We can similarly define internal lattice homomorphism as the following:

Definition 13.3.5.7. (**Internal Lattice homomorphism**) *Suppose* **C** *is a category with finite limits and* L, L' *are two internal lattices with* $\top$ *and* $\bot$. *An arrow* $f : L \longrightarrow L'$ *is said to be an internal lattice*

[10] The diagram is the *internal way* of saying that for $x, y \in L$, $x \leq y \iff x = x \wedge y$.

homomorphism if the following diagrams commute[11]:

$$\begin{tikzcd} L & L \times L \arrow[l, "\wedge"'] \\ L \arrow[u, leftarrow, "f"] & L \times L \arrow[l, "\wedge"] \arrow[u, leftarrow, "f\times f"'] \end{tikzcd} \qquad \begin{tikzcd} L & L \times L \arrow[l, "\vee"'] \\ L \arrow[u, leftarrow, "f"] & L \times L \arrow[l, "\vee"] \arrow[u, leftarrow, "f\times f"'] \end{tikzcd}$$

$f(x \wedge y) = f(x) \wedge f(y)$ $\qquad$ $f(x \vee y) = f(x) \vee f(y)$

$$\begin{tikzcd} 1 \arrow[r, "\top"] \arrow[d, "\top"'] & L \\ L \arrow[ur, "f"'] & \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r, "\perp"] \arrow[d, "\perp"'] & L \\ L \arrow[ur, "f"'] & \end{tikzcd}$$

$f(\top) = \top$ $\qquad$ $f(\perp) = \perp$

Subobject Lattices, Internal & External

We now discuss the two subobject lattices, the external, $\mathrm{Sub}_{\mathbf{E}}(A)$, and the internal, PA. We first begin by showing that $\mathrm{Sub}_{\mathbf{E}}(A)$ is in-fact a Heyting Algebra, which extends the result obtained in Theorem 13.3.2.3, 1.

Proposition 13.3.5.8. *Suppose* $\mathbf{E}$ *is a topos. The lattice* $\mathrm{Sub}_{\mathbf{E}}(A)$ *of subobjects of object* A *is a Heyting algebra where the top and bottom elements are* $\perp : 0 \rightarrowtail A$[12] *and* $\top := 1 : A \rightarrowtail A$*. Moreover, for any* $k : A \to B$ *in* $\mathbf{E}$*, the functor* $(k)^{-1} : \mathrm{Sub}_{\mathbf{E}}(B) \to \mathrm{Sub}_{\mathbf{E}}(A)$ *is a Heyting algebra homomorphism.*

Proof. First note that $\mathrm{Sub}_{\mathbf{E}}(A) \cong \mathrm{Sub}_{\mathbf{E}/A}(1)$ therefore it is enough to talk about the subobjects of identity in slice $\mathbf{E}/A$. To construct the $(\Rightarrow)$ in $\mathrm{Sub}_{\mathbf{E}/A}(1)$, take any two open objects[13] $U \rightarrowtail 1$ and $V \rightarrowtail 1$ in the $\mathrm{Sub}_{\mathbf{E}/A}(1)$. Clearly, U^V is also open, so $U^V \rightarrowtail 1$ in $\mathrm{Sub}_{\mathbf{E}/A}(1)$. The corresponding $U^V \rightarrowtail A$ in $\mathbf{E}$ is the required exponential.
For next result, take an arrow $k : A \to B$ in $\mathbf{E}$. Since change of base functor $k^* : \mathbf{E}/B \longrightarrow \mathbf{E}/A$ is a logical morphism and preserves limits and colimits (Theorem 13.3.4.2), therefore the corresponding arrow $(k)^{-1} : \mathrm{Sub}_{\mathbf{E}}(B) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(A)$ is also structure preserving as it preserves the meet (limit), join (image), top & bottom (by functoriality) and exponents (logical morphism). $\square$

As to what we eluded earlier in remark of Theorem 13.3.2.3, we have proved the external part of it in the Proposition 13.3.5.8. The remaining thing to do is to show that the internal subobject lattice PA is also an internal Heyting algebra. This is exactly what we do now.

Proposition 13.3.5.9. *Suppose* $\mathbf{E}$ *is a topos. The power object* PA *for any object* A *is an internal Heyting algebra. Moreover, for any arrow* $k : A \to B$ *in* $\mathbf{E}$*, the map* $Pk : PB \to PA$ *is an internal Heyting algebra homomorphism.*

[11] The diagrams are the *internal way* of saying that f preserves all structure.
[12] See Corollary 13.3.4.5.
[13] See Definition 13.3.2.5.

Proof. The proof follows the canonical construction of Definition 13.3.3.2. Construct internal meet in PA as the arrow $\bigwedge : PA \times PA \longrightarrow PA$ given by the Yoneda lemma on the natural transformation $\bigwedge_{(-)} : \mathrm{Hom}_{\mathbf{E}}(-, PA \times PA) \Rightarrow \mathrm{Hom}_{\mathbf{E}}(-, PA)$. Similarly for $\bigvee$. Now for internal $(\Rightarrow) : PA \times PA \longrightarrow PA$, we again follow the same construction on the $(\Rightarrow)'$ of the Heyting algebra $\mathrm{Sub}_{\mathbf{E}}(A \times X)$, which gives the following implication arrow in external subobject lattice:

$$(\Rightarrow)' : \mathrm{Sub}_{\mathbf{E}}(A \times X) \times \mathrm{Sub}_{\mathbf{E}}(A \times X) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(A \times X)$$

which can be seen to give rise to a natural transform $(\Rightarrow)'_{(-)} : \mathrm{Hom}_{\mathbf{E}}(-, PA \times PA) \longrightarrow \mathrm{Hom}_{\mathbf{E}}(-, PA)$ which then by Yoneda lemma gives an arrow $(\Rightarrow) : PA \times PA \longrightarrow PA$ which is the required implication for internal Heyting algebra PA. Since top element of $\mathrm{Sub}_{\mathbf{E}}(A)$ is the $1 : A \rightarrowtail A$, which gives the corresponding arrow $1 \to PA$ by the natural isomorphism $\mathrm{Sub}_{\mathbf{E}}(A \times 1) \cong \mathrm{Hom}_{\mathbf{E}}(1, PA)$. Similarly for $\bot$. To see the Pk is an internal Heyting algebra homomorphism, we use generalized elements and argue the $\mathrm{Hom}_{\mathbf{E}}(X, Pk)$ is induces an external Heyting algebra homomorphism between $\mathrm{Sub}_{\mathbf{E}}(B \times X)$ and $\mathrm{Sub}_{\mathbf{E}}(A \times X)$. This can be seen easily since $\mathrm{Sub}_{\mathbf{E}}(B \times X) \cong \mathrm{Hom}_{\mathbf{E}}(X, PB)$ and similarly for PA. Since the square thus formed commutes, therefore we have an external Heyting algebra homomorphism by externalization of Pk, which means Pk is an internal Heyting algebra homomorphism. □

Remark 13.3.5.10. (**Internal logic of a topos is Intuitionistic**) What we have just proved in Propositions 13.3.5.8 & 13.3.5.9 is a very striking fact that in a topos, all the subobjects of an object forms a Heyting algebra instead of a Boolean algebra. This is striking because law of excluded middle (*either there is something or nothing*, more concretely, $x \vee \neg x = \top$ or $\neg\neg x = x$) does not hold in a Heyting algebra. This means that for a subobject S of A in a topos, $S \vee \neg S \neq \top$, which when unraveled means $S \vee (S \Rightarrow 0) \neq A$ since $\neg S := (S \Rightarrow \bot)$ and $\top := 1 : A \rightarrowtail A$ and $\bot : 0 \rightarrowtail A$.
For example, the topos **Sets** is such that the subobject lattice in **Sets** forms a Boolean algebra since $S \cup S^{\complement} = A$ for $S \subset A$. But the fact that this doesn't hold in an arbitrary topos suggests that a topos is a **generalized universe to do sets-like mathematics**.

This concludes the basic properties of topoi. We now study how one can generalize the concept of a topology, and therefore a sheaf, to an arbitrary topos.

13.4 Sheaves in an arbitrary Topos

We have studied two notions of sheaves, one on a topological space X, whose sheaf category is denoted $\mathrm{Sh}(X)$ and the other one on a site $(\mathbf{C}, J)$, whose sheaf category is denoted $\mathrm{Sh}(\mathbf{C}, J)$. Both times we saw that the sheaf category is a reflective subcategory of $\widehat{\mathbf{O(X)}}$ (for $\mathrm{Sh}(X)$, Theorem **??**) and $\widehat{\mathbf{C}}$ (for $\mathrm{Sh}(\mathbf{C}, J)$, Theorem 13.2.5.8). We now generalize the notion of a sheaf to an arbitrary topos. We will see that the same relations of the sheaf category in a topos and the underlying topos holds as one eluded to earlier. We would in-fact see that the notion of a sheaf in a topos is indeed a generalization of sheaves over a site, and hence over a topological space. But the interesting observation would here be that we do not access the "space" itself in the following generalization of sheaves. That is, we assume that our arbitrary topoi acts as if it is a presheaf topos of some notion of "generalized space" and this "space" is completely inaccessible to us[14].

[14]We would later see how to "access" it, via what we would call *points of a topos*.

13.4.1 The Lawvere-Tierney Topology on a Topos

Of-course, to define a sheaf, we would first need to define a notion of a cover. But since we are working with a given topos as if it were a presheaf topoi, so we try to look at what properties are sufficient to give a "topology" on that "generalized space" over whom we consider our topos is actually it's presheaf topos.

To see what are the sufficient conditions to identify a topology on the "generalized space" while only having access to it's presheaf topos, we first look at the example of sheaves over a site $(\mathbf{C}, J)$. That is, we first study the notion of a Grothendieck topology J but not looking at the "space" $\mathbf{C}$, but looking at the effect of J on the presheaf category $\widehat{\mathbf{C}}$.
First note that the presheaf topos has the subobject classifier Ω which maps each object of $\mathbf{C}$ to the corresponding collection of sieves on it. It is also quite easy to see that a sieve S over C is a J-covering sieve if and only if $\bar{S}$ is maximal. Therefore J determines a natural transformation $j : \Omega \Rightarrow \Omega$ where $j_C(S) := \bar{S}$ where $\bar{S} := \{f : \mathrm{dom}\,(f) \to C \mid f^*(S) \in J\mathrm{dom}\,(f)\}$ (See footnote 12). Let's analyze this natural transformation $j : \Omega \Rightarrow \Omega$. First, if $\mathsf{true} : \mathbf{1} \Rightarrow \Omega$ is the subobject classifier of $\widehat{\mathbf{C}}$, then $j \circ \mathsf{true} = \mathsf{true}$ because $\mathsf{true}_C(\star) := S_C^{\max}$ (Section 13.1.1). Next, we note that $j \circ j = j$ and this is obvious. Finally, note that $j_C(S \cap T) = \overline{S \cap T}$ is such that for any $f \in \overline{S \cap T}$, $f^*(S \cap T) \in J\mathrm{dom}\,(f)$ which implies that $f^*(S) \cap f^*(T) \in J\mathrm{dom}\,(f)$ which further means that $f^*(S), f^*(T) \in J\mathrm{dom}\,(f)$ so that $f \in \bar{S} \cap \bar{T}$. Similarly, for $f \in \bar{S} \cap \bar{T}$, we get $f^*(S), f^*(T) \in J\mathrm{dom}\,(f)$ therefore $f^*(S \cap T) = f^*(S) \cap f^*(T) \in J\mathrm{dom}\,(f)$. Therefore $j_C(S \cap T) = j_C(S) \cap j_C(T)$.
Therefore, we are motivated to define the following axioms for a "topology" on a topos $\mathbf{E}$. The topology, is in-fact on the underlying space, but if see j as above on the presheaf topos, then we can safely say that it corresponds to a topology in that underlying generalized space.

Definition 13.4.1.1. (**Lawvere-Tierney Topology**) *Suppose $\mathbf{E}$ is a topos and Ω is it's truth object. An arrow $j : \Omega \longrightarrow \Omega$ is called a Lawvere-Tierney topology on $\mathbf{E}$ if j satisfies the following commutative diagrams:*

LTT.1

$$\begin{tikzcd} \mathbf{1} \arrow[d, tail, "\mathsf{true}"'] \arrow[dr, tail, "\mathsf{true}"] & \\ \Omega \arrow[r, "j"'] & \Omega \end{tikzcd}$$

$j \circ \mathsf{true} = \mathsf{true}$

LTT.2

$$\begin{tikzcd} \Omega \arrow[d, "j"'] \arrow[dr, "j"] & \\ \Omega \arrow[r, "j"'] & \Omega \end{tikzcd}$$

$j \circ j = j$

LTT.3

$$\begin{tikzcd} \Omega \times \Omega \arrow[r, "\wedge"] \arrow[d, "j \times j"'] & \Omega \arrow[d, "j"] \\ \Omega \times \Omega \arrow[r, "\wedge"'] & \Omega \end{tikzcd}$$

$j \circ \wedge = \wedge \circ (j \times j)$

where $\wedge : \Omega \times \Omega \longrightarrow \Omega$ is the internal meet in Ω as in Definition 13.3.3.2.

Remark 13.4.1.2. As mentioned in Definition 13.3.3.2, the truth object in a topos defines an internal meet-semilattice object $(\Omega, \wedge, \mathsf{true} : \mathbf{1} \longrightarrow \mathsf{true})$ where true is the top element. Now, the axiom LTT.1 says that j preserves 0-ary meet/top element in Ω. LTT.3 on the other hand tells us that j preserves the meet in Ω and LTT.2 says that j is idempotent. Hence, in accordance with Definition 13.3.5.7, we can say that a Lawvere-Tierney topology on a topos $\mathbf{E}$ is equivalently **an idempotent internal meet-semilattice endomorphism on** Ω, where Ω is the truth object of $\mathbf{E}$.

Example 13.4.1.3. An easy to see example of this is the arrow $j : \Omega \Rightarrow \Omega$ in $\mathbf{Sets}^{\mathbf{O(X)}^{\mathrm{op}}}$ which is the characteristic arrow of the subobject $J : \mathbf{O(X)}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ which takes an open set $U \subseteq X$ to the collection of all those sieves over U which forms an open cover of U. In particular, this natural transformation j is given by components $j_U : \Omega U \to \Omega U$ which takes a sieve S over U to it's *closure* $\bar{S}$, i.e., $\bar{S}$ is that sieve which contains all those open subsets V of X for which $V \cap S$ forms an open cover of V.

The Closure Operator

Any arrow j on a topos $\mathbf{E}$ gives equivalently an operator which sends each subobject to it's j-*closure*.

Definition 13.4.1.4. (j**-Closure Operator**) *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is any such arrow. Then* j *determines a map* $(\bar{-}) : \mathrm{Sub}_{\mathbf{E}}(A) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(A)$ *for any object* A *given as follows:*

$$\begin{tikzcd}
\mathrm{Sub}_{\mathbf{E}}(A) \arrow[r, dash, "\cong"] \arrow[d, "(\bar{-})"'] & \mathrm{Hom}_{\mathbf{E}}(A,\Omega) \arrow[d, "\mathrm{Hom}_{\mathbf{E}}(A,j)"] \\
\mathrm{Sub}_{\mathbf{E}}(A) \arrow[r, dash, "\cong"] & \mathrm{Hom}_{\mathbf{E}}(A,\Omega)
\end{tikzcd}$$

which, as the above diagram says, maps each subobject of $m : S \rightarrowtail A$ *of* A *to another subobject of* A, $\bar{m} : \bar{S} \rightarrowtail A$ *which is called* j*-closure of* S *and this subobject* $\bar{S}$ *is given by the subobject characterized by the arrow* $j \circ$ *char* m*, that is,*

$$\mathrm{char}\ \bar{m} = j \circ \mathrm{char}\ m.$$

One can see that the closure operator $(\bar{-}) : \mathrm{Sub}_{\mathbf{E}}(A) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(A)$ is natural in A:

Lemma 13.4.1.5. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is any such arrow in* $\mathbf{E}$*. For any arrow* $f : A \longrightarrow B$ *in* $\mathbf{E}$ *and it's corresponding subobject pullback functor* $(f)^{-1} : Sub_{\mathbf{E}}(B) \longrightarrow Sub_{\mathbf{E}}(A)$*, we have that*

$$(f)^{-1}(\bar{S}) = \overline{(f)^{-1}(S)}.$$

for any subobject $S \to B$*.*

Proof. The following diagram directly shows the result:

$$\begin{tikzcd}
(f)^{-1}(S) \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner"] & S \arrow[r] \arrow[d, tail, "m"] \arrow[dr, phantom, "\lrcorner"] & \mathbf{1} \arrow[r, equal] \arrow[d, "\mathsf{true}"] \arrow[dr, phantom, "\lrcorner"] & \mathbf{1} \arrow[d, "\mathsf{true}"] \\
A \arrow[r, "f"'] & B \arrow[r, "\mathrm{char}\ m"'] & \Omega \arrow[r, "j"'] & \Omega
\end{tikzcd}$$

as characteristic arrow for both $(f)^{-1}(\bar{S})$ and $\overline{(f)^{-1}(S)}$ are same. □

One can give the description of a Lawvere-Tierney topology from the closure operator.

Proposition 13.4.1.6. *Suppose* $\mathbf{E}$ *is a topos. An arrow* $j : \Omega \longrightarrow \Omega$ *is a Lawvere-Tierney topology on* $\mathbf{E}$ *if and only if the corresponding* j*-closure operator* $(\bar{-})$ *satisfies for any subobjects* $S, T \rightarrowtail A$ *for any object* A *the following*

$$S \subseteq \bar{S} \quad \bar{S} = \bar{\bar{S}} \quad \overline{S \cap T} = \bar{S} \cap \bar{T}$$

Proof. (L $\implies$ R) Let $j : \Omega \longrightarrow \Omega$ to be a Lawvere-Tierney topology on $\mathbf{E}$. To show that $S \subseteq \bar{S}$, we need a monomorphism $S \rightarrowtail \bar{S}$, which can be seen to exist by the fact that S forms a cone $(m, !_A)$ over the subobject pullback of $\bar{S}$ and the fact that this does forms a cone depends on LTT.1. To show that $S = \bar{\bar{S}}$, we can simply note that from LTT.2, we have $\text{char } \bar{m} = j \circ \text{char } m = (j \circ j) \circ \text{char } m = \text{char } \bar{\bar{m}}$. To show $\overline{S \cap T} = \bar{S} \cap \bar{T}$, we can note the following by the help of LTT.3 ($n : A \to \Omega$ is the characteristic of T):

$$\begin{aligned} \text{char } \overline{m \cap n} &= j \circ \text{char } m \cap n \\ &= j \circ \wedge \langle \text{char } m, \text{char } n \rangle \\ &= \wedge \circ (j \times j) \circ \langle \text{char } m, \text{char } n \rangle \\ &= \wedge \circ \langle j \circ \text{char } m, j \circ \text{char } n \rangle \\ &= \wedge \circ \langle \text{char } \bar{m}, \text{char } \bar{n} \rangle \\ &= \text{char } \bar{m} \wedge \text{char } \bar{n} \end{aligned}$$

which is indeed the required result.
(R $\implies$ L) LTT.2 and LTT.3 essentially follows from following the arguments above in reverse. Whereas for LTT.1, since we have $S \rightarrowtail \bar{S}$ for any subobject S, then for $S = \mathbf{1}$ and $A = \Omega$, that is, if we take the subobject classifier $\mathsf{true} : \mathbf{1} \rightarrowtail \Omega$ as our subobject, then because characteristic arrow of true is 1_Ω, therefore $\bar{\mathbf{1}} = \mathbf{1}$ and so $j \circ 1_\Omega \circ \mathsf{true} = 1 \circ \mathsf{true}$, i.e. $j \circ \mathsf{true} = \mathsf{true}$. □

Closed & Dense Subobjects

Based on what the j-closure of a subobject $m : S \rightarrowtail A$ looks like in relation to S and A, each j determines two classes of subobjects:

Definition 13.4.1.7. (**Closed & Dense Subobjects**) *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is a Lawvere-Tierney topology on* $\mathbf{E}$. *Let* $m : S \rightarrowtail A$ *be any subobject. We then define*

- *The subobject m is* ***closed*** *if* $\bar{S} = S$.
- *The subobject m is* ***dense*** *if* $\bar{S} = A$.

As expected, any Lawvere-Tierney topology $j : \Omega \Rightarrow \Omega$ on $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ determines a Grothendieck topology on $\mathbf{C}$, therefore generalizing it.

Proposition 13.4.1.8. *Every Grothendieck topology* J *on a small category* $\mathbf{C}$ *determines a Lawvere-Tierney topology* $j : \Omega \Rightarrow \Omega$ *in the presheaf topos* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$.

Proof. Let $(\mathbf{C}, J)$ be a site. Define $j : \Omega \Rightarrow \Omega$ to be a natural transformation with components $j_C : \Omega C \to \Omega C$ which takes a sieve to it's J-closure. The fact that j is a Lawvere-Tierney topology in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ can be seen from the beginning discussion above Definition 13.4.1.1. □

13.4.2 j-Sheaves in a topos

A Grothendieck topology J on a small category $\mathbf{C}$ leads to a notion of sheaves over a site as given by Definition 13.2.3.2. Remember that a J-covering sieve S_C is a subpresheaf of $\mathbf{Yon}(C)$ such that the closure of sieve S_C is $S_C^{\max}$. As in the proof of Proposition 13.4.1.8, the Lawvere-Tierney topology j in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ corresponding to J is such that the j-closure of a sieve is maximal if and only if that sieve is a J-cover, that is, a J-cover S_C is a j-dense subobject of $\mathbf{Yon}(C)$. A matching family of a J-cover S_C is simply a natural transformation $S_C \Rightarrow P$ where S_C is viewed as a subpresheaf of $\mathbf{Yon}(C)$. An amalgamation of a matching family $S_C \Rightarrow P$ is hence a natural transformation from $\mathbf{Yon}(C) \Rightarrow P$. The condition when P is a sheaf says *every matching family has a unique amalgamation*, which from above discussion surmounts to the fact that the following commutes:

$$\begin{tikzcd} S_C \arrow[r, Rightarrow, "\forall"] \arrow[d, Rightarrowtail, "\text{(dense)}"'] & P \\ \mathbf{Yon}(C) \arrow[ur, Rightarrow, "\exists!"'] & \end{tikzcd} .$$

This motivates the following definition of a sheaf in any arbitrary topos:

Definition 13.4.2.1. (**Sheaf Object in a Topos**) *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is a Lawvere-Tierney topology on it. An object* F *is called a sheaf* in $\mathbf{E}$ *if for all dense subobjects* $m : A \rightarrowtail E$*, any arrow* $A \to F$ *can be factored uniquely via* m*, that is the following commutes:*

$$\begin{tikzcd} A \arrow[r, "\forall"] \arrow[d, rightarrowtail, "m", "\text{(dense)}"'] & F \\ E \arrow[ur, "\exists!"'] & \end{tikzcd} .$$

In other words, the following is an isomorphism for all dense subobjects m*:*

$$\mathrm{Hom}_{\mathbf{E}}(E, F) \xrightarrow[\cong]{-\circ m} \mathrm{Hom}_{\mathbf{E}}(A, F) \ .$$

Remark 13.4.2.2. The full subcategory of sheaf objects is denoted

$$\mathrm{Sh}_j\mathbf{E}.$$

A weaker condition of the above definition would gives us the following definition which generalizes separated presheaves[15]:

Definition 13.4.2.3. (**Separated Object in a Topos**) *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is a Lawvere-Tierney topology on it. An object* G *in* $\mathbf{E}$ *is called separated if for all dense subobjects* $m : A \rightarrowtail E$*, the following is a monomorphism:*

$$\mathrm{Hom}_{\mathbf{E}}(E, G) \xrightarrowtail{-\circ m} \mathrm{Hom}_{\mathbf{E}}(A, G) \ .$$

Remark 13.4.2.4. The full subcategory of separated objects is denoted

$$\mathrm{Sep}_j\mathbf{E}.$$

[15] To remind, a presheaf is separated if every matching family has an amalgamation (not necessarily unique).

13.4.3 $\mathrm{Sh}_j\mathbf{E}$ is a Topos

As in the case of sheaves over a site, $\mathrm{Sh}_j\mathbf{E}$ is a topos. We begin with proving that $\mathrm{Sh}_j\mathbf{E}$ has finite limits and exponentials.

Lemma 13.4.3.1. *Let* $\mathbf{E}$ *be a topos. The full subcategory* $Sh_j\mathbf{E}$ *has all finite limits and exponentials with any object in* $\mathbf{E}$*.*

Proof. To show finite limits, we just have to show that it has terminal object, equalizers and binary products. The terminal object $\mathbf{1}$ of $\mathbf{E}$ is clearly a sheaf. To show equalizer of any two parallel arrows of sheaves is a sheaf, take any parallel pair and it's equalizer in $\mathbf{E}$

$$\begin{tikzcd} E \arrow[r, tail, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

take any dense subobject $m : S \rightarrowtail C$ and any arrow $k_A : S \to A$ and $k_B : S \to B$. Take any arrow $k_E : S \to E$. We then have the following diagram due to A and B being sheaf objects and m being dense:

$$\begin{tikzcd} S \arrow[rr, "e \circ k_e"] \arrow[dd, tail, "m"', "\text{(dense)}"] & & A \arrow[dd, shift right, "f"'] \arrow[dd, shift left, "g"] \\ & & \\ C \arrow[uurr, dashed, "u_{e \circ k_e}"] \arrow[rr, dashed, shift left, "u_{f \circ e \circ k_e}"] \arrow[rr, dashed, shift right, "u_{g \circ e \circ k_e}"'] & & B \end{tikzcd}$$

It can be seen now that the arrows $f \circ u_{e \circ k_E} = g \circ u_{e \circ k_E}$. Hence, $\exists! v : C \longrightarrow E$ by universality of equalizer such that $e \circ v = u_{e \circ k_E}$. Therefore $e \circ v \circ m = e \circ k_E$ and since e is monic, hence $v \circ m = k_E$. Binary products can also be seen as above. Therefore finite limits exists.

To see about exponentials, we can see that for any sheaf object F and any object B, F^B is a sheaf by the following; Suppose $m : A \rightarrowtail E$ is a dense subobject, therefore we have:

$$\mathrm{Hom}_{\mathbf{E}}(E, F) \cong \mathrm{Hom}_{\mathbf{E}}(A, F)$$

Now, because m is dense, then $m \times 1_B : A \times B \rightarrowtail E \times B$ is dense because $\overline{A \times B} \cong \overline{(\pi)^{-1}(A)} \cong (\pi)^{-1}(\overline{A}) = (\pi)^{-1}(E) = E \times B$ where $\pi : E \times B \longrightarrow E$ is the first projection. Using the density of $m \times 1_B$ and the fact that F is a sheaf, we can now see the following:

$$\begin{tikzcd} \mathrm{Hom}_{\mathbf{E}}(E \times B, F) \arrow[r, dash, "\cong"] \arrow[d, "\mathrm{Hom}_{\mathbf{E}}(m \times 1, F)"'] & \mathrm{Hom}_{\mathbf{E}}(E, F^B) \arrow[d, "\mathrm{Hom}_{\mathbf{E}}(m, F^B)"] \\ \mathrm{Hom}_{\mathbf{E}}(A \times B, F) \arrow[r, dash, "\cong"'] & \mathrm{Hom}_{\mathbf{E}}(A, F^B) \end{tikzcd}$$

where the $\mathrm{Hom}_{\mathbf{E}}(m \times 1, F)$ is an isomorphism and since the square commutes, so $\mathrm{Hom}_{\mathbf{E}}(m, F^B)$ is an isomorphism, proving that F^B is a sheaf. This shows exponentials exists in $\mathrm{Sh}_j\mathbf{E}$. $\square$

Next, the subobject classifier of $\mathrm{Sh}_j\mathbf{E}$ is given by the mono-epi factor (image) of $j : \Omega \longrightarrow \Omega$:

Lemma 13.4.3.2. *Suppose* $\mathbf{E}$ *is a topos,* $j : \Omega \longrightarrow \Omega$ *is an LT topology in* $\mathbf{E}$ *and* Ω *is it's subobject classifier. Then the equalizer of* j *and* 1_Ω*, denoted* Ω_j *classifies the* j*-closed subobjects, that is, for any object* E *in* $\mathbf{E}$*, there is an isomorphism*

$$\mathrm{Hom}_{\mathbf{E}}(E, \Omega_j) \cong \mathit{ClSub}_{\mathbf{E}}(E)$$

where $\mathit{ClSub}_{\mathbf{E}}(E)$ *is the sub-lattice of* j*-closed subobjects of* E*.*

Proof. The object Ω_j is constructed, as stated above, by the equalizer of j and 1_Ω can equivalently be realized by the mono-epic factor of j, i.e. the following commutes:

$$\begin{tikzcd}
\Omega_j \arrow[r, tail, "m"] \arrow[dr, "1_{\Omega_j}"'] & \Omega \arrow[d, two heads, "r"'] \arrow[dr, "j"] & \\
& \Omega_j \arrow[r, tail, "m"'] & \Omega
\end{tikzcd}.$$

This follows from the fact that $j \circ j = j$ and therefore Ω itself forms a cone over the equalizer diagram of 1_Ω and j. Now, the main result follows as from the fact that for any closed subobject $s : A \rightarrowtail E$ of E, we have $\bar{A} = A$, that is

$$\begin{aligned} j \circ \mathrm{char}\ s &= \mathrm{char}\ s \\ m \circ r \circ \mathrm{char}\ s &= \mathrm{char}\ s \end{aligned}$$

and hence $\mathrm{char}\ s$ is factored via the unique arrow $r \circ \mathrm{char}\ s$. $\square$

Now if we ought to show that Ω_j is the subobject classifier of $\mathrm{Sh}_j\mathbf{E}$, then Ω_j must be a sheaf first of all:

Lemma 13.4.3.3. *Suppose* $\mathbf{E}$ *is a topos,* æ $: \Omega \longrightarrow \Omega$ *is an LT topology in* $\mathbf{E}$ *and* Ω_j *is the equalizer of* j *and* 1_Ω*. Then for any dense subobject* $m : A \rightarrowtail E$*, the pullback functor along* m*:*

$$(m)^{-1} : \mathit{ClSub}_{\mathbf{E}}(E) \longrightarrow \mathit{ClSub}_{\mathbf{E}}(A)$$

$$(k : C \rightarrowtail E) \qquad\qquad (m)^{-1}(k)$$

is an isomorphism.

Proof. We only wish to find a map $\tau : \mathrm{ClSub}_{\mathbf{E}}(A) \longrightarrow \mathrm{ClSub}_{\mathbf{E}}(E)$ such that $\tau \circ (m)^{-1} = 1$ and

$(m)^{-1} \circ \tau = 1$. Consider the following candidate for τ:

$$\tau : \mathrm{ClSub}_{\mathbf{E}}(A) \longrightarrow \mathrm{ClSub}_{\mathbf{E}}(E)$$

$$\begin{tikzcd} h : B \rightarrowtail A \arrow[r, mapsto] & B \arrow[d, two heads, "e_{moh}"'] \arrow[dr, tail, "moh"] & \\ & \mathrm{Im}(B) \arrow[r, tail, "u_{moh}"'] & E \end{tikzcd}$$

$$\overline{\mathrm{Im}(B)} \xrightarrow[\overline{u_{moh}}]{} E$$

To show that $\tau \circ (m)^{-1} = 1$, take any closed subobject $k : C \rightarrowtail E$ of E, then $\tau\left((m)^{-1}(C)\right) = \overline{\mathrm{im}\left(()\,(m)^{-1}(C)\right)}$. Now since $\overline{u_{m\circ((m)^{-1}(k))}} : \overline{\mathrm{im}\left(()\,(m)^{-1}(C)\right)} \rightarrowtail E$ is as given below:

$$\begin{aligned} \mathrm{char}\ \overline{\mathrm{im}\left(()\,(m)^{-1}(C)\right)} &= j \circ \mathrm{char}\ \mathrm{im}\left(()\,(m)^{-1}(C)\right) \\ &= j \circ \wedge\langle \mathrm{char}\ m, \mathrm{char}\ k\rangle \\ &= \wedge \circ (j \times j) \circ \langle \mathrm{char}\ m, \mathrm{char}\ k\rangle \\ &= \wedge \circ \langle j \circ \mathrm{char}\ m, j \circ \mathrm{char}\ k\rangle \\ &= \mathrm{char}\ \bar{m} \cap \mathrm{char}\ \bar{k} \\ &= \mathrm{char}\ \bar{A} \cap \mathrm{char}\ \bar{C} \\ &= \mathrm{char}\ E \cap \mathrm{char}\ C \\ &= \mathrm{char}\ C \end{aligned}$$

and therefore $\tau \circ (m)^{-1} = 1$. For $(m)^{-1} \circ \tau = 1$, take a closed subobject $h : B \rightarrowtail A$, so that $(m)^{-1}(\tau(h)) = (m)^{-1}\left(\overline{\mathrm{im}\left(()B\right)}\right) \cong \overline{(m)^{-1}\left(\mathrm{im}\left(()B\right)\right)} = \bar{B} = B$ where the second-to-last equality is obtained via what is called Beck-Chevalley condition, which we hadn't discussed here. Hence $(m)^{-1} \circ \tau = 1$. □

Now we can safely say that Ω_j is a sheaf object:

Corollary 13.4.3.4. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in* $\mathbf{E}$*, then* Ω_j*, the equalizer of* j *and* 1_Ω*, is a* j*-sheaf in* $\mathbf{E}$*.*

Proof. Take any dense subobject $m : A \rightarrowtail E$. Then:

$$\begin{aligned} \mathrm{Hom}_{\mathbf{E}}(E, \Omega_j) &\cong \mathrm{ClSub}_{\mathbf{E}}(E) && \text{By Lemma 13.4.3.2} \\ &\cong \mathrm{ClSub}_{\mathbf{E}}(A) && \text{By Lemma 13.4.3.3} \\ &\cong \mathrm{Hom}_{\mathbf{E}}(A, \Omega_j) && \text{By Lemma 13.4.3.2} \end{aligned}$$

Hence proved. □

Lemma 13.4.3.5. *Suppose* $\mathbf{E}$ *is a topos. If* E *is a sheaf in* $\mathbf{E}$*, then*

$$m : A \rightarrowtail E \text{ is closed in } E \iff A \text{ is also a sheaf.}$$

Proof. (L $\implies$ R) Let E be a sheaf in $\mathbf{E}$ and $m : A \rightarrowtail E$ be closed. Take any dense subobject $k : S \rightarrowtail B$ and let $f : S \to A$ be any arrow. Since E is a sheaf, therefore $\exists! b : B \to E$ such that $b \circ k = m \circ f$. Let's now take the following pullback:

$$\begin{tikzcd} (b)^{-1}(A) \arrow[r, "\pi_2"] \arrow[d, tail, "\pi_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & \bar{A} = A \arrow[d, tail, "\bar{m}=m"] \\ B \arrow[r, "b"'] & E \end{tikzcd}$$

and since $b \circ k = m \circ f$, therefore $\exists! u : S \to (b)^{-1}(A)$ such that $\pi_1 \circ u = k$ and $\pi_2 \circ u = f$. Now $S \subset (b)^{-1}(A) \implies \bar{S} \subset \overline{(b)^{-1}(A)} = (b)^{-1}(\bar{A}) = (b)^{-1}(A) \implies B \subset (b)^{-1}(A)$ where the last implication is drawn from the fact that $B = \bar{S}$ (k is dense). Therefore $\exists g : B \longrightarrow (b)^{-1}(A)$ such that $\pi_1 \circ g = 1$. Hence $(\pi_1 \circ g) \circ k = k \implies \pi_1 \circ u = k = (\pi_1 \circ g) \circ k \implies u = g \circ k$ and so $\pi_2 \circ g : B \to A$ is the required arrow.
(R $\implies$ L) Suppose A is a sheaf. We need to show that $m : A \rightarrowtail E$ is closed, i.e. $\bar{A} = A$, where E itself is a sheaf. Take the trivially dense subobject $d : A \rightarrowtail \bar{A}$. Since A and E are sheaves, then we have that $\exists! u_d : \bar{A} \to A$ such that $u_d \circ d = 1$ and $\exists! u_m : \bar{A} \to E$ such that $u_m \circ d = m$. Now since $(m \circ u_d) \circ d = m \circ 1 = m$ and since u_m is unique such that $u_m \circ d = m$, therefore $u_m = m \circ u_d$. Since the closure of m is such that $\bar{m} \circ d = m$, therefore $\bar{m} = u_m$ as u_m is unique with the property that $u_m \circ d = m$. But $\bar{m}$ is a monic and also $(\bar{m} \circ d) \circ u_d = (m) \circ u_d = \bar{m}$, we get $d \circ u_d = 1$, therefore $\bar{A} \cong A$. $\square$

Lemma 13.4.3.6. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in it. Then the* $\mathsf{true}_j : \mathbf{1} \longrightarrow \Omega_j$ *is the subobject classifier of* $Sh_j\mathbf{E}$ *which is given as composition of* true *with the epic part of the mono-epi factorization of* j *as follows:*

$$\begin{tikzcd} \mathbf{1} \arrow[d, tail, "\mathsf{true}_j := r \circ \mathsf{true}"'] \arrow[dr, tail, "\mathsf{true}"] & \\ \Omega_j \arrow[r, tail, "m"'] & \Omega \end{tikzcd} .$$

Proof. By Lemma 13.4.3.4, we have that Ω_j is a j-sheaf object in $\mathbf{E}$. Now if true_j ought to be the subobject classifier of $\mathrm{Sh}_j\mathbf{E}$, then for each subobject $m : A \rightarrowtail E$ of sheaves (that is A and E are sheaves), we must have a unique arrow $\mathrm{char}\ m : E \to \Omega_j$ such that the following is a pullback:

$$\begin{tikzcd} A \arrow[r] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, tail, "\mathsf{true}_j"] \\ E \arrow[r, dashed, "\mathrm{char}\ m"'] & \Omega_j \end{tikzcd} .$$

But by Lemma 13.4.3.2, A must equivalently be closed! Therefore we would be done if we could show that any subobject of sheaves $m : A \rightarrowtail E$ is always closed. This just follows from Lemma 13.4.3.5. $\square$

We finally can prove that $\mathrm{Sh}_j\mathbf{E}$ is a topos:

Theorem 13.4.3.7. *Suppose* $\mathbf{E}$ *is a topos and let* $j : \Omega \longrightarrow \Omega$ *be any Lawvere-Tierney topology in* $\mathbf{E}$*. Then the full subcategory* $Sh_j\mathbf{E}$ *of sheaf objects in* $\mathbf{E}$ *is a topos.*

Proof. $\mathrm{Sh}_j\mathbf{E}$ has finite limits and exponentials by Lemma 13.4.3.1. Subobject classifier of $\mathrm{Sh}_j\mathbf{E}$ is true_j as showed in Lemma 13.4.3.6. □

13.4.4 The Sheafification Functor $a : \mathbf{E} \longrightarrow \mathrm{Sh}_j\mathbf{E}$

We will now construct a left adjoint to the inclusion functor $i : \mathrm{Sh}_j\mathbf{E} \hookrightarrow \mathbf{E}$ where $\mathbf{E}$ is a topos and $j : \Omega \longrightarrow \Omega$ is a Lawvere-Tierney topology in $\mathbf{E}$. We would achieve our task in two steps. First, we would take any object E of $\mathbf{E}$ to a separated object E' and then we would take a separated object to a sheaf object $\overline{E'}$. In essence, we would need to construct two functors, $L_1 : \mathbf{E} \longrightarrow \mathrm{Sep}_j\mathbf{E}$ and $L_2 : \mathrm{Sep}_j\mathbf{E} \longrightarrow \mathrm{Sh}_j\mathbf{E}$, where L_1 and L_2 both must be left adjoints of the corresponding inclusions. We begin with constructing the separated object E':

From an object E to a separated object E'

To construct such a separated object E', we first note the following definition and the lemmas:

Definition 13.4.4.1. (**Graph of an arrow**) *Suppose* $\mathbf{E}$ *is a topos and* $f : A \longrightarrow B$ *is an arrow in it. The graph of* f *is defined to be the following subobject:*

$$A \xrightarrow{\langle 1,f\rangle} A \times B \ .$$

We also write this subobject as $G(f) := \langle 1, f\rangle$.

Now, we show that any subobject of a separated object is also separated:

Lemma 13.4.4.2. *Suppose* $\mathbf{E}$ *is a topos and* $m : B \rightarrowtail C$ *is a subobject. If* C *is separated, then so is* B.

Proof. Take any dense subobject $k : S \rightarrowtail E$. We then have the following

$$\begin{tikzcd}
\mathrm{Hom}_{\mathbf{E}}(E,C) \arrow[r, "\mathrm{Hom}_{\mathbf{E}}(k,C)"] & \mathrm{Hom}_{\mathbf{E}}(S,C) \\
\mathrm{Hom}_{\mathbf{E}}(E,B) \arrow[u, tail, "\mathrm{Hom}_{\mathbf{E}}(E,m)"] \arrow[r, "\mathrm{Hom}_{\mathbf{E}}(k,B)"'] & \mathrm{Hom}_{\mathbf{E}}(S,B) \arrow[u, tail, "\mathrm{Hom}_{\mathbf{E}}(S,m)"']
\end{tikzcd}$$

where both the left and right vertical arrows are injective. Now, the bottom arrow $\mathrm{Hom}_{\mathbf{E}}(k,B)$ is injective because if for two $x, y \in \mathrm{Hom}_{\mathbf{E}}(E,B)$ we have $x \circ k = y \circ k$, then since $\mathrm{Hom}_{\mathbf{E}}(k,C)$ is injective because C is separated, therefore we will have

$$\begin{aligned}
& m \circ x \circ k = m \circ y \circ k \\
\implies \ & m \circ x = m \circ y && \text{because } - \circ k \text{ of top arrow is injective} \\
\implies \ & x = y
\end{aligned}$$

Hence proved. □

We next show the equivalent conditions for graph of any arrow to be a closed subobject:

Lemma 13.4.4.3. *Suppose* $\mathbf{E}$ *is a topos. Let* C *be any object in* $\mathbf{E}$ *and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in* $\mathbf{E}$. *Then, the following are equivalent:*

1. *C is separated.*
2. *The diagonal $\Delta_C \in Sub_{\mathbf{E}}(C \times C)$ is a closed subobject.*
3. *The following commutes*

$$\begin{tikzcd} C \arrow[r, "\{\cdot\}_C"] \arrow[dr, "\{\cdot\}_C"'] & \Omega^C \arrow[d, "j^C"] \\ & \Omega^C \end{tikzcd} .$$

4. *For any $f : A \longrightarrow C$, the graph f, $G(f)$, is a closed subobject of $A \times C$.*

Proof. ($1 \implies 2$) If C is separated, then because $\Delta_C \subset \overline{\Delta_C}$, and so we have the usual dense subobject $k : C \to \overline{C}$. Therefore, we have $\mathrm{Hom}_{\mathbf{E}}(k, C) : \mathrm{Hom}_{\mathbf{E}}\left(\overline{C}, C\right) \rightarrowtail \mathrm{Hom}_{\mathbf{E}}(C, C)$. Take $\pi_1 \circ \bar{\Delta}_C$ and $\pi_2 \circ \bar{\Delta}_C \in \mathrm{Hom}_{\mathbf{E}}\left(\bar{C}, C\right)$ where $\pi_1, \pi_2 : C \times C \rightrightarrows C$ are the projections. But then $\pi_1 \circ \bar{\Delta}_C \circ k = \pi_2 \circ \bar{\Delta}_C \circ k \implies \pi_1 \circ \bar{\Delta}_C = \pi_2 \circ \bar{\Delta}_C$ as $- \circ k$ is injective as C is separated. Therefore $\bar{\Delta}_C$ forms a cone over $\pi_1, \pi_2 : C \times C \rightrightarrows C$, therefore there exists unique $l : \bar{C} \longrightarrow C$ with $\Delta_C \circ l = \bar{\Delta}_C$ which means that $\bar{\Delta}_C \subset \Delta_C$ so that $\bar{\Delta}_C = \Delta_C$.
($2 \implies 3$) This is trivial because if $\Delta_C : C \times C \rightarrowtail C$ is a closed subobject, then $j \circ \mathrm{char}\ \Delta_C = \mathrm{char}\ \Delta_C$. This commuting diagram gives rise to another commuting diagram obtained by it's P-transpose, which proves the result:

$$\begin{tikzcd} C \times C \arrow[r, "\mathrm{char}\ \Delta_C"] \arrow[dr, "\mathrm{char}\ \Delta_C"'] & \Omega \arrow[d, "j"] \\ & \Omega \end{tikzcd} \xrightarrow{P-\mathrm{Transpose}} \begin{tikzcd} C \arrow[r, "\mathrm{char}\hat{\ }\Delta_C"] \arrow[dr, "\mathrm{char}\hat{\ }\Delta_C"'] & PC \cong \Omega^C \arrow[d, "j^C"] \\ & PC \cong \Omega^C \end{tikzcd} .$$

($2 \implies 4$) For $f : A \to C$, the graph $G(f) = \langle 1, f \rangle : A \rightarrowtail A \times C$ is obtained by the pullback of Δ_C along $f \times 1$. The naturality of the closure operator proves the rest.
($4 \implies 1$) Suppose for $f : A \to C$ the graph $G(f) = \langle 1, f \rangle : A \rightarrowtail A \times C$ is closed. Take a dense subobject $m : S \rightarrowtail B$ and let $b_1, b_2 : B \rightrightarrows C$ be such that $b_1 \circ m = b_2 \circ m$. We wish to prove that $b_1 = b_2$. Next, let's look at the graph of b_1 and $b_1 \circ m$:

$$\begin{tikzcd} S \arrow[r, tail, "m"] \arrow[d, tail, "\langle 1, b_1 \circ m \rangle"'] & B \arrow[r, "b_1"] \arrow[d, tail, "\langle 1, b_1 \rangle"] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, tail, "\Delta_C"] \\ S \times C \arrow[r, tail, "m \times 1"'] & B \times C \arrow[r, "b_1 \times 1"'] & C \times C \end{tikzcd}$$

where Δ_C is closed, $\langle 1, b_1 \rangle$ is then closed and then $\langle 1, b_1 \circ m \rangle$ is also closed. Right and the whole square are pullbacks, therefore left one is, and $m \times 1$ is dense, which means $\overline{\langle 1, b_1 \circ m \rangle} = \langle 1, b_1 \circ m \rangle = \langle 1, b_1 \rangle$. Hence $b_1 \circ m = b_2 \circ m \implies \langle 1, b_1 \rangle = \langle 1, b_2 \rangle \implies b_1 = b_2$. □

We now construct the separated object E' for each object E in a topos $\mathbf{E}$, in the following lemma:

Lemma 13.4.4.4. *Suppose $\mathbf{E}$ is a topos and $j : \Omega \longrightarrow \Omega$ is an LT topology in it. For any object E in $\mathbf{E}$, there is an epimorphism*

$$\theta_E : E \longrightarrow E'$$

where E' is a separated object in $\mathbf{E}$.

Proof. Consider the arrow

$$E \xrightarrow{\{\cdot\}_E} \Omega^E \xrightarrow{r^E} \Omega_j^E$$

where as usual, the Ω_j is the mono-epic factor of j or equivalently the equalizer of 1 and j. Now denote E' as the mono-epic factor of $r^E \circ \{\cdot\}_E$, as shown below:

$$\begin{tikzcd} & E' \arrow[dr, tail] & \\ E \arrow[ur, two heads, "\theta_E"] \arrow[rr, "r^E\circ\{\cdot\}_E"'] & & \Omega_j^E \end{tikzcd}.$$

Now, by Lemma 13.4.4.2, E' is a separated object because Ω_j^E is. □

The left adjoint of $i : \mathbf{Sep}_j\mathbf{E} \hookrightarrow \mathbf{E}$

Remember our aim is to construct first a left adjoint of inclusion $i : \mathrm{Sep}_j\mathbf{E} \hookrightarrow \mathbf{E}$. To this end, we have found a way to form a separated object for any object of $\mathbf{E}$. We now wish to show that this construction of separated object is indeed a left adjoint to inclusion. For which we have to check it's universality. For that, we have the following lemmas

Lemma 13.4.4.5. *Suppose* $\mathbf{E}$ *is a topos. For any object* E *of* $\mathbf{E}$*, there exists an epimorphism* $\theta_E : E \longrightarrow E'$ *such that the kernel pair of* θ_E *is the closure* $\overline{\Delta_E}$ *of the subobject* $\Delta_E : E \rightarrowtail E \times E$*.*

Proof. Section 5.3, p.p. 229, Lemma 5, [**MacMoer**]. □

An immediate corollary of the above lemma proves the universality of the construction in Lemma 13.4.4.4:

Corollary 13.4.4.6. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in it. For each object* E *of* $\mathbf{E}$*, the corresponding epimorphism to a separated object* E'

$$\theta_E : E \longrightarrow E'$$

is universal amongst all arrows from E *to a separated object.*

Proof. If there is an arrow $f : E \to S$ where S is separated, then by Theorem 13.3.4.7 and Lemma 13.4.4.5, the epic $\theta_E : E \longrightarrow E'$ is given as the following coequalizer

$$\begin{tikzcd} & & S \\ \bar{E} \arrow[r, shift left, "\pi_1\circ\bar{\Delta}_E"] \arrow[r, shift right, "\pi_2\circ\bar{\Delta}_E"'] & E \arrow[ur, "f"] \arrow[r, two heads, "\theta_E"'] & E' \end{tikzcd}.$$

Now, because $f \circ \pi_1 \circ \Delta_E = f = f \circ \pi_2 \circ \Delta_E$, therefore $f \circ \pi_1 \circ \bar{\Delta}_E \circ k = f \circ \pi_2 \circ \bar{\Delta}_E \circ k$, where $k : E \rightarrowtail \bar{E}$, so, because $- \circ k : \mathrm{Hom}_{\mathbf{E}}\left(\bar{E}, S\right) \longrightarrow \mathrm{Hom}_{\mathbf{E}}\left(E, S\right)$ is injective because S is separated, therefore, $f \circ \pi_1 \circ \bar{\Delta}_E = f \circ \pi_2 \circ \bar{\Delta}_E$ and hence $\exists! : l : S \to E'$. □

Finally, we can now conclude that indeed, the construction of separated object E' is the left adjoint of inclusion:

Corollary 13.4.4.7. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in it. Then there is an adjunction as given below:*

$$\begin{tikzcd} \mathbf{E} \arrow[r, bend left, "L_1"] \arrow[r, phantom, "\perp"] & Sep_j\mathbf{E} \arrow[l, hook, bend left, "i"] \end{tikzcd}$$

where L_1 *is given by:*

$$\begin{aligned} L_1 : \mathbf{E} &\longrightarrow Sep_j\mathbf{E} \\ E &\longmapsto E' \\ (f : E \to F) &\longmapsto (f' : E' \to F') \end{aligned}$$

where f' *is given by the following universal arrow because of* $\theta_F \circ f$ *forming a cocone over the top coequalizer diagram as in the following:*

$$\begin{tikzcd} \bar{E} \arrow[r, shift left, "\pi_1 \circ \bar{\Delta}_E"] \arrow[r, shift right, "\pi_2 \circ \bar{\Delta}_E"'] & E \arrow[r, two heads, "\theta_E"] \arrow[d, "f"] & E' \arrow[d, dashed, "f'"] \\ \bar{F} \arrow[r, shift left, "p_1 \circ \bar{\Delta}_F"] \arrow[r, shift right, "p_2 \circ \bar{\Delta}_F"'] & F \arrow[r, two heads, "\theta_F"] & F' \end{tikzcd} .$$

Proof. To show that L_1 as above is indeed the left adjoint, take any object E and a separated object G and then take any arrow $f : E \longrightarrow iG$ in $\mathbf{E}$, and then just observe that the following commutes:

$$\begin{tikzcd} E \arrow[d, "\eta_E"'] \arrow[dr, "f"] & \\ iL_1(E) = i(E') \arrow[r, dashed, "\text{By Coroll. 13.4.4.6}"'] & i(G) \end{tikzcd}$$

Hence proved that L_1 is left adjoint of inclusion. □

The left adjoint of $i : \mathbf{Sh}_j\mathbf{E} \hookrightarrow \mathbf{Sep}_j\mathbf{E}$

And finally, we have the left adjoint of the inclusion $i : \mathrm{Sh}_j\mathbf{E} \hookrightarrow \mathrm{Sep}_j\mathbf{E}$:

Lemma 13.4.4.8. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology. Then there is an adjunction as given below:*

$$\begin{tikzcd} Sep_j\mathbf{E} \arrow[r, bend left, "L_2"] \arrow[r, phantom, "\perp"] & Sh_j\mathbf{E} \arrow[l, hook, bend left, "i"] \end{tikzcd}$$

where L_2 is given by:

$$L_2 : Sep_j\mathbf{E} \longrightarrow Sh_j\mathbf{E}$$
$$E \longmapsto \bar{E}$$
$$(f : E \to F) \longmapsto (\bar{f} : \bar{E} \to \bar{F})$$

where the closure of a separated subobject is a sheaf because $\overline{r^E \circ \{\cdot\}_E} : \bar{E} \rightarrowtail \Omega_j^E$ is a closed subobject of the sheaf Ω_j^E and so by Lemma 13.4.3.5, $\bar{E}$ is a sheaf.

Proof. To show that L_2 is indeed the left adjoint of inclusion, take any $f : E \longrightarrow i(F)$ in $\mathrm{Sep}_j\mathbf{E}$ where E is separated and F is a sheaf. We then have the following commuting diagram, which establishes the result:

$$\begin{tikzcd}
E \arrow[d, tail, "\text{(dense)}"'] \arrow[drr, "f"] & & \\
iL_2(E) = i(\bar{E}) = \bar{E} \arrow[rr, dashed, "\text{by sheaf condition of } F"'] & & i(F) \quad \text{(sheaf)}
\end{tikzcd}$$

Hence proved. □

Finally, we have the sheafification functor:

Theorem 13.4.4.9. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in it. Then the full-subcategory* $Sh_j\mathbf{E}$ *is reflective. That is, there is a left adjoint of inclusion:*

$$\begin{tikzcd}
\mathbf{E} \arrow[r, bend left, "a"] \arrow[r, phantom, "\perp"] & Sh_j\mathbf{E} \arrow[l, bend left, "i"]
\end{tikzcd}$$

where $a = L_2 \circ L_1$ as in Lemma 13.4.4.8 and Corollary 13.4.4.7.

Proof. Take any $f : E \longrightarrow i(F)$ in $\mathbf{E}$ where E is any object in $\mathbf{E}$ and F is any sheaf. We then have:

$$\begin{tikzcd}
\text{(sep.)}\ E' \arrow[d, tail, "\text{(dense)}"'] & E \arrow[l, two heads, "\theta_E"'] \arrow[dr, "f"] & \\
\text{(sheaf)}\ a(E) = \bar{E}' \arrow[rr, dashed, "\text{by sheaf cond. of } F"] & & i(F)
\end{tikzcd}.$$

Hence proved. □

The following shows, like all sheafification adjunction studied previously, that the left adjoint a is left-exact:

Proposition 13.4.4.10. *Suppose* $\mathbf{E}$ *is a topos and* $j : \Omega \longrightarrow \Omega$ *is an LT topology in it. Then a in the adjunction $i \vdash a$ of Theorem 13.4.4.9 is left exact.*

Proof. Section 3.3, p.p. 232, [**MacMoer**]. □

13.5 Geometric Morphisms

One of the important aspect that we are witnessing continuously in the above sections is the repetitive rise of left exactness of the left adjoint in the sheafification adjunction. This is a general phenomenon for maps between two topoi, the study of which leads to a natural notion of *points of a topos* and generalization of tensor product. We hence define a geometric morphism between two topoi as an adjoint pair where the left adjoint is left exact:

Definition 13.5.0.1. (**Geometric Morphism**) *Suppose* **E** *and* **F** *are two topoi. An adjunction*

$$\begin{tikzcd} \mathbf{E} \arrow[r, bend left, "f^*"] \arrow[r, phantom, "\perp"] & \mathbf{F} \arrow[l, bend left, "f_*"] \end{tikzcd}$$

where the left adjoint f^* is also left exact (preserves finite limits) is then called a geometric morphism and is denoted as

$$\mathbf{F} \xrightarrow{\quad f \quad} \mathbf{E}\ .$$

The left adjoint f^* is called the inverse-image part and the right adjoint f_* is called the direct-image part of the geometric morphism f.

Example 13.5.0.2. A trivial example is that of direct and inverse image of sheaves. Take two topological spaces X and Y and a continuous map $f : X \to Y$. Then there is the following adjunction

$$\begin{tikzcd} \mathrm{Sh}\,(Y) \arrow[r, bend left, "f^*"] \arrow[r, phantom, "\perp"] & \mathrm{Sh}\,(X) \arrow[l, bend left, "f_*"] \end{tikzcd}$$

where

$$\begin{aligned} f_* : \mathrm{Sh}\,(X) &\longrightarrow \mathrm{Sh}\,(Y) \\ F &\longmapsto F((f)^{-1}(-)) \end{aligned}$$

and

$$\begin{aligned} f^* : \mathrm{Sh}\,(Y) &\longrightarrow \mathrm{Sh}\,(X) \\ F &\longrightarrow \Gamma_{(f^*(\Lambda_F))}. \end{aligned}$$

Note that the f^* in sub-script is the pullback functor. The functors $\Gamma_{(-)}$ and Λ_- in the adjunction $\Lambda_{(-)} \vdash \Gamma_{(-)}$ are as follows:

$$\begin{aligned} \Gamma_{(-)} : \mathbf{Bund}\,(X) &\longrightarrow \mathrm{Sh}\,(X) \\ (p : Y \to X) &\longmapsto (F : \mathbf{O(X)}^{\mathrm{op}} \to \mathbf{Sets}) \\ & \qquad\qquad U \mapsto \{s : U \to Y \mid p \circ s = \iota : U \hookrightarrow X\}. \end{aligned}$$

and

$$\begin{aligned} \Lambda_{(-)} : \mathrm{Sh}\,(X) &\longrightarrow \mathbf{Bund}\,(X) \\ F &\longmapsto (p : \Lambda_F \to X) \end{aligned}$$

where $\Lambda_F := \{\mathrm{germ}_x s \mid \forall s \in FU, \forall U \in \mathfrak{U}\,(x)\,, \forall x \in X\}$. Hence Γ is just the sheaf of cross-sections functor and the Λ is the trivial étale bundle of germs functor.

We have the following characterization of geometric morphisms between two sheaf topoi $\mathrm{Sh}\,(X)$ and $\mathrm{Sh}\,(Y)$:

Lemma 13.5.0.3. *Consider two topological spaces X and Y where Y is Hausdorff and their corresponding sheaf topoi Sh (X) and Sh (Y). Then there is a bijection between geometric morphisms between Sh (X) and Sh (Y) and continuous maps between X and Y.*

Proof. Take any continuous map $f : X \longrightarrow Y$. The above example shows that we get a geometric morphism $f : \mathrm{Sh}\,(X) \longrightarrow \mathrm{Sh}\,(Y)$. For a geometric morphism $f : \mathrm{Sh}\,(X) \longrightarrow \mathrm{Sh}\,(Y)$, where Y is Hausdorff, we wish to find a continuous mapping $\hat{f} : X \to Y$. We first note that because f^* is left exact, therefore all subobjects of $\mathbf{1}$ of $\mathrm{Sh}\,(Y)$ are preserved, that is, a subobject $F \rightarrowtail \mathbf{1}$ in $\mathrm{Sh}\,(Y)$ is mapped to a subobject $f^*F \rightarrowtail \mathbf{1}$ in $\mathrm{Sh}\,(X)$. But a subobject of $\mathbf{1}$ in $\mathrm{Sh}\,(Y)$ is an open set of Y (Proposition 20.6.3.3). Therefore we have the map $f^* : O(Y) \to O(X)$. Now define a function $\hat{f} : X \longrightarrow Y$ which takes $x \in X$ to that $y \in Y$ for which $x \in f^*(V)\ \forall V \in \mathfrak{U}\,(y)$. Let's first show that this function $\hat{f}$ is well defined. If $y_1 \neq y_2 \in Y$ are such that $\exists x \in X$ with $\hat{f}(x) = y_1$ and $\hat{f}(x) = y_2$, then $x \in f^*(V) \forall V \in \mathfrak{U}\,(y_1)$ and similarly for y_2. Since Y is Hausdorff, therefore $\exists$ open $V_1 \ni y_1$ and $V_2 \ni y_2$ such that $V_1 \cap V_2 = \phi$, which means that $x \in f^*(V_1) \cap f^*(V_2) = f^*(V_1 \cap V_2) = \phi$ which is a contradiction, where the first equality holds because f^* preserves limits and $- \cap -$ is the pullback of subobjects. Similarly, if $x \in X$, then $\exists y \in Y$ such that $\hat{f}(x) = y$ because if it's not $\forall y \in Y$, $\exists V_y \in \mathfrak{U}\,(y)$ such that $x \notin f^*(V)$. But then if we collect all such open $V_y \ni y\ \forall y \in Y$ as in $\bigcup_{y \in Y} V_y$, then clearly $\bigcup_{y \in Y} V_y = Y$ and $f^*(Y) = X \ni x$, which is a contradiction. This proves that f^* is well defined. We now wish to show that this function $\hat{f}$ is continuous. For this, take any open $V \subseteq Y$. We have $\left(\hat{f}\right)^{-1}(V) = \{x \in X \mid \hat{f}(x) \in V\} = \{x \in X \mid x \in f^*(V)\} = f^*(V)$ which is clearly open. $\square$

Example 13.5.0.4. There are other examples of geometric morphisms, like the change of base adjunction, sheafification adjunction and global sections adjunction among others. More succinctly, the following adjunctions are geometric morphisms:

$$\begin{tikzcd} \mathbf{E}/A \arrow[r, bend left, "k^* := (k)^{-1}"] \arrow[r, phantom, "\perp"] & \mathbf{E}/B \arrow[l, bend left, "k_* := \prod_k"] \end{tikzcd} \qquad \begin{tikzcd} \mathbf{E} \arrow[r, bend left, "i^* := a"] \arrow[r, phantom, "\perp"] & \mathrm{Sh}_j\mathbf{E} \arrow[l, bend left, "i_* := i"] \end{tikzcd} \qquad \begin{tikzcd} \mathbf{Sets} \arrow[r, bend left, "\Delta(-) := \coprod_{s \in (-)} \mathbf{1}_{\mathbf{E}}"] \arrow[r, phantom, "\perp"] & \mathbf{E} \arrow[l, bend left, "\Gamma(-) := \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, -)"] \end{tikzcd}\ .$$

It is interesting to note that the collection of all topoi and geometric morphisms is a 2-category:

Definition 13.5.0.5. (**2-Category of Topoi & Geometric Morphisms**) *Consider the category* **Topoi** *whose objects are topoi and arrows are geometric morphisms between them. Then,*

- $\mathrm{Hom}_{\mathbf{Topoi}}(\mathbf{E}, \mathbf{F})$ **is a category** *where objects are geometric morphisms* $f : \mathbf{E} \longrightarrow \mathbf{F}$ *and an arrow between two geometric morphisms* f *and* g *is given by a natural transformation* $\eta^* : f^* \Rightarrow g^*$ *or equivalently by a natural transformation* $\eta_* : g_* \Rightarrow f_*$.
- **The 2-category Topoi** *is formed by noticing that any geometric morphism* $g : \mathbf{G} \longrightarrow \mathbf{F}$ *induces a functor* $\mathrm{Hom}_{\mathbf{Topoi}}(g, \mathbf{E})$ *for any topoi* $\mathbf{E}$ *as follows:*

$$\begin{aligned} \mathrm{Hom}_{\mathbf{Topoi}}(g, \mathbf{E}) : \mathrm{Hom}_{\mathbf{Topoi}}(\mathbf{F}, \mathbf{E}) &\longrightarrow \mathrm{Hom}_{\mathbf{Topoi}}(\mathbf{G}, \mathbf{E}) \\ (f : \mathbf{F} \to \mathbf{E}) &\longmapsto (f \circ g : \mathbf{G} \to \mathbf{E}) \\ (\eta : h \to f) &\longmapsto (\eta \circ g : h \circ g \to f \circ g) \end{aligned}$$

where $(\eta \circ g)^* : (h \circ g)^* \to (f \circ g)^*$ *is same as* $g^* \circ \eta^* : g^* \circ h^* \Rightarrow g^* \circ f^*$.

Therefore the 2-category **Topoi** is where objects are topoi $\mathbf{E}, \mathbf{F}, \ldots$, 1-cells are geometric morphisms $f : \mathbf{F} \to \mathbf{E}, \ldots$ and 2-cells are natural transformations between 1-cells $\eta : f \to g$.

13.5.1 Tensor Products

We now study the generalization of tensor products of modules to that of arbitrary contravariant and covariant functors. First, let us note the famous $\otimes$-hom adjunction of modules:

$$\mathrm{Hom}_R(X_S \otimes {}_S Z_R, Y_R) \cong \mathrm{Hom}_S\left(X_S, \underline{\mathrm{Hom}_R}({}_S Z_R, Y_R)\right)$$

where $X_S, {}_S Z_R, Y_R$ is a left S-module, left S right R-module and a right R-module respectively. Moreover, $\underline{\mathrm{Hom}_R}({}_S Z_R, Y_R)$ is the right S-module of R-linear maps ${}_S Z_R \to Y_R$. It is important to note here that the above adjunction is not a geometric morphism. But tensor products would subsequently help us in making new geometric morphisms.
The tensor product above is, in essence, between two functors because the right S-module X_S can be represented as the contravariant functor $F : \mathbf{1}^{\mathrm{op}}{}_{\mathbf{AbGrp}} \to \mathbf{AbGrp}$ where $\mathbf{1}$ is the one object **AbGrp** enriched category and F as the **AbGrp** enriched functor and similarly ${}_S Z_R$ would be a bifunctor given by the **AbGrp** enriched functor $G : \mathbf{1}^{\mathrm{op}}{}_{\mathbf{AbGrp}} \times \mathbf{1}_{\mathbf{AbGrp}} \longrightarrow \mathbf{AbGrp}$.
Now, suppose that we have a small category $\mathbf{C}$ and a co-complete category $\mathbf{E}$ with a functor $A : \mathbf{C} \longrightarrow \mathbf{E}$. What is then the meaning of tensor product of a presheaf $P : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$ and A?

$$\mathbf{C}^{\mathrm{op}} \xrightarrow{P} \mathbf{Sets} \qquad \overset{??}{\bigotimes} \qquad \mathbf{C} \xrightarrow{A} \mathbf{E}$$

To obtain such a general notion of tensor product, we will need to understand the *category of elements* (Definition 13.1.2.1) construction in the very first result, the Theorem 13.1.2.2. We would essentially define the left adjoint L as the tensor product functor $- \otimes_{\mathbf{C}} A$ with A. But elaborating this would make it clear on how one should approach tensor products.
Recall that all colimits can be constructed from coproducts and coequalizers. In particular, take $H : \mathbf{J} \longrightarrow \mathbf{E}$ to be a diagram in $\mathbf{E}$ of index J. Now consider the coproducts $\coprod_{(u:i\to j)\in \mathrm{Ar}(\mathbf{J})} H(\mathrm{dom}\,(u))$ and $\coprod_{i \in \mathrm{Ob}(\mathbf{J})} H(i)$. Consider the following two parallel arrows:

$$\coprod_{(u:i\to j)\in \mathrm{Ar}(\mathbf{J})} H(\mathrm{dom}\,(u)) \underset{\tau}{\overset{\theta}{\rightrightarrows}} \coprod_{i \in \mathrm{Ob}(\mathbf{J})} H(i) \ .$$

The θ is formed because $\coprod_{i\in \mathrm{Ob}(\mathbf{J})} H(i)$ forms a cocone over $\{H(\mathrm{dom}\,(u))\}_{u\in \mathrm{Ar}(\mathbf{J})}$ by injections κ_i and τ is formed because $\coprod_{i\in \mathrm{Ob}(\mathbf{J})} H(i)$ forms a cocone over $\{H(\mathrm{dom}\,(u))\}_{u:i\to j\in \mathrm{Ar}(\mathbf{J})}$ by $\kappa_j \circ H(u)$. The colimit of the diagram H is then given as the coequalizer of θ & τ:

$$\begin{tikzcd} \coprod_{(u:i\to j)\in \mathrm{Ar}(\mathbf{J})} H(\mathrm{dom}\,(u)) \arrow[r, shift left, "\theta"] \arrow[r, shift right, "\tau"'] & \coprod_{i\in \mathrm{Ob}(\mathbf{J})} H(i) \arrow[r, "\phi"] & \varinjlim H \end{tikzcd} .$$

Now, replacing H by $A\circ \pi_P$, we get the following (where $\pi_P : \int_{\mathbf{C}} P \to \mathbf{C}$ is the projection):

$$\begin{tikzcd} \coprod_{(u:(C,p)\to(C',p'))\in \mathrm{Ar}(\int_{\mathbf{C}} P)} A(C) \arrow[r, shift left, "\theta"] \arrow[r, shift right, "\tau"'] & \coprod_{(C,p)\in \mathrm{Ob}(\int_{\mathbf{C}} P)} A(C) \arrow[r, "\phi"] & \varinjlim(A\circ \pi_P) =: L(P) \end{tikzcd} .$$

We simply define $L(P) := \varinjlim(A\circ \pi_P)$ as the tensor product of P with A, $P\otimes_{\mathbf{C}} A$. Then, by the Theorem 13.1.2.2, we have the following adjunction:

$$\begin{tikzcd} \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \arrow[rr, bend left, "L:=-\otimes_{\mathbf{C}} A"] \arrow[rr, phantom, "\perp"] & & \mathbf{E} \arrow[ll, bend left, "R:=\mathrm{Hom}_{\mathbf{E}}(A(\cdot),-)"] \end{tikzcd} .$$

This adjunction in terms of hom-sets is:

$$\mathrm{Hom}_{\mathbf{E}}\left(P\otimes_{\mathbf{C}} A, E\right) \;\text{——}\cong\text{——}\; \mathrm{Hom}_{\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}}\left(P, \mathrm{Hom}_{\mathbf{E}}\left(A(-), E\right)\right) .$$

13.5.2 Points of a Topos

Consider a topological space X. Let $x\in X$ be a point of it. One can alternatively write $x\in X$ as an arrow $x : \mathbf{1}\to X$ in the **Top**. But as we saw below Defn. 13.5.0.1, the fact that we then have a geometric morphism $x : \mathbf{Sets} \longrightarrow \mathrm{Sh}\,(X)$ as follows:

$$\begin{tikzcd} \mathrm{Sh}\,(X) \arrow[rr, bend left, "x^*"] \arrow[rr, phantom, "\perp"] & & \mathrm{Sh}\,(\mathbf{1}) \cong \mathbf{Sets} \arrow[ll, bend left, "x_*"] \end{tikzcd} .$$

Hence we are representing the point x of the *underlying space* of the sheaf topos $\mathrm{Sh}\,(X)$ as a geometric morphism $x : \mathbf{Sets} \longrightarrow \mathrm{Sh}\,(X)$. This becomes the motivation for the following definition:

Definition 13.5.2.1. (**Points of a Topos**) *Let* $\mathbf{E}$ *be a topos. A point* f *of topos* $\mathbf{E}$ *is defined to be a geometric morphism* $f : \mathbf{Sets} \longrightarrow \mathbf{E}$.

However, the more interesting observations lies in trying to characterize the points of a Grothendieck topos.

Points of a Presheaf Topos

We wish to characterize the points of a Grothendieck topos. In order to do so, we begin by studying the points of a presheaf topos $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$. Take a point of $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$

$$f : \mathbf{Sets} \longrightarrow \mathbf{Sets}^{\mathbf{C}^{\text{op}}}.$$

We know that each object in $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$ is a colimit of representables by Proposition 13.1.2.3. But f^* preserves those colimits as it is a left adjoint. Therefore f^* can be studied by studying $f^* \circ \mathbf{Yon}\,(-) : \mathbf{C} \longrightarrow \mathbf{Sets}$ only. Our aim is now to show that for each point f, the functor $- \otimes_{\mathbf{C}} (f^* \circ \mathbf{Yon}\,(-)) : \mathbf{Sets}^{\mathbf{C}^{\text{op}}} \longrightarrow \mathbf{Sets}$ is isomorphic to f^*. To further our discussion, we have to first simplify the definition of tensor product to set valued functors as it will make analysis easier:

Definition 13.5.2.2. (**Tensor Product of Set Functors**) *Let* $\mathbf{C}$ *be a small category and* $R : \mathbf{C}^{\text{op}} \longrightarrow \mathbf{Sets}$ *and* $A : \mathbf{C} \longrightarrow \mathbf{Sets}$*. Then the tensor product* $R \otimes_{\mathbf{C}} A$ *is alternatively given as:*

$$R \otimes_{\mathbf{C}} A := \coprod_{C \in \text{Ob}(\mathbf{C})} RC \times AC / \sim$$

where $(r, a)_C \sim (r', c')_{C'}$ *if and only if* $\exists (r_0, c_0)_{C_0}, (r_1, c_1)_{C_1}, \ldots, (r_n, c_n)_{C_n} \in \coprod_{C \in \text{Ob}(\mathbf{C})} RC \times AC$ *such that*

1. $(r_0, c_0)_{C_0} = (r, a)_C$ *and* $(r_n, c_n)_{C_n} = (r', a')_{C'}$.
2. $\forall 1 \leq k \leq n$, $\exists u_k : C_k \longrightarrow C_{k-1}$ *in* $\mathbf{C}$ *such that*

$$\begin{aligned} Ru_k(r_{k-1}) &= r_k \\ Au_k(a_k) &= a_{k-1} \end{aligned}$$

OR, equivalently, $\exists u_k : C_{k-1} \longrightarrow C_k$ in $\mathbf{C}$ *such that*

$$\begin{aligned} Ru_k(r_k) &= r_{k-1} \\ Au_k(a_{k-1}) &= a_k. \end{aligned}$$

With the above definition, we can see now the following:

Proposition 13.5.2.3. *Suppose* $\mathbf{C}$ *is a small category and* $f : \mathbf{Sets} \longrightarrow \mathbf{Sets}^{\mathbf{C}^{\text{op}}}$ *is a point of the* $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$*. Then, there exists a unique functor* $A = f^* \circ \mathbf{Yon}\,(-) : \mathbf{C} \longrightarrow \mathbf{Sets}$ *such that*

$$f^* \cong - \otimes_{\mathbf{C}} A.$$

Proof. Take any $R : \mathbf{C}^{\text{op}} \longrightarrow \mathbf{Sets}$. We then have $f^*(R)$ and $R \otimes_{\mathbf{C}} A$ both in $\mathbf{Sets}$. But for these two sets in $\mathbf{Sets}$, we have a canonical map:

$$\begin{aligned} e_R : R \otimes_{\mathbf{C}} A &\longrightarrow f^*(R) \\ (r \otimes a)_C &\longmapsto f^*(\eta^r)(a) \end{aligned}$$

because for any $r \in RC$, $\exists!\ \eta^r : \mathbf{Yon}\,(C) \Rightarrow R$ by Yoneda Lemma where $f^*(\eta^r) : f^*(\mathbf{Yon}\,(C)) =: AC \longrightarrow f^*(R)$ takes $a \in AC$ to $f^*(\eta^r)(a)$. The well definiteness of e_R can be checked readily, that is, the fact that $e_R((Rg(r) \otimes a')_{C'}) = e_R((r \otimes Ag(a'))_C)$ can be seen via unraveling of definitions and Yoneda lemma. Now, we wish to show that e_R is an isomorphism. To this extent we just need to show that $e_{\mathbf{Yon}(C)}$ is an isomorphism because every presheaf is a colimit of representables and $-\otimes_{\mathbf{C}} A$ is itself a colimit construction. In order to show this, we have

$$e_{\mathbf{Yon}(C)} : \mathbf{Yon}\,(C) \otimes_{\mathbf{C}} A \longrightarrow f^*(\mathbf{Yon}\,(arg1))$$

But we have that $A = f^* \circ \mathbf{Yon}\,(-) : \mathbf{C} \longrightarrow \mathbf{Sets}$ and because of the fact that $\mathbf{Yon}\,(C) \otimes_{\mathbf{C}} F \cong FC$ for any $F : \mathbf{C} \to \mathbf{Sets}$[16], therefore $\mathbf{Yon}\,(C) \otimes_{\mathbf{C}} A \cong f^*(\mathbf{Yon}\,(C))$, hence $e_{\mathbf{Yon}(C)}$ is an isomorphism. □

Now, by the virtue of tensor product, we already have an adjunction:

$$\begin{tikzcd} \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \arrow[rr, bend left, "-\otimes_{\mathbf{C}} A"] \arrow[rr, phantom, "\perp"] & & \mathbf{Sets} \arrow[ll, bend left, "\mathrm{Hom}_{\mathbf{Sets}}(A(\cdot), -)"] \end{tikzcd} .$$

So if we forcefully assume that $-\otimes_{\mathbf{C}} A$ is left-exact, a condition we then define as *flatness* of A, then, we can safely say that any such flat functor $A : \mathbf{C} \longrightarrow \mathbf{Sets}$ gives a point of the presheaf topos $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ because we then have the geometric morphism as above. This leads to following proposition, which we have just proved:

Proposition 13.5.2.4. *Suppose* $\mathbf{C}$ *is a small category and* $A : \mathbf{C} \longrightarrow \mathbf{Sets}$ *is a* ***flat*** *functor, then* $\exists$ *a unique point of* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$

$$f : \mathbf{Sets} \longrightarrow \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

where $f^* := -\otimes_{\mathbf{C}} A$ *and* $f_* := \mathrm{Hom}_{\mathbf{Sets}}\,(A(\cdot), -)$.

We therefore have the following equivalence:

$$\boxed{\mathrm{Flat}(\mathbf{C}, \mathbf{Sets}) \equiv \mathrm{Hom}_{\mathbf{Topoi}}\left(\mathbf{Sets}, \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}\right)}$$

where $\mathrm{Flat}(\mathbf{C}, \mathbf{Sets})$ is the category of flat functors $A : \mathbf{C} \longrightarrow \mathbf{Sets}$ and natural transformations.

[16] This happens by Adjoint isomorphism and Yoneda Lemma:

$$\begin{aligned} \mathrm{Hom}_{\mathbf{Sets}}\,(\mathbf{Yon}\,(C) \otimes_{\mathbf{C}} A, S) &\cong \mathrm{Hom}_{\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}}\,(\mathbf{Yon}\,(C), \mathrm{Hom}_{\mathbf{Sets}}\,(A(-), S)) \\ &= \mathrm{Nat}\,(\mathrm{Hom}_{\mathbf{C}}\,(-, C), \mathrm{Hom}_{\mathbf{Sets}}\,(A(-), S)) \\ &\cong \mathrm{Hom}_{\mathbf{Sets}}\,(AC, S). \end{aligned}$$

By generalized elements, we have the isomorphism $\mathbf{Yon}\,(C) \otimes_{\mathbf{C}} A \cong AC$.

13.6 Categorical Semantics

What we have seen so far is the fact that a topos is an another framework/universe in which one can do *sets-like mathematics*. But one of the important facets of the category of **Sets** is that one can interpret logical theories in it, for example, the interpretation of theory of abelian groups in sets leads to an abelian group in the usual set-theoretic sense. But we already saw above that one can interpret abelian groups in a category with *enough structure*, which was an internal group object. What we saw there was an example of categorical semantics; interpreting syntactic languages in a category (with enough structure). In this section we would develop this line of thinking more formally. We follow the *Elephant* Section D1.2 [**Elephant**] for the discussion below.

13.6.1 Σ-Structures

Definition 13.6.1.1. (**Σ-Structure in a Category**) *Suppose Σ is a first order signature of a language and* **C** *is a category with finite products. A Σ-structure on* **C** *is a map M from the signature Σ to* **C** *which takes:*

- ***Finite list of Sorts*** $A_1, \dots, A_n \in \Sigma - \text{Sort}$ *to an object*

$$MA_1 \times \cdots \times MA_n$$

  in **C**. *The empty list is mapped to terminal object* **1** *in* **C**.
- ***Function symbol*** $f : A_1 \dots A_n \longrightarrow B \in \Sigma - \text{Fun}$ *to an arrow*

$$Mf : MA_1 \times \cdots \times MA_n \longrightarrow MB$$

  in **C**.
- ***Relation Symbol*** $R \rightarrowtail A_1 \dots A_n \in \Sigma - \text{Rel}$ *to a monomorphism*

$$MR \rightarrowtail MA_1 \times MA_n$$

  in **C**.

The collection of all Σ-structures over a category themselves form a category:

Definition 13.6.1.2. (**Category of Σ-Structures**) *Suppose Σ is a signature and* **C** *is a category with finite products. We can then form a category denoted*

$$\Sigma - \mathbf{Str}\,(\mathbf{C})$$

whose:

- ***Objects*** *are Σ-structures $M, N, \dots$.*
- ***Arrows*** *are Σ-structure homomorphisms between two Σ-structures M & N, which are denoted by:*

$$h : M \longrightarrow N$$

  and defined as a collection of arrows in **C**

$$\{h_A : M_A \longrightarrow N_A\}_{A \in \Sigma - \text{Sort}}$$

  for which the following two conditions hold:

– *The following commutes for each* $f : A_1 \dots A_N \longrightarrow B$ *in* $\Sigma - \mathrm{Fun}$:

$$\begin{tikzcd} MA_1 \times \cdots \times MA_n \arrow[r, "Mf"] \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] & MB \arrow[d, "h_B"] \\ NA_1 \times \cdots \times NA_n \arrow[r, "Nf"] & NB \end{tikzcd} .$$

– *The following commutes for each* $R \rightarrowtail A_1 \dots A_n$ *in* $\Sigma - \mathrm{Rel}$:

$$\begin{tikzcd} MR \arrow[r, tail] \arrow[d, "h_R"] & MA_1 \times \cdots \times MA_n \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] \\ NR \arrow[r, tail] & NA_1 \times \cdots \times NA_n \end{tikzcd} .$$

13.6.2 Terms

The term of some sort of a signature can also be interpreted as an arrow in a category with finite products:

Definition 13.6.2.1. (**Term in a Σ-Structure**) *Suppose Σ is a signature and M is a Σ-structure over a category $\mathbf{C}$ with finite products. Let $\vec{x}.t$ be a term in a context $\vec{x} = \{x_1, \dots, x_n\}$ where $x_i : A_i$, $1 \leq i \leq n$ and $t : B$. Then the same term $\vec{x}.t$ is interpreted in the Σ-structure M as an arrow denoted*

$$MA_1 \times \cdots \times MA_n \xrightarrow{[[\vec{x}.t]]_M} MB$$

generated by the following conditions:

- *If the term $t : B$ is simply a variable of sort B, then t must be some $x_i : A_i$ from the context $\vec{x}$, and therefore the corresponding arrow simply becomes the following projection:*

$$[[\vec{x}.t]]_M := \pi_i : MA_1 \times \cdots \times MA_n \longrightarrow MA_i.$$

- *If the term $t : B$ is actually the term $f(t_1, \dots, t_m) : B$ for some $f \in \Sigma - \mathrm{Fun}$ and $\vec{x}.t_i : C_i$ are other terms, then the arrow $[[\vec{x}.t]]$ would be the composite*[17]*:*

$$MA_1 \times \cdots \times MA_n \xrightarrow{\langle [[\vec{x}.t_1]]_M, \dots, [[\vec{x}.t_m]]_M \rangle} MC_1 \times \dots MC_m \xrightarrow{Mf} MB$$

Example 13.6.2.2. A term $\vec{x}.t : B$ in a signature Σ can be constructed by $f(g(h(a(x_1), b(c(x_2)))))$ where $\vec{x} = \{x_1, x_2\}$ and x_1, x_2 are variables of sorts $A_1, A_2 \in \Sigma{-}\mathrm{Sort}$ respectively and $f, g, h, a, b, c \in \Sigma - \mathrm{Fun}$ with *target* of f being B. Then the corresponding interpretation $[[\vec{x}.t]] : MA_1 \times MA_2 \longrightarrow MB$ would be given by the following composition:

$$MA_1 \times MA_2 \xrightarrow{\langle Ma, Mb \circ Mc \rangle} M\mathrm{cod}\,(a) \times M\mathrm{cod}\,(b) \xrightarrow{Mf \circ Mg \circ Mh} MB \ .$$

[17] Note that the terms $t_i,\ 1 \leq i \leq m$ are also in context $\vec{x}$, so the arrows $[[\vec{x}.t_i]] : MA_1 \times \cdots \times MA_n \longrightarrow MC_i$ are the term arrows in $\mathbf{C}$ for them.

We can also interpret the usual substitution of a term $\vec{x}.t$ by list of terms $\vec{s}$, denoted $t[\vec{s}/\vec{x}]$, in a category by the help of implicit composition in that category.

Proposition 13.6.2.3. *Suppose M is a Σ-structure over a category $\mathbf{C}$ with finite products. Let $\vec{y}.t : C$ be a term of sort C in context $\vec{y},\ y_i : B_i$. Suppose $\vec{s}$ is a list of terms and this list has same length and type as $\vec{y}$ and let $\vec{x},\ x_i : A_i$ be a common context for each term $s_i \in \vec{s}$. Then, the term $\vec{x}.t[\vec{s}/\vec{y}]$ is interpreted as the following arrow given by composition in $\mathbf{C}$:*

$$MA_1 \times \cdots \times MA_n \xrightarrow{\langle [[\vec{x}.s_1]]_M, \ldots, [[\vec{x}.s_m]]_M \rangle} MB_1 \times \cdots \times MB_m \xrightarrow{[[\vec{y}.t]]_M} MC.$$

Proof. Each term is given by a chain of application of function symbols over all or some of the variables present in context. Therefore, the substitution in term $\vec{y}.t$ by terms $\vec{s}$ is again a term in which there are more applications of function symbols on some or all the arguments of the context $\vec{x}$. Now, using Definition 13.6.2.1, specifically the second point, we get the desired result by unrolling the whole chain of application of functions inductively. □

We now see that the homomorphism of Σ-structures preserves the interpretation of terms upto *naturality*:

Proposition 13.6.2.4. *Suppose M & N are two Σ-structures and $h : M \longrightarrow N$ is a Σ-structure homomorphism. If $\vec{x}.t$ is term in Σ where $t : B$ and $x_i : A_i$, then, the following commutes:*

$$\begin{tikzcd} MA_1 \times \cdots \times MA_n \arrow[r, "{[[\vec{x}.t]]_M}"] \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"'] & MB \arrow[d, "h_B"] \\ NA_1 \times \cdots \times NA_n \arrow[r, "{[[\vec{x}.t]]_N}"'] & NB \end{tikzcd}$$

Proof. Denote the term $\vec{x}.t$ as a chain of application of function symbols as in the example above. Since the arrow $[[\vec{x}.t]]_M$ would be the composition of all arrows involved, similarly for $[[\vec{x}.t]]_N$, and since each individual arrow in the composition for M would have corresponding arrows from domain object and target object to that of N for which the natural square would commute by Definition 13.6.1.2, therefore the whole big rectangle will commute. This big rectangle is clearly the one required. □

13.6.3 Formulae

Definition 13.6.2.1 tells us how to interpret terms of a signature. The next step would thus be to interpret a formula of some signature in a category. As we know, formulas of some signature themselves are categorized by several restrictions. These restrictions are atomic, horn, regular, coherent, first order, geometric & infinitary first order formulas. Each of which would thus be interpreted in a category which would have enough structure suitable for their residence in it.

Definition 13.6.3.1. (**Formula in a Σ-Structure**) Suppose Σ is a signature and M is a Σ-structure in a category $\mathbf{C}$ which has finite limits. Let $\vec{x}.\phi$ be a formula in context $\vec{x},\ x_i : A_i,\ 1 \leq i \leq n$. The formula $\vec{x}.\phi$ would be interpreted in M as the subobject

$$[[\vec{x}.\phi]] \rightarrowtail MA_1 \times \cdots \times MA_n$$

in **C** and this subobject $[[\vec{x}.\phi]]$ is generated recursively by the following :

1. **RELATIONS** : If ϕ is simply $R(t_1,\ldots,t_m)$ where $R \in \Sigma - \mathrm{Rel}$ and $t_j : B_j,\ 1 \le j \le m$, then the subobject $[[\vec{x}.\phi]]$ is the following pullback:

$$\begin{tikzcd} {[[\vec{x}.\phi]]} \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & MR \arrow[d, tail] \\ MA_1 \times \cdots \times MA_n \arrow[r, "{\langle [[\vec{x}.t_1]]_M,\ldots,[[\vec{x}.t_m]]_M \rangle}"'] & MB_1 \times \cdots \times MB_m \end{tikzcd} .$$

2. **EQUALITY** : If ϕ is simply $\vec{x}.s = \vec{x}.t$, an equality of terms of some sort B, then the subobject $[[\vec{x}.\phi]]$ is the following equalizer:

$$\begin{tikzcd} {[[\vec{x}.\phi]]} \arrow[r, tail] & MA_1 \times \cdots \times MA_n \arrow[r, shift left, "{[[\vec{x}.s]]_M}"] \arrow[r, shift right, "{[[\vec{x}.t]]_M}"'] & MB \end{tikzcd} .$$

3. **TRUTH** : If ϕ is $\top$, i.e. truth, then the subobject $[[\vec{x}.\phi]]$ is simply the top element of the lattice $\mathrm{Sub}_{\mathbf{C}}(MA_1 \times \cdots \times MA_n)$:

$$[[\vec{x}.\phi]] = \top \rightarrowtail MA_1 \times \cdots \times MA_n \ .$$

4. **BINARY MEETS** : If ϕ is $\psi \wedge \chi$ for other formulas ψ and χ, then the subobject $[[\vec{x}.\phi]]$ is the following pullback:

$$\begin{tikzcd} {[[\vec{x}.\phi]]} \arrow[r, tail] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & {[[\vec{x}.\chi]]} \arrow[d, tail] \\ {[[\vec{x}.\psi]]} \arrow[r, tail] & MA_1 \times \cdots \times MA_n \end{tikzcd} .$$

5. **FALSITY** : If ϕ is $\perp$, then the subobject $[[\vec{x}.\phi]]$ is the bottom element of the lattice $\mathrm{Sub}_{\mathbf{C}}(MA_1 \times \cdots \times MA_n)$:

$$[[\vec{x}.\phi]] = \perp \rightarrowtail MA_1 \times \cdots \times MA_n \ .$$

6. **BINARY JOINS** : If ϕ is $\psi \vee \chi$ for other formulas ψ and χ and **C** is a coherent category[18], then the subobject $[[\vec{x}.\phi]]$ is the join of the two terms as subobjects:

$$\begin{tikzcd} {[[\vec{x}.\psi]] \amalg [[\vec{x}.\chi]]} \arrow[dr, two heads, "e"] \arrow[ddrr, bend right, "!"'] & & {[[\vec{x}.\chi]]} \arrow[ll, "i_2"'] \arrow[dd, tail] \\ & {[[\vec{x}.\phi]]} \arrow[dr, tail, "m"] & \\ {[[\vec{x}.\psi]]} \arrow[uu, "i_1"] \arrow[rr, tail] & & MA_1 \times \cdots \times MA_n \end{tikzcd}$$

[18] A coherent category is a regular category in which each subobject poset have finite joins and change of base functor for any arrow preserves these finite joins. Also remember the arrow factorization property in a regular category.

7. **IMPLICATION** : If ϕ is $\psi \Rightarrow \chi$ for other formulas ψ and χ and $\mathbf{C}$ is a Heyting category[19], then the subobject $[[\vec{x}.\phi]]$ is the implication $[[\vec{x}.\psi]] \Rightarrow [[\vec{x}.\chi]]$ in the Heyting algebra $\mathrm{Sub}_{\mathbf{C}}\,(MA_1 \times \cdots \times MA_n)$:

$$[[\vec{x}.\phi]] = \forall_{m_{[[\vec{x}.\psi]]}}([[\vec{x}.\psi]] \cap [[\vec{x}.\chi]]) \rightarrowtail MA_1 \times \cdots \times MA_n$$

where $\forall_{m_{[[\vec{x}.\psi]]}} : \mathrm{Sub}_{\mathbf{C}}\,([[\vec{x}.\psi]]) \longrightarrow \mathrm{Sub}_{\mathbf{C}}\,(MA_1 \times \cdots \times MA_n)$.

8. **NEGATION** : If ϕ is $\neg\psi$ for some formula ψ and $\mathbf{C}$ is a Heyting category, then the subobject $[[\vec{x}.\phi]]$ is the negation $\neg[[\vec{x}.\psi]]$ in the Heyting algebra $\mathrm{Sub}_{\mathbf{C}}\,(MA_1 \times \cdots \times MA_n)$:

$$[[\vec{x}.\phi]] = ([[\vec{x}.\psi]] \Rightarrow \bot) \longrightarrow MA_1 \times \cdots \times MA_n \ .$$

9. **EXISTENTIAL QUANTIFICATION** : If ϕ is $(\exists y)\psi$ where y is a variable of sort B and $\mathbf{C}$ is a regular category, then the subobject $[[\vec{x}.\phi]]$ is the following image[20]:

$$\begin{tikzcd}
{[[(\vec{x},y).\psi]]} \arrow[rd, two heads, "e"] \arrow[dd, tail] & & \\
& {[[\vec{x}.\phi]]} \arrow[rd, tail, "m"] & \\
MA_1 \times \cdots \times MA_n \times MB \arrow[rr, "\pi"'] & & MA_1 \times \cdots \times MA_n
\end{tikzcd}$$

where π is the unique arrow due to projection onto first n terms.

10. **UNIVERSAL QUANTIFICATION** : If ϕ is $(\forall y)\psi$ where y is a variable of sort B and $\mathbf{C}$ is a Heyting category, then the subobject $[[\vec{x}.\phi]]$ is the following (see footnote 27):

$$[[\vec{x}.\phi]] = \forall_\pi\,([[(\vec{x},y).\phi]]) \rightarrowtail MA_1 \times \cdots \times MA_n$$

11. **INFINITARY JOINS** : If ϕ is $\bigvee_{i\in I}\psi_i$ where ψ_i are other formulas and $\mathbf{C}$ is a geometric category[21], then the subobject $[[\vec{x}.\phi]]$ is the join $\bigcup_{i\in I}[[\vec{x}.\psi_i]]$ in the join-semilattice $\mathrm{Sub}_{\mathbf{C}}\,(MA_1 \times \cdots \times MA_n)$:

$$[[\vec{x}.\phi]] = \textstyle\bigcup_{i\in I}[[\vec{x}.\psi_i]] \rightarrowtail MA_1 \times \cdots \times MA_n \ .$$

12. **INFINITARY MEETS** : If ϕ is $\bigwedge_{i\in I}\psi_i$ where ψ_i are other formulas and $\mathbf{C}$ has arbitrary meets of subobjects, then the subobject $[[\vec{x}.\phi]]$ is given by the meet $\bigcap_{i\in I}[[\vec{x}.\psi_i]]$ in the meet-semilattice $\mathrm{Sub}_{\mathbf{C}}\,(MA_1 \times \ldots MA_n)$:

$$[[\vec{x}.\phi]] = \textstyle\bigcap_{i\in I}[[\vec{x}.\psi_i]] \rightarrowtail MA_1 \times \cdots \times MA_n \ .$$

[19] A category $\mathbf{C}$ is Heyting if it is coherent and for any arrow $f : X \to Y$, the change of base functor

$$f^* : \mathrm{Sub}_{\mathbf{C}}\,(Y) \longrightarrow \mathrm{Sub}_{\mathbf{C}}\,(X)$$

has a right adjoint $\forall_f : \mathrm{Sub}_{\mathbf{C}}\,(X) \longrightarrow \mathrm{Sub}_{\mathbf{C}}\,(Y)$. It follows that in a Heyting category, as expected, each subobject lattice is a Heyting algebra.

[20] $\vec{x}, y$ denotes that we have extended our initial context $\vec{x}$ to now also include y.

[21] A geometric category is just an infinitary coherent category, meaning that each subobject poset have infinitary joins and any change of base functor preserves them.

Remark 13.6.3.2. (**Infinitary first order formulas can just be interpreted in a geometric category**) Suppose **C** is a geometric category. Then **C** is also a Heyting category. Therefore each subobject lattice has infinitary meets, hence the heading. However, the change of base functor may not preserve the infinitary meets, i.e., change of base functor in this geometric category may not preserve non-geometric formulas.

Remark 13.6.3.3. Suppose M is a Σ-structure on a category **C** and **C** has enough structure to interpret a particular formula $\vec{x}.\phi$. We then call ϕ to be interpretable in category **C**.

The substitution property extends from terms to formulas:

Proposition 13.6.3.4. *Suppose Σ is a signature and $\vec{y}.\phi$ is a formula in Σ in context $\vec{y}$ which can be interpreted in a category* $\mathbf{C}$[22] *and M is a Σ-structure on* **C**. *Let $\vec{s}$ be a list of terms with same length and type as that of context $\vec{y}$, where each term $s_i \in \vec{s}$ is in common context $\vec{x}$. Then, the formula $[[\vec{x}.\phi[\vec{s}/\vec{y}]]]$ is given as the following pullback (Note $x_i : A_i,\ 1 \leq i \leq n$, $y_i : B_i,\ 1 \leq i \leq m$):*

$$\begin{tikzcd}
{[[\vec{x}.\phi[\vec{s}/\vec{y}]]]_M} \arrow[rr] \arrow[d, tail] \arrow[drr, phantom, "\lrcorner", very near start] & & {[[\vec{y}.\phi]]_M} \arrow[d, tail] \\
MA_1 \times \cdots \times MA_n \arrow[rr, "{\langle [[\vec{x}.s_1]]_M, \ldots, [[\vec{x}.s_m]]_M \rangle}"'] & & MB_1 \times \cdots \times MB_m
\end{tikzcd}.$$

Proof. Any formula, as Definition 13.6.3.1 instructs, is fundamentally generated from relations and equality. The fact that the formula $\vec{y}.\phi$ is substituted by $\vec{s}$ to give a new formula $\vec{x}.\phi[\vec{s}/\vec{y}]$ just makes the previous formula stated in new terms. We thus have two formulas, $[[\vec{y}.\phi]]$ and $[[\vec{x}.\phi[\vec{s}/\vec{y}]]]$. If $\vec{y}.\phi$ is simply a relation or an equality, then by Proposition 13.6.2.3 one can see that the later subobject $[[\vec{x}.\phi[\vec{s}/\vec{y}]]]$ is just the relevant pullback. Since the result hold for atomic formulas, therefore it will hold for all generated from other constructions. $\square$

Unfortunately, not all formulas in context are natural with respect to Σ-structure homomorphisms:

Proposition 13.6.3.5. *Let* **C** *be atleast a cartesian*[23] *category and let $\vec{x}.\phi$ be a geometric formula in context over Σ which is interpretable in* **C**. *Suppose $h : M \longrightarrow N$ is a homomorphism of Σ-structures. Then there is a commutative square (Note $x_i : A_i,\ 1 \leq i \leq n$):*

$$\begin{tikzcd}
{[[\vec{x}.\phi]]_M} \arrow[r, tail] \arrow[d] & MA_1 \times \cdots \times MA_n \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"] \\
{[[\vec{x}.\phi]]_N} \arrow[r, tail] & NA_1 \times \cdots \times NA_n
\end{tikzcd}.$$

[22] That is, **C** has enough structure to interpret $\vec{y}.\phi$ in itself, as instructed in Definition 13.6.3.1.

[23] One which has all finite limits.

Proof. Again, we wish to see whether there is a commutative square for relations and equality. For equality, this can be seen easily; suppose ϕ is $\vec{x}.s = \vec{x}.t$ where s and t are terms of sort B. Then, we have the following diagram (The two squares on the right commute due to Proposition 13.6.2.4):

$$\begin{tikzcd}
& {[[\vec{x}.\phi]]_N} \arrow[r, tail] & NA_1 \times \cdots \times NA_n \arrow[r, shift left] \arrow[r, shift right] & NB \\
& P \arrow[u] \arrow[r, tail] \arrow[ur, phantom, "\urcorner", very near start] & MA_1 \times \cdots \times MA_n \arrow[u, "h_{A_1}\times\cdots\times h_{A_n}"] \arrow[r, shift left] \arrow[r, shift right] & MB \arrow[u, "h_B"'] \\
{[[\vec{x}.\phi]]_M} \arrow[uur, dashed] \arrow[urr, tail] & & &
\end{tikzcd}$$

where the dotted arrow is obtained from universal property of equalizer $[[\vec{x}.\phi]]_N$. Therefore we have a cone over the pullback P and so there is a unique arrow $[[\vec{x}.\phi]]_M \to P$ through which the subobject $[[\vec{x}.\phi]]_M$ factors through. Clearly, this arrow $[[\vec{x}.\phi]] \to P$ has to be a monic. Hence we have a commutative square as required. Note that we only needed the universal property of the equalizer for the equality, hence the same will hold in case of relations. Since a geometric formula is generated by those basic formulas interpretation of whom preserves finite limits, therefore we have the required commutative square for all geometric formulas interpretable in **C**. □

The above motivates the following definitions:

Definition 13.6.3.6. *Suppose Σ is a signature and $h : M \longrightarrow N$ is a homomorphism of Σ-structures over* **C***. Then,*

1. (**Elementary Morphism**) *The morphism $h : M \longrightarrow N$ is called an elementary morphism if for each* ***first-order formula*** *in context $\vec{x}.\phi$ over Σ where $x_i : A_i,\ 1 \le i \le n$, there is a* ***commutative square*** *as shown:*

$$\begin{tikzcd}
{[[\vec{x}.\phi]]_M} \arrow[r, tail] \arrow[d] & MA_1 \times \cdots \times MA_n \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] \\
{[[\vec{x}.\phi]]_N} \arrow[r, tail] & NA_1 \times \cdots \times NA_n
\end{tikzcd}.$$

2. (**Elementary Embedding**) *The morphism $h : M \longrightarrow N$ is called an elementary morphism if for each* ***first-order formula*** *in context $\vec{x}.\phi$ over Σ where $x_i : A_i,\ 1 \le i \le n$, there is a* ***pullback*** *as shown:*

$$\begin{tikzcd}
{[[\vec{x}.\phi]]_M} \arrow[r, tail] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & MA_1 \times \cdots \times MA_n \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] \\
{[[\vec{x}.\phi]]_N} \arrow[r, tail] & NA_1 \times \cdots \times NA_n
\end{tikzcd}.$$

3. (**Embedding**) *The morphism $h : M \longrightarrow N$ is called an elementary morphism if for each* ***atomic***

formula *in context $\vec{x}.\phi$ over Σ where $x_i : A_i,\ 1 \leq i \leq n$, there is a* ***pullback*** *as shown:*

$$\begin{tikzcd}
{[[\vec{x}.\phi]]_M} \arrow[r, tail] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & MA_1 \times \cdots \times MA_n \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] \\
{[[\vec{x}.\phi]]_N} \arrow[r, tail] & NA_1 \times \cdots \times NA_n
\end{tikzcd} .$$

13.6.4 Theories and Models

We now come to models of a theory over a signature, defining both:

Definition 13.6.4.1. *Suppose Σ is a signature and M be a Σ-structure over a category* **C***. Then,*

1. (**Satisfiability of a Sequent**) *Let $\sigma = \phi \vdash_{\vec{x}} \psi$ be a sequent over Σ interpretable*[24] *in* **C***. Then the sequent σ is defined to be satisfiable in the structure M if*

$$[[\vec{x}.\phi]]_M \leq [[\vec{x}.\psi]]_M$$

in $\mathrm{Sub}_{\mathbf{C}}(MA_1 \times \cdots \times MA_n)$ where $\leq$ is the order induced by the subobject lattice. The satisfiability of the sequent σ in structure M is denoted as:

$$M \vDash \sigma.$$

2. (**Model of a Theory**) *Let $\mathbb{T}$ be a theory over Σ interpretable in* **C**[25]*. Then the structure M over* **C** *is said to be a model of the theory $\mathbb{T}$ if all the axioms/sequents of the theory $\mathbb{T}$ is satisfiable in M. We denote a model M of a theory $\mathbb{T}$ by:*

$$M \models \mathbb{T}.$$

3. (**Category of Models of a Theory**) *Let $\mathbb{T}$ be a theory over a signature Σ. The full-subcategory of $\Sigma - \mathbf{Str}(\mathbf{C})$ where structures(objects) are all the models of the theory $\mathbb{T}$ and structure morphisms between them is called the category of models of theory $\mathbb{T}$ and is denoted as:*

$$\mathbb{T} - \mathbf{Mod}(\mathbf{C}).$$

Remark 13.6.4.2. The subcategory of all the models of a theory $\mathbb{T}$ over a signature Σ on category **C** between whom the structure morphisms are elementary morphisms is denoted as:

$$\mathbb{T} - \mathbf{Mod}(\mathbf{C})_e.$$

Before going further, we need to know particularities about functors between two categories with appropriate structures:

1. If categories **C** and **D** are cartesian, then a **cartesian functor** $F : \mathbf{C} \longrightarrow \mathbf{D}$ is defined to be a functor which preserves finite limits.
2. If categories **C** and **D** are regular, then a **regular functor** $F : \mathbf{C} \longrightarrow \mathbf{D}$ is defined to be a functor which preserves finite limits and regular epimorphisms.

[24] A sequent is *interpretable* if each formula in the sequent is interpretable in **C**.

[25] A theory is *interpretable* in **C** if all the sequents of the theory are interpretable.

3. If categories $\mathbf{C}$ and $\mathbf{D}$ are coherent, then a **coherent functor** $F : \mathbf{C} \longrightarrow \mathbf{D}$ is defined to be a functor which preserves finite limits, regular epimorphisms and finite joins of subobjects.
4. If categories $\mathbf{C}$ and $\mathbf{D}$ are Heyting, then a **Heyting functor** $F : \mathbf{C} \longrightarrow \mathbf{D}$ is defined to be a functor which preserves finite limits, regular epimorphisms and right adjoints $\forall_f$: $\mathrm{Sub}_{\mathbf{C}}(X) \to \mathrm{Sub}_{\mathbf{C}}(Y)$ for any arrow $f : X \to Y$ in $\mathbf{C}$.
5. If categories $\mathbf{C}$ and $\mathbf{D}$ are geometric, then a **geometric functor** $F : \mathbf{C} \longrightarrow \mathbf{D}$ is defined to be a functor which preserves finite limits, regular epimorphisms and infinitary joins of subobjects.

The above defined functors preserves satisfiability of a sequent:

Proposition 13.6.4.3. *Let $T : \mathbf{C} \longrightarrow \mathbf{D}$ be a cartesian (regular, coherent, Heyting, geometric) functor. Let M be a Σ-structure over $\mathbf{C}$ and let σ be a sequent over Σ interpretable in $\mathbf{C}$. If $M \vDash \sigma$, then, $\Sigma - \textbf{\textit{Str}}(T) : \Sigma - \textbf{\textit{Str}}(\mathbf{C}) \longrightarrow \Sigma - \textbf{\textit{Str}}(\mathbf{D})$, which takes each interpretation through T to get an interpretation over $\mathbf{D}$, gives a structure over $\mathbf{D}$ for which σ is still satisfiable, that is,*

$$\Sigma - \textbf{\textit{Struc}}(T)(M) \vDash \sigma.$$

Proof. Take $\mathbf{C}$ to be cartesian (respectively, all else). Suppose $\sigma = \phi \vdash_{\vec{x}} \psi$ is a sequent satisfiable in M and interpretable in $\mathbf{C}$, where $x_i : A_i,\ 1 \le i \le n$. Hence, we have $[[\vec{x}.\phi]]_M \le [[\vec{x}.\psi]]_M$ in the $\mathrm{Sub}_{\mathbf{C}}(MA_1 \times \cdots \times MA_n)$. Since for two subobjects S, T in any $\mathrm{Sub}_{\mathbf{C}}(A)$, we say $S_1 \le S_2$ if and only if $S_1 \cap S_2 = S_1$ where $\cap$ is the meet of two subobjects given by pullback of them (which exists since $\mathbf{C}$ has finite limits (atleast)). Since T preserves limits as it is cartesian (atleast), therefore if $S_1 \cap S_2 = S_1$, then $T(S_1) = T(S_1 \cap S_2) \cong T(S_1) \cap T(S_2)$. Hence, if $[[\vec{x}.\phi]]_M \le [[\vec{x}.\psi]]_M$, then $T([[\vec{x}.\phi]]_M) \le T([[\vec{x}.\psi]]_M)$, which just means that if $M \vDash \sigma$, then $\Sigma - \mathbf{Struc}(T)(M) \vDash \sigma$. ☐

The converse of the above proposition additionally requires T to be conservative:

Proposition 13.6.4.4. *Let $T : \mathbf{C} \longrightarrow \mathbf{D}$ be a cartesian (regular, coherent, Heyting, geometric) functor which is additionally conservative[26]. Let M be a Σ-structure over $\mathbf{C}$ and let σ be a sequent over Σ interpretable in $\mathbf{C}$. If $\Sigma - \textbf{\textit{Struc}}(T)(M) \vDash \sigma$, then, $M \vDash \sigma$.*

Proof. If $\Sigma - \mathbf{Struc}(T)(M) \vDash \sigma$, then $T([[\vec{x}.\phi]]_M) \le T([[\vec{x}.\psi]]_M)$ which means that $T([[\vec{x}.\phi]]_M \cap [[\vec{x}.\psi]]_M) \cong T([[\vec{x}.\phi]]_M)$. But $[[\vec{x}.\phi \wedge \psi]]_M := [[\vec{x}.\phi]]_M \cap [[\vec{x}.\psi]]_M$ (Definition 13.6.3.1). Therefore $T([[\vec{x}.\phi \wedge \psi]]_M) \cong T([[\vec{x}.\phi]]_M)$. By conservativity of T, we have that $[[\vec{x}.\phi]]_M \cong [[\vec{x}.\phi \wedge \psi]]_M$, which is what we wanted. ☐

Example 13.6.4.5. (**The Category of Models of the Theory of Abelian Groups over Sets is AbGrp**) We wish to show the following where $\mathbb{T}$ is the theory of abelian groups over it's canonical signature with one sort G, three function symbols: $f : G, G \to G$, $1 : [] \to G$, $i : G \to G$; and no relation symbols. We wish to show:

$$\mathbb{T} - \mathbf{Mod}(\mathbf{Sets}) \cong \mathbf{AbGrp}$$

First, the theory of abelian groups consists of the following four atomic sequents/axioms:
1. $\top \vdash_{\{x,y,z\}} m(m(x,y),z) = m(x, m(y,z))$

[26]A functor $F : \mathbf{C} \longrightarrow \mathbf{D}$ is conservative if for any $f : C \to C'$ in $\mathbf{C}$, we have that $Ff : FC \to FC'$, then $f : C \to C'$ was an isomorphism.

2. $\top \vdash_{\{x,y\}} m(x,y) = m(y,z)$
3. $\top \vdash_{\{x\}} m(x,1()) = x$
4. $\top \vdash_{\{x\}} m(x,i(x)) = 1()$

To show first that each model defines a unique group, take any model M of $\mathbb{T}$ over **Sets**. Since M is a model, therefore for each axiom $\top \vdash_{\vec{x}} \phi$ of $\mathbb{T}$, the following is true:

$$[[\vec{x}.\top]]_M \leq [[\vec{x}.\phi]]_M.$$

But this means that

$$[[\vec{x}.(\top \wedge \phi)]]_M \cong [[\vec{x}.\top]]_M.$$

Moreover, we know that

$$[[\vec{x}.\top]]_M = MA_1 \times \dots MA_n$$

where $x_i : A_i$. This means that

$$[[\vec{x}.\top]]_M \cap [[\vec{x}.\phi]]_M = [[\vec{x}.(\top \wedge \phi)]]_M = [[\vec{x}.\top]]_M = MA_1 \times \cdots \times MA_n$$

and so

$$[[\vec{x}.\phi]]_M = MA_1 \times \cdots \times MA_n.$$

Therefore, for associativity of $Mm : MG \times MG \longrightarrow MG$:

$$[[\{x,y,z\}.m(m(x,y),z) = m(x,m(y,z))]]_M = MG \times MG \times MG$$

which means Mm is associative for all x, y, z in MG. Hence Mm is associative. Similarly, Mm is commutative, each x in G has inverse and $1() \in MG$ is the identity. Hence (MG, Mm) is an abelian group determined by the model M. For each Σ-structure homomorphism $h : M \to N$, the only component gives rise to a group homomorphism $h_G : MG \to NG$ in **Sets**. This follows because of the natural square of the three function symbols. Similarly, each group homomorphism $h : G \to H$ determines a unique Σ-structure homomorphism because the conditions of group homomorphism are the ones which make the three natural squares commute.

Proposition 13.6.4.4 has an interesting corollary:

Proposition 13.6.4.6. *Let $\mathbb{T}$ be a geometric theory over a signature Σ. Then,*

1. *For any small category $\mathbf{C}$, a Σ-structure M in $\mathbf{Sets}^{\mathbf{C}}$ is a $\mathbb{T}$-model if and only if the Σ-structure $ev_C(M)$ on **Sets** for each $C \in \mathrm{Ob}(\mathbf{C})$ is a $\mathbb{T}$-model. The functor $ev_C(-) : \mathbf{Sets}^{\mathbf{C}} \longrightarrow \mathbf{Sets}$ takes a set-functor to the set obtained by evaluating it at C.*
2. *For any topological space X and for any element $x \in X$, the inverse image $x^* : Sh(X) \longrightarrow \mathbf{Sets}$ of the geometric morphism:*

$$\begin{tikzcd} x : \mathbf{Sets} \arrow[r, bend left, "x_*"] \arrow[r, phantom, "\top"] & Sh(X) \arrow[l, bend left, "x^*"] \end{tikzcd}$$

*is such that for a Σ-structure M in the category $Sh(X)$ is a $\mathbb{T}$-model if and only if the Σ-structure $x^*(M)$ in **Sets** is a $\mathbb{T}$-model. $x^*(M)$ takes the interpreation of M in $Sh(X)$ to an interpretation in **Sets** via composition with x^*.*

Proof. **1.** (L $\implies$ R) Let M be a $\mathbb{T}$-model over the category $\mathbf{Sets}^{\mathbf{C}}$. Now, for a fixed object $C \in \mathrm{Ob}(\mathbf{C})$, we have by composition a Σ-structure over the category **Sets**, $\mathrm{ev}_C(M)$, which takes each sort, function symbols & relation symbols first to it's interpretation in $\mathbf{Sets}^{\mathbf{C}}$ via M and then to **Sets** via evaluating that interpretation at $C \in \mathrm{Ob}(C)$. Note that $\mathrm{ev}_C(-) : \mathbf{Sets}^{\mathbf{C}} \longrightarrow \mathbf{Sets}$ is a geometric functor because limits of set functors is determined point-wise, an epimorphic natural transform is by definition one whose each component is an epimorphism and join of two subobjects in $\mathbf{Sets}^{\mathbf{C}}$ is join of their corresponding components (colimits are computed point-wise). By Proposition 13.6.4.3, since each sequent of $\mathbb{T}$ is satisfied by M as it is a $\mathbb{T}$-model, therefore $\mathrm{ev}_C(M)$ is also a Σ-structure over **Sets** where each sequent of $\mathbb{T}$ is satisfiable. Hence $\mathrm{ev}_C(M)$ is also a $\mathbb{T}$-model, for each $C \in \mathrm{Ob}(\mathbf{C})$.
(R $\implies$ L) It is quite simple to see that $\mathrm{ev}_C(-) : \mathbf{Sets}^{\mathbf{C}} \longrightarrow \mathbf{Sets}$ is conservative, because if for some natural transform $\eta : F \Rightarrow G$ in $\mathbf{Sets}^{\mathbf{C}}$, the function $\mathrm{ev}_C(\eta) := \eta_C : FC \longrightarrow GC$ is an isomorphism for each $C \in \mathrm{Ob}(\mathbf{C})$, then η is also a natural isomorphism as each natural transform is determined by it's components. Therefore by Proposition 13.6.4.4, we have that M is also a $\mathbb{T}$-model.

2. For $x : \mathbf{1} \longrightarrow X$, we have that the induced geometric morphisms $x : \mathrm{Sh}(\mathbf{1}) \cong \mathbf{Sets} \longrightarrow \mathrm{Sh}(X)$ is such that the inverse image (left-adjoint) $x^* : \mathrm{Sh}(X) \longrightarrow \mathbf{Sets}$ is a stalk functor, that is, it takes each sheaf F over X to stalk at x, F_x. Now because x^* is left adjoint so it preserves small colimits (and so coequalizers, hence regular epics), x^* is inverse image of a geomteric morphism so it preserves finite limits. Therefore x^* is a geometric functor. Now $x^* : \mathrm{Sh}(X) \longrightarrow \mathbf{Sets}$ is also conservative because if for some $\eta : F \Rightarrow G$ it s true that $x^*(\eta) : F_x \to G_x$ is an isomorphism for each $x \in X$, then η is a natural isomorphism. We can then follow the same argument as in **1** to conclude the result. $\square$

13.7 Topoi and Logic

We now study some of the interconnections between topoi and logic, studying Cohen's proof of independence of continuum hypothesis from ZFC axioms and a brief introduction to synthetic differential geometry, in between.

13.7.1 Natural Numbers Object in a Topos : $\mathbb{N}_{\mathbf{E}}$

The axioms of set theory demand existence of an infinite set, the set of natural numbers $\mathbb{N}$. In a topos, this axiom is interpreted as the existence of the following object:

Definition 13.7.1.1. (**Natural Numbers Object**) *Suppose* **E** *is a topos. An object* $\mathbb{N}_{\mathbf{E}}$ *in* **E** *is defined to be a natural numbers object if it has two arrows*

$$\mathbf{1} \rightarrowtail^{0} \mathbb{N}_{\mathbf{E}} \xrightarrow{s} \mathbb{N}_{\mathbf{E}}$$

such that for any other object X *with arrows* $\mathbf{1} \xrightarrow{x} X \xrightarrow{f} X$*, there exists a unique arrow* $h : \mathbb{N}_{\mathbf{E}} \longrightarrow X$

such that the following commutes:

$$\begin{tikzcd}
\mathbf{1} \arrow[r, tail, "0"] \arrow[d, equal] & \mathbb{N}_{\mathbf{E}} \arrow[r, "s"] \arrow[d, dashed, "h"'] & \mathbb{N}_{\mathbf{E}} \arrow[d, dashed, "h"] \\
\mathbf{1} \arrow[r, tail, "x"'] & X \arrow[r, "f"'] & X
\end{tikzcd} .$$

Remark 13.7.1.2. (**Natural Numbers Objects are unique upto isomorphism**) Take any other object N in $\mathbf{E}$ for which satisfies the condition of an NNO as in Definition 13.7.1.1 with the following defining arrows: $\mathbf{1} \xrightarrow{n} N \xrightarrow{f} N$. Hence there exists the unique arrows $a : \mathbb{N}_{\mathbf{E}} \longrightarrow N$ and $b : N \longrightarrow \mathbb{N}_{\mathbf{E}}$ which are universal, for which $a \circ b : N \longrightarrow N$ is such that $f \circ a \circ b = a \circ b \circ f$ and $a \circ b \circ x = x$. Now we have another $\mathbf{1} \xrightarrow{x} N \xrightarrow{a \circ b} N$, therefore $\exists$ unique $h : N \to N$ with $h \circ x = x$ and $a \circ b \circ h = h \circ f$. But $a \circ b \circ x = x$ too and $a \circ b$ is also unique, therefore $a \circ b = h$. Moreover, because $h \circ f = a \circ b \circ h$, therefore by uniqueness, $f = h \implies f = a \circ b$. Hence, $h = 1_N = a \circ b$. Similarly, $b \circ a = 1_{\mathbb{N}_{\mathbf{E}}}$.

We first see that each geometric morphism between two topoi in which one has an NNO, implies that the other one has it too:

Lemma 13.7.1.3. *Let* $\mathbf{E}, \mathbf{F}$ *be a topoi where* $\mathbf{E}$ *has a natural numbers object* $\mathbb{N}_{\mathbf{E}}$*. Let there be a following geometric morphism between* $\mathbf{F}$ *and* $\mathbf{E}$*:*

$$\begin{tikzcd}
g : \mathbf{F} \arrow[r, bend left, "g_*"] \arrow[r, phantom, "\top"] & \mathbf{E} \arrow[l, bend left, "g^*"]
\end{tikzcd} .$$

Then, $g^*(\mathbb{N}_{\mathbf{E}})$ *is an NNO for* $\mathbf{F}$*.*

Proof. It can be seen, that for each object X of $\mathbf{F}$ for which there are arrows $\mathbf{1}_{\mathbf{F}} \xrightarrow{x} X \xrightarrow{f} X$, there exists unique arrow $g^*(\mathbb{N}_{\mathbf{E}}) \longrightarrow X$ which is the transpose of the unique arrow $\mathbb{N}_{\mathbf{E}} \longrightarrow g_*(X)$, and hence it makes the corresponding square in $\mathbf{F}$ commute. This also depends on the fact g^* preserves finite limits as it is inverse image of a geometric morphism and so preserves terminals. □

This lemma has very important corollaries, first of which shows that **each presheaf category has an NNO**:

Corollary 13.7.1.4. *For a small category* $\mathbf{C}$*, the presheaf category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *has a natural numbers object.*

Proof. We have the global sections adjunction (a geometric morphism):

$$\begin{tikzcd}
\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \arrow[r, bend left, "\Gamma"] \arrow[r, phantom, "\top"] & \mathbf{Sets} \arrow[l, bend left, "\Delta"]
\end{tikzcd} .$$

Now use Lemma 13.7.1.3. Therefore the NNO for $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is the constant to $\mathbb{N}$ presheaf, $\Delta(\mathbb{N}) : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$. ☐

We now see that **each Grothendieck topos has an NNO**:

Corollary 13.7.1.5. *Let $(\mathbf{C}, J)$ be a site. The sheaf topos $Sh(\mathbf{C}, J)$ has a natural numbers object, given by:*

$$\mathbb{N}_{Sh(\mathbf{C},J)} = \coprod_{n\in\mathbb{N}} \mathbf{1}_{Sh(\mathbf{C},J)}$$

Proof. The site $(\mathbf{C}, J)$ determines a Lawvere-Tierney topology on $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, j [REF]. We also know that there is a geometric morphism (sheafification):

$$\begin{tikzcd} \mathrm{Sh}_j\mathbf{E} \arrow[rr, bend left, "i"] \arrow[rr, phantom, "\top"] & & \mathbf{E} \arrow[ll, bend left, "a"] \end{tikzcd} .$$

Now for any $\mathbf{E}$ being the presheaf category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, we have by Lemma 13.7.1.3 and Corollary 13.7.1.4 that $\mathrm{Sh}(\mathbf{C}, J)$ has an NNO, given by:

$$a(\Delta(\mathbb{N})) : \mathbf{C}^{\mathrm{op}} \longrightarrow \mathbf{Sets}.$$

Now since $\mathbb{N} \cong \coprod_{n\in\mathbb{N}} \mathbf{1}$ and a & Δ are left adjoints of geometric morphisms so they preserve the terminals and the small coproducts to give the desiderata. ☐

13.7.2 The ¬¬ Lawvere-Tierney Topology in a Topos

There is an LT topology in a topos which gives us as it's sheaf topos, a Boolean topos. To get to that result, we would need to understand how the operations in the two Heyting lattices $\mathrm{Sub}_{\mathrm{Sh}_j\mathbf{E}}(F)$ and $\mathrm{Sub}_{\mathbf{E}}(F)$ interacts.

$\mathbf{Sub}_{\mathbf{Sh}_j\mathbf{E}}(F)$ and $\mathbf{Sub}_{\mathbf{E}}(F)$

The structures of the Heyting algebra structures $\mathrm{Sub}_{\mathrm{Sh}_j\mathbf{E}}(F)$ and $\mathrm{Sub}_{\mathbf{E}}(F)$ are comparable (where F is j-sheaf):

Proposition 13.7.2.1. *Let $\mathbf{E}$ be a topos and let $j : \Omega \longrightarrow \Omega$ be a Lawvere-Tierney topology on it. Then, for any closed subobjects S, T of a j-sheaf F in $\mathbf{E}$, the following identities hold in $Sub_{Sh_j\mathbf{E}}(F)$:*

1. $1_j = 1$
2. $S \wedge_j T = S \wedge T$
3. $0_j = \bar{0}$
4. $S \vee_j T = \overline{S \vee T}$
5. $S \Rightarrow_j T = S \Rightarrow T$
6. $\neg_j S = \overline{\neg S}$

where $(-)_j$ denotes corresponding operation in the Heyting algebra $Sub_{Sh_j\mathbf{E}}(F)$.

Proof. 1. Because $1 : F \to F$ is closed, hence 1 is also the top element of $\mathrm{Sub}_{\mathrm{Sh}_j\mathbf{E}}(F)$.
2. The meet of two closed subobjects S, T in $\mathrm{Sub}_{\mathrm{Sh}_j\mathbf{E}}(F)$ would be closed by LTT.3. Hence $S \wedge T = \bar{S} \wedge \bar{T} = \overline{S \wedge T} = S \wedge_j T$.
3. 0 is the bottom element of $\mathrm{Sub}_{\mathbf{E}}(F)$, the closure of 0 would hence be the smallest closed subobject of F.
4. The join of two closed subobjects, $S \vee_j T$, would be the smallest closed subobject containing S and T, which is $\overline{S \vee T}$.
5. If we could show that $S \Rightarrow T$ is closed, then we could argue that: Since $S \Rightarrow_j T$ is the unique subobject in $\mathrm{Sub}_{\mathrm{Sh}_j\mathbf{E}}(F)$ with the property for any closed subobject R that $R \leq S \Rightarrow_j T$ if and only if $R \wedge_j S \leq T$, therefore we need to show that the same is true for $S \Rightarrow T$:

$$\begin{aligned} R \wedge_j S \leq T &\iff R \wedge S \leq T \\ &\iff R \leq S \Rightarrow T. \end{aligned}$$

Therefore we would be done if we could show that $S \Rightarrow T$ is also closed. To this extent, for any subobject W of F in $\mathrm{Sub}_{\mathbf{E}}(F)$, we have:

$$\begin{aligned} W \leq S \Rightarrow T &\iff W \wedge S \Rightarrow T \iff W \wedge \bar{S} \leq \bar{T} \iff \overline{W \wedge \bar{S}} \leq \bar{\bar{T}} \iff \bar{W} \wedge \bar{\bar{S}} \leq \bar{T} \iff \bar{W} \wedge \bar{S} \leq \bar{T} \\ &\iff \bar{W} \wedge S \leq T \iff \bar{W} \leq S \Rightarrow T \iff \bar{\bar{W}} \leq \overline{S \Rightarrow T} \iff \bar{W} \leq \overline{S \Rightarrow T} \iff W \leq \overline{S \Rightarrow T} \end{aligned}$$

where last line follows from the fact that $W \leq \bar{W}$.
6. Negation in a Heyting algebra is simply $\neg A := (A \Rightarrow 0)$. Hence from the above results (particularly 3 & 5), $\neg_j S := (S \Rightarrow_j 0_j) = (S \Rightarrow 0) =: \neg S$. $\square$

To a Boolean Topos from any Topos

A **Boolean topos** is a topos whose internal/external subobject lattice is an internal/external Boolean lattice. That is, PX is an internal Boolean lattice or, equivalently, $\mathrm{Sub}_{\mathbf{E}}(X)$ is an external Boolean lattice for each object X of $\mathbf{E}$. We will now see that, for the negation $\neg : \Omega \longrightarrow \Omega$ of the internal Heyting lattice Ω, the arrow $\neg\neg := \neg \circ \neg : \Omega \longrightarrow \Omega$ is a Lawvere-Tierney topology in $\mathbf{E}$ and it's sheaf topos is in-fact a Boolean topos:

Theorem 13.7.2.2. *Suppose* $\mathbf{E}$ *is a topos. Consider the internal Heyting algebra* Ω*, whose negation operator is* $\neg : \Omega \longrightarrow \Omega$*. Then, the arrow*

$$\neg\neg : \Omega \longrightarrow \Omega$$

is a Lawvere-Tierney topology in $\mathbf{E}$*. Moreover, the following then holds:*

$$\mathit{Sh}_{\neg\neg}\mathbf{E} \text{ is a } \textbf{\textit{Boolean Topos}}.$$

Proof. It's a basic result that for $x, y \in H$ for any Heyting algebra H that the following holds:

$$x \leq \neg\neg x \ , \quad \neg\neg x = \neg\neg\neg\neg x \ , \quad \neg\neg(x \wedge y) = \neg\neg x \wedge \neg\neg y.$$

So the corresponding requirements for the closure operator in $\mathrm{Sub}_{\mathbf{E}}(X)$ is already satisfied. But we wish to show that this is also natural. Since for any $f : X \to Y$ in $\mathbf{E}$, we have by Proposition

13.3.5.8 that $f^* : \mathrm{Sub}_{\mathbf{E}}(Y) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(X)$ is a Heyting algebra homomorphism, therefore the negation is preserved by f^*. Hence the $\neg\neg : \mathrm{Sub}_{\mathbf{E}}(X) \longrightarrow \mathrm{Sub}_{\mathbf{E}}(X)$ is a natural operator. Hence, $\neg\neg$ is an LT topology.

The latter result can be seen to hold because for any subsheaf S of sheaf X, S has to be $\neg\neg$-closed in $\mathbf{E}$, i.e. $\neg\neg S = S$ in $\mathbf{E}$. But we wish to show the result in $\mathrm{Sh}_{\neg\neg}\mathbf{E}$. Now note that we wish to show that $\neg_{\neg\neg}\neg_{\neg\neg}S = S$ (i.e. in $\mathrm{Sh}_j\mathbf{E}$). But by Proposition 13.7.2.1, $\neg_{\neg\neg}S = \neg\left(\bar{S}^{\neg\neg}\right) = \neg\neg\neg S$, so $\neg_{\neg\neg}\neg_{\neg\neg}S = (\neg)^6 S$. Hence, the statement $\neg\neg S = S$ in $\mathbf{E}$ is equivalent to $(\neg)^6 S = S$ in $\mathrm{Sh}_{\neg\neg}\mathbf{E}$. The result then follows. □

The next result would help us in establishing the independence of axiom of choice with continuum hypothesis:

Proposition 13.7.2.3. *Suppose* $\mathbf{E}$ *is a topos which is generated by the subobjects of* $\mathbf{1}$ *and each subobject lattice* $\mathrm{Sub}_{\mathbf{E}}(A)$ *is a complete Boolean algebra. Then,* $\mathbf{E}$ *satisfies axiom of choice.*

Proof. Take an epimorphism $p : X \longrightarrow I$. Since $\mathrm{Sub}_{\mathbf{E}}(I)$ is a complete Boolean algebra, therefore the collection of all subobjects $n : N \rightarrowtail I$ of I which has a section $s : N \longrightarrow X$ with $p \circ s = n$, has a maximal subobject $m : M \rightarrowtail I$. We wish to show that there is a section $I \longrightarrow X$ of p. To do this, we have to argue that the maximal element M of all sections of p must be the I itself. Hence let us assume that this maximal subobject of sections M is not I. Now, since $\mathrm{Sub}_{\mathbf{E}}(I)$ is a Boolean algebra so it has $\neg M$. Now since $\mathrm{Sub}_{\mathbf{E}}(\mathbf{1})$ generates $\mathbf{E}$, therefore there is a monic from a subobject $V \rightarrowtail \mathbf{1}$ with $V \to \neg M$. Now construct the pullback

$$\begin{tikzcd} X' \arrow[rr, tail] \arrow[d, two heads, "p'"'] \arrow[drr, phantom, "\lrcorner", very near start] & & X \arrow[d, two heads, "p"] \\ V \arrow[r, tail, "t"] & \neg M \arrow[r, tail] & I \end{tikzcd} .$$

Again, there is a subobject $W \rightarrowtail \mathbf{1}$ which has a monic $r : W \rightarrowtail X'$. Clearly, $p' \circ r : W \rightarrowtail V$ is a monic. Therefore we also have $t \circ p' \circ r : W \rightarrowtail \neg M$. Hence $W \wedge \neg M = \bot$ and so $W \vee M = W \amalg M$. Since we already have $W \rightarrowtail X$ and $M \to X$ which are sections of p. But M was the greatest subobject of I which has a section of p, therefore our assumption $M \neq I$ is wrong, and hence $M = I$ and hence we have a section $I \rightarrowtail X$. □

The $\neg\neg$ sheaves over a poset forms a topos which follows axiom of choice:

Proposition 13.7.2.4. *Let* $\mathbf{P}$ *be a poset regarded as a category. Then the topos* $\mathrm{Sh}(\mathbf{P}, \neg\neg)$ *satisfies axiom of choice.*

Proof. We have to just show that $\mathrm{Sh}(\mathbf{P}, \neg\neg)$ is generated by $\mathrm{Sub}_{\mathrm{Sh}(\mathbf{P},\neg\neg)}(\mathbf{1})$ because $\mathrm{Sh}(\mathbf{P}, \neg\neg)$ has complete subobject lattices as it is a Grothendieck topos. Now because each presheaf is a colimit of representables (Proposition 13.1.2.3) and so by the sheafification geometric morphism $i \vdash a$, any sheaf F is such that $i(F) \cong \varinjlim D$ where $D : I \to \mathbf{Sets}^{\mathbf{P}^{op}}$ is a diagram of representables. Now, $F \cong a \circ i(F) = a(\varinjlim D) \cong \varinjlim a(D)$ where last isomorphism comes from the fact that a is the left adjoint. Hence, $\mathrm{Sh}(\mathbf{P}, \neg\neg)$ is generated by the sheafification of representables. But since $\mathbf{Yon}(p) \rightarrowtail \mathbf{1}_{\mathbf{Sets}^{\mathbf{P}^{op}}}$, and because a is left-exact (geometric morphism), therefore $a(\mathbf{Yon}(p)) \rightarrowtail \mathbf{1}_{\mathrm{Sh}(\mathbf{P},\neg\neg)}$. □

Dense Topology & $\neg\neg$

Let's first revisit the dense topology:

Definition 13.7.2.5. (**Dense Topology on a Poset**) *Let* $\mathbf{P}$ *be a poset regarded as a category. We first define a subset* $D_p \subseteq \{q \in \mathrm{Ob}(\mathbf{P}) \mid q \leq p\}$ *to be* ***dense below*** p *if for any* $r \leq p$, $\exists\, q \in D_p$ *such that* $q \leq r$*. We then define a Grothendieck topology* J *of* $\mathbf{P}$ *given as (for any* $p \in \mathrm{Ob}(\mathbf{P})$*):*

$$J(p) := \{D_p \mid D_p \text{ is dense below } p\}\,.$$

Extending the above definition to any category is obvious:

Definition 13.7.2.6. (**Dense Topology on a Category**) *Let* $\mathbf{C}$ *be a small category. The dense topology on* $\mathbf{C}$ *is defined as follows: for any object* C *of* $\mathbf{C}$*,*

$$JC := \{S \mid \text{for any } f : D \to C\ ,\ \ \exists\, g : E \to D \text{ such that } f \circ g \in S\}\,.$$

We will now see that dense topology on the $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is exactly the $\neg\neg$ topology on it:

Proposition 13.7.2.7. *Suppose* $\mathbf{C}$ *is a small category. Then, the* $\neg\neg$ *Lawvere-Tierney topology on* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is equivalent to the dense topology on* $\mathbf{C}$*.*

Proof. First note that in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, for a subobject $A \rightarrowtail E$, it's negation $(\neg A) \rightarrowtail E$ is given as follows:

$$(\neg A)(C) := \{x \in EC \mid \forall\, f : \mathrm{dom}\,(f) \to C\ ,\ \ Ef(x) \notin A(\mathrm{dom}\,(f))\}.$$

Therefore $\neg\neg A$ would be:

$$(\neg\neg A)(C) := \{x \in EC \mid \forall f : \mathrm{dom}\,(f) \to C\ ,\ \ \exists\, g : \mathrm{dom}\,(g) \to \mathrm{dom}\,(f) \text{ such that } E(f \circ g)(x) \in A(\mathrm{dom}\,(g))\}\,.$$

Secondly, for any given site $(\mathbf{C}, J)$, the closure operation of LT topology on $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ induced by Grothendieck topology J is given by: for $A \rightarrowtail E$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$,

$$x \in \bar{A}(C) \iff \left\{f : \mathrm{dom}\,(f) \to C \mid \bar{A}(f)(x) \in E(\mathrm{dom}\,(f))\right\} \in JC.$$

In particular, if we let $(\mathbf{C}, J)$ to be a dense topology, then the above condition would be:

$$x \in \bar{A}(C) \iff \forall\, f : \mathrm{dom}\,(f) \to C\ ,\ \exists\, g : \mathrm{dom}\,(g) \to \mathrm{dom}\,(f) \text{ such that } E(f \circ g)(x) \in A(\mathrm{dom}\,(g))$$

and this is same as that of $\neg\neg A$. □

13.7.3 Axiom of Choice in a Topos

Axiom of choice says that for a collection of non-empty sets $\{X_i\}_{i \in I}$, the set $\prod_{i \in I} X_i$ is also non-empty. This condition can also be stated equivalently as: A surjective function $p : X \longrightarrow I$ has a section $s : I \longrightarrow X$ so that $p \circ s = 1_I$. We hence define the following:

Definition 13.7.3.1. (**Axiom of Choice**) *Suppose* $\mathbf{E}$ *is a topos. Then* $\mathbf{E}$ *is said to follow axiom of choice if for each epimorphism*

$$X \xrightarrow{\ \ p\ \ } \!\!\!\!\rightarrow Y$$

has a section

$$Y \xrightarrow{\ \ s\ \ } X\ .$$

That is, for every epis p*, there is an arrow* s *as above such that* $p \circ s = 1_Y$*.*

There is a weaker property for AC, called the internal axiom of choice.

Definition 13.7.3.2. (**Internal Axiom of Choice**) *Suppose* $\mathbf{E}$ *is a topos. Consider the functor for any object* E *of* $\mathbf{E}$:

$$\begin{aligned}(-)^E : \mathbf{E} &\longrightarrow \mathbf{E}\\ X &\longmapsto X^E\\ (f : X \to Y) &\longmapsto (f^E : X^E \to Y^E).\end{aligned}$$

The topos $\mathbf{E}$ *is said to follow the internal axiom of choice if the above functor* $(-)^E$ *for any object* E*, preserves epimorphisms.*

Remark 13.7.3.3. (**AC** $\implies$ **IAC**) If $p : X \to Y$ is any epimorphism in $\mathbf{E}$ which has a section $s : Y \to X$, then $p^E : X^E \to Y^E$ and $s^E : Y^E \to X^E$ are such that $p^E \circ s^E = (p \circ s)^E = (1_Y)^E = 1_{Y^E}$, because $(-)^E$ is a functor, and therefore s^E is a section of p^E. Hence axiom of choice implies internal axiom of choice.

An interesting property of **Sets** is that the terminal object $\mathbf{1}$ generates the whole category. Here, a collection of objects $\mathcal{G}$ of a category $\mathbf{C}$ is said to **generate** $\mathbf{C}$ if and only if for any two non-equal parallel pair of arrows $f \neq g : A \rightrightarrows B$ there exists an object $G \in \mathcal{G}$ and an arrow $u : G \to A$ such that $f \circ u \neq g \circ u$. A topos $\mathbf{E}$ is said to be **well-pointed** if the terminal $\mathbf{1}$ generates the $\mathbf{E}$. A topos $\mathbf{E}$ is additionally said to be **non-degenerate** if $\mathbf{0} \ncong \mathbf{1}$.
Clearly **Sets** is a non-degenerate, well-pointed topos. Moreover:

Lemma 13.7.3.4. *Suppose* $\mathbf{E}$ *is a non-degenerate topos. Then* $\mathbf{E}$ *is also a well-pointed topos if and only if the functor*

$$\mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, -) : \mathbf{E} \longrightarrow \mathbf{Sets}$$

is faithful.

Proof. (L $\implies$ R) Suppose $\mathbf{E}$ is well-pointed and non-degenerate. If we have $\mathrm{Hom}_{\mathbf{E}}(1, A) \cong \mathrm{Hom}_{\mathbf{E}}(1, B)$, then since $\mathbf{1}$ generates $\mathbf{E}$, therefore $\mathrm{Hom}_{\mathbf{E}}(X, A) \cong \mathrm{Hom}_{\mathbf{E}}(X, B)$ because if $f \neq g : X \rightrightarrows A$, then $\exists\ u : \mathbf{1} \to X$ (because $\mathbf{0} \ncong \mathbf{1}$) such that $f \circ u \neq g \circ u : \mathbf{1} \longrightarrow A \in \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, A)$. Since $\mathrm{Hom}_{\mathbf{E}}(X, A) \cong \mathrm{Hom}_{\mathbf{E}}(X, B)$ for any object X of $\mathbf{E}$, hence by generalized elements, $A \cong B$, proving that the functor $\mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, -)$ is injective over hom-sets.
(R $\implies$ L) If $\mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, -)$ is faithful, then if we take any two non-equal parallel arrows $f \neq g : A \rightrightarrows B$, because $\mathbf{0} \ncong \mathbf{1}$, then we can conclude that $\mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, f) \neq \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, g) : \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, A) \rightrightarrows \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, B)$, which means that for any $u : \mathbf{1} \to A$, $\mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, f)(u) \neq \mathrm{Hom}_{\mathbf{E}}(\mathbf{1}, g)(u) \implies f \circ u \neq g \circ u$. Hence $\mathbf{1}$ generates $\mathbf{E}$, so $\mathbf{E}$ is well-pointed. □

13.7.4 Independence of Continuum Hypothesis : The Cohen Topos

We now prove that there is a Boolean topos (a model of set theory) in which the continuum hypothesis doesn't hold. We show the entire construction in the theorem below:

Theorem 13.7.4.1. (***Independence of Continuum Hypothesis***) *There is a Boolean topos satisfying the axiom of choice in which continuum hypothesis doesn't hold.*

Proof. **Act 1:** *Requirement in* **Sets** *to follow CH*
We first understand what we need in **Sets** in order for it to follow CH. The continuum hypothesis says that there is no set whose cardinality is between $\mathbb{N}$ and $P\mathbb{N} = \mathbb{R}$. If this doesn't hold, then we must have a set X such that $\mathbb{N} \rightarrowtail X \rightarrowtail P\mathbb{N}$ where the subobjects are strict, that is, there are no epimorphisms $\mathbb{N} \to X$ and $X \to P\mathbb{N}$. We will construct a Boolean topos (a model of set theory) where we would indeed have an object X with the above mentioned monics. In order to construct this Boolean topos, let us begin with the usual **Sets** where we take a set B with cardinality strictly greater than that of $P\mathbb{N}$. We would use[27] B to *force* some other unique set to be in between the nno and it's power object in a so constructed Boolean topos. We would then conclude that this so constructed topos will not follow CH.

Act 2: *Construction of the Cohen Poset* **P**
To make this Boolean topos, let's first analyze our requirement of $g : B \rightarrowtail P\mathbb{N}$. We can equivalently state it by the power adjunction:

$$\hat{g} : B \times \mathbb{N} \longrightarrow \Omega \cong \mathbf{2}$$
$$(b, n) \longmapsto \begin{cases} 0 & \text{if } n \in g(b) \\ 1 & \text{if } n \notin g(b) \end{cases}$$

where Ω is the subobject classifier of **Sets** . If g ought to be a monic, then we must have that

$$b \neq b' \implies g(b) \neq g(b') \iff \exists\, n \text{ such that } \hat{g}(b, n) \neq \hat{g}(b', n).$$

or the contrapositive:

$$\hat{g}(b, n) = \hat{g}(b', n) \forall\, n \implies b = b'$$

These conditions are particularly important as we would try to reach this condition for two $b \neq b'$. Now, define a **condition** as a tuple (F_p, p) where $F_p \subseteq B \times \mathbb{N}$ is finite and $p : F_p \longrightarrow \mathbf{2}$. What we wish to do is to get closer and closer to the whole $B \times \mathbb{N}$ gradually, this means that as F_p gets bigger and bigger, we wish to get the corresponding p closer and closer to $\hat{g}$. In order to argue this more concretely, we construct the following poset, called the Cohen poset:

$$P := \{(F_p, p) \mid F_p \subseteq B \times \mathbb{N} \text{ is finite } \& \; p : F_p \to \mathbf{2}\}$$

and the order $\leq$ in P is given by:

$$p \leq q \iff F_q \subseteq F_p \;\&\; p|_{F_q} = q.$$

For the following we regard the Cohen poset P as a category **P**.

Act 3: *Transferring* $B \times \mathbb{N}$ *from* **Sets** *to* $\mathbf{Sets}^{\mathbf{P}^{\mathrm{op}}}$

[27]Note that the idea here is to use B to force another set to be in between them in some other model of set theory

We now use the adjunction[28] given in Definition 13.2.6.2, as follows:

$$\begin{tikzcd} \mathbf{Sets}^{\mathbf{P}^{\mathrm{op}}} \arrow[rr, bend left, "\Gamma"] \arrow[rr, phantom, "\top"] & & \mathbf{Sets} \arrow[ll, bend left, "\Delta"] \end{tikzcd} .$$

The left adjoint Δ takes $B \times \mathbb{N}$ to $\Delta(B \times \mathbb{N})$ but since Δ is left-exact (footnote 35), hence $\Delta(B \times \mathbb{N}) \cong \Delta B \times \Delta\mathbb{N}$. Now, as we pointed to earlier, we wish to get a p as close to $\hat{g}$ as possible, so we consider the following subobject of $\Delta(B \times \mathbb{N})$:

$$A \rightarrowtail \Delta(B \times \mathbb{N})$$

which takes a condition p of $\mathbf{P}$ to:

$$\begin{aligned} A : \mathbf{P}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ p &\longmapsto \{(b, n) \in B \times \mathbb{N} \mid p(b, n) = 0\} . \end{aligned}$$

We now observe the following:

Act 4: *A is a closed subobject of $\Delta(B \times \mathbb{N})$ with respect to $\neg\neg$ topology on* $\mathbf{Sets}^{\mathbf{P}^{\mathrm{op}}}$
The proof of the above statement is as follows: We just wish to show $\neg\neg A \subseteq A$ in the lattice $\mathrm{Sub}_{\mathbf{P}}(\Delta(B \times \mathbb{N}))$ as the other side is trivial. As the proof of Proposition 13.7.2.7 showed, we have $(b, n) \in \neg\neg A(p)$ iff $\forall q \leq p$, $\exists r \leq q$ such that $\Delta(B \times \mathbb{N})(r \leq p)((b, n)) \in A(r)$ or $(b, n) \in A(r)$ or $r(b, n) = 0$. Assume $(b, n) \notin A(p)$. Hence $p(b, n) \neq 0$. But then we are not sure whether $p(b, n) = 1$ or $p(b, n)$ is undefined. For the former, if $p(b, n) = 1$, then clearly for any $r \leq p$, $r(b, n) = 1$ because $r|_{F_p} = p$ and so $(b, n) \notin \neg\neg A(p)$. For the latter, if $p(b, n)$ is undefined, then for some $q \leq p$ we would have $q(b, n) = 1$, and then for any $r \leq q$, $r(b, n) = 1$ to conclude that $(b, n) \notin A(p)$. Both cases suggest $(b, n) \notin \neg\neg A(p)$. Hence proved the fact that if $(b, n) \notin A(p) \implies (b, n) \notin \neg\neg A(p)$ and it's contrapositive gives the required result.

Act 5: *Yielding an arrow* $\Delta B \longrightarrow \Omega_{\neg\neg}^{\Delta\mathbb{N}}$
Since A is a closed subobject of $\Delta(B \times \mathbb{N})$, therefore the characteristic arrow $\Delta(B \times \mathbb{N}) \longrightarrow \Omega$ factors via the LT Topology $\neg\neg : \Omega \longrightarrow \Omega$ and $\neg\neg : \Omega \longrightarrow \Omega$ itself factors via the $\Omega_{\neg\neg}$ which is the subobject classifier of $\mathrm{Sh}(\mathbf{P}, \neg\neg)$, as shown below:

$$\begin{tikzcd} \Delta(B \times \mathbb{N}) \arrow[r, "\mathrm{char}\ A"] \arrow[d, "\mathrm{char}\ A"'] & \Omega \\ \Omega \arrow[ur, "\neg\neg"] \arrow[r, two heads, "r"'] & \Omega_{\neg\neg} \arrow[u, tail, "m"'] \end{tikzcd} .$$

Therefore, we have an arrow

$$f := r \circ \mathrm{char}\ A : \Delta B \times \Delta\mathbb{N} \longrightarrow \Omega_{\neg\neg}$$

[28] As alluded to earlier, this is actually a geometric morphism.

and hence it's P-transpose would be:

$$\hat{f} : \Delta B \longrightarrow \Omega_{\neg\neg}^{\Delta\mathbb{N}}.$$

Act 6: $\hat{f} : \Delta B \longrightarrow \Omega_{\neg\neg}^{\Delta\mathbb{N}}$ *is a monomorphism in* $\mathbf{Sets}^{\mathbf{P}^{\mathrm{op}}}$
To show the above, we just have to show that each component $\hat{f}_p : \Delta B(p) = B \longrightarrow \Omega_{\neg\neg}^{\Delta\mathbb{N}}(p)$ is a monomorphism. To this extent, take any two non-equal elements $b \neq b'$ from B. We wish to show that $\hat{f}_p(b) \neq \hat{f}_p(b')$. First, we note that $\Omega_{\neg\neg}^{\Delta\mathbb{N}}$, by Proposition 13.1.3.1, is given as follows:

$$\begin{aligned} \Omega_{\neg\neg}^{\Delta\mathbb{N}} : \mathbf{P}^{\mathrm{op}} &\longrightarrow \mathbf{Sets} \\ p &\longmapsto \mathrm{Nat}\left(\mathrm{Hom}_{\mathbf{P}}\left(-,p\right) \times \Delta\mathbb{N}, \Omega_{\neg\neg}\right). \end{aligned}$$

Therefore, $\hat{f}_p(b)$ is a natural transformation as:

$$\hat{f}_p(b) : \mathrm{Hom}_{\mathbf{P}}\left(-,p\right) \times \Delta\mathbb{N} \Rightarrow \Omega_{\neg\neg}.$$

To fulfill our aim, we must show that $\hat{f}_p(b) \neq \hat{f}_p(b')$. Again, as both are natural transforms, so we would be done if we would show that for each $q \leq p$, $(\hat{f}_p(b))_q \neq (\hat{f}_p(b'))_q$. Again,

$$\begin{aligned} (\hat{f}_p(b))_q : (\mathrm{Hom}_{\mathbf{P}}\left(-,p\right) \times \Delta\mathbb{N})(q) \cong \mathrm{Hom}_{\mathbf{P}}\left(q,p\right) \times \Delta\mathbb{N}(q) \cong \{\star\} \times \mathbb{N} \cong \mathbb{N} &\longrightarrow \Omega_{\neg\neg}(q) \\ n &\longmapsto \{r \in \mathbf{P} \mid r \leq q \;\&\; r(b,n) = 0\}. \end{aligned}$$

Now consider $(\hat{f}_p(b))_q(n)$ as above. If $r \in (\hat{f}_p(b))_q(n)$, then $r(b,n) = 0$ and so all $t \leq r$ is in $(\hat{f}_p(b))_q(n)$. Since all conditions F_r are finite, hence for large enough n, neither (b,n) nor (b',n) would be defined for $r : F_r \to \mathbf{2}$. One can now easily construct some $t \leq r$ with $t(b,n) = 0$ and $t(b,n) = 1$ and so $t \in (\hat{f}_p(b))_q(n)$ but $t \notin (\hat{f}_p(b))_q(n)$. Hence we are done.

Act 7: *Transferring monic* $\hat{f} : \Delta B \longrightarrow \Omega_{\neg\neg}^{\Delta\mathbb{N}}$ *from* $\mathbf{Sets}^{\mathbf{P}^{\mathrm{op}}}$ *to a monic in* $\mathrm{Sh}\left(\mathbf{P}, \neg\neg\right)$ *via sheafification*
Since sheafification is inverse image of a geometric morphism, therefore it would trivially preserve the monic $\hat{f}$. In particular, we would have the following monic:

$$a(\hat{f}) : a(\Delta B) \rightarrowtail a(\Omega_{\neg\neg}^{\Delta\mathbb{N}})$$

and because $a(X^Y) \cong a(X)^{a(Y)}$, therefore $a(\Omega_{\neg\neg}^{\Delta\mathbb{N}}) \cong (a(\Omega_{\neg\neg}))^{a(\Delta\mathbb{N})} \cong \Omega_{\neg\neg}^{a(\Delta\mathbb{N})}$, we can rewrite above as (denoting $\widehat{(-)} := a(\Delta(-))$):

$$\widehat{\hat{f}} : \widehat{B} \rightarrowtail \Omega_{\neg\neg}^{\widehat{\mathbb{N}}}.$$

Act 7: *Conclusion*
We have finally proved that $\widehat{B}$ is a subobject of $\Omega_{\neg\neg}^{\widehat{\mathbb{N}}} \cong P(\widehat{\mathbb{N}})$ where $P(\widehat{\mathbb{N}})$ is the power object of the nno $\widehat{\mathbb{N}}$ in $\mathrm{Sh}\left(\mathbf{P}, \neg\neg\right)$. One can also show that (but we won't here for space concerns! Refer to Section 6.3, pp 284 of **[MacMoer]** instead) the cardinal inequality of our choice of B in **Sets** as $\mathbb{N} < P\mathbb{N} < B$ is preserved in $\mathrm{Sh}\left(\mathbf{P}, \neg\neg\right)$ as

$$\widehat{\mathbb{N}} < \widehat{P\mathbb{N}} < \widehat{B}$$

and this, combined with $\widehat{B} < P(\widehat{\mathbb{N}})$, gives us that in the Boolean Grothendieck topos (which is a model of set theory!) $\mathrm{Sh}\,(\mathbf{P}, \neg\neg)$, we would have the following cardinal inequality:

$$\boxed{\widehat{\mathbb{N}} < \widehat{P\mathbb{N}} < P(\widehat{\mathbb{N}}) \quad \text{in } \mathrm{Sh}\,(\mathbf{P}, \neg\neg).}$$

Therefore, $\mathrm{Sh}\,(\mathbf{P}, \neg\neg)$ is a Boolean Grothendieck topos in which continuum hypothesis fails as above but axiom of choice holds by Proposition 13.7.2.4. $\square$

Remark 13.7.4.2. (**The Cohen Topos**) The Boolean Grothendieck Topos of $\neg\neg$-sheaves over the Cohen poset $\mathbf{P}$, $\mathrm{Sh}\,(\mathbf{P}, \neg\neg)$, is usually called the Cohen topos.

13.7.5 Integers in a Topos

The concept of Dedekind cuts is usually used to generate irrationals from rationals. One can essentially extend the idea on to a sheaf topos $\mathrm{Sh}\,(X)$ where X is some topological space.
Let's first construct the integers from naturals:

From $\mathbb{N}_{\mathbf{E}}$ to $\mathbb{Z}_{\mathbf{E}}$

In the usual category **Sets**, we have the $\mathbb{N}$. One can construct all integers by collecting all pairs of naturals whose difference between them is same, that is, $\mathbb{Z}$ can be constructed as the following quotient set:

$$\mathbb{Z} := \{(n, m) \mid n, m \in \mathbb{N}\}/\sim$$

where $\sim$ is the following equivalence relation,

$$(n, m) \sim (n', m') \iff n + m' = m + n'.$$

The relation $\sim$ essentially collects all the pairs (n, m) whose difference $(n - m)$ are same. More categorically, we can represent above as an universal construction in **Sets** as follows:

1. Construct the kernel pair of $+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ in **Sets**:

$$\begin{tikzcd} E \arrow[r, "b"] \arrow[d, "a"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbb{N} \times \mathbb{N} \arrow[d, "+"] \\ \mathbb{N} \times \mathbb{N} \arrow[r, "+"'] & \mathbb{N} \end{tikzcd} .$$

2. Then construct the following coequalizer where $\pi_1, \pi_2 : \mathbb{N} \times \mathbb{N} \rightrightarrows \mathbb{N}$ are the product projections:

$$\begin{tikzcd} E \arrow[r, shift left, "\langle \pi_1 \circ a, \pi_2 \circ b \rangle"] \arrow[r, shift right, "\langle \pi_2 \circ a, \pi_1 \circ b \rangle"'] & \mathbb{N} \times \mathbb{N} \arrow[r, two heads] & \mathbb{Z} \end{tikzcd} .$$

The fact that this universal description is equivalent to the previous set-theoretic one can be seen via the observation that E is the collection of 4-tuples (n, m', n', m) with $n + m' = m + n'$ with a and b being the respective projections, and the arrows $\pi_1 \circ a$ takes (n, m', n', m) to n, $\pi_2 \circ b$ takes it to m, $\pi_2 \circ a$ takes it to m' and $\pi_1 \circ b$ takes it to n'.
With this, we are motivated to state the following:

Definition 13.7.5.1. **(Integer Object in a Topos)** *Let* $\mathbf{E}$ *be a topos and* $\mathbb{N}_{\mathbf{E}}$ *being the natural numbers object in* $\mathbf{E}$*. Then, we define the integer object* $\mathbb{Z}_{\mathbf{E}}$ *in* $\mathbf{E}$ *as the following coequalizer:*

1. *Let* E *be the following pullback:*

$$\begin{tikzcd} E \arrow[r, "b"] \arrow[d, "a"'] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbb{N}_{\mathbf{E}} \times \mathbb{N}_{\mathbf{E}} \arrow[d, "+"] \\ \mathbb{N}_{\mathbf{E}} \times \mathbb{N}_{\mathbf{E}} \arrow[r, "+"'] & \mathbb{N}_{\mathbf{E}} \end{tikzcd}$$

2. *And then define the* $\mathbb{Z}_{\mathbf{E}}$ *is defined as:*

$$\begin{tikzcd} E \arrow[r, shift left, "\langle \pi_1 \circ a, \pi_2 \circ b \rangle"] \arrow[r, shift right, "\langle \pi_2 \circ a, \pi_1 \circ b \rangle"'] & \mathbb{N}_{\mathbf{E}} \times \mathbb{N}_{\mathbf{E}} \arrow[r, two heads] & \mathbb{Z}_{\mathbf{E}} \end{tikzcd} .$$

Chapter 14

Language of ∞-Categories

Contents

In this chapter we give an overview of the language of ∞-categories. The ultimate goal of this chapter is to define the ∞-category of ∞-groupoids and state the Yoneda lemma in ∞-categorical setting. Therefore, if you believe that ∞-groupoids are spaces (homotopy hypothesis), then we would construct the ∞-category of spaces by the end of this chapter. In the process, we will learn about the techniques invovled in manipulating ∞-categories. In this chapter, we will only work with the category of compactly generated spaces and by **Top** we mean category of compactly generated spaces.

14.1 Simplicial sets

The goal of simplicial sets is to obtain a combinatorial approximation of topological spaces. We describe simplicial sets as a presheaf over simplex category. One can motivate themselves why this definition is the correct definition for combinatorially handling topological spaces by reading the review paper cite[Bergner's simplicial set paper]. We give some basic properties of such objects. We also give a general important result about presheaves, "every presheaf is a colimit of representable presheaves". This is of fundamental importance in the development of ∞-categories.

Remark 14.1.0.1. (*Notations*)

1. As we will be frequently constructing and dealing with presheaf category over a category $\mathbf{C}$, therefore instead of only denoting it by $\mathbf{PSh}(\mathbf{C})$, we will also be denoting it by $\widehat{\mathbf{C}}$, depending on the convenience of the given situation.
2. We will denote objects of an ordinary 1-category $\mathbf{C}$ by lowercase alphabets like $a, b, c, .., x, y, z$, morphisms in $\mathbf{C}$ by lowercase letters like $f, g, h, \dots$ and functors $\mathbf{C} \to \mathbf{D}$ by uppercase alphabets like $A, B, C, ..., X, Y, Z$ and also by lowercase alphabets like $f, g, h, \dots$.
3. Let $\mathbf{C}$ be a category and $c \in \mathbf{C}$ be an object. We denote by $h_c : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$ the contravariant hom-functor given by $a \mapsto \mathrm{Hom}_{\mathbf{C}}(a, c)$.
4. We will denote ∞-categories by $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$.

Let us also as a reminder put the usual Yoneda lemma.

Lemma 14.1.0.2. *(Yoneda lemma) Let $\mathbf{A}$ be a category, $a \in \mathbf{A}$ be an object and F be a presheaf over $\mathbf{A}$. Then, there is a natural isomorphism*

$$\mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, F) \xrightarrow{\cong} F(a).$$

Proof. Consider the following maps

$$\begin{aligned} \varphi : \mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, F) &\longrightarrow F(a) \\ \alpha &\longmapsto \alpha_a(\mathrm{id}_a). \end{aligned}$$

and

$$\begin{aligned} \xi : F(a) &\longrightarrow \mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, F) \\ x &\longmapsto \beta : h_a \to F \end{aligned}$$

where β is defined as follows. Consider any $a' \in \mathbf{A}$ and any $f \in h_a(a')$. Define $\beta_{a'} : h_a(a') \to F(a')$ by mapping as $f \mapsto F(f)(x)$.

One then sees that $\xi \circ \varphi = \mathrm{id}$ by naturality of α and that $\varphi \circ \xi = \mathrm{id}$ by funtoriality of F. □

Corollary 14.1.0.3. *Let $\mathbf{A}$ be a category and define* Yon $: A \to \widehat{\mathbf{A}}$ *to be the Yoneda functor given by $a \mapsto h_a$ and $a \to b \mapsto h_a \to h_b$. Then,* Yon *is fully-faithful.*

Proof. We have by Yoneda lemma (Lemma 14.1.0.2) that $\mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, h_b) \cong h_b(a) = \mathrm{Hom}_{\mathbf{A}}(a, b)$. □

14.1.1 Extension of functor by colimits-I

The following is a fundamental result in category theory.

Theorem 14.1.1.1. *Let $\mathbf{A}$ be a small category and $\mathbf{C}$ be a locally small category with all small colimits. Then, for any functor*

$$f : \mathbf{A} \to \mathbf{C}$$

we get two functors

$$f_! : \widehat{\mathbf{A}} \longrightarrow \mathbf{C}$$

and

$$\begin{aligned} f^* : \mathbf{C} &\longrightarrow \widehat{\mathbf{A}} \\ c &\longmapsto \mathrm{Hom}_{\mathbf{C}}(f(-), c) \end{aligned}$$

such that

1. $f_!$ *is left adjoint of* f^*

$$\begin{tikzcd} \widehat{\mathbf{A}} \arrow[r, shift left=2, "f_!"] \arrow[r, phantom, "\perp"] & \mathbf{C} \arrow[l, shift left=2, "f^*"] \end{tikzcd} .$$

   The functor $f_!$ *is called the extension of* f *by colimits.*

2. *Denoting* $\mathrm{Yon} : \mathbf{A} \hookrightarrow \widehat{\mathbf{A}}$ *to be the Yoneda embedding, we get that the following commutes upto a unique natural isomorphism*

$$\begin{tikzcd} \widehat{\mathbf{A}} \arrow[r, "f_!"] \arrow[dr, phantom, "\Searrow"] & \mathbf{C} \\ \mathbf{A} \arrow[u, hook, "\mathrm{Yon}"] \arrow[ur, "f"'] & \end{tikzcd} .$$

   That is, $f_!(h_a) \cong f(a)$ *where* $h_a = \mathrm{Hom}_{\mathbf{A}}(-, a)$.

Proof. The main heart of the proof is the fact that every presheaf is a colimit of representables (Proposition 1.1.8 of cite[Cis]).

1. We define $f_! : \widehat{\mathbf{A}} \to \mathbf{C}$ as follows. Pick any $X \in \widehat{\mathbf{A}}$. Consider the functor $\varphi_X : \mathbf{A}/X \to \widehat{\mathbf{A}}$ from the category of elements of X given by $(a, s) \mapsto h_a$ and on maps by $f : (a, s) \to (b, t)$ by $h_f : h_a \to h_b$. By Proposition 1.1.8 of cite[Cis], we have $\varinjlim_{(a,s)} h_a = X$. Define $f_!(X) = \varinjlim_{(a,s)} f(a)$ which exists in $\mathbf{C}$ as $\mathbf{C}$ has all small colimits. Consequently, we obtain the following natural isomorphisms by Yoneda lemma (Lemma 14.1.0.2) and limit preserving properties of contravariant homs

$$\begin{aligned} \mathrm{Hom}_{\mathbf{C}}(f_!(X), c) &\cong \mathrm{Hom}_{\mathbf{C}}\left(\varinjlim_{(a,s)} f(a), c\right) \cong \varprojlim_{(a,s)} \mathrm{Hom}_{\mathbf{C}}(f(a), c) \\ &\cong \varprojlim_{(a,s)} \mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, f^*(c)) \cong \mathrm{Hom}_{\widehat{\mathbf{A}}}\left(\varinjlim_{(a,s)} h_a, f^*(c)\right) \\ &\cong \mathrm{Hom}_{\widehat{\mathbf{A}}}(X, f^*(c)). \end{aligned}$$

2. Since $\mathrm{Hom}_{\mathbf{C}}(f_!(h_a), c) \cong \mathrm{Hom}_{\widehat{\mathbf{A}}}(h_a, f^*c) \cong f^*(c)(a) = \mathrm{Hom}_{\mathbf{C}}(f(a), c)$ for all objects $c \in \mathbf{C}$, therefore the result follows by Corollary 14.1.0.3. □

The above theorem will take a central place in constructions of this chapter. Indeed, let us point the following applications of this theorem before stating its proof.

Corollary 14.1.1.2. *Let* $\mathbf{A}$ *be a small category and* $\mathbf{C}$ *be a category with small colimits. If* $F : \widehat{\mathbf{A}} \to \mathbf{C}$ *is a colimit preserving functor. Then*

1. *there exists* $f : \mathbf{A} \to \mathbf{C}$ *such that* $f_! \cong F$ *naturally,*
2. F *has a right adjoint.*

Proof. We first define the required f. For any object $a \in \mathbf{A}$, define $f(a) := F(h_a)$. Then by Theorem 14.1.1.1, 1, it follows that $f_!(h_a) \cong f(a) = F(h_a)$. Since we know that every presheaf is presented as a colimit of representable functors indexed by its category of elements, thus we get that for any $X \in \widehat{\mathbf{A}}$, $F(X) \cong f_!(X)$ naturally. The second conclusion also follows from Theorem 14.1.1.1, 2. $\square$

The following result will be important in order to define simplicial mapping spaces.

Proposition 14.1.1.3. *For any small category $\mathbf{A}$, the presheaf category $\widehat{\mathbf{A}}$ is Cartesian closed where the internal hom object is defined by*

$$\underline{\mathrm{Hom}}\,(X,Y)(a) := \mathrm{Hom}_{\widehat{\mathbf{A}}}\,(X \times h_a, Y)$$

and on morphisms as

$$\underline{\mathrm{Hom}}\,(X,Y)(f) := \mathrm{Hom}_{\widehat{\mathbf{A}}}\,(X \times h_f, Y).$$

Proof. We wish to show that $\underline{\mathrm{Hom}}\,(-,-)$ acts as internal hom object in $\widehat{\mathbf{A}}$. This can be seen by establishing following natural bijections

$$\mathrm{Hom}_{\widehat{\mathbf{A}}}\,(T \times X, Y) \cong \mathrm{Hom}_{\widehat{\mathbf{A}}}\,(T, \underline{\mathrm{Hom}}\,(X,Y))$$

This follows from contemplating the functor $f_X : \mathbf{A} \to \widehat{\mathbf{A}}$ for all $X \in \widehat{\mathbf{A}}$ given by $a \mapsto X \times \mathrm{Yon}(-)$, together with Theorem 14.1.1.1. Indeed, first observe that $\widehat{\mathbf{A}}$ is locally small with all small colimits. Second, observe from the proof of Theorem 14.1.1.1 that for any $Z \in \widehat{\mathbf{A}}$, we have that $f_{X!} : \widehat{\mathbf{A}} \to \widehat{\mathbf{A}}$ takes any $Z \in \widehat{\mathbf{A}}$ and maps it to $f_{X!}(Z) = \varinjlim_{(a,s)} f_X(a)$ where (a,s) varies over $\mathbf{A}/Z$, the category of elements of Z. Since $f_X(a) = X \times h_a$ and filtered colimits commute with finite limits, therefore we get a natural isomorphism $f_{X!}(Z) \cong X \times Z$. Consequently, the adjunction of Theorem 14.1.1.1 completes the proof. $\square$

14.1.2 Categories $\boldsymbol{\Delta}$ and sSet

Consider the category of all finite sets $[n]$ with $n+1$ elements with linear order $0 < 1 < \cdots < n$ and mappings being the non-decreasing maps. Denote this category by $\boldsymbol{\Delta}$ and call it the *simplex category*. A *simplicial set* is then a presheaf over $\boldsymbol{\Delta}$. Let the category of all simplicial sets be denoted **sSet**.

There are two important class of maps in $\boldsymbol{\Delta}$.

Definition 14.1.2.1 (Face and degeneracy maps). For each $n \in \mathbb{N}$, we have $n+1$ face maps

$$\begin{aligned} d^i : [n-1] &\to [n] \\ j &\mapsto \begin{cases} j & \text{if } j \le i \\ j+1 & \text{if } j > i \end{cases} \end{aligned}$$

where $0 \le i \le n$ and n degeneracy maps

$$\begin{aligned} s^i : [n] &\to [n-1] \\ j &\mapsto \begin{cases} j & \text{if } j \le i \\ j-1 & \text{if } j > i \end{cases} \end{aligned}$$

where $0 \le i \le n-1$.

Remark 14.1.2.2. By Yoneda embedding, we will allow ourselves to abuse the notation by writing $d^i : [n-1] \to [n]$ as the unique map $d^i : \Delta^{n-1} \to \Delta^n$ and $s^i : [n] \to [n-1]$ as the unique map $s^i : \Delta^n \to \Delta^{n-1}$ in **sSet**. For a simplicial set X we may thus interpret $x \in X_n$ as $x : \Delta^n \to X$. The above give maps which we denote as

$$\begin{aligned} d_i : X_n &\longrightarrow X_{n-1} \\ x &\longmapsto x \circ d_i \end{aligned}$$

for each $0 \leq i \leq n$ also called the face maps and

$$\begin{aligned} s_i : X_{n-1} &\longrightarrow X_n \\ x &\longmapsto x \circ s_i \end{aligned}$$

for each $0 \leq i \leq n-1$ also called degeneracy maps.

It is quite easy to observe, but very important for applications, the following relations satisfied by face and degeneracy maps. All these are immediate from definition given above.

Proposition 14.1.2.3. *The following relations hold in* $\mathbf{\Delta}$ *(and thus in* **sSet** *):*

1. *If $i < j$, then $d^j d^i = d^i d^{j-1}$.*
2. *If $i \leq j$, then $s^j s^i = s^i s^{j+1}$.*
3. *We have*

$$s^j d^i = \begin{cases} d^i s^{j-1} & \text{if } i < j \\ \mathrm{id} & \text{if } i = j, j+1 \\ d^{i-1} s^j & \text{if } j+1 < i. \end{cases}$$

□

Consequently, for a simplicial set, dual relations hold.

Proposition 14.1.2.4. *Let X be a simplicial set. Then the face and degeneracy maps of X satisfies the following relations:*

1. *If $i < j$, then $d_i d_j = d_{j-1} d_i$.*
2. *If $i \leq j$, then $s_i s_j = s_{j+1} s_i$.*
3. *We have*

$$d_i s_j = \begin{cases} s_{j-1} d_i & \text{if } i < j \\ \mathrm{id} & \text{if } i = j, j+1 \\ s_j d_{i-1} & \text{if } j+1 < i. \end{cases}$$

□

Remark 14.1.2.5. One may keep the following picture in mind while working with a simplicial set X:

$$\begin{tikzcd}
{[0]} \arrow[r, shift left=3, "d^0"] \arrow[r, shift left=1, "d^1"'] & {[1]} \arrow[l, shift left=2, "s^0"] \arrow[r, shift left=4, "d^0"] \arrow[r, shift left=2, "d^1"] \arrow[r, "d^2"'] & {[2]} \arrow[l, shift left=2, "s^0"] \arrow[l, shift left=4, "s^1"] \arrow[r, shift left=2, "d^i"] & \cdots \arrow[l, shift left=2, "s^i"]
\end{tikzcd}.$$

$$\begin{tikzcd}
X_0 \arrow[r, shift right=2, "s_0"'] & X_1 \arrow[l, shift right=3, "d_0"'] \arrow[l, shift right=1, "d_1"] \arrow[r, shift right=2, "s_0"'] \arrow[r, shift right=4, "s_1"'] & X_2 \arrow[l, shift right=4, "d_0"'] \arrow[l, shift right=2, "d_1"] \arrow[l, "d_2"] \arrow[r, shift right=2, "s_i"'] & \cdots \arrow[l, shift right=2, "d_i"']
\end{tikzcd}$$

Remark 14.1.2.6. There is a functor

$$\begin{aligned} |-| : \mathbf{\Delta} &\longrightarrow \mathbf{Top} \\ [n] &\longmapsto |\Delta^n| \end{aligned}$$

where $|\Delta^n|$ is the standard topological n-simplex in $\mathbb{R}^{n+1}$ and for $f : [n] \to [m]$ in $\mathbf{\Delta}$, we have

$$\begin{aligned} |f| : |\Delta^n| &\to |\Delta^m| \\ (t_0, \dots, t_n) &\longmapsto \left(\sum_{i \in f^{-1}(0)} i, \dots, \sum_{i \in f^{-1}(m)} i \right). \end{aligned}$$

Example 14.1.2.7 (*Singular chains*)**.** An important example of a simplicial set is that of $\mathrm{Sing}(X)$ defined as

$$\begin{aligned} \mathrm{Sing}(X) : \mathbf{\Delta}^{op} &\to \mathbf{Set} \\ [n] &\mapsto \mathrm{Hom}_{\mathbf{Top}}(|\Delta^n|, X). \end{aligned}$$

The main point is that $\mathrm{Sing}(X)$ as a simplicial set knows all about the homotopy type of space X. Consequently, we will denote $X([n]) := \mathrm{Sing}(X)_n = \mathrm{Hom}_{\mathbf{Top}}(|\Delta^n|, X)$.

Example 14.1.2.8 (*Nerve of a category*)**.** Let $\mathbf{C}$ be a category. Define the nerve of $\mathbf{C}$ as

$$\begin{aligned} N\mathbf{C} : \mathbf{\Delta}^{op} &\to \mathbf{Set} \\ [n] &\mapsto \mathrm{Hom}_{\mathbf{Cat}}([n], \mathbf{C}) \end{aligned}$$

where $[n]$ is a category as it is a poset. Consequently, $N\mathbf{C}_n$ is the set of all n-composable arrows of $\mathbf{C}$.

Theorem 14.1.2.9. *Nerve construction is a fully-faithfull embedding of categories into simplicial sets.*

14.1.3 Operations on simplicial sets

*Define product, coproducts, subspaces, unions, quotients, limits, colimits and mapping objects.***TODO!!**

Example 14.1.3.1 (*Standard Δ^n, boundaries $\partial_i \Delta^n$, $\partial \Delta^n$ and horn Λ^n_i*)**.** Denote $\Delta^n = N[n]$ to be the nerve of the category $[n]$. That is,

$$\Delta^n([m]) = \Delta^n_m = \mathrm{Hom}_{\mathbf{\Delta}}([m], [n]) = h_{[n]}([m])$$

which are exactly all representable presheaves over $\mathbf{\Delta}$. These are the combinatorial analogues of the topological n-simplex $|\Delta^n|$ and we tend to think about Δ^n using the intuition gained from the topological one. There are some important simplicial subsets of Δ^n.

Let $E \subseteq [n]$ be a totally ordered subset. Define $\Delta^E = NE$ to be a simplicial subset of Δ^n. From this we derive the following simplicial subsets of Δ^n. The first is the *i^{th}-boundary of* Δ^n for $0 \le i \le n$ given by

$$\partial_i \Delta^n = \bigcup_{i \notin E \subsetneq [n]} \Delta^E \cong \Delta^{n-1}.$$

The second is the *boundary of* Δ^n given by

$$\partial\Delta^n = \bigcup_{E\subsetneq[n]} \Delta^E = \bigcup_{i=0}^{n} \partial_i\Delta^n$$

The third is the i^{th}-horn of Δ^n denoted Λ_i^n given by

$$\Lambda_i^n = \bigcup_{i\in E\subsetneq[n]} \Delta^E.$$

Remark 14.1.3.2. Let $n \geq 1$, $0 \leq i \leq n$ and $m \geq 0$. Note that we have

$$(\partial_i\Delta^n)_m = \begin{Bmatrix}\text{Order preserving maps } f : [m] \to \\ [n] \text{ which are not surjective and} \\ i \notin \mathrm{Im}\,(f).\end{Bmatrix}$$

$$(\partial\Delta^n)_m = \begin{Bmatrix}\text{Order preserving maps } f : [m] \to \\ [n] \text{ which are not surjective.}\end{Bmatrix}$$

$$(\Lambda_i^n)_m = \begin{Bmatrix}\text{Order preserving maps } f : [m] \to \\ [n] \text{ which are not surjective and} \\ i \in \mathrm{Im}\,(f).\end{Bmatrix}.$$

For two simplicial sets, we can define the internal hom using the Proposition 14.1.1.3.

Definition 14.1.3.3 (Homotopy & mapping complex). Let S, T be a simplicial set. Then $\underline{\mathrm{Hom}}\,(S, T)$ denotes the following simplicial set

$$[n] \mapsto \mathrm{Hom}_{\mathbf{sSet}}\,(S \times \Delta^n, T).$$

An n-simplex of $\underline{\mathrm{Hom}}\,(S, T)$ is defined to be an n-homotopy from S to T. A 1-homotopy H is also referred to as a homotopy from $H|_{S\times\{0\}} =: f$ to $H|_{S\times\{1\}} =: g$.

14.1.4 Basic properties

Do results and exercises from various sources to showcase the combinatorial arguments used while working with simplicial sets.

We discuss some properties of simplicial sets which would be useful later on. Some of these might be taken as exercises on combinatorial manipulations with simplicial sets.

We first begin with a simple observation.

Definition 14.1.4.1 (**n-degenerate**). A simplicial set X is said to be n-degenerate if for all $m > n$, all m-simplices in X_m are degenerate.

Example 14.1.4.2. Each standard simplicial sets Δ^n are n-degenerate. Indeed, its m-simplices for $m > n$ are

$$\Delta^n_m = \{\text{Order preserving maps } f : [m] \to [n]\}.$$

But as $m > n$, therefore every such f is necessarily non-injective. It follows that each simplex in Δ^n_m is in the image of $s_i : X_{m-1} \to X_m$ for some $0 \le i \le m-1$.

For similar reasons, the i^{th}-boundary $\partial_i \Delta^n$, boundary $\partial \Delta^n$ and horns Λ^n_i for $0 \le i \le n$ are all $n-1$-degenerate.

Lemma 14.1.4.3. *Let X be an n-degenerate simplicial set and Y be a simplicial set. Then any collection of functions $\{\varphi_m : X_m \to Y_m\}_{0 \le m \le n}$ such that for any $f : [k] \to [l]$ in $\mathbf{\Delta}$ with $0 \le k, l \le n$, the following square commutes*

$$\begin{tikzcd} X_l \arrow[r, "\varphi_l"] \arrow[d, "f^*"'] & Y_l \arrow[d, "f^*"] \\ X_k \arrow[r, "\varphi_k"'] & Y_k \end{tikzcd},$$

the collection $\{\varphi_m\}_{0 \le m \le n}$ lifts to a unique map of simplicial sets $\varphi : X \to Y$.

Remark 14.1.4.4. As a consequence of Lemma 14.1.4.3, in order to give a map of simplicial sets from an n-degenerate simplicial set S, it suffices to construct the required map only on m-simplices for $0 \le m \le n$.

Proof of Lemma 14.1.4.3. Let $\{\varphi_m\}_{0 \le m \le n}$ be as given. We wish to define $\varphi_{m'}$ for $m' > n$. We proceed by induction on m'. Suppose $\varphi_{m'-1}$ is given to us. Since we have the following diagram

$$\begin{tikzcd} X_{m'-1} \arrow[r, "\varphi_{m'-1}"] \arrow[d, "s_i"'] & Y_{m'-1} \arrow[d, "s_i"] \\ X_{m'} \arrow[r, "\varphi_{m'}"'] & Y_{m'} \end{tikzcd}$$

and that every element of $X_{m'}$ is in image of s_i (guaranteed by Proposition 14.1.2.4, 3), it follows that there is a unique choice of $\varphi_{m'}$ to fit in the above diagram, as required. □

Lemma 14.1.4.5. *Let $n \ge 1$ and $0 \le i \le n$. Then,*

1. *$\partial_i \Delta^n$ has exactly 1 non-degenerate $n-1$-simplex,*

2. *$\partial\Delta^n$ has exactly $n+1$ non-degenerate $n-1$-simplices,*
3. *Λ^n_i has exactly n non-degenerate $n-1$-simplices.*

Proof. These three items are immediate from Remark 14.1.3.2. □

The following is an important adjunction which will be consistently used in later sections. Moreover, it is generally good to keep in mind all the time while working with simplicial sets so that one transfer intution from topological spaces to that of simplicial sets, as we would need to do time and time again (for example when dealing with homotopy of simplicial sets).

Theorem 14.1.4.6 (*Geometric realization*)**.** *The singular functor* Sing : **Top** → **sSet** *has a left adjoint* $|-| : \mathbf{sSet} \to \mathbf{Top}$

$$\begin{tikzcd} \mathbf{sSet} \arrow[r, shift left, "|-|"] \arrow[r, phantom, "\perp"] & \mathbf{Top} \arrow[l, shift left, "\mathrm{Sing}"] \end{tikzcd}.$$

The functor $|-|$ is called the geometric realization and for a simplicial set X, we have

$$|X| = \coprod_{n\geq 0} X_n \times |\Delta^n| \,/ \sim$$

where $\sim$ is generated by $(f^*x, t) \sim (x, |f|t)$ for all $f : [n] \to [m]$ in $\mathbf{\Delta}$, $x \in X_m$ and $t \in |\Delta^n|$.

Proof. The main idea is that any map $X \to \mathrm{Sing}(Y)$ is a natural transform of presheaves. One observes that the naturality conditions on this morphism is equivalently represented in terms of a map $|X| \to Y$. **TODO**. □

Example 14.1.4.7. As an example, one can show that

$$\Delta^1 \times \Delta^1 \cong \Delta^2 \cup_{\Delta^1} \Delta^2.$$

Observe that even though Δ^1 has all simplices of dimension ≥ 2 as degenerate, yet $\Delta^1 \times \Delta^1$ has two non-degenerate 2-dimensional simplices.

A consequence of the above isomorphism is that the geometric realization of $\Delta^1 \times \Delta^1$ is exactly I^2, the unit square, which is the product of the geometric realization of Δ^1 with itself. Indeed, this is an instantiation of the general result in Theorem 14.1.4.9.

Remark 14.1.4.8. It is immediate to observe from Theorem 14.1.4.6 that geometric realization of Δ^n is exactly the standard topological n-simplex. Similarly, $\partial\Delta^n$ and $|\Lambda^n_i|$ are homeomorphic to exactly the pictures that we used in our mind to understand them.

Finally, the main result is as follows.

Theorem 14.1.4.9. *Let X, Y be simplicial sets. Then the natural map*

$$|X \times Y| \to |X| \times |Y|$$

is a homeomorphism.

We show that nerve of a category satisfies lifting property which would prove to be useful later on while discussing ∞-categories.

Proposition 14.1.4.10. *Let $X = N\mathbf{C}$ be thr nerve of a small category $\mathbf{C}$. Then for any $0 < i < n$, the following lifting problem is uniquely filled:*

$$\begin{tikzcd} \Lambda^n_i \arrow[r] \arrow[d, hook] & X \\ \Delta^n \arrow[ur, dashed, "!"'] & \end{tikzcd} .$$

Proposition 14.1.4.11. *Let S be a simplicial set and X a Kan complex. Then $\underline{\mathrm{Hom}}\,(S, X)$ is a Kan complex, denoted* $\mathrm{Map}(S, X)$.

14.1.5 Eilenberg-Zilber categories

This is an abstraction of the type of combinatorial proofs that we would like to make in the simplex category $\mathbf{\Delta}$. Indeed, as we would have to work with surjections and injections in $\mathbf{\Delta}$ primarily, which interacts with the the size of $[n]$ (which we would define to be n in a minute) as an injection would only increase the size and a surjection would only decrease it, therefore we need a systematic toolset to work with these things. In particular, if we denote $\mathbf{\Delta}_+$ to be the subcategory of all injective maps, $\mathbf{\Delta}_-$ to be the subcategory of all surjective maps in $\mathbf{\Delta}$ and $d : \mathrm{Ob}(\mathbf{\Delta}) \to \mathbb{N}$ the size map, then we have the following properties about the tuple $(\mathbf{\Delta}, \mathbf{\Delta}_+, \mathbf{\Delta}_-, d)$:

1. All bijections are in both $\mathbf{\Delta}_+$ and $\mathbf{\Delta}_-$.
2. The dimension map d takes bijective cardinals to the same natural.
3. Let $\sigma : [n] \to [m]$ in $\mathbf{\Delta}$ which is not bijective. If σ is injective (i.e. in $\mathbf{\Delta}_+$), then $d(a) < d(b)$ and if σ is surjective (i.e. in $\mathbf{\Delta}_-$), then $d(a) > d(b)$.
4. Any map $\sigma : [n] \to [m]$ in $\mathbf{\Delta}$ factors as a surjection followed by an injection.
5. For any surjective map $\sigma : [n] \to [m]$ in $\mathbf{\Delta}$, there exists a section $\pi : [m] \to [n]$, i.e. such that $\pi\sigma = \mathrm{id}_{[n]}$.
6. If $\sigma : [n] \to [m]$ is surjective, then the set of sections of σ uniquely determines the map σ.

Remark 14.1.5.1. We need this abstraction of properties of $\mathbf{\Delta}$ so that with the same techniques, we can work with bisimplicial sets, which is important if one wishes to consider simplicial objects in **sSet**.

These considerations about $\mathbf{\Delta}$ motivates the following definition.

Definition 14.1.5.2. (**Eilenberg-Zilber categories**) A category $\mathbf{A}$ is said to be an Eilenberg-Zilber category (or much simply, *EZ cateogry*), if there exists subcategories $\mathbf{A}_+, \mathbf{A}_-$ and a function $d : \mathrm{Ob}(\mathbf{A}) \to \mathbb{N}$ which satisfies the following axioms:

1. All isomorphisms of $\mathbf{A}$ are in both $\mathbf{A}_+$ and $\mathbf{A}_-$.
2. If a, a' are isomorphic objects in $\mathbf{A}$, then $d(a) = d(a')$.
3. Let $\sigma : a \to a'$ not be an isomorphism. If σ is in $\mathbf{A}_+$, then $d(a) < d(a')$. If σ is in $\mathbf{A}_-$, then $d(a) > d(a')$.
4. For any map $\sigma : a \to a'$ in $\mathbf{A}$, there exists unique factorization of σ into maps $p : a \to c$ in $\mathbf{A}_-$ and $i : c \to a'$ in $\mathbf{A}_+$

$$\begin{tikzcd} & c \arrow[dr, "i"] & \\ a \arrow[ur, "p"] \arrow[rr, "\sigma"'] & & a' \end{tikzcd} .$$

5. If $\sigma : a \to a'$ is a map in $\mathbf{A}_-$, then there exists a section $\pi : a' \to a$, i.e. $\pi\sigma = \mathrm{id}_a$.
6. If σ, σ' are two maps in $\mathbf{A}_-$ such both of them has the same collection of sections, then $\sigma = \sigma'$.

The main thrust of this section is to discuss presheaves over an EZ category $\mathbf{A}$, keeping in mind the prototypical case of $\mathbf{A} = \mathbf{\Delta}$. Indeed, the main result and its corollaries will serve first as a practice for the type of arguments we shall need later and also as a tool to be consistently used later in constructions with simplicial sets (which are, presheaves over $\mathbf{\Delta}$).

Example 14.1.5.3. By previous discussion, it is clear that the simplex category $\mathbf{\Delta}$ is an EZ category.

Let $\mathbf{A}$ be an EZ category and X a presheaf over $\mathbf{A}$. Then the category of elements of X, $\mathbf{A}/X$, is an EZ category again. Indeed, define $(\mathbf{A}/X)_+$ exactly as those pairs (a, s) where $a \in \mathbf{A}_+$ and $(\mathbf{A}/X)_-$ exactly as those pairs (a, s) where $a \in \mathbf{A}_-$. Further for an object (a, s), define $d(a, s) = d(a)$, where the latter d is coming from the EZ structure on $\mathbf{A}$.

14.1.6 Kan complexes and homotopy groups

In the beginning section, we showed how simplicial sets can be viewed as combinatorial version of usual spaces. However, in order to do homotopy theory in this combinatorial setting, we need to isolate a class of simplicial sets which are right for this kind. Indeed, these will be Kan complexes.

We wish to define $\pi_n(X, x)$ where $x \in X_0$ is a 0-simplex of a simplicial set. To this end, we immediately run into problems as $\pi_0(X, x)$ should be the equivalence class of all those 0-simplices which are boundaries of a 1-simplex. But it is immediately clear that this is not an equivalence relation! Indeed, consider Δ^1. Then the above relation is not symmetric as we have $0 \to 1$ as a 1-simplex but there is no $1 \to 0$ in Δ^1_1. Similarly, if we try to prove transitivity of the above relation, we land up in the following situation. Let $x \to y, y \to z$ be two 1-simplices in X. Then, we wish to find a 1-simplex $x \to z$ in X. Note that given $x \to y$ and $y \to z$, we have a map $\Lambda^2_1 \to X$ and we wish to wish to fill the following diagram

$$\begin{tikzcd} \Lambda^2_1 \arrow[r] \arrow[d, hook] & X \\ \Delta^2 \arrow[ur, dashed] & \end{tikzcd} .$$

So we need those simplicial sets, where the above dotted arrow always exists.

Definition 14.1.6.1 (Kan complex). A simplicial set X is a Kan complex if for any $n \geq 0$ and any $0 \leq i \leq n$, the following lifting diagram is filled

$$\begin{tikzcd} \Lambda^n_i \arrow[r] \arrow[d, hook] & X \\ \Delta^n \arrow[ur, dashed] & \end{tikzcd} .$$

Example 14.1.6.2. Clearly, Δ^n are not Kan complexes for $n \geq 0$. Moreover, $N\mathbf{C}$ is a Kan complex if and only if $\mathbf{C}$ is a groupoid. This is immediate.

The prototypical (and in some sense, the only example) of Kan complexes are those obtained by Theorem 14.1.4.6.

Proposition 14.1.6.3. *Let X be a topological space. Then* $\mathrm{Sing}(X)$ *is a Kan complex.*

Proof. Let $n \geq 0$, $0 \leq i \leq n$. Then a map $\Lambda^n_i \to \mathrm{Sing}(X)$ is equivalent to a map $|\Lambda^n_i| \to X$ by adjunction of Theorem 14.1.4.6. Consequently, we wish to fill the following diagram

$$\begin{tikzcd} {|\Lambda^n_i|} \arrow[r] \arrow[d, hook] & X \\ {|\Delta^n|} \arrow[ur, dashed] & \end{tikzcd}.$$

But this is immediate as $|\Lambda^n_i| \hookrightarrow |\Delta^n|$ is a retraction. ☐

Another example of a Kan complex is the mapping complex.

Definition 14.1.6.4 (Mapping complex). Let X, Y be two spaces. The mapping complex, denoted $\mathrm{Map}(X,Y)$, is the one whose n-simplices are

$$[n] \mapsto \mathrm{Hom}_{\mathbf{Top}}\left(X \times |\Delta^n|, Y\right).$$

Note that $\mathrm{Map}(X,Y)$ has 0-simplices as continuous maps, 1-simplices as homotopies and so on. The name is justified by the following result.

Corollary 14.1.6.5. *Let X, Y be spaces. Then the mapping complex* $\mathrm{Map}(X,Y)$ *is a Kan complex.*

Proof. As all spaces are compactly generated, therefore we have $\mathrm{Hom}_{\mathbf{Top}}\left(X \times |\Delta^n|, Y\right) \cong \mathrm{Hom}_{\mathbf{Top}}\left(|\Delta^n|, Y^X\right)$. Consequently, $\mathrm{Map}(X,Y) \cong \mathrm{Sing}(Y^X)$, which is a Kan complex by Proposition 14.1.6.3. ☐

One can now check that the relation mentioned in the beginning on X_0 is indeed an equivalence relation and would thus yield the definition of $\pi_0(X, x)$, however, we would like to define all homotopy groups in one go. We first define the notion of two simplicies being homotopic.

Definition 14.1.6.6 (Homotopic rel ∂). Let X be a Kan complex and $\sigma, \tau \in X_n$ be two n-simplicies. Then σ and τ are homotopic rel ∂ if there exists an $n+1$-simplex $\psi \in X_{n+1}$ such that $\sigma|_{\partial\Delta^n} = \tau|_{\partial\Delta^n}$ and $\partial_n \psi = \sigma$, $\partial_{n+1}\psi = \tau$ and for all $0 \leq i \leq n-1$, $\partial_i \psi = \sigma d^i s^{n-1} = \tau d^i s^{n-1}$. We can diagramatically represent these conditions as the commutativity of the the diagrams

$$\begin{tikzcd} X & \Delta^n \arrow[l, "\tau"'] \\ \Delta^n \arrow[u, "\sigma"] & \partial\Delta^n \arrow[l, hook'] \arrow[u, hook] \end{tikzcd},$$

$$\begin{tikzcd} \Delta^{n+1} \arrow[r, "\psi"] & X & \Delta^{n+1} \arrow[l, "\psi"'] \\ \Delta^n \arrow[u, "d^n"] \arrow[ur, "\sigma"'] & & \Delta^n \arrow[u, "d^{n+1}"'] \arrow[ul, "\tau"] \end{tikzcd}$$

and for $0 \leq i \leq n-1$

$$\begin{tikzcd} \Delta^{n+1} \arrow[r, "\psi"] & X & \Delta^n \arrow[l, "\tau"'] \\ \Delta^n \arrow[u, "d^i"] \arrow[rr, "s^{n-1}"'] & & \Delta^{n-1} \arrow[u, "d^i"'] \end{tikzcd}$$

in **sSet**.

Indeed, we see that this generalizes our previous notion of homotopy relative to boundary as follows.

Lemma 14.1.6.7. *Let $X = \mathrm{Sing}(Y)$ be the Kan complex associated to a space Y. Then two n-simplices $\sigma, \tau \in X_n$ are homotopic rel ∂ if and only if $\sigma, \tau : |\Delta^n| \to Y$ are homotopic relative to $|\partial\Delta^n| \hookrightarrow |\Delta^n|$ in the classical sense.*

Proof. Let $\sigma, \tau : \Delta^n \to X$ be two n-simplices. These are homotopic rel ∂ if the diagrams in Definition 14.1.6.6 commutes in **sSet**. By the adjunction of Theorem 14.1.4.6, this is equivalent to commutativity of the following diagrams in **Top**

$$\begin{tikzcd}
Y & {|\Delta^n|} \arrow[l, "\tau"'] \\
{|\Delta^n|} \arrow[u, "\sigma"] & {|\partial\Delta^n|} \arrow[l, hook'] \arrow[u, hook]
\end{tikzcd},$$

$$\begin{tikzcd}
{|\Delta^{n+1}|} \arrow[r, "\psi"] & X & {|\Delta^{n+1}|} \arrow[l, "\psi"'] \\
{|\Delta^n|} \arrow[u, "d^n"] \arrow[ur, "\sigma"'] & & {|\Delta^n|} \arrow[u, "d^{n+1}"'] \arrow[ul, "\tau"]
\end{tikzcd}$$

and for $0 \le i \le n-1$

$$\begin{tikzcd}
{|\Delta^{n+1}|} \arrow[r, "\psi"] & Y & {|\Delta^n|} \arrow[l, "\tau"'] \\
{|\Delta^n|} \arrow[u, "d^i"] \arrow[rr, "s^{n-1}"'] & & {|\Delta^{n-1}|} \arrow[u, "d^i"']
\end{tikzcd}.$$

But this is equivalent to the data of a homotopy $H : |\Delta^n| \times I \to Y$ rel boundary. Indeed, since we want $H_0 = \sigma$, $H_1 = \tau$ and $H_t|_{|\partial\Delta^n|} = \sigma|_{|\partial\Delta^n|} = \tau|_{\partial\Delta^n}$ for all $t \in I$, therefore we naturally get that H factors through $|\Delta^{n+1}|$. This can be visualized for a 3-simplex immediately. $\square$

Finally, when X is a Kan complex, then this is an equivalence relation. This is where combinatorial relations between face and degeneracy maps of a simplicial set as provided in Proposition 14.1.2.3 becomes handy.

Proposition 14.1.6.8. *Let X be a Kan complex. The homotopy rel ∂ is an equivalence relation on each X_n for $n \geq 1$.*

Proof. We first show reflexivity. Consider $\sigma : \Delta^n \to X$ an n-simplex. We claim that $\psi = \sigma s^n : \Delta^{n+1} \to X$ works as the homotopy from σ to σ. Indeed, we see that $\psi d^n = \sigma s^n d^n = \sigma$ and $\psi d^{n+1} = \sigma s^n d^{n+1} = \sigma$ by Proposition 14.1.2.3. Similarly, for $0 \le i \le n-1$, we have $\psi d^i = \sigma s^n d^i = \sigma d^i s^{n-1}$ by the same result. This establishes reflexivity.

We now show symmetry. Let $\sigma \sim \tau$ for some $\sigma, \tau : \Delta^n \to X$ with $\sigma|_{\partial\Delta^n} = \tau|_{\partial\Delta^n}$. Then, there exists $\psi : \Delta^{n+1} \to X$ with $\psi d^n = \sigma$, $\psi d^{n+1} = \tau$ and $\psi d^i = \sigma d^i s^{n-1} = \tau d^i s^{n-1}$ for $0 \le i \le n-1$. We wish to show that $\tau \sim \sigma$. We will use the fact that X is a Kan complex (so that all horns can be filled). In particular, we will construct a horn $\kappa : \Lambda^{n+2}_n \to X$, filling which will give us the required homotopy from τ to σ. Indeed, let κ be obtained from ψ by adding degeneracies such that $\kappa d^{n+2} = \psi$ and $\kappa d^{n+1} = \sigma s^n$, that is, the degenerate $n+1$-simplex obtained by repeating the n^{th}-vertex of σ. For $0 \leq i \leq n-1$, we keep $\kappa d^i = \psi d^i s^{n-1}$. Filling the horn κ yields an $n+2$-simplex $\tilde{\kappa} : \Delta^{n+2} \to X$ whose n^{th}-face is exactly a homotopy from τ to σ. Indeed, denote $\phi = \tilde{\kappa} d^n$. Then, $\phi d^n = \tilde{\kappa} d^n d^n = \tilde{\kappa} d^{n+1} d^n = \sigma$. Similarly, $\phi d^{n+1} = \tau$. These follow from the observation that $d^n d^n$ is the unique n-simplex of $\tilde{\kappa}$ not containing the vertices n and $n+1$ in $\tilde{\kappa}$. For $0 \leq i \leq n-1$, by Proposition 14.1.2.3, we have $\phi d^i = \tilde{\kappa} d^n d^i = \tilde{\kappa} d^i d^{n-1}$, which in turn is $\kappa d^i d^{n-1} = \psi d^i s^{n-1} d^{n-1} = \psi d^i = \sigma d^i s^{n-1}$, as needed. This shows symmetry.

Finally, we wish to show transitivity. Let $\psi, \psi' \in X_{n+1}$ be homotopies $\sigma \sim \tau$ and $\tau \sim \eta$ respectively for some $\sigma, \tau, \eta \in X_n$. We wish to construct a homotopy $\psi'' \in X_{n+1}$ between σ and η. Indeed, we obtain a horn $\kappa : \Lambda^{n+2}_{n+2} \to X$ whose n^{th} and $n+1^{\text{th}}$ boundaries are ψ and ψ' respectively and the rest boundaries are required degeneracies. Filling this horn up by the Kan condition gives $\tilde{\kappa}$ and its $n+2^{\text{th}}$-boundary is the required homotopy. $\square$

Consequently, we define homotopy groups of a Kan complex as follows.

Definition 14.1.6.9 (**Homotopy groups of a Kan complex**)**.** Let X be a Kan complex and $x \in X_0$ be a base point. Then for $n \geq 1$ define

$$\pi_n(X, x_0) = \{\sigma \in X_n \mid \sigma|_{\partial\Delta^n} = c_{x_0}\} / \sim$$

where $\sim$ is the homotopy rel ∂. For $n = 0$, define

$$\pi_0(X) = X_0 / \sim$$

where

$$x \sim y \iff \exists \gamma \in X_1 \text{ s.t. } \gamma d^1 = x \ \& \ \gamma d^0 = y.$$

We now show that $\pi_n(X, x)$ is indeed a group.

Construction 14.1.6.10 (*Composition and group operation on* $\pi_n(X, x)$)**.** Let X be a Kan complex and $x \in X$ be a point in it (i.e a 0-simplex). Let $\sigma, \tau \in X_n$ be two n-simplices such that $\sigma|_{\partial\Delta^n} = \tau|_{\partial\Delta^n} = c_x$. We construct the composition $\sigma \cdot \tau$ of σ and τ as the following n-simplex. Construct the following horn $\kappa : \Lambda^{n+1}_n \to X$ whose $n-1^{\text{th}}$-boundary is σ, $n+1^{\text{th}}$-boundary is τ[1] and for $0 \leq i \leq n-2$, $\kappa d^i = c_x$. It follows from horn-filling condition of X that we get an $n+1$-simplex $\psi : \Delta^{n+1} \to X$ extending the horn κ. Consequently, we define the *concatenation* $\sigma \cdot \tau$ as the n^{th}-boundary of ψ, that is,

$$\sigma \cdot \tau = \psi d^n.$$

We claim that the operation

$$\begin{aligned} \cdot : \pi_n(X, x) \times \pi_n(X, x) &\longrightarrow \pi_n(X, x) \\ ([\alpha], [\beta]) &\longmapsto [\alpha \cdot \beta] \end{aligned}$$

[1]If one is thinking of paths, then it is important to note that $\sigma \cdot \tau$ is the one where τ is traversed first and then σ. One has to let go of the past notation because here simplices are not merely going to be paths, homotopies and so on, but rather more general objects like arrows of a category, 2-arrows and so on, so to be consistent with the notion of composition, it is best that we change the order in which we concatenate paths.

is a well-defined function, that is, $[\alpha \cdot \beta]$ only depends on $[\alpha]$ and $[\beta]$.

Indeed, suppose $[\alpha] = [\alpha']$ and $[\beta] = [\beta']$ in $\pi_n(X, x)$. Then, we have a homotopy rel ∂ denoted $\psi \in X_{n+1}$ from α to α' and $\chi \in X_{n+1}$ from β to β'. We wish to construct $\phi \in X_{n+1}$ which is a homotopy rel ∂ from $\alpha \cdot \beta$ to $\alpha' \cdot \beta'$. As usual, we obtain this homotopy by constructing a higher horn and filling it by Kan condition.

To correctly denote the simplex to be constructed, we first observe that if $\delta \sim \epsilon$ is a homotopy rel ∂ of two n-simplices with $\delta|_{\partial\Delta^n} = \epsilon|_{\partial\Delta^n} = x$, then $\delta \cdot x \sim \epsilon$. Indeed, consider an $n+2$-horn whose $n+2$-boundary is the composition simplex $\delta \cdot x$, $n+1$-boundary is ϵs_0 and all i-boundaries for $0 \leq i \leq n-1$ are degeneracies of δ. Filling this yields the n-boundary as the required homotopy.

We now show that if $\beta \sim \beta'$, then $\alpha \cdot \beta \sim \alpha \cdot \beta'$. This would complete the proof. Indeed, this is immediate by considering an $n+2$-horn given by $\kappa : \Lambda^{n+2}_{n+1} \to X$ such that κd^{n+2} is the homotopy $\beta \sim \beta'$, κd^n is the composition $\alpha \cdot \beta$, κd^{n-1} is the composition $\alpha \cdot \beta'$ and the κd^i for $0 \leq i \leq n-2$ are all x.

Consequently, we get that $\cdot$ is a well defined operation on $\pi_n(X, x)$. We will later show that $\cdot$ makes $\pi_n(X, x)$ into a group. Moreover, we will show that $\pi_n(X, x) \cong \pi_n(|X|, x)$ and $\pi_n(\mathrm{Sing}(Y), y) \cong \pi_n(Y, y_0)$!

14.1.7 The fundamental group of a Kan complex

Let X be a Kan complex and $x_0 \in X$ be a 0-simplex. The fundamental group $\pi_1(X, x_0)$ is explicitly given by the following

$$\begin{aligned}\pi_1(X, x_0) &= \{\sigma : \Delta^1 \to X \mid \sigma|_{\partial\Delta^1} = x_0\}/\sim \\ &= \{\sigma \in X_1 \mid d_0(\sigma) = d_1(\sigma) = x_0\}/\sim\end{aligned}$$

where $\sigma \sim \tau$ if and only if there exists a homotopy rel ∂ denoted $H : \Delta^2 \to X$, from σ to τ. That is, $d_1(H) = \sigma$, $d_2(H) = \tau$ and $d_0(H) = s_0(d_0(\sigma)) = s_0(x_0) = \mathrm{id}_{x_0}$. This is a group where the operation is

$$\begin{aligned}\cdot : \pi_1(X, x_0) \times \pi_1(X, x_0) &\longrightarrow \pi_1(X, x_0) \\ ([\sigma], [\tau]) &\longmapsto [\sigma \cdot \tau]\end{aligned}$$

where $\sigma \cdot \tau$ is a 1-simplex which is their composition (Construction 14.1.6.10), obtained by the 1-boundary of the 2-simplex δ obtained by filling the horn

$$\kappa : \Lambda^2_1 \to X$$

whose $\partial_0\Lambda^2_1 = \sigma$ and $\partial_1\Lambda^2_1 = \tau$. More precisely, we define κ_1 as follows (this is sufficient by Example 14.1.4.2 and Lemma 14.1.4.3):

$$\begin{aligned}\kappa_1 : (\Lambda^2_1)_1 &\longrightarrow X_1 \\ \{0,1\} &\longmapsto \sigma \\ \{1,2\} &\longmapsto \tau \\ \{1,1\} &\longmapsto s_0(d_0(\sigma)).\end{aligned}$$

In this section, following the usual terminology, we will write $\sigma \cdot \tau$ as $\tau * \sigma$ for $\sigma, \tau \in X_1$ with $d_0(\sigma) = d_1(\sigma) = x_0$.

We now prove some basic results about $\pi_1(X, x_0)$. It is a good exercise to show that $\pi_1(X, x_0)$ is a group.

Theorem 14.1.7.1. *Let X be a Kan complex and $x_0 \in X_0$. Then $\pi_1(X, x_0)$ is a group.*

Proof. We first show that for any three $\alpha, \beta, \gamma \in X_1$ with their boundaries being x_0, we have $(\alpha * \beta) * \gamma \sim \alpha * (\beta * \gamma)$. Indeed, let H be the witness of composition $\alpha * \beta$, K that of $\beta * \gamma$, L that of $(\alpha * \beta) * \gamma$ and P that of $\alpha * (\beta * \gamma)$. Now consider the following horn $\kappa : \Lambda^3_2 \to X$ given by following on non-degenerate 2-simplices:

$$\begin{aligned} \kappa_2 : (\Lambda^3_2)_2 &\longrightarrow X_2 \\ \{0,1,2\} &\longmapsto H \\ \{1,2,3\} &\longmapsto K \\ \{0,2,3\} &\longmapsto L. \end{aligned}$$

By Kan condition, the above horn is filled and its 2-boundary yields a 2-simplex $\chi \in X_2$ such that $d_0(\chi) = \beta * \gamma$, $d_1(\chi) = (\alpha * \beta) * \gamma$ and $d_2(\chi) = \alpha$. We now construct the required homotopy by filling another 3-horn. Indeed, consider a horn $\lambda : \Lambda^3_1 \to X$ such that $d_0(\lambda) = s_1(\beta * \gamma)$, $d_2(\lambda) = P$ and $d_3(\lambda) = \chi$. This fills to give its 1-boundary as the required homotopy.

The identity element being c_{x_0} and the existence of inverses are also immediate results of horn filling and is thus omitted. $\square$

A Kan complex is path-connected if $\pi_0(X) = 0$.

14.1.8 ∞-categories

14.1.9 Theorem of Boardman-Vogt

14.2 Classical homotopical algebra

14.2.1 Model categories

We discuss now a general setup in which one can "do" homotopy theory.

Definition 14.2.1.1 (**Model categories**). Let $\mathbf{C}$ be a category and $W, C, F \subseteq \mathbf{C}$ be subcategories of $\mathbf{C}$ which are called weak equivalences ($\approx$), cofibrations ($\rightarrowtail$) and fibrations ($\twoheadrightarrow$) respectively. We call $W \cap C$ weak/acyclic cofibrations and $W \cap F$ weak/acyclic fibrations. Then, the tuple $(\mathbf{C}, W, C, F)$ is a model category if it satisfies the following axioms:

1. The category $\mathbf{C}$ has all finite limits and colimits.
2. Weak equivalences satisfies 2 out-of 3 property.
3. For any $f : X \to Y$ in $\mathbf{C}$, we have two factorizations

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, tail, "\approx"'] & & Y \\ & Z \arrow[ur, two heads] & \end{tikzcd} \qquad \begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, tail] & & Y \\ & Z \arrow[ur, two heads, "\approx"'] & \end{tikzcd}$$

one a weak cofibration followed by fibration and another one a cofibration followed by a weak fibration.

4. We have

$$rlp(W \cap C) = F$$
$$C = llp(W \cap F)$$

where for a subcategory $S \subseteq \mathbf{C}$, the collection $rlp(S)$ denotes the collection of all maps $X \to Y$ in $\mathbf{C}$ satisfying right lifting property wrt S, i.e., such that for *any* commutative square

$$\begin{tikzcd} Z \arrow[r] \arrow[d, "\in S"'] & X \arrow[d, "f"] \\ W \arrow[r] \arrow[ur, dashed, "h"] & Y \end{tikzcd}$$

where left vertical arrow is in S, there exists a lift $h : W \to X$ as shown which makes all diagrams commute. Similarly, one defines $llp(S)$.

Definition 14.2.1.2 (Cofibrant/fibrant objects). An object X in a model category $\mathbf{C}$ is cofibrant (fibrant) if the unique map from initial object $\emptyset \to X$ (to terminal object $X \to \text{pt.}$) is a cofibration (fibration).

Chapter 15

Algèbre Commutative Dérivée

Part VI

Special Topics

Chapter 16

Commutative Algebra

Contents

In this chapter, we collect topics from contemporary commutative algebra. The most need of all this material comes from algebraic goemetry. In particular, in the following, we list out the topics that we would need for our treatment of basic algebraic geometry.

1. Dimension theory : For dimension of schemes, Hauptidealsatz, local complete intersection, etc.
2. Integral dependence : For proper maps between affine varieties, normalization, finiteness of integral closure, certain DVRs of dimension 1, etc.
3. Field theory : For birational classification of varieties, primitive element theorem, basic algebra in general, etc.
4. Completions : Local analysis of singularities, formal schemes, complete local rings, Cohen structure theorem, Krull's theorem, etc.
5. Valuation rings : For curves and their non-singular points (DVRs) and various equivalences, Dedekind domains, etc.
6. Multiplicities : For intersections in projective spaces, intersection multiplicity, Hilbert polynomials, flat families, studying singularities in an algebraic variety etc.
7. Kähler differentials : For differential forms on schemes, this will be used consistently in further topics.
8. Depth and Cohen-Macaulay : For local complete intersections, blowing up, etc.
9. Tor and Ext functors : They are tools for other algebraic notions, generizable to global algebra, tor dimension, etc.
10. Projective modules : For vector bundles, projective dimension and Ext, pd + depth = dim for regular local rings, etc.
11. Flatness : Family of schemes varying continuously, smooth and étalé maps, etc.
12. Lifting properties - Étale, unramified and smooth morphisms : These are used heavily for the corresponding scheme maps, and beyond.

Notation 16.0.0.1. Let R be a ring and $f(x) \in R[x]$ be a polynomial. We will denote $c_n(f) \in R$ to be the coefficient of x^n in $f(x)$. If $f(x,y) \in R[x,y]$, then we will denote $c_{n,m}(f) \in R$ to be the coefficient of $x^n y^m$ in $f(x,y)$. We may also write $c_{x^n}(f)$ for $c_n(f)$ and $c_{x^n y^m}(f)$ for $c_{n,m}(f)$ if it makes statements more clear.

Remark 16.0.0.2. We will consistently keep using the geometric viewpoint given by the theory of schemes (see Chapter 1) in discussing the topics below, as a viewpoint to complement the algebraic viewpoint. This will also showcase the usefulness of scheme language.

16.1 General algebra

We discus here general results about prime ideals, modules and algebras.

16.1.1 Nakayama lemma

Let R be a ring. Denote the set of all units of R as $R^\times$. The *Jacobson radical* is the ideal $\mathfrak{r}$ of R formed by the intersection of all maximal ideals of R. A *finitely generated R-module M* is a module which has a finite collection of elements $\{x_1, \dots, x_n\} \subset M$ such that for any $z \in M$, there are $r_1, \dots, r_n \in R$ so that $z = r_1 x_1 + \cdots + r_n x_n$. More concisely, if there is a surjection R-module homomorphism $R^n \twoheadrightarrow M$. Let's begin with a simple observation.

Lemma 16.1.1.1. *Let R be a ring and M be a simple R-module. Then*

$$M \cong R/\mathfrak{m}$$

where $\mathfrak{m} \leq R$ is a maximal ideal.

Proof. As M is simple, therefore for any non-zero $f \in M$, we have $Rf = M$. It follows that the map $\varphi : R \to M$ mapping $r \mapsto rf$ is surjective. Thus, $M \cong R/\mathrm{Ker}\,(\varphi)$. If $\mathrm{Ker}\,(\varphi)$ is not maximal, then $R/\mathrm{Ker}\,(\varphi)$ has a non-trivial submodule, a contradiction to simplicity of M. □

We then have the following results about $\mathfrak{r}$.

Proposition 16.1.1.2. *Let R be a ring and let $\mathfrak{r}$ denotes it Jacobson radical. Then,*

1. *$x \in \mathfrak{r}$ if and only if $1 - xy \in R^\times$ for any $y \in R$.*
2. *(Nakayama lemma) Let M be a finitely generated R-module. If $\mathfrak{q} \subseteq \mathfrak{r}$ is an ideal of R such that $\mathfrak{q}M = M$, then $M = 0$.*
3. *Let M be a finitely generated module and $\mathfrak{q} \subseteq \mathfrak{r}$. Let $N \leq M$ be a submodule of M such that $M = N + \mathfrak{q}M$, then $M = N$.*
4. *If R is a local ring and M, N are two finitely generated modules, then*
$$M \otimes_R N = 0 \iff M = 0 \text{ or } N = 0.$$

Proof. 1. (L ⇒ R) Suppose there is $y \in R$ such that $1 - xy \notin R^\times$. Since each non-unit element is contained in a maximal ideal by Zorn's lemma, therefore $1 - xy \in \mathfrak{m}$ for some maximal ideal. Since $x \in \mathfrak{r}$, therefore $x \in \mathfrak{m}$. Hence $xy, 1 - xy \in \mathfrak{m}$, which means that $1 \in \mathfrak{m}$, a contradiction.
(R ⇒ L) Suppose $1 - xy \in R^\times$ for all $y \in R$ and $x \notin \mathfrak{r}$. Then, again by Zorn's lemma we have $x \in R^\times$. Hence let $y = x^{-1}$ to get that $1 - xy = 1 - 1 = 0 \in R^\times$, a contradiction.

2. Suppose $M \neq 0$. Since M is finitely generated, therefore there is a submodule $N \subset M$ such that M/N is simple (has no proper non-trivial submodule). By Lemma 16.1.1.1, $M/N \cong R/\mathfrak{m}$. Then, $\mathfrak{m}R \neq R$ which is same as $\mathfrak{m}M \neq M$. Since $\mathfrak{q} \subseteq \mathfrak{r} \subseteq \mathfrak{m}$, hence $\mathfrak{q}M \neq M$, a contradiction.

3. Apply 2. on M/N.

4. The only non-trivial part is L ⇒ R. Since $(M \otimes_R N)/\mathfrak{m}(M \otimes_R N) = M/\mathfrak{m}M \otimes_{R/\mathfrak{m}} N/\mathfrak{m}N$, therefore we have $M/\mathfrak{m}M \otimes_{R/\mathfrak{m}} N/\mathfrak{m}N = 0$. Since $R/\mathfrak{m}$ is a field therefore $M/\mathfrak{m}M = 0$ WLOG. Hence, $M = \mathfrak{m}M$ and since R is local, therefore $\mathfrak{r} = \mathfrak{m}$. We conclude by Nakayama. □

Note that if $\mathfrak{a}$ is an ideal, then $1 + \mathfrak{a}$ is a multiplicative set. In-fact this is quite a special multiplicative set because of the following.

Lemma 16.1.1.3. *Let R be a ring and $\mathfrak{a} \leq R$. For the multiplicative set $S = 1 + \mathfrak{a}$, $S^{-1}\mathfrak{a}$ is in Jacobson radical of $S^{-1}R$.*

Proof. Pick $x/s \in S^{-1}A$ and $a/t \in S^{-1}\mathfrak{a}$. It is equivalent to show that $1 - \frac{xa}{st}$ is a unit in $S^{-1}A$ by Proposition 16.1.1.2, 1. Indeed, $st = 1 + a'$ for some $a' \in \mathfrak{a}$. Thus $1 - \frac{xa}{st} = \frac{1+a'-ax}{st}$ where numerator is in S, hence a unit as required. □

Using this observation, we can find a single annihilator for certain modules.

Lemma 16.1.1.4. *Let $\mathfrak{a} \leq R$ be an ideal such that $\mathfrak{a}M = M$ for some finitely generated R-module M. Then, there exists $x \in R$ such that $xM = 0$.*

Proof. Let $S = 1 + \mathfrak{a}$, $N = S^{-1}M$, $B = S^{-1}A$ and $\mathfrak{b} = S^{-1}\mathfrak{a}$. Then $N = \mathfrak{b}N$ and hence by Nakayama, we get $N = 0$. As N is finitely generated, therefore there exists an element $x \in S$ such that $xM = 0$, as required. □

Here's a simple, but important example.

Lemma 16.1.1.5. *Let M, N be A-modules and N be finitely generated. Let $\mathfrak{a} \leq A$ be in Jacobson radical and $\varphi : M \to N$ be an A-linear map. If $\bar{\varphi} : M/\mathfrak{a}M \to N/\mathfrak{a}N$ is surjective, then φ is surjective.*

Proof. We wish to show that $N = \varphi(M)$. By Nakayama, it is sufficient to show that $N = \varphi(M) + \mathfrak{a}N$. Pick any $n \in N$. By hypothesis, there exists $m \in M$ such that $\varphi(m) - n = \sum_i a_i n_i \in \mathfrak{a}N$. The result follows. □

Corollary 16.1.1.6. *Let $(R, \mathfrak{m}, \kappa)$ be a local ring and $\varphi : M \to N$ be an R-module homomorphism and N be finitely generated. If $\varphi \otimes \mathrm{id} : M \otimes \kappa \to N \otimes \kappa$ is surjective, then φ is surjective.* □

Here's another application which is quite interesting arithmetically.

Proposition 16.1.1.7. *Let R be a ring whose underlying abelian group is free of finite rank, A be a ring whose underlying abelian group is finitely generated and $\varphi : R \to A$ be a ring homomorphism. Then the following are equivalent:*

1. *φ is an isomorphism.*
2. *For each $\mathfrak{p} \in \mathrm{Spec}\,(\mathbb{Z})$, the homomorphism*
$$\varphi_{\mathfrak{p}} : R \otimes_{\mathbb{Z}} \kappa(\mathfrak{p}) \to A \otimes_{\mathbb{Z}} \kappa(\mathfrak{p})$$
is an isomorphism.

Proof. By functoriality of tensor products, it is immediate that (1. $\Rightarrow$ 2.). For the converse, we proceed as follows. We first show that φ is surjective. Let $M = \mathrm{CoKer}\,(\varphi)$. We have the short exact sequence of finitely generated $\mathbb{Z}$-modules
$$R \xrightarrow{\varphi} A \longrightarrow M \longrightarrow 0\ .$$
Suppose φ is not surjective. Then $M \neq 0$. Consequently, there exists a prime $\mathfrak{p} \in \mathrm{Spec}\,(\mathbb{Z})$ such that $M_{\mathfrak{p}} \neq 0$. Localizing the above sequence at this prime, we obtain by exactness the following sequence of $\mathbb{Z}_{\mathfrak{p}}$-modules:
$$R_{\mathfrak{p}} \xrightarrow{\varphi_{\mathfrak{p}}} A_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \longrightarrow 0\ .$$
By Corollary 16.1.1.6 applied on the local ring $\mathbb{Z}_{\mathfrak{p}}$, it follows that $\varphi_{\mathfrak{p}}$ as above is surjective. It follows by exactness of the sequence above that $M_{\mathfrak{p}} = 0$, a contradiction.

To see injectivity of φ, observe that $\mathrm{Ker}\,(\varphi)$ is a submodule of R. Since R is free of finite rank, therefore by Proposition 16.23.0.8, it follows that $K := \mathrm{Ker}\,(\varphi)$ is a free module of finite rank. Consequently, the following exact sequence is of free $\mathbb{Z}$-modules:
$$0 \longrightarrow K \xrightarrow{\iota} R \xrightarrow{\varphi} A \longrightarrow 0$$
As $\mathbb{Z}_{\mathfrak{o}} = \mathbb{Q} = \kappa(\mathfrak{o})$, thus by exactness of localization, we get the following exact sequence
$$0 \longrightarrow K \otimes \mathbb{Q} \xrightarrow{\iota \otimes \mathrm{id}} R \otimes \mathbb{Q} \xrightarrow{\varphi \otimes \mathrm{id}} A \otimes \mathbb{Q} \longrightarrow 0\ .$$

By hypothesis, $\varphi \otimes \mathrm{id}$ is an isomorphism. Conseqeuntly, $\mathrm{rank}K + \mathrm{rank}A = \mathrm{rank}R$, but the isomorphism implies that $\mathrm{rank}R = \mathrm{rank}A$. It follows at once that $\mathrm{rank}K = 0$ and since K is free, therefore $K = 0$, as required. □

As another simple application, here's a lemma which is a good exercise as well.

Lemma 16.1.1.8. *Let $(R, \mathfrak{m})$ be a local domain and $k = R/\mathfrak{m}$ and $K = Q(R)$. Denote $Q(M)$ to be the fraction module of M; $Q(M) = M \otimes_R R_\circ$. If M is a finitely generated R-module such that*

$$\dim_k M/\mathfrak{m}M = \dim_K Q(M) = n,$$

then $M \cong R^n$.

16.1.2 Localization

We next consider localization of rings and R-modules. Take any multiplicative set $S \subset R$ which contains 1. Then, localizing an R-module M on S is defined as

$$S^{-1}M := \{m/s \mid m \in M, s \in S\}.$$

where $m/s = n/t$ if and only if $\exists u \in S$ such that $u(mt - ns) = 0$. We have that $S^{-1}M$ is an R-module where addition $m/s + n/t = (mt + ns)/st$. In the case when $M = R$, we get a ring structure on $S^{-1}R$ as well where multiplication is given by $m/s \cdot n/t := mn/st$. There is a natural map $M \to S^{-1}M$ which maps $m \mapsto m/1$ and it may not be an injection if $\exists m \in M$ and $s \in S$ such that $s \cdot M = 0$.

Lemma 16.1.2.1. *Let $S \subset R$ be a multiplicative set in a ring R and M be an R-module. Then,*

$$S^{-1}M \cong S^{-1}R \otimes_R M.$$

Proof. One can do this by directly checking the universal property of tensor product of $S^{-1}R$ and M over R for $S^{-1}M$. We have the map $\varphi : S^{-1}R \times M \to S^{-1}M$ given by $(r/s, m) \mapsto rm/s$. Now for any bilinear map $f : S^{-1}R \times M \to N$, we can define the map $\tilde{f} : S^{-1}M \to N$ given by $\tilde{f}(m/s) := f(1/s, m)$. Clearly, $\tilde{f}$ is well-defined and $\tilde{f}\varphi = f$. Moreover, if $g : S^{-1}M \to N$ is such that $g\varphi = f$, then $g(m/s) = f(1/s, m) = \tilde{f}(m/s)$. Hence $\tilde{f}$ is unique with this property. □

Lemma 16.1.2.2. *Localization w.r.t a multiplicative set $S \subset R$ is an exact functor on* **Mod**(R).

Proof. Let $0 \to M' \to M \to M'' \to 0$ be an exact sequence of R-modules. Then we have the localized sequence $S^{-1}M' \to S^{-1}M \to S^{-1}M''$. Since $S^{-1}0 = 0$, therefore this is left exact. Exactness at middle follows from exactness at middle of the first sequence. The right exactness can be seen by right exactness of tensor product functor $S^{-1}R \otimes_R -$ and by Lemma 16.1.2.1. □

Lemma 16.1.2.3. *Let R be a ring and $S \subset R$ be a multiplicative set. Then*

$$\begin{aligned}\{\text{prime ideals of } R \text{ not intersecting } S\} &\xrightarrow{\cong} \{\text{prime ideals of } S^{-1}R\}\\ \mathfrak{p} &\longmapsto S^{-1}\mathfrak{p}\end{aligned}$$

Proof. Trivial. □

Next we see an important property of modules, that is their "local characteristic". This means that one can check whether an element of a module is in a submodule by checking it locally at each prime, as the following lemma suggests. This has geometric significance in algebraic geometry (M induces and is induced by a quasi-coherent sheaf over $\mathrm{Spec}\,(R)$, see **??**).

Lemma 16.1.2.4. *Let M be an R-module. Then,*

1. *$M \neq 0$ if and only if there exists a point $\mathfrak{p} \in \mathrm{Spec}\,(R)$ such that $M_{\mathfrak{p}} \neq 0$.*
2. *If $N \subset M$ is a submodule and $0 \neq x \in M$, then $x \in N$ if and only if $x \in N_{\mathfrak{p}} \subseteq M_{\mathfrak{p}}$ for each point $\mathfrak{p} \in \mathrm{Spec}\,(R)$.*

Proof. 1. (L $\Rightarrow$ R) Since $\exists x \in M$ which is non-zero, therefore consider the annihilator ideal $\mathrm{Ann}(x) = \{r \in R \mid rx = 0\}$ of R. Then, this ideal is contained in a maximal ideal $\mathfrak{m}$ of R by Zorn's lemma. Hence consider $M_{\mathfrak{m}}$, which contains $x/1$. Now if there exists $r \in R \setminus \mathfrak{m}$ such that $rx = 0$, then $r \in \mathrm{Ann}(x)$, but since $\mathfrak{m} \supseteq \mathrm{Ann}(x)$, hence we have a contradiction.
(R $\Rightarrow$ L) Let $\mathfrak{p} \in \mathrm{Spec}\,(R)$ be such that $x/r \in M_{\mathfrak{p}}$ and $x/r \neq 0$. Since $M_{\mathfrak{p}}$ is an R-module, therefore $r \cdot (x/r)$ is well-defined in $M_{\mathfrak{p}}$. Hence $(rx)/r = x/1 \in M_{\mathfrak{p}}$. If $x/1 = 0$ in $M_{\mathfrak{p}}$, therefore $\varphi_{\mathfrak{p}}(x) = 0$ and hence $x = 0$ as $\varphi_{\mathfrak{p}}$ is injective. Thus, $x/r = 0$ in $M_{\mathfrak{p}}$, a contradiction. Therefore $x/1 \neq 0$ and hence $x \neq 0$ in M.

2. This follows from using 1. on the module $(N+Rx)/N$. We do this by observing the following chain of equivalences, whose key steps are explained below:

$$x \in N \iff N + Rx = N \iff (N + Rx)/N = 0 \iff ((N + Rx)/N)_{\mathfrak{p}} \forall \mathfrak{p} \in \mathrm{Spec}\,(R) \iff$$
$$(N + Rx)_{\mathfrak{p}}/N_{\mathfrak{p}} = 0 \forall \mathfrak{p} \in \mathrm{Spec}\,(R) \iff (N + Rx)_{\mathfrak{p}} = N_{\mathfrak{p}} \forall \mathfrak{p} \in \mathrm{Spec}\,(R) \iff$$
$$N_{\mathfrak{p}} + (Rx)_{\mathfrak{p}} = N_{\mathfrak{p}} \forall \mathfrak{p} \in \mathrm{Spec}\,(R) \iff (Rx)_{\mathfrak{p}} \subseteq N_{\mathfrak{p}} \forall \mathfrak{p} \in \mathrm{Spec}\,(R) \iff \varphi_{\mathfrak{p}}(x) = x/1 \in N_{\mathfrak{p}} \forall \mathfrak{p} \in \mathrm{Spec}\,(R).$$

For two submodules $N, K, L \subset M$ where $L \subseteq N$ and $\mathfrak{p} \in \mathrm{Spec}\,(R)$, we get $(N/L)_{\mathfrak{p}} = N_{\mathfrak{p}}/L_{\mathfrak{p}}$ by exactness of localization (Lemma 16.1.2.2) on the exact sequence

$$0 \to L \to N \to N/L \to 0.$$

Finally $(N + K)_{\mathfrak{p}} = N_{\mathfrak{p}} + K_{\mathfrak{p}}$ in $M_{\mathfrak{p}}$ is true by direct checking and where we use the primality of $\mathfrak{p}$. □

Remark 16.1.2.5. (*Few life hacks*) The above proof tells us few ways how one can approach the problems in ring theory. Note especially that $x \in N$ if and only if $N + Rx = N$, which quickly turns a set-theoretic relation into an algebraic one, where we can now use various constructions as we did, like localization.

The following is the universal property for localization.

Proposition 16.1.2.6. *Let R be a ring and S be a multiplicative set. If $\varphi : R \to T$ is a ring homomorphism such that $\varphi(S) \subseteq T^{\times}$ where $T^{\times}$ is the unit group of T, then there exists a unique map $\tilde{\varphi} : S^{-1}R \to T$ such that the following commutes*

$$\begin{tikzcd} R \arrow[r, "\varphi"] \arrow[d, "i"'] & T \\ S^{-1}R \arrow[ur, dashed, "\tilde{\varphi}"'] & \end{tikzcd}.$$

Proof. Pick any ring map $\varphi : R \to T$. Take any map $f : S^{-1}R \to T$ which makes the above commute. We claim that $f(r/s) = \varphi(r)\varphi(s)^{-1}$. Indeed, we have that $f(r/1) = \varphi(r)$ for all $r \in R$. Further, for any $s \in S$, we have $f(1/s) = 1/f(s/1) = 1/\varphi(s) = \varphi(s)^{-1}$. Consequently, we get for any $r/s \in S^{-1}R$ the following

$$f\left(\frac{r}{s}\right) = f\left(\frac{r}{1}\cdot\frac{1}{s}\right) = f\left(\frac{r}{1}\right)\cdot f\left(\frac{1}{s}\right) = \varphi(r)\varphi(s)^{-1}.$$

This proves uniqueness. Clearly, this is a ring homomorphism. This completes the proof. $\square$

Remark 16.1.2.7. As Proposition 16.1.2.6 is the universal property of localization, therefore the construction $S^{-1}R$ is irrelevant; the property above completely characterizes localization upto a unique isomorphism.

Lemma 16.1.2.8. *Let R be a ring and $f \in R \setminus \{0\}$. Then,*

$$R_f \cong \frac{R[x]}{\langle fx - 1\rangle}.$$

In particular, R_f is a finite type R-algebra.

Proof. We shall use Proposition 16.1.2.6. We need only show that $R[x]/\langle fx-1\rangle$ satisfies the same universal property as stated in Proposition 16.1.2.6. Indeed, we first have the map $i : R \to R[x]/\langle fx-1\rangle$ given by $r \mapsto r + \langle fx - 1\rangle$. Let $\varphi : R \to T$ be any map such that $\varphi(f) \in T^\times$. We claim that there exists a unique map $\tilde{\varphi} : R[x]/\langle fx-1\rangle \to T$ such that $\tilde{\varphi}\circ i = \varphi$. Indeed, take any map $g : R[x]/\langle fx-1\rangle \to T$ such that $g \circ i = \varphi$. Thus, for all $r \in R$, we have $g(r + \langle fx-1\rangle) = \varphi(r)$. As $fx + \langle fx-1\rangle = 1 + \langle fx-1\rangle$, therefore we obtain that $g(f+\langle fx-1\rangle)\cdot g(x+\langle fx-1\rangle) = \varphi(f)\cdot g(x+\langle fx-1\rangle) = 1$. Hence, we see that $g(x+\langle fx-1\rangle) = \varphi(f)^{-1}$. Hence for any element $p(x) + \langle fx-1\rangle$, we see that $f(p(x)+\langle fx-1\rangle) = p(\varphi(f)^{-1})$. This makes g unique well-defined ring homomiorphism. This completes the proof. $\square$

The following is a simple but important application of technique of localization.

Lemma 16.1.2.9. *Let R be a ring. Then the nilradical of R, $\mathfrak{n}$, the ideal consisting of nilpotent elements is equal to the intersection of all prime ideals of R:*

$$\mathfrak{n} = \bigcap_{\mathfrak{p}\in \mathrm{Spec}(R)} \mathfrak{p}.$$

Proof. Take $x \in \bigcap_{\mathfrak{p}\in\mathrm{Spec}(R)}\mathfrak{p}$. We then have $x \in \mathfrak{p}$ for each $\mathfrak{p} \in \mathrm{Spec}\,(R)$. Hence if for each $n \in \mathbb{N}$ we have that $x^n \neq 0$, then we get that $S = \{1, x, x^2, \dots\}$ forms a multiplicative system. Considering the localization $S^{-1}R$, we see that it is non-zero. Therefore $S^{-1}R$ has a prime ideal, which corresponds to a prime ideal $\mathfrak{p}$ of R which does not intersects S, by Lemma 16.1.2.3. But this is a contradiction as x is in every prime ideal.
Conversely, take any $x \in \mathfrak{n}$ and any prime ideal $\mathfrak{p} \in \mathrm{Spec}\,(R)$. Since $x^n = 0$ for some $n \in \mathbb{N}$, therefore $x^n \in \mathfrak{p}$ for each $\mathfrak{p} \in \mathrm{Spec}\,(R)$. Hence it follows from primality of each $\mathfrak{p}$ that $x \in \mathfrak{p}$. $\square$

We next give two results which are of prominent use in algebraic geometry. The first result says that finite generation of a module can be checked locally.

Lemma 16.1.2.10. *Let M be an R-module and suppose $f_i \in R$ are elements such that $\sum_{i=1}^n Rf_i = R$. Then, the following are equivalent:*

1. *M is a finitely generated R-module.*
2. *M_{f_i} is a finitely generated R_{f_i}-module for all $i = 1, \dots, n$.*

Proof. (1. ⇒ 2.) This is simple, as finite generation is preserved under localization.
(2. ⇒ 1.) Let M_{f_i} be generated by $m_{ij}/(f_i)^{n_{ij}}$ for $j = 1, \dots, n_i$. Let $N \leq M$ be a submodule generated by m_{ij} for each $j = 1, \dots, n_i$ and for each $i = 1, \dots, n$. Clearly, N is a finitely generated R-module. Moreover, N_{f_i} for each $i = 1, \dots, n$ is equal to M_{f_i}. We wish to show that $(M/N)_{\mathfrak{p}} = 0$. To this, end, let f_i be such that $f_i \notin \mathfrak{p}$. As $(M/N)_{\mathfrak{p}} = \varinjlim_{f \notin \mathfrak{p}} (M/N)_f$, so it suffices to show that there is a cofinal system of $f \notin \mathfrak{p}$ such that $(M/N)_f = 0$. Indeed, as $(M/N)_{f_i} = M_{f_i}/N_{f_i} = 0$, so we need only show that for any basic open $D(g) \subseteq D(f_i)$, we have $(M/N)_g = 0$. As by Lemma 1.2.1.4, 2 we have that $g^n = rf_i$ for some $r \in R$, therefore we deduce that $(M/N)_g = 0$ as $(M/N)_{f_i} = 0$. It follows that $(M/N)_{\mathfrak{p}} = 0$ for all primes $\mathfrak{p}$ and hence $M/N = 0$ by Lemma 16.1.2.4, 1, hence $M = N$ and M is finitely generated. □

The second result gives a partial analogous result as to Lemma 16.1.2.10 did, but for algebras. This is again an important technical tool used often in algebraic geometry.

Lemma 16.1.2.11. *Let A be a ring and B be an A-algebra. Suppose $f_1, \dots, f_n \in B$ are such that $\sum_{i=1}^n Bf_i = B$. If for all $i = 1, \dots, n$, B_{f_i} is a finitely generated A-algebra, then B is a finitely generated A-algebra.*

Proof. Let B_{f_i} be generated by

$$\left\{ \frac{b_{ij}}{f_i^{n_j}} \right\}_{j=1,\dots,M_i}$$

as an A-algebra, for each $i = 1, \dots, n$. Further, we have $c_1, \dots, c_n \in B$ such that $c_1 f_1 + \dots + c_n f_n = 1$. We claim that $S = \{b_{ij}, f_i, c_i\}_{i,j}$ is a finite generating set for B.

Let C be the sub-algebra of B generated by S. Pick any $b \in B$. We wish to show that $b \in C$. Fix an $i = 1, \dots, n$. Observe that the image of b in the localized ring B_{f_i} is generated by some polynomial with coefficients in A and indeterminates replaced by

$$\left\{ \frac{b_{ij}}{f_i^{n_j}} \right\}_{j=1,\dots,M_i}.$$

We may multiply b by $f_i^{N_i}$ for N_i large enough so that $f_i^{N_i} b$ is then represented by a polynomial with coefficients in A evaluated in f_i and b_{ij} for $j = 1, \dots, M_i$. Consequently, $f_i^{N_i} b \in C$, for each $i = 1, \dots, n$. Observe that $f_1, \dots, f_n$ in C generates the unit ideal in C. By Lemma 16.23.0.2, 2, we see that $f_1^{N_1}, \dots, f_n^{N_n}$ also generates the unit ideal in C. Hence, we have $d_1, \dots, d_n \in C$ such that $1 = d_1 f_1^{N_1} + \dots + d_n f_n^{N_n}$. Multiplying by b, we obtain $b = d_1 f_1^{N_1} b + \dots + d_n f_n^{N_n}$ where by above, we now know that each term is in C. This completes the proof. □

An observation which is of importance in the study of varieties is the following.

Lemma 16.1.2.12. *Let R be an integral domain. Then*

$$\bigcap_{\mathfrak{m}<R} R_{\mathfrak{m}} = R$$

where the intersection runs over all maximal ideals $\mathfrak{m}$ of R and the intersection is carried out in the fraction field $R_{\langle 0\rangle}$.

Proof. We already have that

$$R \hookrightarrow R_{\mathfrak{m}}$$

for any maximal ideal $\mathfrak{m} < R$. Thus,

$$R \hookrightarrow \bigcap_{\mathfrak{m}<R} R_{\mathfrak{m}}.$$

Thus it would suffice to show that $\bigcap_{\mathfrak{m}<R} R_{\mathfrak{m}} \hookrightarrow R$. Indeed, consider the following map

$$\begin{aligned} \bigcap_{\mathfrak{m}<R} R_{\mathfrak{m}} &\longrightarrow R \\ [f_{\mathfrak{m}}/g_{\mathfrak{m}}] &\longmapsto f_{\mathfrak{m}}g_{\mathfrak{m}'} \end{aligned}$$

where $f_{\mathfrak{m}}/g_{\mathfrak{m}} = f_{\mathfrak{m}'}/g_{\mathfrak{m}'}$ for two maximal ideals $\mathfrak{m}, \mathfrak{m}'$ in R. Thus, $f_{\mathfrak{m}}g_{\mathfrak{m}'} = f_{\mathfrak{m}'}g_{\mathfrak{m}}$. Hence the above map is well-defined and is injective as $f_{\mathfrak{m}}g_{\mathfrak{m}'} = 0$ implies $f_{\mathfrak{m}} = 0$ as $g_{\mathfrak{m}'} \neq 0$. The result follows. $\square$

One may wonder when localization and Hom commutes. It does when one of the modules is finitely presented.

Proposition 16.1.2.13. *Let M, N be R-modules where M is finitely presented and $S \subseteq R$ be a multiplicative set. Then,*

$$S^{-1}(\mathrm{Hom}_R(M,N)) \cong \mathrm{Hom}_{S^{-1}R}\left(S^{-1}M, S^{-1}N\right).$$

Proof. Consider the map

$$\begin{aligned} \theta_M : S^{-1}(\mathrm{Hom}_R(M,N)) &\longrightarrow \mathrm{Hom}_{S^{-1}R}\left(S^{-1}M, S^{-1}N\right) \\ \frac{\varphi}{s} &\longmapsto \frac{m}{t} \mapsto \frac{\varphi(m)}{st}. \end{aligned}$$

We claim that θ_M is an isomorphism. To this end, first observe that if M is free, then it is immediate from standard Hom identities. Now consider a finite presentation $R^m \to R^n \to M \to 0$ of M. Localizing at S we get a finite presentation $(S^{-1}R)^m \to (S^{-1}R)^n \to S^{-1}M \to 0$ of $S^{-1}M$ as an $S^{-1}R$-module. As Hom is left exact and localization is exact, then we get the following commutative diagram where rows are exact:

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Hom}_{S^{-1}R}\left(S^{-1}M, S^{-1}N\right) \arrow[r] & \mathrm{Hom}_{S^{-1}R}\left((S^{-1}R)^n, S^{-1}N\right) \arrow[r] & \mathrm{Hom}_{S^{-1}R}\left((S^{-1}R)^m, S^{-1}N\right) \\ 0 \arrow[r] & S^{-1}\mathrm{Hom}_R(M,N) \arrow[r] \arrow[u, "\theta_M"] & S^{-1}\mathrm{Hom}_R(R^n,N) \arrow[r] \arrow[u, "\theta_{R^n}", "\cong"'] & S^{-1}\mathrm{Hom}_R(R^m,N) \arrow[u, "\theta_{R^m}", "\cong"'] \end{tikzcd}.$$

By five-lemma, θ_M is an isomorphism, as required. $\square$

Local rings

A ring R is said to be *local* if there is a unique maximal ideal of R. In such a case we denote it by $(R, \mathfrak{m})$. We first study tangent and cotangent space to certain type of regular local rings, which are important in the study of rational points.

Definition 16.1.2.14. (**Zariski (co)tangent space**) Let $(R, \mathfrak{m})$ be a local ring. Then, we define the Zariski *cotangent space* of $(R, \mathfrak{m})$ to be $T^*R = \mathfrak{m}/\mathfrak{m}^2$ and the Zariski *tangent space* to be its dual $TR = \mathrm{Hom}_k\left(\mathfrak{m}/\mathfrak{m}^2, k\right)$.

Remark 16.1.2.15. The Zariski cotangent space T^*R is a κ-vector space where $\kappa = R/\mathfrak{m}$ is the residue field. Indeed, the scalar multiplication is given by

$$\begin{aligned} \kappa \times T^*R &\longrightarrow T^*R \\ (c + \mathfrak{m}, x + \mathfrak{m}^2) &\longmapsto cx + \mathfrak{m}^2 \end{aligned}$$

where $c \in R$ and $x \in \mathfrak{m}$. Indeed, this is well-defined as can be seen by a simple check. Consequently, the tangent space $TR = \mathrm{Hom}_k\left(\mathfrak{m}/\mathfrak{m}^2, k\right)$ is also a κ-vector space.

Definition 16.1.2.16. (**Regular local ring**) Let $(A, \mathfrak{m})$ be a local ring with $k = A/\mathfrak{m}$ being the residue field. Then A is said to be regular if $\dim_k \mathfrak{m}/\mathfrak{m}^2 = \dim A$.

There is an important geometric lemma that one should keep in mind about certain local rings.

Definition 16.1.2.17. (**Rational local k-algebras**) Let k be a field. A local k-algebra $(R, \mathfrak{m})$ is said to be rational if its residue field $\kappa = R/\mathfrak{m}$ is isomorphic to the field k.

Rational local k-algebras have a rather simple tangent space.

Proposition 16.1.2.18. *Let $(A, \mathfrak{m}_A)$ be a rational local k-algebra. Then,*

$$TA \cong \mathrm{Hom}_{k,\mathrm{loc}}\left(A, k[\epsilon]\right)$$

where $k[\epsilon] := k[x]/x^2$ is the ring of dual numbers and $\mathrm{Hom}_{k,\mathrm{loc}}\left(A, k[\epsilon]\right)$ denotes the set of all local k-algebra homomorphisms.

Proof. Pick any k-algebra homomorphism $\varphi : A \to k[\epsilon]$. Denote by $\mathfrak{m}_\epsilon = \langle \epsilon \rangle \lneq k[\epsilon]$ the unique maximal ideal of $k[\epsilon]$. Since

$$k[\epsilon]/\mathfrak{m}_\epsilon \cong k,$$

therefore $k[\epsilon]$ is a rational local k-algebra as well. By Lemma 16.23.0.7, we may write $A = k \oplus \mathfrak{m}_A$ and $k[\epsilon] = k \oplus \mathfrak{m}_\epsilon$. We now claim that the datum of a local k-algebra homomorphism $\varphi : A \to k[\epsilon]$ is equivalent to datum of a k-linear map of k-modules $\theta : \mathfrak{m}_A/\mathfrak{m}_A^2 \to k$.

Indeed, we first observe that for any $\varphi : A \to k[\epsilon]$ as above, we have $\varphi(\mathfrak{m}_A) \subseteq \mathfrak{m}_\epsilon$. Thus, $\varphi(\mathfrak{m}_A^2) \subseteq \mathfrak{m}_\epsilon^2 = 0$. Thus, we deduce that for any such φ, $\mathrm{Ker}\left(\varphi\right) \supseteq \mathfrak{m}_A^2$. It follows from universal property of quotients that any such φ is in one-to-one correspondence with k-algebra homomorphisms

$$\tilde{\varphi} : A/\mathfrak{m}_A^2 \cong k \oplus (\mathfrak{m}_A/\mathfrak{m}_A^2) \longrightarrow k[\epsilon].$$

As $\varphi(\mathfrak{m}_A) \subseteq \mathfrak{m}_\epsilon$, therefore $\tilde{\varphi}(\mathfrak{m}_A/\mathfrak{m}_A^2) \subseteq \mathfrak{m}_\epsilon$. Thus, we obtain a k-linear map of k-modules

$$\theta : \mathfrak{m}_A/\mathfrak{m}_A^2 \longrightarrow k \cong \mathfrak{m}_\epsilon$$

where $\mathfrak{m}_\epsilon \cong k$ as k-modules. It suffices to now show that from any such θ, one can obtain a unique k-algebra map $\tilde{\varphi} : k \oplus (\mathfrak{m}_A/\mathfrak{m}_A^2) \to k[\epsilon]$, which furthermore sets up a bijection between all such $\tilde{\varphi}$ and θ.

Indeed, from k-linear map θ, we may construct the following k-algebra map

$$\begin{aligned} \tilde{\varphi} : k \oplus (\mathfrak{m}_A/\mathfrak{m}_A^2) &\longrightarrow k[\epsilon] \\ (k + \bar{m}) &\longmapsto k + \theta(\bar{m})\epsilon. \end{aligned}$$

Then we observe that $\tilde{\varphi}$ is a k-algebra homomorphism as

$$\begin{aligned} \tilde{\varphi}((k_1 + \bar{m}_1)(k_2 + \bar{m}_2)) &= \tilde{\varphi}(k_1k_2 + k_1\bar{m}_2 + k_2\bar{m}_1 + \bar{m}_1\bar{m}_2) \\ &= k_1k_2 + k_1\theta(\bar{m}_2)\epsilon + k_2\theta(\bar{m}_1)\epsilon + \theta(\bar{m}_1\bar{m}_2)\epsilon \\ &= k_1k_2 + k_1\theta(\bar{m}_2)\epsilon + k_2\theta(\bar{m}_1)\epsilon \\ &= (k_1 + \theta(\bar{m}_1)\epsilon) \cdot (k_2 + \theta(\bar{m}_2)\epsilon) \\ &= \tilde{\varphi}(k_1 + \bar{m}_1) \cdot \tilde{\varphi}(k_2 + \bar{m}_2). \end{aligned}$$

Hence, from θ one obtain $\tilde{\varphi}$ back, thus setting up a bijection and completing the proof. □

In general, restriction and then extension of scalars wont yield the same module back. The following gives a criterion when this happens.

Lemma 16.1.2.19. *Let $\varphi : A \to B$ be an A-algebra such that $B \otimes_A B \cong B$. If M is a B-module and M_A is the A-module by restriction, then*

$$M_A \otimes_A B \cong M.$$

Proof. Immediate since

$$M_A \otimes_A B \cong (M \otimes_B B)_A \otimes_A B \cong M \otimes_B (B \otimes_A B) \cong M \otimes_B B \cong M.$$

□

16.1.3 Structure theorem

Let M be a finitely generated R-module. We can understand the structure of such modules completely in terms of the ring R, when R is a PID (so that it's UFD). This is the content of the structure theorem. We first give the following few propositions which is used in the proof of the structure theorem but is of independent interest as well, in order to derive a usable variant of structure theorem. The following theorem tells us a direct sum decomposition exists for any finitely free torsion module over a PID.

Proposition 16.1.3.1. *Let M be a finitely generated torsion module over a PID R. If $\operatorname{Ann}(M) = \langle c \rangle$ where $c = p_1^{k_1} \dots p_r^{k_r}$ and $p_i \in R$ are prime elements, then*

$$M \cong M_1 \oplus \cdots \oplus M_r$$

where $M_i = \{x \in M \mid p_i^{r_i} x = 0\} \leq M$ for all $i = 1, \dots, r$, that is, where $\operatorname{Ann}(M_i) = \langle p_i^{r_i} \rangle$ for all $i = 1, \dots, r$.

The next result tells us that we can further write each of the above M_is as a direct sum decomposition of a special kind.

Proposition 16.1.3.2. *Let M be a finitely generated torsion module over a PID R. If $\mathrm{Ann}M = \langle p^r\rangle$ where $p \in R$ is a prime element, then there exists $r_1 \geq r_2 \geq \cdots \geq r_k \geq 1$ such that*

$$M \cong R/\langle p^{r_1}\rangle \oplus \cdots \oplus R/\langle p^{r_k}\rangle.$$

The structure theorem is as follows.

Theorem 16.1.3.3. *(Structure theorem) Let R be a PID and M be a finitely generated R-module. Then there exists an unique $n \in \mathbb{N}\cup\{0\}$ and $q_1, \dots q_r \in R$ unique upto units such that $q_{i-1}|q_i$ for all $i = 2, \dots, r$ and*

$$M \cong R^n \oplus R/\langle q_1\rangle \oplus \cdots \oplus R/\langle q_r\rangle.$$

The most useful version of this is the following:

Corollary 16.1.3.4. *Let M be a finitely generated torsion module over a PID R. Then, there exists k-many prime elements $p_1, \dots, p_k \in R$, $n_j \in \mathbb{N}$ for each $j = 1, \dots, k$ and $1 \leq r_{1j} \leq \cdots \leq r_{n_j j} \in \mathbb{N}$ for each $j = 1, \dots, k$ such that*

$$M \cong \bigoplus_{j=1}^{k} \left(R/\langle p_j^{r_{1j}}\rangle \oplus \cdots \oplus R/\langle p_j^{n_j j}\rangle\right).$$

Proof. This is a consequence of Propositions 16.1.3.1 and 16.1.3.2. □

This is the famous structure theorem for finitely generated modules over a PID. Note that the ring $\mathbb{Z}$ is PID and any abelian group is a $\mathbb{Z}$-module. Thus, we can classify finitely generated abelian groups using the structure theorem.

Example 16.1.3.5. An example of a module which is not finitely generated is the polynomial module $R[x]$ over a ring R. Indeed, the collection $\{1, x, x^2, \dots\}$ will make it free but not finitely generated.

Example 16.1.3.6. Classification of all abelian groups of order $360 = 2^3 \cdot 3^2 \cdot 5$, for example, can be achieved via structure theorem. Indeed using Corollary 16.1.3.4, we will get that there are 6 total such abelian groups given by

- $\left(\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3\mathbb{Z}} \oplus \frac{\mathbb{Z}}{3\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$
- $\left(\frac{\mathbb{Z}}{2^2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3\mathbb{Z}} \oplus \frac{\mathbb{Z}}{3\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$
- $\left(\frac{\mathbb{Z}}{2^3\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3\mathbb{Z}} \oplus \frac{\mathbb{Z}}{3\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$
- $\left(\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3^2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$
- $\left(\frac{\mathbb{Z}}{2^2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3^2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$
- $\left(\frac{\mathbb{Z}}{2^3\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{3^2\mathbb{Z}}\right) \oplus \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)$

16.1.4 UFDs

16.1.5 Gauss' lemma

Add results surrounding primitive polynomials and Gauss' lemma here from notebook.

Spectra of polynomial rings over UFDs

We now calculate the prime spectra of polynomial rings over UFDs. For that, we need the following two lemmas.

Lemma 16.1.5.1. *Let R be a UFD and $I \leq R[x]$ be an ideal containing two elements with no common factors. Then I contains a non-zero constant from R.*

Proof. Indeed, let $f, g \in R[x]$ be two elements with no common factors. Let Q denote the fraction field of R. We first claim that $f, g \in Q[x]$ have no common factor as well. Indeed, suppose $h(x) \in Q[x]$ is a common factor of $f(x)$ and $g(x)$. It follows from the result on primitive polynomials that we can write $h(x) = ch_0(x)$ where $c \in Q$ and $h_0(x) \in R[x]$ is primitive. Hence, we see that $h_0(x) \in R[x]$ is a polynomial such that $h_0|f$ and $h_0|g$ in $Q[x]$. Again, by general results in UFD, we then conclude that $h_0|f$ and $h_0|g$ in $R[x]$. As f and g have no common factor, therefore $h_0(x) \in R[x]$ is a unit. Hence $h(x) \in Q[x]$ is a unit. Thus, there is no common factor of $f(x)$ and $g(x)$ in $Q[x]$ if there is none in $R[x]$.

Hence, $f(x), g(x)$ in $Q[x]$ have gcd 1, where $Q[x]$ is a PID. Consequently, $f(x)$ and $g(x)$ generates the unit ideal in $Q[x]$. It follows that there exists $p(x), q(x) \in Q[x]$ such that

$$1 = p(x)f(x) + q(x)g(x).$$

By theorem on primitive polynomials, we may write $p(x) = \frac{a}{b}p_0(x)$ and $q(x) = \frac{c}{d}q_0(x)$ where $a/b, c/d \in Q$ and $p_0(x), q_0(x) \in R[x]$ are primitive. The above equation hence becomes

$$\begin{aligned} 1 &= \frac{a}{b}p_0(x)f(x) + \frac{c}{d}q_0(x)g(x) \\ &= \frac{adp_0(x)f(x) + bcq_0(x)g(x)}{bd}, \end{aligned}$$

which thus yields

$$bd = adp_0(x)f(x) + bcq_0(x)g(x)$$

where RHS is in $I \leq R[x]$ because $ad, p_0, bc, q_0 \in R[x]$ and $f, g \in I$ and LHS is in R. Hence $I \cap R$ is not zero. □

Lemma 16.1.5.2. *Let R be a PID and $f, g \in R[x]$ be non-zero polynomials such that f and g have no common factors. Then,*

1. *any prime ideal $\mathfrak{p} \lneq R[x]$ containing f and g is maximal,*
2. *any maximal ideal $\mathfrak{m} \lneq R[x]$ containing f and g is of the form $\langle p, h(x) \rangle$ where $p \in R$ is prime and $h(x)$ is prime modulo p,*
3. *there are only finitely many maximal ideals of $R[x]$ containing f and g.*

Proof. 1. : Let $\mathfrak{p} \lneq R[x]$ be a prime ideal containing f and g. Observe by Lemma 16.1.5.1 that there exists $b \in R \setminus 0$ such that $b \in \mathfrak{p} \cap R$, that is, $\mathfrak{p} \cap R \neq 0$. As R is a PID and $\mathfrak{p} \cap R$ is a prime ideal of R, therefore $\mathfrak{p} \cap R = pR$ for some prime element $p \in \mathfrak{p} \cap R$. We wish to show that $R[x]/\mathfrak{p}$ is a field.

Indeed, we see that (note $\langle p, \mathfrak{p}\rangle = \mathfrak{p}$ as $p \in \mathfrak{p}$)

$$\begin{aligned}\frac{R[x]}{\mathfrak{p}} &\cong \frac{\frac{R[x]}{pR[x]}}{\frac{\langle p,\mathfrak{p}\rangle}{pR[x]}} = \frac{\frac{R[x]}{pR[x]}}{\frac{\mathfrak{p}}{pR[x]}}\\ &\cong \frac{\frac{R}{pR}[x]}{\overline{\mathfrak{p}}}\end{aligned}$$

where $\overline{\mathfrak{p}} = \pi(\mathfrak{p})$ where $\pi : R[x] \twoheadrightarrow \frac{R}{pR}[x]$ is the quotient map. As R is a PID and pR is a non-zero prime ideal, therefore it is maximal. Consequently, R/pR is a field and hence $\frac{R}{pR}[x]$ is a PID. Suppose $\overline{\mathfrak{p}} = 0$, then f and g have a common factor given by $p \in R$, which is not possible. Consequently, $\overline{\mathfrak{p}}$ is a proper prime ideal of $\frac{R}{pR}[x]$ by correspondence theorem. But in PIDs, non-zero prime ideals are maximal ideals, hence we obtain that $\frac{R}{pR}[x]/\overline{\mathfrak{p}}$ is a field, as required.

2. : Let $\mathfrak{m} \lneq R[x]$ be a maximal ideal of $R[x]$ containing f and g. Hence, from Lemma 16.1.5.1 and R being a PID, there exists $p \in R$ a prime such that $\mathfrak{m} \cap R = pR$. Hence R/pR is a field as R is a PID and pR a non-zero prime ideal (so maximal). Consequently, we have a quotient map

$$\pi : R[x] \twoheadrightarrow \frac{R[x]}{pR[x]} \cong \frac{R}{pR}[x]$$

As $p \in \mathfrak{m}$, therefore by correspondence thereom $\pi(\mathfrak{m}) = \overline{\mathfrak{m}}$ is a maximal ideal of $\frac{R}{pR}[x]$. As R/pR is a field, therefore $\frac{R}{pR}[x]$ is a PID. Hence, $\overline{\mathfrak{m}} = \langle \overline{h(x)}\rangle$ for some $h(x) \in R[x]$ such that $\overline{h(x)}$ is irreducible (so it generates a maximal ideal). Again, by correspondence theorem we have $\pi^{-1}(\overline{\mathfrak{m}}) = \mathfrak{m} = h(x)R[x] + pR[x] = \langle p, h(x)\rangle$, as required.

3. : We will use notations of proof of 2. above. Take any maximal ideal $\mathfrak{m} = \langle p, h(x)\rangle \lneq R[x]$ which contains $f(x)$ and $g(x)$, $p \in R$ is prime and $h(x)$ is irreducible modulo p. As R is a PID, so it is a UFD, hence $R[x]$ is a UFD by Gauss' lemma. Hence, writing $f(x)$ and $g(x)$ as product of prime factors in $R[x]$, we observe that there exists distinct primes $p(x), q(x) \in R[x]$ such that $p(x), q(x) \in \mathfrak{m}$. Replacing f by p and g by q, we may assume f and g are irreducible (or prime) in $R[x]$.

By Lemma 16.1.5.1, there exists $b \in R \setminus 0$ such that $b \in \mathfrak{m} \cap R$. As the proof of 2. above shows, $p|b$ in R. As R is a PID, so it is a UFD, hence there are only finitely many choices for p.

Now, going modulo prime p, we see that $\overline{f(x)}, \overline{g(x)} \in \overline{\mathfrak{m}} \lneq \frac{R}{pR}[x]$ has a common factor in $\frac{R}{pR}[x]$, given by $\overline{h(x)}$ as $\overline{\mathfrak{m}} = \langle \overline{h(x)}\rangle$ (by proof of 2.). As $\overline{h(x)}$ generates a maximal ideal in $\frac{R}{pR}[x]$, therefore $\overline{h(x)}$ is a prime element of $\frac{R}{pR}[x]$, which has to divide $\overline{f(x)}$ and $\overline{g(x)}$. As $\frac{R}{pR}[x]$ is a PID, therefore there are only finitely many choices for $\overline{h(x)}$, and since $\mathfrak{m} = \pi^{-1}(\langle \overline{h(x)}\rangle)$, therefore every choice of p as above, yields finitely many choices for $\mathfrak{m}$.

Consequently, there are finitely many choices for p and once p is fixed, there are only finitely many choices for the ideal $\overline{\mathfrak{m}}$. As $\mathfrak{m} = \pi^{-1}(\overline{m})$, therefore there are finitely many maximal ideals containing f and g. $\square$

We now classify Spec $(R[x])$ for a UFD R.

Theorem 16.1.5.3. *Let R be a PID. Any prime ideal $\mathfrak{p} \lneq R[x]$ is of one of the following forms*

1. $\mathfrak{p} = \mathfrak{o}$,

2. $\mathfrak{p} = \langle f(x) \rangle$ *for some irreducible* $f(x) \in R[x]$,
3. $\mathfrak{p} = \langle p, h(x) \rangle$ *for some prime* $p \in R$ *and* $h(x) \in R[x]$ *irreducible modulo* p *and this is also a maximal ideal.*

Proof. Indeed, pick any prime ideal $\mathfrak{p} \lneq R[x]$. If $\mathfrak{p}$ is 0, then it is prime as $R[x]$ is a domain. We now have two cases. If $\mathfrak{p}$ is principal, then $\mathfrak{p} = \langle f(x) \rangle$ for some $f(x) \in R[x]$. As $\langle f(x) \rangle$ is prime therefore $f(x)$ is a prime element. As $R[x]$ is a UFD by Gauss' lemma, therefore $f(x)$ is also irreducible. Consequently, $\mathfrak{p} = \langle f(x) \rangle$ where $f(x)$ is irreducible.

On the other hand if $\mathfrak{p}$ is not principal, there exists $f(x), g(x) \in \mathfrak{p}$ such that $f(x) \not| g(x)$ and $g(x) \not| f(x)$. As $R[x]$ is a UFD and $\mathfrak{p}$ is prime, therefore there exists prime factors of f and g which are in $\mathfrak{p}$. Replacing f and g by these prime factors, we may assume f and g are distinct irreducibles in $\mathfrak{p}$. Consequently, by Lemma 16.1.5.2, we see that $\mathfrak{p} = \langle p, h(x) \rangle$ for some prime $p \in R$ and $h(x)$ irreducible modulo p. Moreover by Lemma 16.1.5.2 we know that $\mathfrak{p}$ in this case is maximal. □

We now portray their use in the following.

Lemma 16.1.5.4. *Let F be an algebraically closed field. Then,*
1. *every non-constant polynomial* $f(x, y) \in F[x, y]$ *has at least one zero in* F^2,
2. *every maximal ideal of* $F[x, y]$ *is of the form* $\mathfrak{m} = \langle x - a, y - b \rangle$ *for some* $a, b \in F$.

Proof. 1. : Take any polynomial $f(x, y) \in F[x, y]$. Going modulo y, we see that $\overline{f(x, y)} \in F[x, y]/\langle y \rangle = F[x]$. If $\overline{f(x, y)} = 0$, then $(a, 0)$ is a root of $f(x, y)$ for any $a \in F$. if $\overline{f(x, y)} \neq 0$, then since F is algebraically closed, therefore we may write $\overline{f(x, y)} = (x - a_1) \dots (x - a_n)$. Consequently, any $(a_i, 0)$ is a zero of $f(x, y)$. Hence, in any case, $f(x, y)$ has a root in F^2. 2. : Let $R = F[x]$. We know that R is a PID. Take any maximal ideal $\mathfrak{m} \lneq R[y] = F[x, y]$. Then by Theorem 16.1.5.3, we have that either $\mathfrak{m} = \langle f(x, y) \rangle$ where $f(x, y)$ is irreducible or $\mathfrak{m} = \langle p(x), h(x, y) \rangle$ where $p(x) \in R$ is prime and $h(x, y)$ is irreducible modulo $p(x)$.

In the former, we claim that $\langle f(x, y) \rangle$ is not maximal. Indeed, by item 1, we have that $f(x, y)$ has a zero in F^2, say (a, b). Dividing $f(x, y)$ by $y - b$ in $R[y]$, we obtain $f(x, y) = h(x, y)(y - b) + k(x)$, where $k(x) \in R$. Consequently, $k(a) = 0$. Hence, $k(x) = (x - a)l(x)$. Thus, we have $f(x, y) = h(x, y)(y - b) + (x - a)l(x)$, showing $f(x, y) \in \langle x - a, y - b \rangle$. By Theorem 16.1.5.3 above, we know that $\langle x - a, y - b \rangle \in R[y]$ is a maximal ideal and we also know that it contains $f(x, y)$. We hence need only show that $\langle f(x, y) \rangle \subsetneq \langle x - a, y - b \rangle$. Indeed, observe that $x - a \notin \langle f(x, y) \rangle$ as if it is, then $f(x, y) | x - a$. But then $f(x, y)$ is in R, hence $y - b \notin \langle f(x, y) \rangle$. So in either case, $\langle f(x, y) \rangle$ is properly contained in $\langle x - a, y - b \rangle$, showing that $\langle f(x, y) \rangle$ cannot be maximal. Thus, no maximal ideal of $R[y]$ can be of the form $\langle f(x, y) \rangle$.

In the latter, where $\mathfrak{m} = \langle p(x), h(x, y) \rangle$ where $p(x) \in R$ is prime and $h(x, y)$ is irreducible modulo $p(x)$, we first see that $p(x) = x - a$ for some $a \in F$ as $R = F[x]$ and only primes of $F[x]$ are of this type. Let $\pi : R \twoheadrightarrow \frac{R}{p(x)R}[y] \cong \frac{R}{\langle x - a \rangle}[y] \cong F[y]$ be the quotient map by the ideal $p(x)R[y]$. Then we see that by correspondence theorem, $\pi(\mathfrak{m}) = \overline{\mathfrak{m}} = \langle \overline{h(x, y)} \rangle$ is a prime ideal of $F[y]$. Hence, $\overline{\mathfrak{m}} = \langle k(y) \rangle$ for some $k(y) \in F[y]$. Further, since $\overline{\mathfrak{m}}$ is prime and F algebraically closed, therefore $k(y) = y - b$. Thus, we see that modulo $p(x)$ we have $h(x, y) = k(y) = y - b$. We then see that $\mathfrak{m} = \pi^{-1}(\overline{\mathfrak{m}}) = \pi^{-1}(\langle \overline{y - b} \rangle) = \langle p(x), y - b \rangle = \langle x - a, y - b \rangle$, as required. □

Another example gives us finiteness of intersection of two algebraic curves over an algebraically closed field.

Proposition 16.1.5.5. *Let F be an algebraically closed field and $f, g \in F[x, y]$ be two polynomials with no common factors. Then, $Z(f) \cap Z(g)$ is a finite set, that is, f and g intersects at finitely many points in $\mathbb{A}^2_F$.*

Proof. We first show that for any $h(x, y) \in F[x, y]$, $h(a, b) = 0$ for some $(a, b) \in F^2$ if and only if $h \in \langle x - a, y - b\rangle$. Clearly, ($\Leftarrow$) is immediate. For ($\Rightarrow$), we proceed as follows. Going modulo $y - b$ in $F[x, y]$, we obtain $\overline{h(x, y)} \in F[x, y]/\langle y - b\rangle \cong F[x]$. Observe that $\langle y - b\rangle$ is the kernel of the map $F[x, y] \to F[x]$ taking $y \mapsto b$, hence $\overline{h(x, y)} = \overline{h(x, b)}$. As F is algebraically closed, therefore we may write

$$\overline{h(x, b)} = \overline{h(x, y)} = (x - c_1) \dots (x - c_n)$$

for $c_i \in F$. As, $h(a, b) = 0$, therefore $(x - a)|\overline{h(x, b)}$. Hence, for some i, we must have $c_i = a$. This allows us to write

$$h(x, y) - (x - a)k(x) \in \langle y - b\rangle$$

for some $k(x) \in F[x]$. It follows that for some $q(x, y) \in F[x, y]$ we have

$$h(x, y) - (x - a)k(x) = (y - b)q(x, y)$$

Thus, $h(x, y) \in \langle x - a, y - b\rangle$. This completes the proof of the claim above.

Now, using above claim $f(a, b) = 0 = g(a, b)$ if and only if $f, g \in \langle x - a, y - b\rangle$. By Lemma 16.1.5.2, as f and g have no common factors, therefore there are finitely many maximal ideals containing f and g. Further, by Lemma 16.1.5.4, we know that each such maximal ideal is of the form $\langle x - a, y - b\rangle$. Hence, there are only finitely many maximal ideals containing f and g, each of which looks like $\langle x - a, y - b\rangle$. Hence, by above claim, there are finitely many points $(a, b) \in F^2$ such that $f(a, b) = 0 = g(a, b)$. □

16.1.6 Finite type k-algebras

We discuss basic theory of finite type k-algebras, that is, algebras of form $k[x_1, \dots, x_n]/I$.

Recall that for a field k, we denote by $k[x]$ the polynomial ring in one variable and we denote the rational function field $k(x)$ to be the field obtained by localizing at prime $\mathfrak{o}$. Further if K/k is a field extension and $\alpha \in K$, then $k[\alpha]$ is a subring of K generated by $\alpha \in K$ and it contains k. Whereas, $k(\alpha)$ is a field extension $k \hookrightarrow k(\alpha) \hookrightarrow K$. The following lemma shows that if K is algebraic, then $k(\alpha) = k[\alpha]$.

Lemma 16.1.6.1. *Let k be a field and K/k be an algebraic extension. If $\alpha_1, \dots, \alpha_n \in K$, then $k[\alpha_1, \dots, \alpha_n] = k(\alpha_1, \dots, \alpha_n)$.*

Proof. The proof uses a standard observation in field theory. First, let $f_1(x) \in k[x]$ be the minimal polynomial of α_1. Consequently, by a standard result in field theory, $k[\alpha_1] = k[x]/f_1(x)$ is a field. Thus $k[\alpha_1] = k(\alpha_1)$. Now observe that $K/k(\alpha_1)$ is an algebraic extension. Consequently, the same argument will yield $k(\alpha_1)[\alpha_2]$ to be a field. By above, we thus obtain $k(\alpha_1)[\alpha_2] = k[\alpha_1][\alpha_2] = k[\alpha_1, \alpha_2]$ to be a field. Consequently, $k[\alpha_1, \alpha_2] = k(\alpha_1, \alpha_2)$. One completes the proof now by induction. □

Lemma 16.1.6.2. *Let k be a field and K/k be an algebraic extension. Then the homomorphism*

$$\begin{aligned} k[x_1, \dots, x_n] &\longrightarrow k(\alpha_1, \dots, \alpha_n) \\ x_i &\longmapsto \alpha_i \end{aligned}$$

has kernel which is a maximal ideal generated by n elements.

Proof. (*Sketch*) Use the proof of Lemma 16.1.6.1 to obtain that for each $1 \leq i \leq n$, we have that $k(\alpha_1, \dots, \alpha_{i-1})[\alpha_i] \cong k(\alpha_1, \dots, \alpha_{i-1})[x_i]/p_i(\alpha_1, \dots, \alpha_{i-1}, x_i)$ and divide an element $p \in k[x_1, \dots, x_n]$ in the kernel inductively by p_i and replacing p_i by remainder, starting at $i = n$. □

16.1.7 Primary decomposition

This is a basic topic which allows us to talk about irreducible components of schemes. The main motivating analogy here is the following: What prime ideals are to prime numbers is what primary ideals are to prime powers.

Definition 16.1.7.1 (Primary ideals). An ideal $\mathfrak{q} \leq R$ is primary if $R/\mathfrak{q} \neq 0$ and every zero-divisor is a nilpotent.

Example 16.1.7.2. Examples are $\langle p^n \rangle \leq \mathbb{Z}$ for primes p, $\langle x^n, y \rangle \leq k[x, y]$. Indeed, $\mathbb{Z}/\langle p^n \rangle$ has zero-divisors exactly multiples of p, which are nilpotent. Similarly, as $k[x, y]/\langle x^n, y \rangle \cong k[x]/\langle x^n \rangle$ has zero-divisors exactly multiples of x, which are again nilpotent.

We will now directly state the main results of primary decomposition. For details, see cite[AMD].

Proposition 16.1.7.3. *Let R be a ring.*
1. *If $\mathfrak{q} \leq R$ is primary, then $\sqrt{\mathfrak{q}}$ is a prime.*
2. *If for some $\mathfrak{a} \leq R$, $\sqrt{\mathfrak{a}}$ is maximal, then $\mathfrak{q}$ is primary.*

Due to above result, its beneficial to introduce the following terminology.

Definition 16.1.7.4 ($\mathfrak{p}$-primary ideals). Let $\mathfrak{p} \leq R$ be a prime ideal. A primary ideal $\mathfrak{q} \leq R$ is said to be $\mathfrak{p}$-primary if $\sqrt{\mathfrak{q}} = \mathfrak{p}$.

Definition 16.1.7.5 (Primary decomposition). Let R be a ring and $\mathfrak{a} \leq R$ be an ideal. We call $\mathfrak{a}$ *decomposable* if $\mathfrak{a} = \bigcap_{i=1}^n \mathfrak{q}_i$ where each $\mathfrak{q}_i$ is a $\mathfrak{p}_i$-primary ideal for primes $\mathfrak{p}_i$. This is called a *minimal primary decomposition* if each $\mathfrak{p}_i$ is distinct and $\mathfrak{q}_i \not\supseteq \bigcap_{j \neq i} \mathfrak{q}_j$ for each i. Each of the $\mathfrak{p}_i$ is called *prime belonging to* $\mathfrak{a}$. The minimal primes amongst $\{\mathfrak{p}_1, \dots, \mathfrak{p}_n\}$ are called *isolated primes belonging to* $\mathfrak{a}$. The remaining are called *embedded primes of* $\mathfrak{a}$. Given a primary decomposition of $\mathfrak{a} = \bigcap_{i=1}^n \mathfrak{q}_i$, the primary ideals corresponding to isolated primes belonging to $\mathfrak{a}$ are called *isolated primary components* and primary ideals corresponding to embedded primes belonging to $\mathfrak{a}$ are called *embedded primary components*.

Lemma 16.1.7.6. *Every decomposable ideal admits a minimal primary decomposition.*

There are three important results about primary decomposition. The first says that decomposable ideals have unique isolated primes. Second says that any prime containing a decomposable ideal $\mathfrak{a}$ contains a minimal prime belonging to $\mathfrak{a}$. Finally, the isolated primary components are unique. Here are the formal statements.

Theorem 16.1.7.7 (Weak uniqueness of minimal primary decomposition)**.** *Let R be a ring, $\mathfrak{a} \leq R$ be a decomposable ideal and $\mathfrak{a} = \bigcap_{i=1}^{n} \mathfrak{q}_i$ be a minimal primary decomposition where $\mathfrak{q}_i$ are $\mathfrak{p}_i$-primary ideals for primes $\mathfrak{p}_i$.*

1. *The isolated primes belonging to $\mathfrak{a}$ are uniquely determined by $\mathfrak{a}$.*
2. *If $\mathfrak{p} \supseteq \mathfrak{a}$ is a prime containing $\mathfrak{a}$, then $\mathfrak{p}$ contains an isolated prime belonging to $\mathfrak{a}$. Consequently, the minimal primes of $A/\mathfrak{a}$ are in bijection with isolated primes belonging to $\mathfrak{a}$.*
3. *The isolated primary components of $\mathfrak{a}$ are uniquely determined by $\mathfrak{a}$.*

There is a correspondence result for primary ideals under localization.

Proposition 16.1.7.8. *Let $S \subseteq R$ be a multiplicative set and $\mathfrak{q} \leq R$ be a $\mathfrak{p}$-primary ideal. If $S \cap \mathfrak{p} = \emptyset$, then $S^{-1}\mathfrak{q}$ is an $S^{-1}\mathfrak{p}$-primary ideal of $S^{-1}R$. Consequently, there is a bijection between primary ideals $\mathfrak{q} \leq R$ not intersecting S and primary ideals of $S^{-1}R$.*

The above discussion is for ideals which admit a primary decomposition. The question remains that for which rings does every ideal admit a primary decomposition. One such answer is given by noetherian rings.

Theorem 16.1.7.9 (Lasker-Noether)**.** *Let R be a noetherian ring. Then every ideal $\mathfrak{a} \leq R$ admits a primary decomposition.*

16.2 Graded rings & modules

We now study a very important class of rings, which have an extra structure of having their additive abelian group being graded by $\mathbb{Z}$[1]. These include polynomial algebras and quotient of polynomial algebras by homogeneous ideals. In particular, they are the algebraic counterparts of projective varieties. These will also be essential while discussing dimension theory.

Definition 16.2.0.1 (**Graded rings & homogeneous ideals**). A ring S is said to be graded if the additive subgroup of S has a decomposition

$$S = \bigoplus_{d \geq 0} S_d$$

where $S_d \subseteq S$ is a subgroup which is called the subgroup of degree d homogeneous elements, such that for all $d, e \geq 0$, we have

$$S_d \cdot S_e \subseteq S_{d+e}.$$

An ideal $\mathfrak{a} \leq S$ is said to be homogeneous if the additive subgroup of $\mathfrak{a}$ has a decomposition

$$\mathfrak{a} = \bigoplus_{d \geq 0} \mathfrak{a} \cap S_d.$$

Remark 16.2.0.2. Hence, if S is a graded ring, then for all $d \geq 0$, the abelian group S_d is an S_0-module. Moreover, as $S = \bigoplus_{d \geq 0} S_d$, therefore S is an S_0-algebra.

Polynomial rings $S = k[x_0, \dots, x_n]$ are graded rings where S_d is the abelian subgroup of all degree d homogeneous monomials. We will see more examples once we show how to construct quotients and localizations of graded rings. But first we see some important properties of homogeneous ideals.

Proposition 16.2.0.3. *Let S be a graded ring and $\mathfrak{a} \leq S$ be any ideal. Then,*

1. *$\mathfrak{a}$ is homogeneous if and only if there exists $G \subseteq S$ a subset of homogeneous elements such that G generates $\mathfrak{a}$.*
2. *Let $\mathfrak{a}, \mathfrak{b}$ be two homogeneous ideals of S. Then $\mathfrak{a} + \mathfrak{b}, \mathfrak{a} \cdot \mathfrak{b}$ and $\sqrt{\mathfrak{a}}$ are again homogeneous ideals.*
3. *The homogeneous ideal $\mathfrak{a}$ is prime if and only if for any two homogeneous $f, g \in \mathfrak{a}$ it follows that $fg \in \mathfrak{a}$ implies either $f \in \mathfrak{a}$ or $g \in \mathfrak{a}$.*

We now define the notion of graded map of graded rings.

Definition 16.2.0.4 (**Map of graded rings**). Let S, T be graded rings. A ring homomorphism $\varphi : S \to T$ is said to be a graded map if for all $d \geq 0$ we get $\varphi|_{S_d} : S_d \to T_d$. That is, φ preserves degree.

16.2.1 Constructions on graded rings

We now do familiar constructions on graded rings, like quotients, fraction fields and localizations.

[1] we choose to not work in excessive generality; $\mathbb{Z}$-grading is sufficient for us.

Homogeneous localization

Definition 16.2.1.1 (Homogeneous localization). Let S be a graded ring and $T \subseteq S$ be a multiplicative set consisting only of homogeneous elements of S. Then we define $(T)^{-1}S$ to be the degree 0 abelian group of the graded ring $T^{-1}S$. Hence $(T)^{-1}S$ is a commutative ring with 1.

Remark 16.2.1.2. Indeed, $T^{-1}S$ is a graded ring where an element $f/g \in T^{-1}S$ for homogeneous f has degree $\deg f - \deg g$. It is immediate to see that this is well-defined and satisfies $(T^{-1}S)_d \cdot (T^{-1}S)_e \subseteq (T^{-1}S)_{d+e}$, making $T^{-1}S$ a graded ring.

The following is a discussion on localization of a graded ring S at a homogeneous prime ideal $\mathfrak{p}$. Let T denote the multiplicative subset of S consisting of all homogeneous elements not contained in $\mathfrak{p}$. Then $T^{-1}S$ is a graded ring whose degree d-elements are a/f where $a \in S_{d+e}$ and $f \in T$ of degree e. These form an additive abelian group where $a/f + b/g = ag + bf/fg$ where $a \in S_{d+k}, b \in S_{d+l}$ and $f, g \in T$ are of degree k and l respectively. Indeed, then $ag + bf \in S_{d+k+l}$ and $fg \in T$ of degree $k + l$. Consequently, we define

$$S_{(\mathfrak{p})} := (T^{-1}S)_0$$

where $(T^{-1}S)_0$ is the degree 0 elements in the localization $T^{-1}S$. We call this the *homogeneous localization* of the graded ring S at the homogeneous prime ideal $\mathfrak{p}$. Thus $S_{(\mathfrak{p})} = (S_\mathfrak{p})_0$, i.e. homogeneous localization just picks out degree 0 elements from the usual localization. Note that the usual localization $T^{-1}S$ is a graded ring where grading is given by subtracting the degree of numerator by degree of denominator.

Lemma 16.2.1.3. *Let S be a graded ring and $\mathfrak{p}$ be a homogeneous prime ideal of S. Then, the homogeneous localization $S_{(\mathfrak{p})}$ is a local ring.*

Proof. Consider the set $\mathfrak{m} := (\mathfrak{p} \cdot T^{-1}S) \cap S_\mathfrak{p}$. Then, $\mathfrak{m}$ is a maximal ideal of $S_\mathfrak{p}$ as any element not in $\mathfrak{m}$ in $S_\mathfrak{p}$ is a fraction f/g where $\deg f = \deg g$ and $f \notin \mathfrak{p}$ and thus it is invertible. Consequently, $S_\mathfrak{p}$ is local. $\square$

Remark 16.2.1.4. Note that if S is a graded domain, then $S_{(\langle 0 \rangle)}$ yields a field whose elements are of the form f/g where $\deg f = \deg g$ and f, g g is a non-zero homogeneous element of S. This field is called the *homogeneous fraction field* of graded domain S. This is a subfield of usual fraction field $S_{\langle 0 \rangle}$.

Remark 16.2.1.5. Let S be a graded ring and $g \in S$ be a homogeneous element. The *homogeneous localization of S at g* is defined to be the following subring of S_g:

$$S_{(g)} := \{f/g^n \in S_g \mid f \text{ is homogeneous with } \deg f = n \deg g,\ n \in \mathbb{N}\} \leq S_g.$$

Let S be a graded ring. Then an S-module M is said to be *graded S-module* if $M = \bigoplus_{d \in \mathbb{Z}} M_d$ where $M_d \leq M$ is a subgroup of M such that $S_d \cdot M_e \subseteq M_{d+e}$. Then, for a homogeneous element $g \in S$, we denote by $M_{(g)}$ the following submodule of M_g:

$$M_{(g)} := \{m/g^n \mid m \text{ is homogeneous with } \deg m = n \deg g,\ n \in \mathbb{N}\} \leq M_g.$$

The following is an important structural result of homogeneous localization of a graded ring at a homogeneous element.

Proposition 16.2.1.6. *Let S be a graded ring. Fix $f \in S_1$, a degree 1 homogeneous element of S. Consider the homogeneous localization $S_{(f)}$.*

1. *We have an isomorphism of abelian groups for all $e \in \mathbb{Z}$:*

$$(S_f)_e \cong f^e S_{(f)}.$$

Consequently,

$$S_f \cong \bigoplus_{e \in \mathbb{Z}} f^e S_{(f)}.$$

2. *We have an isomorphism of graded rings:*

$$S_f \cong S_{(f)}\left[t, \frac{1}{t}\right].$$

This is also an isomorphism of $S_{(f)}$-algebras.

Proof. 1. Fix $e \in \mathbb{Z}$ and consider the map

$$\begin{aligned} (S_f)_e &\longrightarrow f^e S_{(f)} \\ \frac{g}{f^n} &\longmapsto f^e \cdot \frac{g}{f^{n+e}}. \end{aligned}$$

As $f \in S_1$, this is well-defined. It is immediate that this is an isomorphism.

2. Consider the map

$$\begin{aligned} S_{(f)}\left[t, \frac{1}{t}\right] &\longrightarrow S_f \\ t &\longmapsto f. \end{aligned}$$

This is injective as if $p(f, 1/f) = 0$ in S_f for some $p(t, 1/t) \in S_{(f)}\left[t, \frac{1}{t}\right]$, then by item 1, it follows that $p(t, 1/t) = 0$. Furthermore, this is surjective by construction. This completes the proof. □

For each graded S-module M, one can attach a sequence of graded modules.

Definition 16.2.1.7. (**Twisted modules**) Let S be a graded ring and M a graded S-module. Then, define

$$M(l) := \bigoplus_{d \in \mathbb{Z}} M_{d+l}$$

to be the *l-twisted graded module of M*.

An important lemma with regards to localization of a graded ring at a positive degree element is as follows, it will prove its worth in showing that projective spectrum of a graded ring is a scheme (see Lemma **??**).

Lemma 16.2.1.8. *Let S be a graded ring and $f \in S_d$, $d > 0$. Then we have a bijection*

$$D_+(f) \cong \operatorname{Spec}\left(S_{(f)}\right)$$

where $D_+(f) \subseteq \operatorname{Spec}(S)$ is the set of all homogeneous prime ideals of S which does not contain f and does not contain S_+.

Proof. Consider the following map

$$\begin{aligned} \varphi : D_+(f) &\longrightarrow \operatorname{Spec}\left(S_{(f)}\right) \\ \mathfrak{p} &\longmapsto (\mathfrak{p} \cdot S_f)_0, \end{aligned}$$

that is, the degree zero elements of the prime ideal $\mathfrak{p} \cdot S_f$ of S_f. Indeed, $\varphi(\mathfrak{p})$ is a prime ideal of $S_{(f)}$. Further, if $(\mathfrak{p} \cdot S_f)_0 = (\mathfrak{q} \cdot S_f)_0$ for $\mathfrak{p}, \mathfrak{q} \in D_+(f)$, then for any $g \in \mathfrak{p}$, one observes via above equality that $g \in \mathfrak{q}$. Consequently, $\mathfrak{p} = \mathfrak{q}$. Thus φ is injective. For surjectivity, pick any prime ideal $\mathfrak{p} \in \operatorname{Spec}\left(S_{(f)}\right)$. We will construct a prime ideal $\mathfrak{q} \in D_+(f)$ such that $\varphi(\mathfrak{q}) = \mathfrak{p}$. Indeed, let $K = \{g \in S \mid g \text{ is homogeneous } \& \ \exists n > 0 \text{ s.t. } g/f^n \in \mathfrak{p}\}$ and consider the ideal

$$\mathfrak{q} = \langle K \rangle.$$

We thus need to check the following statements to complete the bijection:

1. $\mathfrak{q}$ is not the unit ideal of S,
2. $\mathfrak{q}$ is homogeneous in S,
3. $\mathfrak{q}$ is prime in S,
4. $\mathfrak{q}$ doesn't contain f,
5. $(\mathfrak{q} \cdot S_f)_0 = \mathfrak{p}$.

Statement 4 tells us that $\mathfrak{q}$ doesn't contain S_+. Statement 1 follows from a degree argument; if $1 \in \mathfrak{q}$, then $1 = a_1 g_1 + \cdots + a_m g_m$ for $g_i \in K$ and $a_i \in S$, but 1 is a degree 0 element whereas the minimum degree of the right is atleast > 0. Statement 2 is immediate as $\mathfrak{q}$ is generated by homogeneous elements. For statement 3, it is enough to check for homogeneous elements $h, k \in S$ that $hk \in \mathfrak{q} \implies h \in \mathfrak{q}$ or $k \in \mathfrak{q}$. This is immediate, after observing that any homogeneous element of $\mathfrak{q}$ is in K because K is the set of all homogeneous elements of S of positive degree which is not a power of f. Statements 4 and 5 are immediate checks. $\square$

16.3 Noetherian modules and rings

Let R be a ring. An R-module M is said to be *noetherian* if it satisfies either of the following equivalent properties:

1. Every increasing chain of submodules of M eventually stabilizes.
2. Every non-empty family of submodules of M has a maximal element.
3. Every submodule is finitely generated.

We prove the equivalence of 1 and 3 as in Proposition 16.3.0.3. But before, let us see that noetherian hypothesis descends to submodules and to quotients:

Lemma 16.3.0.1. *Let R be a ring and M be a noetherian R-module.*

1. *If N is a submodule of M, then N is noetherian.*
2. *If M/N is a quotient of M, then M/N is noetherian.*

Proof. 1. Take any submodule of M which is in N, then it is a submodule of N which is finitely generated.
2. Take any submodule of M/N, which is of the form K/N where $K \subseteq M$ is a submodule of M containing N. Hence K is finitely generated and so is N. Thus K/N is finitely generated. □

We also have that a finitely generated module over noetherian ring necessarily has to be noetherian, so every submodule is also finitely generated, which is not usually the case. This is another hint why having noetherian hypothesis can greatly ease calculations.

Lemma 16.3.0.2. *Let R be a noetherian ring and let M be an R-module. Then M is a noetherian module if and only if M is finitely generated.*

Proof. The only non-trivial side is R ⇒ L. Since M is finitely generated, therefore there is a surjection $f : R^n \twoheadrightarrow M$ where R^n is noetherian as R is noetherian (you may like to see it as a consequence of Corollary 16.3.0.5). Now take an increasing chain of submodules $N_0 \subseteq N_1 \subseteq \dots$ of M. This yields an increasing chain of ideals $f^{-1}(N_0) \subseteq f^{-1}(N_1) \subseteq \dots$, which stabilizes as R is noetherian. Applying f to the chain again we get that $N_0 \subseteq N_1 \subseteq \dots$ stabilizes. □

Here's the proof of equivalence as promised.

Proposition 16.3.0.3. *Let R be a ring. An R-module M is noetherian if and only if every submodule of M is finitely generated.*

Proof. (L $\implies$ R) Suppose R-module M is noetherian and let $S \subseteq M$ be a submodule of M. Note S is also noetherian. This means that any subcollection of submodules of S has a maximal element. Let such a subcollection be the collection of all finitely generated submodules of S, which clearly isn't empty as $\{0\}$ is there. This would have a maximal element, say N. If $N = S$, we are done. If not, then take $x \in S \setminus N$ and look at $N + Rx \subset S$. Clearly this is a submodule of S strictly containing N and is also finitely generated as N is too. This contradicts the maximality of N. Hence every submodule of M is finitely generated.
(R $\implies$ L) Let every submodule of M be finitely generated. We wish to show that this makes M into a noetherian module. So take any ascending chain of submodules $S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots$. Consider the union $S = \cup_{i=0}^{\infty} S_i$. S is also a submodule because for any $x, y \in S$, since $\{S_i\}$ is an ascending chain, there exists S_i such that $x, y \in S_i$, and so $x + y \in S_i \subseteq S$. By hypothesis,

$S = \langle x_1, \dots, x_k \rangle$. Let S_{n_i} be the smallest submodule containing x_i. Then $S_{\max n_i}$ is a member of the chain which contains each of the x_is, which thus means that the $S_{\max n_i}$ is generated by x_is because if it didn't then S would have either a smaller or a larger generating set, contradicting the generation by $x_1, \dots, x_k$. Hence the chain stabilizes after $S_{\max n_i}$. □

The reason one dwells with the noetherian hypothesis is reflected in the following properties enjoyed by it. Given a short exact sequence of modules, it is possible to figure out whether the middle module is noetherian or not by checking the same for the other two:

Proposition 16.3.0.4. *Let $0 \longrightarrow M' \xrightarrow{f} M \xrightarrow{g} M'' \longrightarrow 0$ be a short exact sequence of R-modules. Then, the module M is noetherian if and only if M' and M'' are noetherian.*

Proof. (L $\implies$ R) Let M be noetherian. Then if we consider any ascending chain in M' or M'', then we get an ascending chain in M because of the maps f and g. Remember inverse image of an injective and direct image of a surjective module homomorphism of a submodule is also a submodule.
(R $\implies$ L) Consider an ascending chain of submodules $S_0 \subseteq S_1 \subseteq \dots$ in M. We then have two more ascending chains $\{(f)^{-1}(S_i)\}$ and $\{g(S_i)\}$ in M' and M'' respectively. Since these are noetherian, therefore for both of them $\exists k \in \mathbb{N}$ such that these two chains stabilizes after k. Now, we wish to show that $\{S_i\}$ also stabilizes after k. For this, we just need to show that $S_{k+1} \subseteq S_k$. Hence take any $m \in S_{k+1}$. We have $g(m) \in g(S_k)$, therefore $\exists s \in S_k$ such that $g(m) = g(s) \implies g(m-s) = 0$ in M''. Since the sequence is exact, therefore $\exists m' \in M'$ such that $f(m') = m - s$, or, $m - s \in \operatorname{im}(f)$. Since $m \in S_{k+1}$ and $s \in S_k \subseteq S_{k+1}$, therefore $m - s \in S_{k+1}$. Hence $m - s \in \operatorname{im}(f) \cap S_{k+1}$ and since $\operatorname{im}(f) \cap S_{k+1} = \operatorname{im}(f) \cap S_k$, therefore $m - s \in S_k$ and thus $m \in S_k$. This proves $S_{k+1} \subseteq S_k$, proving $S_k = S_{k+1} = \dots$. □

An easy consequence of the above is that direct sum of finitely many noetherian modules is again noetherian:

Corollary 16.3.0.5. *Suppose $\{M_i\}_{i=1}^n$ be a collection of noetherian R-modules. Then $\bigoplus_{i=1}^n M_i$ is also a noetherian R-module.*

Proof. Since the sum $\bigoplus_{i=1}^n M_i$ sits at the middle of the following short exact sequence:

$$0 \longrightarrow M_1 \xrightarrow{f} \bigoplus_{i=1}^n M_i \xrightarrow{g} \bigoplus_{i=2}^n M_i \longrightarrow 0$$

where f is given by $m \longmapsto (m, 0, \dots, 0)$ and g is given by $(m_1, \dots, m_n) \longmapsto (m_2, \dots, m_n)$. The fact that this is indeed exact is simple to see. One can next use induction to complete the proof. □

An important result in the theory of noetherian rings is the following, which gives us few more (but highly important) examples of noetherian rings in nature. In particular it tells us that the one of the major class of rings which are studied in algebraic geometry, polynomial rings over algebraically closed fields, are noetherian.

Theorem 16.3.0.6. *(Hilbert basis theorem) Let R be a ring. If R is noetherian, then*
1. *$R[x_1, \dots, x_n]$ is noetherian,*
2. *$R[[x_1, \dots, x_n]]$ is noetherian.*

Proof. 1. We need only show that if R is noetherian then so is $R[x]$. Pick any ideal $I \leq R[x]$. We wish to show it is finitely generated. We go by contradiction, let I not be finitely generated.

Let $f_1 \in I$ be the smallest degree non-constant polynomial[2] and denote $I_1 = \langle f_1 \rangle$. Let $f_2 \in I \setminus I_1$ be the smallest degree non-constant polynomial and denote $I_2 = \langle f_1, f_2 \rangle$. Inductively, we define $I_n = \langle f_1, \dots, f_n \rangle$ where $f_n \in I \setminus I_{n-1}$ is of least degree non-constant. As I is not finitely generated, therefore for all $n \in \mathbb{N}$, $I_n \lneq I$. Let $f_n(x) = a_n x^m +$ other terms for each $n \in \mathbb{N}$ so that $a_n \in R$ represents the coefficient of the leading term of $f_n(x)$. Consequently, we obtain a sequence $\{a_n\} \subseteq R$. Let $J = \langle a_1, \dots, a_n, \dots \rangle$. As R is noetherian, therefore there exists $n \in \mathbb{N}$ such that $J = \langle a_1, \dots, a_n \rangle$. It follows that for some $r_1, \dots, r_n \in R$ we have

$$a_{n+1} = r_1 a_1 + \dots + r_n a_n.$$

We claim that $I = \langle f_1, \dots, f_n \rangle =: I_n$.

If not then $f_{n+1} \in I \setminus I_n$ is of least degree non-constant. We will now show that $f_{n+1} \in I_n$, thus obtaining a contradiction. Indeed, we have by the way of choice of f_{n+1} that $\deg f_{n+1} \geq \deg f_i$ for each $i = 1, \dots, n$. Consequently the polynomial

$$g = \sum_{i=1}^{n} r_i f_i \cdot x^{\deg f_{n+1} - \deg f_i}$$

has the property that its degree is equal to $\deg f_{n+1}$ and the coefficient of its leading term is equal to f_{n+1}. It follows that the polynomial $g - f_{n+1} \in I$ has degree strictly less than that of f_{n+1}. By minimality of f_{n+1}, it follows that $g - f_{n+1} \in I_n$. Note that by construction $g \in I_n$. Hence $f_{n+1} \in I_n$, as required.

2. **TODO** : *Write it from your exercise notebook.* □

Any localization of noetherian ring is again noetherian.

Proposition 16.3.0.7. *Let R be a noetherian ring and $S \subset R$ be a multiplicative set. Then $S^{-1}R$ is a noetherian ring.*

Proof. Any ideal of R is $S^{-1}I$ where $I \subseteq R$ is an ideal by exactness of localization (Lemma 16.1.2.2). As I is finitely generated as an R-module, therefore $S^{-1}I$ is finitely generated as an $S^{-1}R$-module, as needed. □

Lemma 16.3.0.8. *Let R be a ring with $\langle f_1, \dots, f_n \rangle = R$. If each R_{f_i} is noetherian, then R is noetherian.*

Proof. Pick any ideal $I \subseteq R$. We wish to show it is finitely generated. By exactness of localization (Lemma 16.1.2.2), we get $I_{f_i} \subseteq R_{f_i}$ is an ideal, thus finitely generated as R_{f_i}-module. By Lemma 16.1.2.10, I is finitely generated as an R-module. □

Corollary 16.3.0.9. *Let R be a ring. Then, R is noetherian if and only if R_f is noetherian for all $f \in R$.* □

[2]this exists by well-ordering by degree.

16.3.1 Dimension 0 noetherian rings

We'll see that they are equivalent to the following.

Definition 16.3.1.1 (Artinian rings). A ring R is said to be artinian if it satisfies descending chain condition for its ideals.

The following are important properties of artinian rings.

Proposition 16.3.1.2. *Let R be an artinian ring. Then,*

1. *every prime is maximal,*
2. *there are finitely many maximals,*
3. *the Jacobson radical is a nilpotent ideal.*

There are two characterizations of artinian rings to keep in mind.

Theorem 16.3.1.3. *Let R be a ring. Then the following are equivalent:*

1. *R is artinian.*
2. *R is noetherian and $\dim R = 0$.*
3. *R is product of finitely many artinian local rings.*

If we allow noetherian hypothesis on R, then we also get the following equivalences.

Theorem 16.3.1.4. *Let R be a noetherian ring. Then the following are equivalent:*

1. *R is artinian.*
2. *$\mathrm{Spec}\,(R)$ is finite discrete.*

For a finitely generated k-algebra, we furthermore have the following.

Proposition 16.3.1.5. *Let k be a field and A be a finitely type k-algebra. Then the following are equivalent:*

1. *A is artinian.*
2. *A is a finite k-algebra.*

16.4 $\mathrm{Supp}\,(M)$, $\mathrm{Ass}\,(M)$ and primary decomposition

Let R be a ring and M be a finitely generated R-module. In the classical case when R is a field and M is then a finite dimensional R-vector space, if $x \in M$ then if even a single element of R annihilate x, then all elements of R annihilate x. This luxury is not enjoyed when R is a ring because not all elements of R may be invertible. What one does then is to study the associated annihilating ideals corresponding to each element of M. The global version of this idea is exactly the concept of *annihilator ideal of M*, i.e. $\mathfrak{a}_M := \{r \in R \mid rM = 0\}$. A module M is then called *faithful* if $\mathfrak{a}_M = 0$. The following exposition is taken from cite[LocalAlgebra].

Now, if we have an R-module M, then we get an ideal of R. This gives us a closed subset of $\mathrm{Spec}\,(R)$ (see Section 1.2). A basic question that then arises is what is the relationship between the module M and the closed set $V(\mathfrak{a}_M) \hookrightarrow \mathrm{Spec}\,(R)$. The following answers that.

Lemma 16.4.0.1. *Let R be a ring and M be a finitely generated R-module. If $\mathfrak{p} \in \mathrm{Spec}\,(R)$ and $\mathfrak{a}_M = \mathrm{Ann}(M)$ be the annihilator ideal, then the following are equivalent:*

1. $M_{\mathfrak{p}} \neq 0$.
2. $\mathfrak{p} \in V(\mathfrak{a}_M)$.

Proof. If we can show that $\mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = (\mathfrak{a}_M)_{\mathfrak{p}}$, then we have the following equivalence

$$M_{\mathfrak{p}} \neq 0 \iff \mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) \neq R_{\mathfrak{p}} \iff (\mathfrak{a}_M)_{\mathfrak{p}} \lneq R_{\mathfrak{p}} \iff \mathfrak{a}_M \subseteq \mathfrak{p}$$

where last equivalence follows from a modified version of Lemma 16.1.2.3. Hence we reduce to showing that $\mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = (\mathfrak{a}_M)_{\mathfrak{p}}$. It is easy to see that $\mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) \supseteq (\mathfrak{a}_M)_{\mathfrak{p}}$. Let $r/s \in \mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}})$. We wish to show that $r/s \in (\mathfrak{a}_M)_{\mathfrak{p}}$. Since M is finitely generated, therefore let $\{m_1, \dots, m_n\}$ be a generating set of M. We thus reduce to showing that $r/s \cdot m_i/1 = 0$ for each $i = 1, \dots, n$. This is exactly the data provided by the fact that $r/s \in \mathrm{Ann}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}})$. □

The above lemma hence gives us a closed subset of $\mathrm{Spec}\,(R)$ attached to each finitely generated R-module M. This has a name.

Definition 16.4.0.2. (**Support of a module**) Let R be a ring and M be a finitely generated R-module. Let $\mathfrak{a}_M$ be the annihilator ideal of M. Then, the support of the module M is defined to be the closed set $\mathrm{Supp}\,(M) := V(\mathfrak{a}_M) \hookrightarrow \mathrm{Spec}\,(R)$. By Lemma 16.4.0.1, it is equivalently given by the set of all those points $\mathfrak{p} \in \mathrm{Spec}\,(R)$ such that $M_{\mathfrak{p}} \neq 0$.

We then define prime ideals associated to an R-module.

Definition 16.4.0.3. (**Associated prime ideals**) Let R be a noetherian ring and M be an R-module. A prime ideal $\mathfrak{p} \in \mathrm{Spec}\,(R)$ is said to be associated to M if there exits $m \in M$ such that

$$\mathfrak{p} = \{r \in R \mid rm = 0\}.$$

The subspace of $\mathrm{Spec}\,(R)$ of all prime ideals associated to M is denoted $\mathrm{Ass}\,(M) \hookrightarrow \mathrm{Spec}\,(R)$.

One can have the following alternate definition of an associated prime ideal.

Lemma 16.4.0.4. *Let R be a noetherian ring and M be an R-module. Then,*

$$\mathfrak{p} \in \mathrm{Ass}\,(M) \iff \exists N \leq M \text{ such that } N \cong R/\mathfrak{p}.$$

Proof. L $\Rightarrow$ R is easy, just consider the map $R \to M$ given by $r \mapsto rm$ where $m \in M$ corresponds to $\mathfrak{p}$. Conversely, take any $0 \neq n \in N$. Then $\mathfrak{p} = \{r \in R \mid rn = 0\}$ as if $r \in R$ is such that $rn = 0$ and $n = s + \mathfrak{p}$, then $rn = rs + \mathfrak{p} = \mathfrak{p}$, that is $rs \in \mathfrak{p}$ and since $s \notin \mathfrak{p}$, therefore $r \in \mathfrak{p}$. Conversely, if $r \in \mathfrak{p}$ then for all $n \in N$, $rn = 0$. □

One can show that $\mathrm{Ass}\,(M)$ is finite for cases of interest.

Proposition 16.4.0.5. *Let R be a noetherian ring and M be a finitely generated R-module. Then* $\mathrm{Ass}\,(()M)$ *is finite.*

Moreover, localization behaves very nicely with associated primes.

Proposition 16.4.0.6. *Let R be a noetherian ring and M be an R-module. For a multiplicative set $S \subseteq R$ such that $S \cap \mathfrak{p} = \emptyset$, we have that $S^{-1}\mathfrak{p} \in \mathrm{Ass}\,(S^{-1}M)$ if and only if $\mathfrak{p} \in \mathrm{Ass}\,(M)$.*

The above proposition will allow us to define associated points of a coherent module over a locally noetherian scheme.

So, for an R-module M, we get two subspaces of Spec (R), one is the closed subspace called support Supp (M) and the other is Ass (M). Support will be used later, but the concept of associated prime ideals of M have a deeper connection with the ring R. They are not unrelated.

Lemma 16.4.0.7. *Let M be an R-module. Then* $\mathrm{Ass}\,(M) \hookrightarrow \mathrm{Supp}\,(M) \hookrightarrow \mathrm{Spec}\,(R)$.

Proof. For $\mathfrak{p} \in \mathrm{Ass}\,(M)$ let $m \in M$ such that its annihilator is $\mathfrak{p}$. Then, for any $r \in \mathfrak{a}_M$, $rm = 0$ and hence $r \in \mathfrak{p}$. Thus $\mathfrak{p} \in V(\mathfrak{a}_M) = \mathrm{Supp}\,(M)$. □

We wish to show the following result from which primary decomposition follows.

Theorem 16.4.0.8. *Let R be a noetherian ring and M be a finitely generated R-module. Then there exists an injective map*

$$M \longrightarrow \prod_{\mathfrak{p} \in \mathrm{Ass}(M)} E_{\mathfrak{p}}$$

where for each $\mathfrak{p} \in \mathrm{Ass}\,(M)$, $E_{\mathfrak{p}}$ is an R-module where $\mathrm{Ass}\,(E_{\mathfrak{p}})$ is a singleton given by $\{\mathfrak{p}\}$. We call such submodules $\mathfrak{p}$-primary.

This result clearly tells us that points of Ass (M) are somewhat special. Let us investigate.

Lemma 16.4.0.9. *Let R be a noetherian ring and M be a finite R-module*[3]*. Then,*

1. *If $N \subseteq M$ is a submodule, then $\mathrm{Ass}\,(N) \subseteq \mathrm{Ass}\,(M)$.*
2. *If $N \subseteq M$ is a submodule, then $\mathrm{Supp}\,(N) \subseteq \mathrm{Supp}\,(M)$.*
3. *If $N \subseteq M$ is a submodule, then $\mathrm{Ass}\,(N) \subseteq \mathrm{Ass}\,(M) \subseteq \mathrm{Ass}\,(N) \cup \mathrm{Ass}\,(M/N)$.*
4. *For any point $\mathfrak{p} \in \mathrm{Spec}\,(R)$, we have $\mathfrak{a}_{R/\mathfrak{p}} := \mathrm{Ann}(R/\mathfrak{p}) = \mathfrak{p}$. Thus, $\mathrm{Supp}\,(R/\mathfrak{p}) = V(\mathfrak{p})$ is an irreducible closed subset of $\mathrm{Spec}\,(R)$.*
5. *For any point $\mathfrak{p} \in \mathrm{Spec}\,(R)$, we have $\mathrm{Ass}\,(R/\mathfrak{p}) = \{\mathfrak{p}\}$. Thus, $\mathrm{Ass}\,(R/\mathfrak{p})$ is exactly the generic point of $\mathrm{Supp}\,(R/\mathfrak{p})$.*
6. *For all $\mathfrak{p} \in \mathrm{Spec}\,(R)$, there exists a maximal submodule $N \subseteq M$ such that $\mathfrak{p} \notin \mathrm{Ass}\,(N)$.*
7. *For all $\mathfrak{p} \in \mathrm{Ass}\,(M)$, there exists a maximal submodule $N \subsetneq M$ such that $\mathfrak{p} \notin \mathrm{Ass}\,(N)$ and none of these maximal submodules are isomorphic to $R/\mathfrak{p}$.*

Proof. Note that by Lemma 16.3.0.2, M is a Noetherian module.

1. If $\mathfrak{p} \in \mathrm{Ass}\,(N)$, then for some $n \in N$, $\mathfrak{p} = \{r \in R \mid rn = 0\}$. Result follows as $n \in M$.
2. If $\mathfrak{p} \in \mathrm{Supp}\,(N)$, then $\mathfrak{p} \supseteq \mathfrak{a}_N$. Result follows as $\mathfrak{a}_N \supseteq \mathfrak{a}_M$.
3. By 1, we need only show $\mathrm{Ass}\,(M) \subseteq \mathrm{Ass}\,(N) \cup \mathrm{Ass}\,(M/N)$. Pick $\mathfrak{p} \in \mathrm{Ass}\,(M)$. By the Lemma 16.4.0.4 and it's proof, the submodule E of M containing of all elements of M who have annihilator as $\mathfrak{p}$ is isomorphic to $R/\mathfrak{p}$. If $E \cap N = \emptyset$, then M/N has a submodule isomorphic to $R/\mathfrak{p}$ and hence $\mathfrak{p} \in \mathrm{Ass}\,(M/N)$. Otherwise if $E \cap N \neq \emptyset$, then pick $x \in E \cap N$. Since $x \in E$, so annihilator of x is $\mathfrak{p}$ and thus $\mathfrak{p} \in \mathrm{Ass}\,(E \cap N)$. By another use of Lemma 16.4.0.4, there is a submodule $F \subseteq E \cap N$ which is isomorphic to $R/\mathfrak{p}$. It follows that N has a submodule isomorphic to $R/\mathfrak{p}$. By a final use of Lemma 16.4.0.4, we conclude that $\mathfrak{p} \in \mathrm{Ass}\,(N)$.
4. $\mathrm{Ann}(R/\mathfrak{p}) = \{r \in R \mid r(R/\mathfrak{p}) = \mathfrak{p}\}$. It follows from primality of $\mathfrak{p}$ that $\mathrm{Ann}(R/\mathfrak{p}) = \mathfrak{p}$.

[3] this is just another name for finitely generated R-modules.

5. As above, this reduces to primality of $\mathfrak{p}$.
6. The set of all submodules N of M satisfying $\operatorname{Ass}(N) \notin p$ has a maximal element as M is a noetherian module.
7. If $\mathfrak{p} \in \operatorname{Ass}(M)$, then the maximal N obtained from 5 cannot be M. The other fact follows from 4.

□

The primary decomposition now is a corollary of the main theorem.

Corollary 16.4.0.10 (Primary decomposition theorem)**.** *Let R be a noetherian ring and M be a finitely generated R-module. If $N \leq M$ is a submodule, then we can write*

$$N = \bigcap_{\mathfrak{p} \in \operatorname{Ass}(M/N)} Q(\mathfrak{p})$$

where $Q(\mathfrak{p})$ is a $\mathfrak{p}$-primary submodule of M, that is, $\operatorname{Ass}(Q(\mathfrak{p})) = \{\mathfrak{p}\}$.

With the above investigation, we are now ready to prove Theorem 16.4.0.8.

Proof of Theorem 16.4.0.8. **TODO.** □

16.5 Tensor, symmetric & exterior algebras

16.5.1 Results on tensor products

We collect some important results on tensor products in this section which are used all over the text. The following results are immediate corollaries of definition of tensor product, but are of immense use in general.

Proposition 16.5.1.1. *Following are some basic properties of tensor products.*

1. *Tensor product is associative and commutative upto isomorphism.*
2. *If $\{M_\lambda\}$ is a family of R-modules and N is an R-module, then*
$$\left(\bigoplus_{\lambda} M_\lambda\right) \otimes_R N \cong \bigoplus_{\lambda} M_\lambda \otimes_R N.$$
3. *Let $\varphi : R \to S$ be a ring homomorphism and M, N be two R-modules. Then the scalar extended modules $M \otimes_R S$ and $N \otimes_R S$ satisfy the following*
$$(M \otimes_R S) \otimes_S (N \otimes_R S) \cong (M \otimes_R N) \otimes_R S.$$
4. *Let R be a ring and M be an R-module. If $I, J \leq R$ are two ideals, then*
$$R/I \otimes_R R/J \cong R/I + J$$
as rings.
5. *If R, S are two rings, then*
$$R \otimes_S S[x] \cong R[x]$$
as rings.

Proof. **TODO.** □

The following is a helpful lemma showing that tensor product commutes with direct limits in all positions.

Lemma 16.5.1.2. *Let M_i, N_i bet R_i-modules where I is directed set and $\{M_i\}, \{N_i\}$ and $\{R_i\}$ are directed systems of modules and rings. Let $M := \varinjlim_{i\in I} M_i$, $N := \varinjlim_{i\in I} N_i$ and $R := \varinjlim_{i\in I} R_i$. Then,*
$$\varinjlim_{i\in I}(M_i \otimes_{R_i} N_i) \cong M \otimes_R N$$
as R-modules.

Proof. We will construct R-linear maps $f : \varinjlim_{i\in I}(M_i \otimes_{R_i} N_i) \longleftrightarrow M \otimes_R N : g$ which will be inverses to each other. We first construct f as follows. For each $i \in I$, we have
$$f_i : M_i \otimes_{R_i} N_i \to M \otimes_{R_i} N \to M \otimes_R N$$

given by $(m_i \otimes n_i) \mapsto ((m_i) \otimes (n_i)) \mapsto ((m_i) \otimes (n_i))$. Note that M, N are R_i-modules canonically. By universal property of $\varinjlim_{i\in I}$, we obtain f as above. To construct g, we need only construct an R-bilinear map

$$\begin{aligned} M \times N &\longrightarrow \varinjlim_{i\in I}(M_i \otimes_{R_i} N_i) \\ ((m_i)_{i\in I}, (n_i)_{i\in I}) &\longmapsto ((m_i \otimes n_i)_{i\in I}). \end{aligned}$$

This can be said to be R-bilinear, thus yielding a map g as required. It is straightforward to see they are inverses to each other. □

The following says that localization commutes with tensor products.

Lemma 16.5.1.3. *Let M, N be two R-modules and $S \subseteq R$ be a multiplicative set. Then,*

$$S^{-1}(M \otimes_R N) \cong S^{-1}M \otimes_{S^{-1}R} S^{-1}N.$$

Proof. We may write by Lemma 16.1.2.1 the following

$$\begin{aligned} S^{-1}M \otimes_{S^{-1}R} S^{-1}N &\cong (M \otimes_R S^{-1}R) \otimes_{S^{-1}R} (S^{-1}R \otimes_R N) \\ &\cong M \otimes_R (S^{-1}R \otimes_{S^{-1}R} (S^{-1}R \otimes_R N)) \\ &\cong M \otimes_R (N \otimes_R S^{-1}R) \\ &\cong (M \otimes_R N) \otimes_R S^{-1}R \\ &\cong S^{-1}(M \otimes_R N). \end{aligned}$$

This completes the proof. □

The following is important for calculations of tensor of quotient maps.

Proposition 16.5.1.4. *Let R be a ring and $f : M \to N$, $g : M' \to N'$ be two surjective R-linear maps. Then*

$$\mathrm{Ker}\,(f \otimes g) = \mathrm{id} \otimes j(M \otimes \mathrm{Ker}\,(g)) + i \otimes \mathrm{id}(\mathrm{Ker}\,(f) \otimes M')$$

where $i : \mathrm{Ker}\,(f) \hookrightarrow M$ and $j : \mathrm{Ker}\,(g) \hookrightarrow M'$ are inclusions.

Next, we discuss the notion of fiber of a map of rings. This is easily understood in the scheme language.

Definition 16.5.1.5 (Fiber at a prime ideal). Let $\varphi : R \to S$ be a ring homomorphism and let $\mathfrak{p} \lneq R$ be a prime ideal. Then the fiber of φ at $\mathfrak{p}$ is defined to be $S \otimes_R \kappa(\mathfrak{p})$.

One of the fundamental observation about fiber at a prime ideal is that it is indeed the fiber of the corresponding map of schemes (see Proposition 1.6.5.1), so that the notation makes sense.

Remark 16.5.1.6 (Extension of primes to polynomial ring)**.** If $\mathfrak{p}$ is a prime of A, then the extension $\mathfrak{p} \otimes_A A[x]$ is isomorphic to a prime of $A[x]$. Indeed, the following map shows that it is isomorphic to prime $\mathfrak{p}A[x]$:

$$\begin{aligned} \mathfrak{p} \otimes_A A[x] &\longrightarrow \mathfrak{p}A[x] \\ a \otimes x &\longmapsto ax. \end{aligned}$$

However, if $\mathfrak{m}$ is maximal in A, then $\mathfrak{m}A[x]$ may not be maximal. Indeed, 0 is maximal in a field F, but $0 \otimes_F F[x] = 0$ is not in $F[x]$.

Remark 16.5.1.7. Let $f : A \to B$ be a ring homomorphism and N be a B-module. Let N_A be the restriction of scalars to A and $N_B = N_A \otimes_A B$. It is in general not true that $N_B \cong N$; in-fact, N_B is larger than N. Indeed we do have a map $g : N \to N_B$ mapping $n \mapsto n \otimes 1$ which is injective as can be seen from universal property of tensor products applied to the bilinear map $\pi : N_A \times B \to N_A$, the projection map. Even more is true; we see that the map $p : N_B \to N$ mapping $n \otimes b \mapsto bn$ provides a splitting of the following s.e.s.:

$$\begin{tikzcd} 0 \arrow[r] & N \arrow[r, "g"] & N_B \arrow[r] \arrow[l, dashed, bend left, "p"] & \mathrm{CoKer}\,(g) \arrow[r] & 0 \end{tikzcd}.$$

Thus, we have that in-fact N is a direct summand of N_B.

16.5.2 Determinants

Fix a commutative ring R with unity for the remainder of this section. We shall show in this section that there exists a unique determinant map over $M_n(R)$. This will motivate further notions discussed in later sections.

We begin by defining a multilinear map over $M_n(R)$.

Definition 16.5.2.1. (**Multilinear map over** $M_n(R)$) Let $n \in \mathbb{N}$ and consider $M_n(R)$. An n-linear map over $M_n(R)$ is a function

$$D : M_n(R) \longrightarrow R$$

which is linear in each row. That is, if A_i denotes the i^{th}-row of matrix A and $c \in R$, then for each $i = 1, \dots, n$, we have

$$\begin{aligned} D(A_1, \dots, A_{i-1}, cA_i + B_i, A_{i+1}, \dots, A_n) &= cD(A_1, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_n) \\ &\quad + D(A_1, \dots, A_{i-1}, B_i, A_{i+1}, \dots, A_n). \end{aligned}$$

We may abbreviate the above by simply writing $D(cA_i + B_i) = cD(A_i) = D(B_i)$.

Example 16.5.2.2. The map

$$\begin{aligned} D : M_n(R) &\longrightarrow R \\ A &\longmapsto cA_{1k_1}A_{2k_2} \dots A_{nk_n} \end{aligned}$$

is an n-linear map where $c \in R$ is a constant and $1 \le k_i \le n$ are n integers.

We first see that linear combination of n-linear maps is again n-linear.

Lemma 16.5.2.3. *Let $D_1, \dots, D_r$ be n-linear maps and $c_1, \dots, c_r \in R$. Then $c_1D_1 + \dots + c_rD_r$ is an n-linear map.*

Proof. By induction, we may assume $r = 2$. Now this is a straightforward check. □

We now come more closer to determinants by defining the following type of n-linear maps.

Definition 16.5.2.4. (**Alternating & determinant maps**) An n-linear map $D : M_n(R) \to R$ is said to be alternating if

1. $D(A) = 0$ if $A_i = A_j$ for any $i \neq j$,
2. $D(\sigma_{ij}(A)) = -D(A)$ where σ_{ij} swaps rows A_i and A_j.

An alternating n-linear map $D : M_n(R) \to R$ is said to be determinant if $D(I_n) = 1$.

Proposition 16.5.2.5. *If $D : M_n(R) \to R$ is an n-linear map such that $D(A) = 0$ whenever $A_i = A_{i+1}$ for some $1 \leq i \leq n$, then D is alternating.*

Proof. Let $A \in M_n(R)$ and $1 \leq i \neq j \leq n$ be such that $A_i = A_j$. We first wish to show that $D(\sigma_{ij}(A)) = -D(A)$. We may assume $j > i$. We go by strong induction over $j - i$. We first show this for $j = i+1$. Indeed, we then have $D(\sigma_{i,i+1}(A)) = D(A_{i+1}, A_i)$. Writing $0 = D(A_{i+1} + A_i, A_i + A_{i+1}) = D(A_{i+1}, A_i) + D(A_i, A_{i+1})$. Thus we get $D(A_{i+1}, A_i) = -D(A_i, A_{i+1})$.

In the inductive case, suppose $D(\sigma_{ij}(A)) = -D(A)$ for all $j - i \leq k$. We wish to show that if $j - i = k + 1$, then the same holds. As $\sigma_{i,i+k+1}(A) = \sigma_{i+k,i+k+1} \circ \sigma_{i,i+k} \circ \sigma_{i+k,i+k+1}(A)$, therefore we are done.

To get that $D(A) = 0$ for A such that $A_i = A_j$ for some $j > i$, we may simply swap rows till they are adjacent, which will be zero by our hypothesis. □

We now define the main candidate for the determinant function over $M_n(R)$.

Definition 16.5.2.6. (E_j) Let $D : M_{n-1}(R) \to R$ be an $n-1$-linear map. For each $1 \leq j \leq n$, define the following map

$$\begin{aligned} E_j : M_n(R) &\longrightarrow R \\ A &\longmapsto \sum_{i=1}^{n} (-1)^{i+j} A_{ij} D(A[i|j]). \end{aligned}$$

Further denote $D_{ij}(A) := D(A[i|j])$.

Theorem 16.5.2.7. *Let $n \in \mathbb{N}$ and $D : M_{n-1}(R) \to R$ be an alternating $n-1$-linear map. For each $1 \leq j \leq n$, the map $E_j : M_n(R) \to R$ defined as above is an alternating n-linear map. If moreover D is a determinant map, then so is each E_j.*

Proof. Fix $1 \leq j \leq n$. We first wish to show that E_j is n-linear. As $D_{ij} : M_n(R) \to R$ is linear in every row except i. Thus $A \mapsto A_{ij} D_i j(A)$ is n-linear. It follows from Lemma 16.5.2.3 that E_j is n-linear.

To show that E_j is alternating, it would suffice from Proposition 16.5.2.5 to show that $E_j(A) = 0$ if A has any two adjacent rows equal, say $A_k = A_{k+1}$. This one checks directly by the definition of E_j.

To see that E_j is determinant if D is determinant is also easy to see. □

We now show the uniqueness of determinants and alternating n-linear maps (upto the value on I_n).

Theorem 16.5.2.8. *Let $D : M_n(R) \to R$ be an alternating n-linear map over $M_n(R)$. Then,*

1. *D is given explicitly on $A \in M_n(R)$ by*

$$D(A) = \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \dots A_{n\sigma(n)} \right) D(I),$$

hence D is unique upto its value over I,

2. *if D is determinant map, then it is uniquely given by*

$$D(A) = \det A := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \dots A_{n\sigma(n)},$$

3. *any alternating map D on $M_n(R)$ is thus uniquely determined by its value on I as*

$$D(A) = (\det A) \cdot D(I).$$

Proof. The proof is straightforward but tedious. **TODO**. □

Corollary 16.5.2.9. *Let $n \in \mathbb{N}$.*

1. *If $A, B \in M_n(R)$, then $\det(AB) = \det(A) \cdot \det(B)$.*
2. *If $B \in M_n(R)$ is obtained by $B_j = A_j + cA_i$ for some fixed $1 \le i, j \le n$ and rest of the rows of B are identical to A, then $\det(B) = \det(A)$.*
3. *If $M \in M_{r+s}(R)$ is given by*

$$M = \begin{bmatrix} A_{r\times r} & B_{r\times s} \\ 0 & C_{s\times s} \end{bmatrix}$$

then $\det(M) = \det(A) \cdot \det(C)$.

4. *For each $1 \le j \le n$, we have*

$$\det(A) = E_j(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A[i|j]).$$

Proof. (*Sketch*) For 1. we can contemplate

$$\begin{aligned} D : M_n(R) &\longrightarrow R \\ A &\mapsto \det(AB). \end{aligned}$$

One claims that D is an n-linear alternating map. Then apply Theorem 16.5.2.8, 3.

2. Follows by multilinearity of det.

3. As elementary row operations only change determinant upto sign and restricting an $r+s$-linear alternating map to first r or last s entries keeps it r-linear and s-linear alternating respectively, therefore the result follows.

4. Follows from Theorem 16.5.2.7 and Theorem 16.5.2.8. □

Construction 16.5.2.10. (*Adjoint of a matrix*) Let $A \in M_n(R)$ be a square matrix. By Corollary 16.5.2.9, the sum $E_j(A) = \det(A)$ for each $1 \le j \le n$

$$\det(A) = \sum_{i=1}^{n} A_{ij} (-1)^{i+j} \det(A[i|j]).$$

Hence, let us define $C_{ij} := (-1)^{i+j}\det(A[i|j])$ as the ij^{th}-*cofactor of* A. Consequently, we get a matrix $(\text{Adj}A)_{ij} = C_{ji}$, called the *adjoint matrix*. Hence, we may rewrite the determinant as

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} A_{ij}C_{ij}\\ &= \sum_{i=1}^{n}(\text{Adj}A)_{ji}A_{ij}.\end{aligned}$$

Thus,

$$\det(A)I = \text{Adj}(A)\cdot A.$$

This also allows us to write that in the case when A is invertible, we have

$$A^{-1} = \frac{1}{\det A}\text{Adj}(A).$$

As similar matrices have same determinant, therefore each linear operator on a finite dimensional vector space has a unique determinant. Thus determinants are invariants of linear operators upto similarity.

16.5.3 Multilinear maps

We now put the previous discussion in a more abstract framework where we work with modules over a commutative ring with 1. We first recall that the rank of a finitely generated module is the size of the smallest generating set. Further recall that a finitely generated free R-module V has a well-defined rank and the smallest generating set is moreover a basis of V (i.e. linearly independent set of generators).

For this section, we would hence fix a commutative ring R with 1.

Definition 16.5.3.1. (**r-linear forms over a module**) Let V be an R-module. An r-linear form L over V is a function

$$L : V^r = V \times \cdots \times V \longrightarrow R$$

such that for any $c \in R$, $\beta_i \in V$ and $(\alpha_1, \ldots, \alpha_r) \in V^r$, we have

$$L(\alpha_1, \ldots, c\alpha_i + \beta_i, \ldots, \alpha_n) = cL(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n) + L(\alpha_1, \ldots, \beta_i, \ldots, \alpha_n)$$

for any $1 \le i \le r$. An r-linear form is usually also called an r-tensor. A 2-linear form/tensor is also usually called a bilinear form. Note that an r-linear form may not be linear. Denote the R-module of all r-linear forms by $M^r(V)$.

Remark 16.5.3.2. Let $f_1, \ldots, f_r \in V^* = \text{Hom}_R(V, R) = M^1(V)$ be a collection of linear functionals. We then obtain $L \in M^r(V)$ given by

$$L(\alpha_1, \ldots, \alpha_r) = f_1(\alpha_1) \cdot \cdots \cdot f_r(\alpha_r).$$

Example 16.5.3.3. We give some examples.

1. Let $V = R^n$ be a free R-module of rank n. Then for a fixed matrix $A \in M_n(R)$, the map
$$\begin{aligned} V \times V &\longrightarrow R \\ (x,y) &\longmapsto x^t A y \end{aligned}$$
is a bilinear form over V.
2. Let $V = R^n$ be a free R-module of rank n. Then we obtain the following n-linear form
$$\begin{aligned} \det : V^n &\longrightarrow R \\ (\alpha_1, \dots, \alpha_n) &\longmapsto \det(A) \end{aligned}$$
where $A \in M_n(R)$ whose i^{th}-row is α_i. Hence, determinant is an n-tensor/n-linear form over V.

Remark 16.5.3.4. (General expression of an r-linear form) Let $L \in M^r(V)$ be an r-form over an R-module V where V is a free module of rank n. Further denote $e_1, \dots, e_n$ be a basis of V. For any $(\alpha_1, \dots, \alpha_r) \in V^r$, we may write $\alpha_i = \sum_{j=1}^n A_{ij} e_j$. Hence we have $A \in M_{r\times n}(R)$. This yields by n-linearity of L that
$$\begin{aligned} L(\alpha_1, \dots, \alpha_r) &= \sum_{j_r=1}^n \cdots \sum_{j_1=1}^n A_{1j_1} \dots A_{rj_r} L(e_{j_1}, \dots, e_{j_r}) \\ &= \sum_{J=\{j_1,\dots,j_r\}} A_J L(e_J) \end{aligned}$$
where $J \in X$ where X is the set of all r-tuples with entries in $\{1, \dots, n\}$. There are therefore n^r terms in the above sum.

Definition 16.5.3.5. (**Tensor product of linear forms**) Let M be an R-module. We then define
$$\begin{aligned} - \otimes - : M^r(V) \times M^s(V) &\longrightarrow M^{r+s}(V) \\ (L, M) &\longmapsto L \otimes M \end{aligned}$$
where $L \otimes M : V^{r+s} \to R$ is given by $(\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s) \mapsto L(\alpha_1, \dots, \alpha_r) M(\beta_1, \dots, \beta_s)$.

Remark 16.5.3.6. We have following observations about tensor of forms:
1. $L \otimes (T+S) = L \otimes T + L \otimes S$,
2. $(L \otimes T) \otimes N = L \otimes (T \otimes N)$,
3. $c(L+T) \otimes S = cL \otimes S + cT \otimes S$,
4. $L \otimes T \neq T \otimes L$.

We now come to an important theorem about $M^r(V)$

Theorem 16.5.3.7. *Let V be a free R-module of rank n and $B = \{e_1, \dots, e_n\} \subseteq V$ be a basis of V. Let X denote the set of all r-tuples with entries in $\{1, \dots, n\}$. Then,*
1. *the R-module $M^r(V)$ is free of rank n^r,*
2. *a basis of $M^r(V)$ is given by $f_J = f_{j_1} \otimes \dots \otimes f_{j_r}$ where $B^* = \{f_1, \dots, f_n\} \subseteq V^* = M^1(V)$ is the dual basis of B, where $J = \{j_1, \dots, j_r\}$ varies over all elements of X.*

Proof. (*Sketch*) We claim that $\{f_J\}_{J\subseteq X}$ forms a basis of $M^r(V)$. Pick any $(\alpha_1,\ldots,\alpha_r)\in V^r$, then by Remark 16.5.3.4, we first have $\alpha_i=\sum_{j=1}^n f_j(\alpha_i)e_j$. Consequently,

$$\begin{aligned}L(\alpha_1,\ldots,\alpha_r) &= \sum_{J=\{j_1,\ldots,j_r\}} L(e_{j_1},\ldots,e_{j_r})\cdot f_{j_1}\otimes\ldots\otimes f_{j_r}(\alpha_1,\ldots,\alpha_r)\\ &= \sum_{J=\{j_1,\ldots,j_r\}} L(e_J)f_{j_1}\otimes\ldots\otimes f_{j_r}(\alpha_1,\ldots,\alpha_r).\end{aligned}$$

Thus, $\{f_J\}_{J\subseteq X}$ spans $M^r(V)$. For linear independence, take any combination

$$\sum_{J\subseteq X} c_J f_J = 0.$$

On the LHS, apply e_I to get $c_I=0$ for each $I\subseteq X$. □

Definition 16.5.3.8. (**Alternating r-linear forms**) Let V be an R-module. An r-linear form $L\in M^r(V)$ is said to be alternating if

1. $L(\alpha_1,\ldots,\alpha_r)=0$ if $\alpha_i=\alpha_j$ for $i\neq j$,
2. $L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})=\operatorname{sgn}(\sigma)L(\alpha_1,\ldots,\alpha_r)$ for all $\sigma\in S_r$.

The collection of all alternating r-linear forms is denoted by $\Lambda^r(V)$ and its a submodule of $M^r(V)$. Note that the second axiom follows from 1, but is important to keep it in mind.

Observe that $\Lambda^1(V)=M^1(V)=V^*$.

Remark 16.5.3.9. Consider $V=R^n$, a free R-module of rank n. We saw earlier that $\det\in M^n(V)$ is an n-linear form.. Theorem 16.5.2.8 shows that det is moreover an unique alternating form with $\det(e_1,\ldots,e_n)=1$. Thus, $\det\in\Lambda^n(V)\subseteq M^n(V)$ is the unique alternating n-linear form over V such that $\det(e_1,\ldots,e_n)=1$, i.e. $\Lambda^n(V)$ is a free R-module of rank 1.

Construction 16.5.3.10. Let V be an R-module. We now construct an R-linear map $\pi_r: M^r(V)\to\Lambda^r(V)$. For each $L\in M^r(V)$, define $L_\sigma\in M^r(V)$ given by $L_\sigma(\alpha_1,\ldots,\alpha_r)=L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)})$ for $(\alpha_1,\ldots,\alpha_r)\in V^r$. Consequently, we claim that the following map is well-defined:

$$\begin{aligned}\pi_r: M^r(V) &\longrightarrow \Lambda^r(V)\\ L &\longmapsto \sum_{\sigma\in S_r}\operatorname{sgn}(\sigma)L_\sigma.\end{aligned}$$

Indeed, we have to show that $\pi_r L$ is an alternating form. Let $(\alpha_1,\ldots,\alpha_r)\in V^r$ be such that $\alpha_i=\alpha_j$ for $i\neq j$. We wish to show that $\pi_r L(\alpha_1,\ldots,\alpha_r)=0$. Let $\tau=(ij)$ be the transposition swapping i and j. First observe that the map $S_r\to S_r$ given by $\sigma\mapsto\tau\sigma$ is a bijection. Consequently, if we let $\sigma_1,\ldots,\sigma_{\frac{n!}{2}}$ to be any $\frac{n!}{2}$ elements of S_r, then the rest $\frac{n!}{2}$ are given by $\tau\sigma_i$, $i=1,\ldots,n!/2$.

Consequently,

$$\begin{aligned}
\pi_r L(\alpha_1, \dots, \alpha_r) &= \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) \\
&= \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) \\
&= \sum_{i=1}^{\frac{n!}{2}} \operatorname{sgn}(\sigma_i) L(\alpha_{\sigma_i(1)}, \dots, \alpha_{\sigma_i(r)}) + \sum_{i=1}^{\frac{n!}{2}} \operatorname{sgn}(\tau\sigma_i) L(\alpha_{\tau\sigma_i(1)}, \dots, \alpha_{\tau\sigma_i(r)}) \\
&= \sum_{i=1}^{\frac{n!}{2}} \operatorname{sgn}(\sigma_i) L(\alpha_{\sigma_i(1)}, \dots, \alpha_{\sigma_i(r)}) + \sum_{i=1}^{\frac{n!}{2}} -\operatorname{sgn}(\sigma_i) L(\alpha_{\sigma_i(1)}, \dots, \alpha_{\sigma_i(r)}) \\
&= 0.
\end{aligned}$$

Hence, π_r is indeed an R-linear map from $M^r(V)$ into $\Lambda^r(V)$.

Finally note that if $L \in \Lambda^r(V)$, then $\pi_r L = r!L$ as $L_\sigma = \operatorname{sgn}(\sigma)L$ for any $\sigma \in S_r$.

Example 16.5.3.11. Let $V = R^n$ be the free R-module of rank n. Let $e_1, \dots, e_n \in V$ be the standard R-basis of V. Further, let $f_1, \dots, f_n \in M^1(V)$ be the associated dual basis. Note that for any $\alpha \in V$, we have $\alpha = f_1(\alpha)e_1 + \dots f_n(\alpha)e_n$. Then, we get an n-form

$$L = f_1 \otimes \dots \otimes f_n \in M^n(V).$$

Consequently we obtain an alternating n-form given by

$$\pi_r L = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(f_{\sigma(1)} \otimes \dots \otimes f_{\sigma(n)}).$$

Observe that for any $(\alpha_1, \dots, \alpha_n) \in V^n$, we obtain

$$\begin{aligned}
\pi_r L(\alpha_1, \dots, \alpha_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(f_{\sigma(1)} \otimes \dots \otimes f_{\sigma(n)}\right)(\alpha_1, \dots, \alpha_n) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(f_{\sigma(1)}(\alpha_1) \cdot \dots \cdot f_{\sigma(n)}(\alpha_n)\right).
\end{aligned}$$

This is exactly the determinant of the $n \times n$ matrix over R given by $A = (f_j(\alpha_i))$. That is, $\pi_r L = \det$.

The following properties of π_r will become important later on.

Proposition 16.5.3.12. *Let V be an R-module and $L \in M^r(V)$ and $M \in M^s(V)$ be r and s-forms over V respectively. Then,*

$$\pi_{r+s}(\pi_r(L) \otimes \pi_s(M)) = r!s!\pi_{r+s}(L \otimes M).$$

Proof. ***TODO*** *: Magnum tedium.* □

The above has a very nice and useful corollary.

Corollary 16.5.3.13. *Let V be a free R-module of rank n with $f_1, \dots, f_n \in V^*$ be a dual basis of V^*. Let $I \in X_r$ and $J \in X_s$ where X_r and X_s are the sets of r and s combinations of $\{1, \dots, n\}$, respectively, such that I and J are disjoint ($i_k \neq j_l$ for any $1 \leq k \leq r$, $1 \leq l \leq s$). Denote $D_I = \pi_r(f_I)$ and $D_J = \pi_s(f_J)$ where $f_I = f_{i_1} \otimes \dots \otimes f_{i_r} \in M^r(V)$ and $f_J = f_{j_1} \otimes \dots \otimes f_{j_s} \in M^s(V)$. Then,*

$$\pi_{r+s}(D_I \otimes D_J) = r!s!D_{I \amalg J}.$$

Proof. Follows immediately from Proposition 16.5.3.12 □

We now come to the main result about alternating forms.

Theorem 16.5.3.14. *Let V be a free module of rank n over R.*

1. *If $r > n$, then $\Lambda^r(V) = 0$.*
2. *If $0 \leq r \leq n$, then rank of $\Lambda^r(V)$ is ${}^nC_r$.*

Proof. (*Sketch*) Using Remark 16.5.3.4, statement 1 is straightforward. For 2, observe that we can write for $(\alpha_1, \dots, \alpha_r) \in V^r$, $r \leq n$ as follows

$$L(\alpha_1, \dots, \alpha_r) = \sum_{J=\{j_1,\dots,j_r\} \in X} L(e_J)(f_{j_1} \otimes \dots \otimes f_{j_r})(\alpha_1, \dots, \alpha_r)$$

where X is the set of all r-permutations of $\{1, \dots, n\}$ (as for any repeatitions, the corresponding term is 0). Now, partitioning the set X into classes in which permutations represent the same combination, we obtain an indexing set $\hat{X}$ of size ${}^nC_r$. Again, by the fact that L is alternating, we observe $\mathrm{sgn}(\sigma)L(e_{j_1}, \dots, e_{j_r}) = L(e_{j_{\sigma(1)}}, \dots, e_{j_{\sigma(r)}})$. Consequntly we may write the above sum as

$$L(\alpha_1, \dots, \alpha_r) = \sum_{J=\{j_1,\dots,j_r\} \in \hat{X}} L(e_{j_1}, \dots, e_{j_r}) \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) \left(f_{j_{\sigma(1)}} \otimes \dots \otimes f_{j_{\sigma(r)}}\right)(\alpha_1, \dots, \alpha_r).$$

Therefore denote for each $J \in \hat{X}$ the following

$$D_J = \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) \left(f_{j_{\sigma(1)}} \otimes \dots \otimes f_{j_{\sigma(r)}}\right).$$

One can observe that the D_J for each $J \in \hat{X}$ can alternatively be written as

$$D_J = \pi_r(f_{j_1} \otimes \dots \otimes f_{j_r}).$$

The above shows that D_J is in $\Lambda^r(V)$ and that it spans $\Lambda_r(V)$. The claim now is that these are also linearly independent. Indeed, that follows immediately by using the fact that f_js are dual basis of e_js. □

We can now abstractly obtain the determinant of a linear operator $T : V \to V$ on a free R-module V of rank n.

Corollary 16.5.3.15. *Let V be a free R-module of rank n and $T : V \to V$ be an R-linear operator. Then,*

1. *rank of $\Lambda^n(V) = 1$,*
2. *there exists a unique $c_T \in R$ such that for all $L \in \Lambda^n(V)$,*

$$L \circ T = c_T L.$$

This c_T is defined to be the determinant of the operator T.

Proof. Statement 1 follows from Theorem 16.5.3.14. For statement 2, one need only observe that $L \circ T$ is again an alternating n-tensor and then use statement 1. □

16.5.4 Exterior algebra over characteristic 0 fields

Let us first make the exterior algebra over characteristic 0 fields, before moving to arbitrary ring.

Definition 16.5.4.1. (**Wedge product**) Let K be a field of characteristic 0 and V be an R-vector space. For any $r, s \in \mathbb{N}$, define

$$\begin{aligned}\Lambda^r(V) \times \Lambda^s(V) &\longrightarrow \Lambda^{r+s}(V)\\ (L, M) &\longmapsto L \wedge M := \frac{1}{r!s!}\pi_{r+s}(L \otimes M).\end{aligned}$$

Observe that $D_I \wedge D_J = \frac{1}{r!s!}\pi_{r+s}(\pi_r(f_I) \otimes \pi_s(f_J)) = \frac{r!s!}{r!s!}\pi_{r+s}(f_I \otimes f_J)$ and the latter is either 0 if I and J have a common index or $D_{I \amalg J}$ if they are distinct. This follows from Proposition 16.5.3.12.

In the following result, we see that wedge product is a anti-commutative, distributive and associative operation.

Proposition 16.5.4.2. *Let V be a K-vector space over a field K of characteristic 0.*

1. *Let $\omega, \eta \in \Lambda^k(V), \phi \in \Lambda^l(V)$. Then, wedge product is distributive as*

$$(\omega + \eta) \wedge \phi = \omega \wedge \phi + \eta \wedge \phi,$$

2. *Let $\omega \in \Lambda^k(V), \eta \in \Lambda^l(V)$. Then, wedge product is anti-commutative as*

$$\omega \wedge \eta = (-1)^{kl}\eta \wedge \omega,$$

3. *Let $\omega \in \Lambda^k(V), \eta \in \Lambda^l(V), \phi \in \Lambda^m(V)$. Then, wedge product is associative as*

$$(\omega \wedge \eta) \wedge \phi = \omega \wedge (\eta \wedge \phi).$$

Proof. We need only check these identities on the basis elements $\{D_I\}$ of each $\Lambda^r(V)$.

1. Let $\omega = D_I, \eta = D_J$ and $\varphi = D_M$. Then,

$$\begin{aligned}(D_I + D_J) \wedge D_M &= \pi_{k+l}((D_I + D_J) \otimes D_M) = \pi_{k+l}(D_I \otimes D_M + D_J \otimes D_M)\\ &= \pi_{k+l}(D_I \otimes D_M) + \pi_{k+l}(D_J \otimes D_M) = D_I \wedge D_M + D_J \wedge D_M\end{aligned}$$

as required.

2. **TODO**. □

Using above, we come to the following definition.

Definition 16.5.4.3. (**Exterior algebra**) Let V be a K-vector space where K is a field of characteristic 0. Then the exterior algebra over V is

$$\begin{aligned}\Lambda(V) &= K \oplus \Lambda^1(V) \oplus \Lambda^2(V) \dots\\ &= K \oplus \bigoplus_{k \geq 1} \Lambda^k(V)\end{aligned}$$

where the product is given by wedge product which by Proposition 16.5.4.2 is associative, unital, distributive but non-commutative. This is also sometimes called the Grassmann algebra over V.

Remark 16.5.4.4. Observe that if V is of dimension n, then

$$\Lambda(V) = K \oplus \bigoplus_{k=1}^{n} \Lambda^k(V)$$

as all the higher forms are automatically 0. Consequently, the dimension of $\Lambda(V)$ by Theorem 16.5.3.14 is seen to be

$$\begin{aligned}\dim_K \Lambda(V) &= 1 + \sum_{k=1}^{n} {}^nC_k \\ &= \sum_{k=0}^{n} {}^nC_k \\ &= 2^n.\end{aligned}$$

Remark 16.5.4.5. Let V be a K-vector space of dimension n, where K is of characteristic 0. The exterior algebra $\Lambda(V)$ is a graded K-algebra of dimension 2^n over K. Indeed, the grading is correct as if $\omega \in \Lambda^k(V), \eta \in \Lambda^l(V)$, then $\omega \wedge \eta \in \Lambda^{k+l}(V)$.

16.5.5 Tensor, symmetric & exterior algebras

We now define the three algebras TM, SM and $\wedge M$ associated to a module M over R without any restriction imposed as earlier.

Definition 16.5.5.1 (TM, SM **and** $\wedge M$)**.** Let R be a ring and M be an R-module.

1. The tensor algebra over M is defined to be

$$TM = \bigoplus_{n\geq 0} T^n M$$

where $T^n M = M \otimes \ldots \otimes M$ n-times and $T^0 M = R$. This is a non-commutative graded R-algebra where the multiplication is given by tensor product.

2. The symmetric algebra over M is defined to be the quotient

$$\operatorname{Sym} M = TM/I = \bigoplus_{n\geq 0} S^n M$$

where I is the two-sided graded ideal of TM given by

$$I = \langle x \otimes y - y \otimes x | x, y \in M \rangle.$$

This is a commutative graded R-algebra where $\operatorname{Sym}^n M$ denotes $T^n M / I \cap T^n M$. To emphasize the base ring R, we sometimes write $\operatorname{Sym}_R(M)$ as well. Note that there is a canonical R-linear map $M \to \operatorname{Sym}_R(M)$ given by the composite $M \to TM \twoheadrightarrow \operatorname{Sym}(M)$ where $M \to TM$ is the inclusion in the first factor.

3. The exterior algebra over M is defined to be the quotient

$$\wedge M = TM/J = \bigoplus_{n\geq 0} \wedge^n M$$

where J is the two-sided graded ideal of TM given by

$$J = \langle x \otimes x \mid x \in M \rangle.$$

This is a skew-commutative[4] graded R-algebra where $\wedge^n M$ denotes $T^n M / J \cap T^n M$.

Symmetric algebra

We begin by discussing the universal property of symmetric algebra.

Proposition 16.5.5.2. *Let R be a ring and M be an R-module. Then the R-algebra $\mathrm{Sym}_R(M)$ satisfies the following universal property: for any commutative R-algebra S and an R-linear map $f : M \to S$, there exists a unique R-linear map of algebras $\tilde{f} : \mathrm{Sym}_R(M) \to S$ such that the following commutes:*

$$\begin{tikzcd} \mathrm{Sym}_R(M) \arrow[r, dashed, "\tilde{f}"] & S \\ M \arrow[u] \arrow[ur, "f"'] & \end{tikzcd}.$$

Thus, we have a natural bijection

$$\mathrm{Hom}_{\mathbf{Mod}(R)}(M, S) \cong \mathrm{Hom}_{\mathbf{Alg}(R)}(\mathrm{Sym}_R(M), S).$$

Using the above property, we have following easy conclusions.

Lemma 16.5.5.3 (Base change)**.** *Let R be a ring and $R \to S$ be an R-algebra. If M is an R-module, then we have an isomorphism of graded rings:*

$$\mathrm{Sym}_R(M) \otimes_R S \cong \mathrm{Sym}_R(M \otimes_R S)$$

Lemma 16.5.5.4. *Let R be a ring and M, M' be R-modules. Then*

$$\mathrm{Sym}_R(M \oplus M') \cong \mathrm{Sym}_R(M) \otimes_R \mathrm{Sym}_R(M').$$

Exterior algebra

The following are three important properties of exterior powers of modules.

Theorem 16.5.5.5. *Let R be a ring.*

1. *[Free modules]. The exterior power $\wedge^k(R^n)$ is a free module of rank ${}^nC_k$ with basis elements $\{e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}\}$ where $\{i_j\}_{j=1,\dots,k}$ is an increasing sequence from the set $1, \dots, n$ and $e_{i_j} = \delta_{i_j} \in R^n$.*
2. *[Tensor product]. Let $f : R \to S$ be a ring map and M be an R-module. Then,*

$$(\wedge^k M) \otimes_R S \cong \wedge^k (M \otimes_R S).$$

3. *[Binomial formula]. Let M, N be two R-modules. Then,*

$$\wedge^k(M \oplus N) \cong \sum_{i=0}^{k} \wedge^i M \otimes_R \wedge^{k-i} N.$$

[4] as J contains $x \otimes y + y \otimes x$ by opening $(x+y) \otimes (x+y) \in J$.

16.6 Field theory

We cover some basic material on Galois theory.

16.6.1 Finite extensions, algebraic extensions & compositum

Recall that a field extension K/F is said to be *finite* if K/F is a finite dimensional F-vector space and then we denote $[K : F] := \dim_F K$. It is said to be *algebraic* if for every $\alpha \in K$, there exists $p(x) \in F[x]$ such that $p(\alpha) = 0$, that is, the inclusion $F \hookrightarrow K$ is integral. Let $I = \{p(x) \in F[x] \mid p(\alpha) = 0\} \leq F[x]$ be an ideal. The generating element $m_{\alpha,F}(x)$ of I is called the *minimal polynomial* of $\alpha \in K$. Note that this is irreducible as I is a prime ideal as it is kernel of a map.

The main basic result connecting algebraic and finite extensions is that finitely generated algebraic extensions are equivalent to finite extensions. This is immediate from Proposition 16.7.1.9, but we give an elementary proof. We first begin by elementary observations.

Theorem 16.6.1.1. *Let K/F be a field extension and $\alpha \in K$.*

1. *If K/F is finite, then it is algebraic.*
2. *If $K/L/F$ are extensions, then*
$$[K : F] = [K : L] \cdot [L : F]$$
where $[K : L]$ or $[L : F]$ is infinity if and only if $[K : F]$ is infinity.
3. *If $\alpha_1, \dots, \alpha_n$ are algebraic over F, then $F(\alpha_1, \dots, \alpha_n) = F[\alpha_1, \dots, \alpha_n]$.*
4. *We have $[F(\alpha) : F] = \deg m_{\alpha,F}$.*
5. *The extension $F(\alpha_1, \dots, \alpha_n)/F$ is algebraic if and only if $\alpha_1, \dots, \alpha_n$ are algebraic over F.*
6. *K/F is a finite-type algebraic extension if and only if K/F is finite.*
7. *If K/L and L/F are both algebraic, then K/F is algebraic.*
8. *The set of all algebraic elements in K over F forms a subfield of K containing F denoted $K^{\mathrm{alg}/F}$.*

Proof. 1. Pick any element $x \in K$ and consider $\{1, x, x^2, \dots\}$. Finiteness of K/F makes sure that there is a finite subset of above which is linearly depenedent.

2. Take bases of K/L and L/F and consider their pairwise product. One sees that this new collection is linearly independent and its F-span is K.

3. As $F[\alpha]$ is a field as it is isomorphic to $F[x]/\langle m_{\alpha,F}(x)\rangle$ and $m_{\alpha,F}(x)$ is irreducible. By universal property of quotients, we get $F[\alpha] = F(\alpha)$. By induction, we wish to show that $F(\alpha_1, \dots, \alpha_{n-1})[\alpha_n] = F(\alpha_1, \dots, \alpha_{n-1})(\alpha_n) = F(\alpha_1, \dots, \alpha_{n-1}, \alpha_n)$, which completes the proof.

4. We have $F(\alpha) = F[\alpha] = \frac{F[x]}{m_{\alpha,F}(x)}$ and this is of dimension $\deg m_{\alpha,F}(x)$ over F.

5. Forward is immediate. For converse, proceed by induction. Clearly, $F(\alpha_1)/F$ is algebraic as it is finite. Composition of finite is finite, so $F(\alpha_1, \dots, \alpha_n)/F$ is finite, thus algebraic.

6. Forward is the only non-trivial side. Let $K = F(\alpha_1, \dots, \alpha_n)$ and by algebraicity, α_i are algebraic. Now $F(\alpha_1)/F$ is finite as algebraic. By induction, we get the result.

7. Pick $\alpha \in K$ and consider $m_{\alpha,L}(x) \in L[x]$ as $m_{\alpha,L}(x) = x^n + c_{n-1}x^{n-1} + \dots + c_1 x + c_0$, $c_i \in L$. Then, consider $F(c_0, \dots, c_{n-1}) \subseteq L$. As L/F is algebraic, thus $c_i \in L$ are algebraic and thus by previous, we get $F(c_0, \dots, c_{n-1})/F$ is algebraic and finite. As $F(c_0, \dots, c_{n-1})(\alpha)/F(c_0, \dots, c_{n-1})$ is algebraic as it is finite, thus $F(c_0, \dots, c_{n-1}, \alpha)/F$ is algebraic as it is composite of two finite extensions.

8. Indeed, pick any two algebraic elements $\alpha, \beta \in K$ over F. Then $F(\alpha, \beta)$ is an algebraic extension over F and thus $F(\alpha, \beta) \subseteq K^{\mathrm{alg}/F}$. $\square$

Next, we define compositum, the smallest field containing two subfields.

Definition 16.6.1.2 (**Compositum of fields**). Let F, K be two fields in a field L. Then compositum of F and K in L is the smallest field in L containing both F and K. This is denoted by $F \cdot K$.

The following are the main results for compositum. We will see more later when needed.

Theorem 16.6.1.3. *Let K/F be a field extension and $K_1, K_2 \subseteq K$ be two subfields containing F. Then,*

1. *If $K_1 = F(\alpha_1, \dots, \alpha_n)$ and $K_2 = F(\beta_1, \dots, \beta_m)$, then $K_1 \cdot K_2 = F(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$.*
2. *If K_1/F and K_2/F are algebraic, then $K_1 \cdot K_2/F$ is algebraic.*
3. *If K_1/F and K_2/F are finite, then $K_1 \cdot K_2/F$ is finite.*
4. *If $[K_1 : F]$ and $[K_2 : F]$ are coprime, then $[K_1 \cdot K_2 : F] = [K_1 : F] \cdot [K_2 : F]$.*
5. *We have $[K_1 \cdot K_2 : F] \leq [K_1 : F] \cdot [K_2 : F]$.*

Proof. 1. It is clear that $K_1 \cdot K_2 \supseteq F(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$ since $K_1 \cdot K_2$ contains both K_1, K_2 and F. For the converse, as $K_1 \cdot K_2$ is the smallest field containing both K_1 and K_2 therefore $K_1 \cdot K_2 \subseteq F(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$.

2. Let L be the algebraic closure of F in $K_1 \cdot K_2$. By hypothesis, $L \supseteq K_1, K_2$. Thus $L \supseteq K_1 \cdot K_2$.

3. By Theorem 16.6.1.1, 6, $K_1 = F(\alpha_1, \dots, \alpha_n)$ and $K_2 = F(\beta_1, \dots, \beta_m)$ where α_i, β_j are algebraic elements over F. By item 1, $K_1 \cdot K_2 = F(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$ is a finitely generated algebraic extension, thus finite, as required.

4. Since we have

$$\begin{aligned} [K_1 \cdot K_2 : F] &= [K_1 \cdot K_2 : K_1][K_1 : F] \\ &= [K_1 \cdot K_2 : K_2][K_2 : F]. \end{aligned}$$

By hypothesis, $[K_1 \cdot K_2 : F]$ is a multiple of $[K_1 : F] \cdot [K_2 : F]$. Thus we redude to showing $[K_1 \cdot K_2 : F] \leq [K_1 : F] \cdot [K_2 : F]$. Note by above equations, it suffices to show that

$$[K_1 \cdot K_2 : K_1] \leq [K_2 : F].$$

To this end, let $\alpha_1, \dots, \alpha_n \in K_2$ be an F-basis of K_2. It thus suffices to show that K_1-span of $\alpha_1, \dots, \alpha_n$ is whole of $K_1 \cdot K_2$, that is, we wish to show

$$L := K_1 \cdot \alpha_1 + \dots + K_1 \cdot \alpha_n = K_1 \cdot K_2.$$

Note that it suffices to show that L is a field containing both K_1 and K_2. Indeed, the fact that L contains K_2 is immediate as L contains F and $\alpha_1, \dots, \alpha_n$. Further L contains K_1 as L contains 1 since L contains K_2 and that it is a K_1-vector space. Thus, $L \supseteq K_1, K_2$. We thus reduce to showing that L is a field.

To this end, observe that if $l \in L$, then $l = c_1\alpha_1 + \dots + c_n\alpha_n$ for $c_i \in K_1$. Now, $l \in K_2(c_1, \dots, c_n)$. Thus $l^{-1} \in K_2(c_1, \dots, c_n) = K_2[c_1, \dots, c_n]$, that is, l^{-1} is a polynomial in c_i with coefficients in K_2. But any element of K_2 is an F-linear combination of $\alpha_1, \dots, \alpha_n$. As $K_1 \supseteq F$, therefore l^{-1} is a linear combination of $\alpha_1, \dots, \alpha_n$ with coefficients in K_1 (powers of c_i multiplied by elements of F, so in K_1). Thus, $l^{-1} \in L$, as needed. The fact that L is multiplicatively closed is immediate. This completes the proof.

5. Follows from proof of item 4 above. $\square$

We now see that a finite algebra over a domain which is a domain induces a finite extension of fraction fields.

Lemma 16.6.1.4. *Let $B \hookrightarrow A$ be a finite B-algebra where both A, B are domains. Then $Q(A)$ is a finite extension of $Q(B)$.*

Proof. Let $\alpha_1, \dots, \alpha_n \in A$ be a generating set of A as a B-module and let $\varphi : B \hookrightarrow A$ be the structure map of the finite B-algebra structure on A. Now let $S = B - \{0\}$. Now we get a map $S^{-1}\varphi : Q(B) \hookrightarrow S^{-1}A$. This is a finite map since $S^{-1}A$ as the $Q(B)$ span of $\alpha_1, \dots, \alpha_n$ in $S^{-1}A$ is $S^{-1}A$. To complete the proof, we need only show that the natural inclusion $S^{-1}A \hookrightarrow Q(A)$ given by $\frac{a}{b} \mapsto \frac{a}{b}$ is a finite map. We see something stronger: $Q(A) = S^{-1}A$. Indeed, this is true because $S^{-1}A$ is a field containing A as $S^{-1}A$ is a domain which is finite over the field $Q(B)$, so that by Lemma 16.7.1.14, we get that $S^{-1}(A)$ is a field. As it contains A, so it also contains $Q(A)$. This completes the proof. $\square$

16.6.2 Maps of field extensions

There are some important results which allow us to extend a field homomorphism from a smaller field to a bigger field. These come in handy while discussing splitting fields and algebraic closures.

Proposition 16.6.2.1 (Extension-I)**.** *Let $\varphi : F \to F'$ be a field isomorphism. Let $p(x) \in F[x]$ be an irreducible polynomial and let $\varphi(p(x)) \in F'[x]$ be the irreducible polynomial in the image. If α is a root of $p(x)$ and β is a root of $\varphi(p(x))$, then there exists a field isomorphism $\tilde{\varphi} : F(\alpha) \to F'(\beta)$ mapping $\alpha \mapsto \beta$ and extending φ:*

$$\begin{tikzcd} F(\alpha) \arrow[r, "\tilde{\varphi}", "\cong"'] & F'(\beta) \\ F \arrow[u, hook] \arrow[r, "\cong", "\varphi"'] & F' \arrow[u, hook] \end{tikzcd}.$$

Proof. Since $F(\alpha) = F[x]/p(x)$ and $F'(\beta) = F'[x]/\varphi(p(x))$, therefore we need only construct an isomorphism between them via φ which takes $\bar{x}$ to $\bar{x}$ (as $\bar{x}$ in $F(\alpha)$ is the root of $p(x)$ in $F(\alpha)$ and similarly for $F(\beta)$).

Indeed, consider the map

$$\begin{aligned} \varphi : F[x] &\to F'[x] \\ x &\mapsto x. \end{aligned}$$

Then, we get $\tilde{\varphi} : \frac{F[x]}{\varphi^{-1}(\varphi(p(x)))} \xrightarrow{\cong} \frac{F'[x]}{\varphi(p(x))}$. This completes the proof. $\square$

Corollary 16.6.2.2. *If $p(x) \in F[x]$ is irreducible and $\alpha \neq \beta$ are two roots, then there is an isomorphism*

$$\begin{aligned} F(\alpha) &\longrightarrow F(\beta) \\ \alpha &\longmapsto \beta \end{aligned}$$

which is id *on F.*

Proof. Use $\varphi = \mathrm{id}_F$ with $F' = F$ on Proposition 16.6.2.1 to get the result. $\square$

We next show that transcendental elements are mapped to transcendental elements under a field homomorphism.

Proposition 16.6.2.3. *Let $\varphi : F \to F'$ be a morphism of fields. If K/F is a field extension, $\psi : K \to F'$ is a morphism extending φ, then the following are equivalent:*

1. *$\alpha \in K$ is transcendental over F,*
2. *$\psi(\alpha) \in F'$ is transcendental over $\varphi(F) \subseteq F'$.*

Proof. The main observation is that for transcendental element $\alpha \in K$ over F, we have that $F[\alpha]$ is isomorphic to polynomial ring $F[x]$. Using this, we consider the restriction $\psi : F(\alpha) \to F'$. Note that $\alpha \in F(\alpha)$ is transcendental over F if and only if $\text{Ker}\,(\psi) = 0$. Further $\psi(\alpha)$ is transcendental over $\psi(F)$ if and only if $\text{Ker}\,(\psi) = 0$. We win. $\square$

16.6.3 Splitting fields & algebraic closure

Given a polynomial, we will now construct the smallest field where that polynomial splits into linear factors. We will then see that splitting fields are exactly what are called normal extensions.

Definition 16.6.3.1 (Splitting field). Let $f(x) \in F$ be a field and $f(x) \in F[x]$ be a polynomial. The splitting field of $f(x)$ over F is the smallest field extension K/F such that $f(x) \in K[x]$ is product of linear factors, that is, K is the smallest field containing all roots of $f(x)$.

Theorem 16.6.3.2. *Splitting field exists.*

Proof. Let $f(x) \in F$ be a field and $f(x) \in F[x]$ be a polynomial. We wish to construct the smallest field K/F containing all roots of F. We induct over $\deg f(x) = n$. If $n = 1$, then $K = F$ will do. Suppose for every polynomial $g(x)$ of degree $n-1$ or lower has a splitting field, which we denote by K_g/F. Pick $f(x) \in F[x]$ be of degree n. We wish to construct the splitting field of $f(x)$. We have two cases. If $f(x)$ is reducible, then $f(x) = g(x)h(x)$ where $\deg g, \deg h < n$. We thus have splitting fields K_g and K_h for g and h respectively. We claim that $K_g \cdot K_h$ is a splitting field of $f(x)$. Indeed, $K_g \cdot K_h$ contains all roots of $f(x)$ so splitting field is a subfield of $K_g \cdot K_h$. But since splitting field of $f(x)$ also contains roots of $g(x)$ and $h(x)$, it follows that it must contain K_g and K_h and thus $K_g \cdot K_h$ as well. Hence splitting field is exactly $K_g \cdot K_h$.

On the other hand if $f(x)$ is irreducible, then let $K = \frac{F[x]}{\langle f(x) \rangle}$ which is a finite extension of F. Now, K has atleast one root of $f(x)$, namely $\bar{x}$, which we label as $\alpha \in K$. Thus, we have that $f(x) = (x - \alpha)g(x)$ in $K[x]$. Thus $g(x) \in K[x]$ is of degree $n-1$. Hence by inductive hypothesis, there exists a field $L_g/K/F$ such that $g(x)$ splits into linear factor/L_g contains all roots of $g(x)$. Thus $L_g(\alpha)$ contains all roots of $f(x)$. We claim that $L_g(\alpha)$ is contains a splitting field of $f(x)$. Indeed, we may take intersection of all sub-fields of $L_g(\alpha)$ which contains all roots of $f(x)$. Such a collection is non-empty as $L_g(\alpha)$ contains all roots of $f(x)$. As intersection of subfields is a subfield, we win the induction step. $\square$

We now show that splitting fields are unique upto isomorphism.

Proposition 16.6.3.3 (Extension-II). *Let $\varphi : F \to F'$ be a field isomorphism and $f(x) \in F[x]$ be a polynomial. Let $\varphi(f(x)) \in F'[x]$ be the image of $f(x)$ under φ. Then, φ lifts to an isomorphism $\tilde{\varphi} : K \to$*

K' where K/F is the splitting field of $f(x)$ and K'/F' is the splitting field of $\varphi(f(x))$:

$$\begin{tikzcd} K \arrow[r, "\tilde{\varphi}", "\cong"'] & K' \\ F \arrow[u, hook] \arrow[r, "\cong", "\varphi"'] & F' \arrow[u, hook] \end{tikzcd}.$$

Proof. We will induct on degree of $f(x)$. If $\deg f(x) = 1$, then F has the root of f and thus we may take $\tilde{\varphi}$ to be φ itself. Let $\deg f = n$ and suppose that for any polynomial of degree $n-1$ or lower over any extension of F, we have the required map. Let $f(x) = p(x)g(x)$ where $p(x) \in F[x]$ is an irreducible factor of $f(x)$. Thus $\deg p(x) \leq n-1$. Now, let α be a root of $p(x)$ and α' be a root of $\varphi(p(x))$. Thus by Extension-I (Proposition 16.6.2.1), it follows that we have an extension $\chi : F(\alpha) \to F'(\alpha')$ which extends φ. Now consider $h(x) = f(x)/x - \alpha$ in $F(\alpha)[x]$. Then, $h(x)$ has degree $n-1$ over $F(\alpha)$, so by inductive hypothesis, we get an extension $\tilde{\varphi} : K_h \to K'_h$ where $K_h/F(\alpha)$ and $K'_h/F'(\alpha')$ are splitting fields of $h(x)$ and $\chi(h(x))$ respectively. We claim that K_h is the splitting field of $f(x)$. Indeed, K_h has all roots of $f(x)$, so it contains the splitting field. But roots of $h(x)$ are just those of $f(x)$ except α, so K_h is the splitting field of $f(x)$. This completes the proof. □

Algebraic closure

We now discuss some basic properties of algebraic closure. Note that there is a subtlety to the definition of an extension being algebraically closed.

Definition 16.6.3.4 (Algebraically closed fields & extensions). A field K is algebraically closed if every polynomial in $K[x]$ has a root. An extension K/F is called an algebraically closed extension if K/F is algebraic and K is algebraically closed. In this case, K is called the algebraic closure of F.

Remark 16.6.3.5. The linguistic subtlety here is that $\mathbb{C}/\mathbb{Q}$ is not algebraically closed extension as it is not algebraic. But $\bar{\mathbb{Q}}/\mathbb{Q}$ is an algebraically closed extension.

We will omit the statement that an algebraic closed extension of any field exists as it can be found in any standard book. We however state the following important results about equivalence conditions for a field to be algebraically closed.

Theorem 16.6.3.6. *Let F be a field. Then the following are equivalent:*
1. *F is algebraically closed.*
2. *Only irreducible polynomial in $F[x]$ are linear.*
3. *If K/F is algebraic, then $K = F$.*

Proof. The only non-trivial part is that of 3. ⇒ 1. Indeed, pick any $f(x) \in F[x]$. Then, consider the splitting field K/F of $f(x)$. As K/F is finite, therefore K/F is algebraic and thus by hypothesis we have $K = F$. It follows that F has all roots of F, as required. □

16.6.4 Separable, normal extensions & perfect fields

Let us begin with definitions.

Definition 16.6.4.1 (**Separable polynomials & extensions**). A polynomial $f(x) \in F[x]$ is said to be separable if $f(x)$ has no repeated roots. That is, there doesn't exists $\alpha \in \bar{F}$ such that $(x - \alpha)^2 | f(x)$. An extension K/F is said to be separable if it is algebraic and for all $\alpha \in K$, the minimal polynomial $m_{\alpha,F}(x) \in F[x]$ is separable.

Definition 16.6.4.2 (**Normal extensions**). An extension K/F is said to be normal if it is algebraic and for all $\alpha \in K$, the minimal polynomial $m_{\alpha,F}(x) \in F[x]$ has all roots in K and is thus a product of linear factors in $K[x]$.

Remark 16.6.4.3. Note that if K/F is normal, then K contains the splitting field of all $f(x) \in F[x]$. Thus every splitting field of some $f(x) \in F[x]$ is an intermediate extension of K/F.

Definition 16.6.4.4 (**Frobenius & perfect fields**). Let K be a field of characteristic $p > 0$. Then the Frobenius is the field map $\mathrm{Fr} : K \to K$ mapping $x \mapsto x^p$. A field K is perfect if either $\mathrm{char}(K) = 0$ or the Frobenius $\mathrm{Fr} : K \to K$ is an isomorphism.

Basic properties

For finite normal extensions, we essentially have the following as the most important observation.

Theorem 16.6.4.5. *Let K/F be a finite normal extension. If $\alpha \in K$ and $Z(m_{\alpha,F}(x)) \subseteq K$ is the set of all F-conjugates of α, then* $\mathrm{Aut}\,(K/F)$ *acts on $Z(m_{\alpha,F}(x))$ transitively.*

We prove this using the following statements.

Proposition 16.6.4.6. *Let K/F be an algebraic extension and $\alpha \in K$. Then,*

1. *For any $\sigma \in \mathrm{Aut}\,(K/F)$, $\sigma(\alpha) \in K$ is an F-conjugate of α.*
2. *If $\beta \in \bar{K}$ is an F-conjugate of α, then there exists a map*

$$\sigma : K \longrightarrow \bar{K}$$

   such that $\sigma(\alpha) = \beta$, $\sigma|_F = \mathrm{id}$ and $\sigma(\alpha) = \beta$.
3. *If K/F is a finite normal extension and $\sigma : K \to \bar{K}$ is a field homomorphism such that $\sigma|_F = \mathrm{id}_F$, then $\sigma(K) = K$. That is, if $\sigma : K \to \bar{K}$ is an F-homomorphism, then $\sigma \in \mathrm{Aut}\,(K/F)$.*

Proof. 1. Apply σ on $m_{\alpha,F}(\alpha) = 0$ to get the desired result.

2. By Extension-I (Proposition 16.6.2.1), we have an extension of $\mathrm{id} : F \to F$ denoted $\chi : F(\alpha) \to F(\beta)$. By a generalization of Extension-II (Proposition 16.6.3.3) which gives us the same result but for splitting fields of arbitrary collection, we get an extension of χ to $\tilde{\sigma} : \bar{K} \to \bar{K}$ extending χ. Defining $\sigma = \tilde{\sigma}|_K : K \to \bar{K}$, we get that σ extends id_F and $\sigma(\alpha) = \beta$, as required.

3. Pick any $\alpha \in K$. We first wish to show that $\sigma(\alpha) \in K$. By item 1, $\sigma(\alpha) \in \bar{K}$ is an F-conjugate of α. As the minimal polynomial $m_{\alpha,F}(x) \in F[x]$ splits linearly in K, this shows that $\sigma(\alpha) \in K$,

hence showing that $\sigma(K) \subseteq K$. To show equality, we need only show that $[K : \sigma(K)] = 1$. Indeed, since

$$[K : F] = [K : \sigma(K)] \cdot [\sigma(K) : F] < \infty$$

and since $\sigma : K \to \sigma(K)$ is an F-isomorphism, therefore $[K : F] = [\sigma(K) : F]$. It follows that $[K : \sigma(K)] = 1$, as required. $\square$

Theorem 16.6.4.7 is now immediate.

Proof of Theorem 16.6.4.7. Pick any two root $\beta \in Z(m_{\alpha,F}(x))$. It suffices to show that there exists $\sigma \in \mathrm{Aut}\,(K/F)$ which maps $\alpha \mapsto \beta$. Indeed, by Proposition 16.6.4.6, 2 & 3, we have such an F-automorphism. $\square$

Characterization of normality and separability

Our goal is to study two questions. First is to understand the relationship between splitting fields and normal extensions (we will see that they are equivalent). Second is to understand the relationship between separability and the automorphisms of the extension.

Understanding these two problems will give us the tool which will allow us to show when a field extension is separable or normal, which will come in handy while doing Galois theory.

Let us begin by the first question.

Theorem 16.6.4.7. *Let K/F be an extension. Then the following are equivalent:*
1. *K/F is a splitting field of some $S \subseteq F[x]$.*
2. *K/F is a normal extension.*

Another important characterization of normal extensions in the finite setting is the following.

Theorem 16.6.4.8. *Let K/F be a finite extension. Then the following are equivalent:*
1. *K/F is a normal extension.*
2. *For every $\sigma \in \mathrm{Hom}_F\left(K, \bar{F}\right)$, we have $\sigma(K) = K$ where note that $\bar{F} = \bar{K}$.*

Proof. (1. $\Rightarrow$ 2.) This is the content of Proposition 16.6.4.6, 3.

(2. $\Rightarrow$ 1.) Pick any $\alpha \in K$. We wish to show that every F-conjugate β of α in $\bar{F} = \bar{K}$ is in K. Indeed, by Proposition 16.6.4.6, 2, it follows that there exists $\sigma : K \to \bar{F}$ such that $\sigma(\alpha) = \beta$. By our hypothesis, $\sigma(K) = K$, thus, $\sigma \in \mathrm{Aut}\,(K/F)$. Hence, $\beta \in K$, as required. $\square$

We now build towards answering the second question.

Definition 16.6.4.9 (Separable degree). Let K/F be a finite extension. Then the separable degree of K/F is defined to be

$$[K : F]_s = \left|\mathrm{Hom}_F\left(K, \bar{F}\right)\right|$$

where $\mathrm{Hom}_F\left(K, \bar{F}\right)$ is finite in size since K/F is finite.

There's a tower law for separable degree as well.

Proposition 16.6.4.10. *Let $L/K/F$ be field extensions and L/F be finite. Then,*

$$[L : F]_s = [L : K]_s \cdot [K : F]_s.$$

The following is an easy lemma.

Lemma 16.6.4.11. *Let K/F be a finite extension. Then*

$$[K : F]_s \leq [K : F].$$

Proof. For $K = F(\alpha)$, this is immediate as any $\sigma \in \mathrm{Hom}_F\left(K, \bar{F}\right)$ takes α to some F-conjugate of α. Thus, $[K : F]_s = \#$ conjugates of α in $\bar{F} \leq \deg m_{\alpha,F}(x) = [K : F]$. Now proceed by induction via tower law (Proposition 16.6.4.10). □

Theorem 16.6.4.12. *Let K/F be a field extension. Then the following are equivalent:*
1. *$[K : F]_s = [K : F]$.*
2. *K/F is a separable extension.*

We can now prove that composition of separable extensions is separable.

Lemma 16.6.4.13. *Let L/K and K/F be separable extensions. Then L/F is separable.*

Proof. We have $[L : F]_s = [L : K]_s \cdot [K : F]_s$ by tower law (Proposition 16.6.4.10). By Theorem 16.6.4.12 we have $[L : F]_s = [L : K] \cdot [K : F] = [L : F]$ and thus we conclude that L/F is separable. □

Another important criterion for separability of a polynomial is to check its derivatives. This is useful in positive characteristic settings.

Lemma 16.6.4.14. *Let $f(x) \in F[x]$ be a polynomial where F is a field. If $f(x)$ is irreducible, then the following are equivalent.*
1. *$f(x)$ is separable.*
2. *$f'(x) \neq 0$.*

Proof. (1. $\Rightarrow$ 2.) If $f'(x)$ is zero, then $f(x)$ and $f'(x)$ will have a common root, which implies that $f(x)$ has a repeated root, a contradiction.

(2. $\Rightarrow$ 1.) Suppose $f(x)$ is inseparable, that is, it has a repeated root. This is equivalent to stating that there is a non-trivial common factor of $f'(x)$ and $f(x)$, say $p(x)$, which we may assume to be the gcd of $f(x)$ and $f'(x)$. As $f(x)$ is irreducible and $p(x)|f(x)$, therefore $p(x) = f(x)$. But $p(x)|f'(x)$, so $f(x)|f'(x)$. This is not possible as $\deg f' \leq \deg f - 1$. □

Using the above theorems, we obtain the following useful criterion usually used in induction steps and allows us to reduce to checking the separability and normality for a single element.

Proposition 16.6.4.15. *Let K/F be a field extension and $\alpha \in K$ be an algebraic element. If the minimal polynomial $m_{\alpha,F}(x) \in F[x]$*
1. *is a separable polynomial, then $F(\alpha)/F$ is a separable extension,*
2. *has all roots in $F(\alpha)$, then $F(\alpha)/F$ is a normal extension.*

Proof. 1. Note that since $m_{\alpha,F}(x)$ is separable, we get

$$[F(\alpha) : F]_s = |S(\mathrm{id}, F(\alpha)/F)| = \#\text{conjugates of } \alpha = \deg m_{\alpha,F}(x) = [F(\alpha) : F].$$

By Theorem 16.6.4.12, we win.

2. We claim that $F(\alpha)/F$ is the splitting field of $m_{\alpha,F}(x)$ in this case. Indeed, $F(\alpha)/F$ is the smallest field containing F and α. By hypothesis, it contains all the roots of $m_{\alpha,F}(x)$, of which α is one. It follows that $F(\alpha)/F$ is the smallest field containing all roots of $m_{\alpha,F}(x)$, as required. □

One can further define the separable closure of algebraic extensions.

Definition 16.6.4.16 (**Separable closure**)**.** Let K/F be an algebraic extension. Consider the set of elements

$$L = \{\alpha \in K \mid \alpha \text{ is separable over } F\}.$$

Then L is a field and L/F is said to be the separable closure of F in K.

Remark 16.6.4.17. Indeed, separable closure L of F in K is a field as if $\alpha, \beta \in L$ then $F(\alpha, \beta)/F$ is a separable extension by Proposition 16.6.4.15, 1 (applied twice). It follows that $F(\alpha, \beta) \subseteq L$ and thus L contains $\alpha \pm \beta, \alpha \cdot \beta$ and α^{-1}, β^{-1}.

Perfect fields

There are essentially two main results here. The first one saying any finite field is perfect and the second saying some important equivalent criterion to be perfect.

Theorem 16.6.4.18 (Finite fields are perfect)**.** *Let $\mathbb{F}_{p^n}$ be a finite field of characteristic p. Then $\mathbb{F}_{p^n}$ is perfect.*

Theorem 16.6.4.19 (Perfect equivalence theorem)**.** *Let F be a field. Then the following are equivalent:*

1. *F is a perfect field.*
2. *Every algebraic extension of F is separable.*
3. *Every irreducible polynomial in $F[x]$ is separable.*

16.6.5 Galois extensions

For simplicity, let us only work with finite Galois extensions.

Definition 16.6.5.1 (**Galois extensions & Galois group**)**.** An extension K/F is Galois if it is finite, separable and normal. That is, for all $\alpha \in K$, the minimal polynomial $m_{\alpha,F}(x) \in F[x]$ has all roots in K and each of them is distinct. The Galois group of a Galois extension K/F, denoted $\mathrm{Gal}\,(K/F)$, is defined to be the automorphism group $\mathrm{Aut}\,(K/F)$.

Let us first see that every splitting field of a separable polynomial is a Galois extension over the base.

Proposition 16.6.5.2. *Let F be a field and $f(x) \in F[x]$ be a separable polynomial. Let K/F be the splitting field of $f(x)$ over F. Then K/F is a Galois extension and* $\mathrm{Gal}\,(K/F)$ *is called the Galois group of the polynomial $f(x)$.*

Proof. We first establish that K/F is Galois. Indeed K/F is finite as it is a splitting field of a polynomial. As it is a splitting field, so it is normal (Theorem 16.6.4.7). To show separability, it suffices to show that the separable degree $[K:F]_s = [K:F]$ (Theorem 16.6.4.12). To this end, we first have $K = F(\alpha_1, \dots, \alpha_n)$ for $\alpha_i \in K$ elements algebraic over F. Consequently, by the tower law for separable degree (Proposition 16.6.4.10), we obtain

$$[K:F]_s = [K:F(\alpha_1, \dots, \alpha_{n-1})]_s \cdot \dots \cdot [F(\alpha_1, \alpha_2) : F(\alpha_1)]_s \cdot [F(\alpha_1) : F]_s.$$

By Proposition 16.6.4.15, it suffices to show that $m_{\alpha_i, F(\alpha_1, \dots, \alpha_{i-1})}(x) \in F(\alpha_1, \dots, \alpha_{i-1})[x]$ is a separable polynomial for each i. Indeed, since $f(\alpha_i) = 0$, thus $m_{\alpha_i, F(\alpha_1, \dots, \alpha_{i-1})}(x) | f(x)$ in $F(\alpha_1, \dots, \alpha_{i-1})[x]$. As $f(x)$ is separable, and $\overline{F(\alpha_1, \dots, \alpha_{i-1})} = \overline{F}$, it follows that $m_{\alpha_i, F(\alpha_1, \dots, \alpha_{i-1})}(x)$ is separable, as required. $\square$

There's a converse to the above result as well.

Proposition 16.6.5.3. *Let K/F be a Galois extension. Then there exists $f(x) \in F[x]$ a separable polynomial whose splitting field is K.*

Proof. As K/F is Galois, therefore finite and hence we may write $K = F(\alpha_1, \dots, \alpha_n)$ for $\alpha_i \in K$ such that no α_i and α_j are conjugate for $i \neq j$ (by normality of K/F, this is possible). As K/F is separable, therefore each $m_{\alpha_i, F}(x) \in F[x]$ is a distinct separable polynomial. Let $f(x) = \prod_{i=1}^n m_{\alpha_i, F}(x)$. This is a separable polynomial as no α_i are conjugates. Moreover, $f(x)$ splits into linear factors over K. It follows that the splitting field of $f(x)$, denoted L, is contained in K. As L contains each of the α_i and F, it follows that $L = K$, as required. $\square$

Thus, for the purposes of clarity, we summarize the above two results in the following corollary.

Corollary 16.6.5.4. *Let K/F be a field extension. Then the following are equivalent.*

1. *K/F is a Galois extension.*
2. *There is a separable polynomial $f(x) \in F[x]$ whose splitting field is K.*

Proof. Follows from Proposition 16.6.5.2 and 16.6.5.3. $\square$

We have the following equivalent criterion to be Galois.

Theorem 16.6.5.5. *Let K/F be a finite extension. Then the following are equivalent:*

1. *K/F is a Galois extension.*
2. $|\mathrm{Aut}\,(K/F)| = [K:F]$.

An extremely important result to keep in mind is the following, telling us that a fixed field by a finite subgroup of the automorphism group always gives a Galois extension(!)

Theorem 16.6.5.6. *Let K be a field and $G \leq \mathrm{Aut}\,(K)$ be a finite subgroup. Then,*

1. *The extension K/K^G is a Galois extension.*
2. *The Galois group of K/K^G is equal to G;*

$$\mathrm{Gal}\left(K/K^G\right) = G.$$

Théorème fondamental de la théorie de Galois

Theorem 16.6.5.7 (Fundamental theorem)**.** *Let K/F be a Galois extension with Galois group $G = \mathrm{Gal}\,(K/F)$. Then the maps*

$$\begin{tikzcd} \{L \mid K/L/F \text{ is an intermediate extension}\} \arrow[d, shift left, "\mathrm{Gal}(K/-)"] \\ \{H \mid H \leq G \text{ is a subgroup}\} \arrow[u, shift left, "K^{(-)}"] \end{tikzcd}$$

establish a bijection. Moreover, we have the following:

1. *For any intermediate $K/L/F$, the extension K/L is a Galois extension.*
2. *Both the maps above are antitone, i.e. they reverse the order.*
3. *For any intermediate extension $K/L/F$, the following are equivalent:*
   (a) *L/F is a Galois extension.*
   (b) *$\mathrm{Gal}\,(K/L)$ is a normal subgroup of G and in this case,*
   $$\mathrm{Gal}\,(L/F) \cong \frac{G}{\mathrm{Gal}\,(K/L)}.$$
4. *For any intermediate extension $K/L/F$*[5] *we have a bijection (where $\bar{F}$ is an algebraic closure of F containing K)*
   $$[L : F]_s = \mathrm{Hom}_F\left(L, \bar{F}\right) = \{\sigma : L \to \bar{F} \mid \ \sigma|_F = \mathrm{id}_F\} \cong \frac{G}{\mathrm{Gal}\,(K/L)}$$
   where the RHS is the set of cosets of $\mathrm{Gal}\,(K/L) \leq G$.
5. *For any two intermediate extensions $K/L_1, L_2/F$ with $H_i = \mathrm{Gal}\,(K/L_i)$, we have*
   (a) *$\mathrm{Gal}\,(K/L_1 \cdot L_2) = H_1 \cap H_2$ in G,*
   (b) *$\mathrm{Gal}\,(K/L_1 \cap L_2) = \langle H_1, H_2\rangle$ in G.*

16.6.6 Consequences of Galois theory

We now portray several consequences of Galois theory (not just fundamental theorem, but field theory in general as well). We begin from observing that finite fields are Galois theoretically quite simple.

For mental clarity, we mention below the topics we cover in this section.

- Galois group of finite fields
- Primitive element theorem
- Compositum & Galois closure
- Norm & trace of a finite separable extension
- Norm & trace in general
- Galois group of ≤ 4 degree polynomials
- Solvability by radicals
- Linearly disjoint extensions

[5] even if L/F is not Galois, i.e. $\mathrm{Gal}\,(K/L)$ is not normal.

Galois group of finite fields

The important result in finite fields is that any finite extension of a finite field is a Galois extension.

Theorem 16.6.6.1. *Let $F = \mathbb{F}_{p^m}$ be a finite field of characteristic p. Let K/F be an algebraic extension. Then the following are equivalent.*

1. *K/F is a finite extension.*
2. *K/F is a Galois extension.*

Proof. (1. $\Rightarrow$ 2.) As K/F is a finite dimensional F-vector space, say of dimension n, therefore K is the finite field $\mathbb{F}_{p^{nm}}$. As $\mathbb{F}_{p^{nm}}$ is by definition the splitting field of $x^{p^{nm}} - x \in \mathbb{F}_p[x]$ which is separable as its derivative is -1 and $x^{p^{nm}} - x$ has no roots in common with -1. It follows by Corollary 16.6.5.4 that $\mathbb{F}_{p^{nm}}/\mathbb{F}_p$ is a Galois extension. As $\mathbb{F}_{p^n}$ is an intermediate extension, therefore by fundamental theorem (Theorem 16.6.5.7), it follows that $\mathbb{F}_{p^{nm}}/\mathbb{F}_{p^n}$ is a Galois extension. (2. $\Rightarrow$ 1.) A Galois extension is always finite. □

Next, we show that the Galois group of any finite extension of a finite field is cyclic.

Proposition 16.6.6.2. *Let $\mathbb{F}_{p^m}$ be a characteristic p finite field. If $K/\mathbb{F}_{p^m}$ is a finite extension of degree n, then $K/\mathbb{F}_{p^m}$ is a Galois extension with Galois group*

$$\mathrm{Gal}\,(K/\mathbb{F}_{p^m}) \cong \mathbb{Z}/n\mathbb{Z}.$$

Proof. We have seen by Theorem 16.6.6.1 that K/F is a Galois extension. By Theorem 16.6.5.5, it is further clear that $|\mathrm{Gal}\,(K/\mathbb{F}_{p^m})| = n$. It hence suffices to show that there exists an element of order n in $\mathrm{Gal}\,(K/\mathbb{F}_{p^m})$. Indeed, consider the following automorphism

$$\begin{aligned} \sigma : K &\longrightarrow K \\ \alpha &\longmapsto \alpha^{p^m}. \end{aligned}$$

We show that σ is of order n in $\mathrm{Gal}\,(K/\mathbb{F}_{p^n})$. Indeed if $\sigma^k(\alpha) = \alpha^{p^{mk}} = \alpha$ for α the generating element of the multiplicative cyclic group of order $p^{nm} - 1$ of $\mathbb{F}_{p^n}$, then we conclude that $n = k$, as required. This completes the proof. □

Corollary 16.6.6.3. *Let F be a finite field and $f(x) \in F[x]$ be a polynomial. If α is a root of $f(x)$, then $F(\alpha)$ is the splitting field of $f(x)$.*

Proof. As $F(\alpha)/F$ is an extension of degree $\deg f(x)$, therefore by Theorem 16.6.6.1, it follows that $F(\alpha)/F$ is Galois, thus it has all conjugates of α and thus is a field containing all roots of $f(x)$. Clearly, $F(\alpha)$ is the smallest field containing all roots of $f(x)$, thus, $F(\alpha)$ is the splitting field of $f(x)$. □

Primitive element theorem

An important theorem in Galois theory is the observation that a finite separable extension is always simple. In particular, every Galois extension is a singly generated field extension.

Theorem 16.6.6.4 (Primitive element theorem)**.** *Let K/F be a finite separable extension. Then there exists $\alpha \in K$ such that $K = F(\alpha)$.*

Proof. Omitted. □

Compositum & Galois closure

We now study how Galois extensions behave with compositums. One calls it the *sliding lemma* as it says that Galois extensions slides through arbitrary extensions.

Proposition 16.6.6.5 (Sliding lemma)**.** *Let K/F be a Galois extension and F'/F be an arbitrary extension such that $K, F' \subseteq \Omega$ where Ω is some large field. Then,*

1. *The extension $K \cdot F'/F'$ is a Galois extension.*
2. *There is an injective group homomorphism*

$$\operatorname{Gal}\left(K \cdot F'/F'\right) \hookrightarrow \operatorname{Gal}\left(K/F\right)$$

whose image is $\operatorname{Gal}\left(K/F' \cap K\right)$. *That is,*

$$\operatorname{Gal}\left(K \cdot F'/F'\right) \cong \operatorname{Gal}\left(K/F' \cap K\right).$$

Proof. 1. We first observe by primitive element theorem (Theorem 16.6.6.4) that $K = F(\alpha)$ for some $\alpha \in K$. We hence have $K \cdot F' = F'(\alpha)$. As α is algebraic over F and $F \subseteq F'$, thus, $F'(\alpha)/F'$ is algebraic. As $F'(\alpha)$ is finitely generated as well, thus $F'(\alpha)/F'$ is finite, as required.

Next, we show that $F'(\alpha)/F'$ is separable. Indeed, by Proposition 16.6.4.15, 1, it suffices to show that $m_{\alpha,F'}(x)$ is a separable polynomial in $F'[x]$. As $m_{\alpha,F'}(x)|m_{\alpha,F}(x)$ and the latter is separable, hence $m_{\alpha,F'}(x)$ is separable.

Finally, we wish to show that $F'(\alpha)/F'$ is normal. Again by Proposition 16.6.4.15 and the fact that $m_{\alpha,F'}(x)|m_{\alpha,F}(x)$ where the latter has all roots in $F(\alpha) \subseteq F'(\alpha)$, we conclude the proof.

2. Consider the map

$$\begin{aligned} \varphi : \operatorname{Gal}\left(K \cdot F'/F'\right) &\longrightarrow \operatorname{Gal}\left(K/F\right) \\ \sigma &\longmapsto \sigma|_K . \end{aligned}$$

This is well-defined since $K = F(\alpha)$, so σ restricted to $F(\alpha)$ maps inside $F(\alpha)$ as all F'-conjugates of α are F-conjugates of α. Now φ can easily be seen to be an injective group homomorphism. We need only find its image now. Indeed, we first claim that for every $\sigma \in \operatorname{Gal}\left(K \cdot F'/F'\right)$, the element $\sigma|_K$ fixes $F' \cap K$. Indeed, σ fixes F' and $K = F(\alpha)$. Thus $F' \cap K \subseteq F' \cap F$, the latter of which is fixed. By item 1 and fundamental theorem (Theorem 16.6.5.7), $K/F' \cap K$ is Galois. Thus, $\varphi : \operatorname{Gal}\left(K \cdot F'/F'\right) \to \operatorname{Gal}\left(K/F' \cap K\right)$. We need only show that it is surjective. To this end, we need only show that $\operatorname{Gal}\left(K/F' \cap K\right) = \operatorname{Im}\left(\varphi\right)$. By fundamental theorem (Theorem 16.6.5.7), it suffices to show that their fixed fields are same. Let $\operatorname{Gal}\left(K/F' \cap K\right) = H_1$ and $\operatorname{Im}\left(\varphi\right) = H_2$, so that $H_2 \le H_1$. We already have by fundamental theorem that $K^{H_2} \ge K^{H_1} = F' \cap K$. On the other hand, if $x \in K^{H_2}$, then $x \in K \cap (K \cdot F')^{\operatorname{Gal}(K \cdot F'/F')} = K \cap F'$, as required. □

The following tells us that compositum and intersections of Galois is Galois.

Proposition 16.6.6.6 (Compositum & intersection of Galois)**.** *Let K_1/F and K_2/F be Galois extensions where $K_1, K_2 \subseteq \Omega$ for some large field Ω. Then,*

1. *Extension $K_1 \cdot K_2/F$ is Galois.*
2. *Extension $K_1 \cap K_2/F$ is Galois.*

3. *There is an injective group homomorphism*

$$\varphi : \mathrm{Gal}\,(K_1 \cdot K_2/F) \hookrightarrow \mathrm{Gal}\,(K_1/F) \times \mathrm{Gal}\,(K_2/F)$$
$$\sigma \mapsto (\sigma|_{K_1}, \sigma|_{K_2})$$

whose image is

$$\begin{aligned}\mathrm{Im}\,(\varphi) &= \{(\sigma, \tau) \mid \sigma|_{K_1 \cap K_2} = \tau|_{K_1 \cap K_2}\} \\ &= \mathrm{Gal}\,(K_1/F) \times_{\mathrm{Gal}(K_1 \cap K_2/F)} \mathrm{Gal}\,(K_2/F)\,.\end{aligned}$$

Hence, in particular, if $K_1 \cap K_2 = F$, then

$$\mathrm{Gal}\,(K_1 \cdot K_2/F) \cong \mathrm{Gal}\,(K_1/F) \times \mathrm{Gal}\,(K_2/F)\,.$$

Proof. 1. By Lemma 16.6.4.13 and sliding lemma (Proposition 16.6.6.5), we deduce that $K_1 \cdot K_2/F$ is a separable extension. By primitive element theorem (Theorem 16.6.6.4) or otherwise, we may deduce that $K_1 \cdot K_2/F$ is finite as well. We need only show that $K_1 \cdot K_2/F$ is normal. To this end, we show that $K_1 \cdot K_2$ is a splitting field of some polynomial in $F[x]$. Indeed, consider $K_1 = F(\alpha)$ and $K_2 = F(\beta)$ by primitive element theorem (Theorem 16.6.6.4) so that $K_1 \cdot K_2 = F(\alpha, \beta)$. As $K_i = F(\alpha_i)$ are normal over F, therefore F_i is splitting field of polynomial $f_i(x) \in F[x]$, for $i = 1, 2$. Thus, we claim that $f_1 \cdot f_2 \in F[x]$ has splitting field $K_1 \cdot K_2$. Indeed, $f_1 \cdot f_2$ splits in $K_1 \cdot K_2$ as both f_1 and f_2 splits in it. Thus if K is the splitting field of $f_1 \cdot f_2$, then $K \subseteq K_1 \cdot K_2$. As $K \supseteq K_i$ for each $i = 1, 2$ since K_i are splitting fields of f_i and f_i splits in K, thus we also have $K \supseteq K_1, K_2$ and thus $K \supseteq K_1 \cdot K_2$. It follows that $K = K_1 \cdot K_2$ and thus $K_1 \cdot K_2$ is normal by Theorem 16.6.4.7, as required.

2. Observe that $K_1 \cap K_2$ is finite and separable over F. We now show that it is normal as well. Indeed, for any $\alpha \in K_1 \cap K_2$, we have $m_{\alpha,F}(x) \in F[x]$ is such that it has all roots in K_1 and K_2 since both are Galois over F. It follows that $m_{\alpha,F}(x)$ has all roots in $K_1 \cap K_2$, showing that $K_1 \cap K_2$ is normal, as required.

3. Injectivity is immediate. For surjectivity, use sliding lemma (Proposition 16.6.6.5) in conjunction with a size argument via Theorem 16.6.5.5. □

We now show that any finite separable extension admits a Galois closure.

Lemma 16.6.6.7. *Let K/F be a finite separable extension. Then there exists a Galois extension L/F such that $L \supseteq K$ which is smallest with respect to containing K.*

Proof. We first show that there exists a Galois extension of F containing K. Indeed, consider $K = F\alpha_1 + \cdots + F\alpha_n$ and let $m_{\alpha_i,F}(x) \in F[x]$ be minimal polynomial of α_i. As K is separable, each of $m_{\alpha_i,F}(x)$ is a separable polynomial in $F[x]$. Thus let K_i/F be the splitting field of $m_{\alpha_i,F}(x)$. By Proposition 16.6.5.2, it follows that K_i/F are all Galois. By compositum of Galois (Proposition 16.6.6.6), we deduce that $L = K_1 \cdot \cdots \cdot K_n$ is a Galois extension of F which contains K as it contains $\alpha_1, \dots, \alpha_n$. Thus we have found a Galois extension of F containing K, as required.

We now wish to show that there is a smallest Galois extension of F containing K. Indeed, consider $E = \bigcap_{L/A/K/F} A$ where A/F is a Galois extension containing K. By fundamental theorem (Theorem 16.6.5.7), it follows that there are only finitely many intermediate extensions of

L/F, thus finitely many such A. Thus E is Galois by intersection of Galois (Proposition 16.6.6.6). Clearly, by construction E is the smallest field extension of F containing K and is Galois. This completes the proof. □

The above lemma allows us to define the following.

Definition 16.6.6.8 (Galois closure of a finite separable extension). Let K/F be a finite separable extension. Then the smallest extension L/F containing K such that L/F is Galois is called the Galois closure of K/F. Lemma 16.6.6.7 shows that it always exists.

Norm & trace of a finite separable extension

Let K/F be an extension. A main technique in field theory is to construct non-trivial elements in K not in F. To this end one of the important set of tools available are those provided by norm & trace of a finite separable extension.

Definition 16.6.6.9 (Norm & Trace). Let K/F be a finite separable extension. Consider a fixed algebraic closure $\bar{F}$ of F. Define

$$N_{K/F}(\alpha) = \prod_{\sigma \in \mathrm{Hom}_F\left(K, \bar{F}\right)} \sigma(\alpha)$$

and

$$\mathrm{Tr}_{K/F}(\alpha) = \sum_{\sigma \in \mathrm{Hom}_F\left(K, \bar{F}\right)} \sigma(\alpha)$$

which we respectively call the norm and trace of α in K/F. Note that $\mathrm{Hom}_F\left(K, \bar{F}\right)$ is finite by Lemma 16.6.4.11.

We can give an alternate definition norm and trace.

Lemma 16.6.6.10. *Let K/F be a finite separable extension. Let $L/K/F$ be the Galois closure of K/F and let $\{\sigma_1, \dots, \sigma_k\} \in \mathrm{Gal}\,(L/F)$ be distinct coset representatives of $\mathrm{Gal}\,(L/K)$ in $\mathrm{Gal}\,(L/F)$. Then*

$$N_{K/F}(\alpha) = \prod_{i=1}^{k} \sigma_i(\alpha)$$

and

$$\mathrm{Tr}_{K/F}(\alpha) = \sum_{i=1}^{k} \sigma_i(\alpha).$$

If K/F is Galois, then $N_{K/F}(\alpha) = \prod_{\sigma \in \mathrm{Gal}(K/F)} \sigma(\alpha)$ and $\mathrm{Tr}_{K/F}(\alpha) = \sum_{\sigma \in \mathrm{Gal}(K/F)} \sigma(\alpha)$.

Proof. By fundamental theorem 16.6.5.7, 4, we have a bijection of sets (which is an isomorphism of groups if K/F is Galois by fundamental theorem):

$$\mathrm{Hom}_F\left(K, \bar{F}\right) \cong \frac{\mathrm{Gal}\,(L/F)}{\mathrm{Gal}\,(L/K)}.$$

The result now follows from definition of norm and trace. □

We now state some basic properties of these two functions.

Proposition 16.6.6.11. *Let K/F be a finite separable extension. Let $L/K/F$ be the Galois closure of K/F.*

1. *For any $\alpha \in K$, $N_{K/F}(\alpha) \in F$ and $\mathrm{Tr}_{K/F}(\alpha) \in F$.*
2. *For any $\alpha, \beta \in K$, we have*
$$N_{K/F}(\alpha\beta) = N_{K/F}(\alpha)N_{K/F}(\beta)$$
and
$$\mathrm{Tr}_{K/F}(\alpha + \beta) = \mathrm{Tr}_{K/F}(\alpha) + \mathrm{Tr}_{K/F}(\beta).$$
3. *If $K = F(\sqrt{D})$ for some $D \in F$, then for $a, b \in F$ we have*
$$N_{K/F}(a + b\sqrt{D}) = a^2 - b^2 D$$
and
$$\mathrm{Tr}_{K/F}(a + b\sqrt{D}) = 2a.$$

Proof. For item 1, since these are coefficients of $m_{\alpha,F}(x)$, so they are in F. Item 2 follows immediately from Lemma 16.6.6.10. For item 3, observe that there is only one other conjugate of $\alpha = a + b\sqrt{D}$ (as minimal polynomial is quadratic) given by $\bar{\alpha} = a - b\sqrt{D}$. Now use Lemma 16.6.6.10. □

Lemma 16.6.6.12. *Let K/F be a finite separable extension of degree n and $\alpha \in K$. Then*

1. *Element α acting by left multiplication on K is an F-linear transformation, which we denote by $T_\alpha : K \to K$.*
2. *The minimal polynomial of element $\alpha \in K$, denoted $m_{\alpha,F}(x)$ is same as the minimal polynomial of the F-linear map $T_\alpha : K \to K$, denoted $p(x) \in F[x]$.*
3. *The norm $N_{K/F}(\alpha)$ and trace $\mathrm{Tr}_{K/F}(\alpha)$ are respectively the determinanat and trace of the F-linear map T_α.*

Proof. 1. Indeed, $T_\alpha : K \to K$ is given by $x \mapsto \alpha x$ which F-linear as $T_\alpha(x + cy) = \alpha(x + cy) = \alpha x + c\alpha y = T_\alpha(x) + cT_\alpha(y)$ where $c \in F$.

2. As $m_{\alpha,F}(x)$ is irreducible, we need only show that $p(x)|m_{\alpha,F}(x)$. Note that $m_{\alpha,F}(T_\alpha) = 0$ since for any $z \in K$, we have
$$m_{\alpha,F}(T_\alpha)(z) = m_{\alpha,F}(\alpha)z = 0.$$
Hence $p(x)|m_{\alpha,F}(x)$, as required.

3. Let $m_{\alpha,F}(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$ in $F[x]$ and $[K : F] = n$. By item 2, the minimal polynomial $p(x)$ of T_α is also $m_{\alpha,F}(x)$. Determinant of T_α is the product of all eigenvalues (with repetitions) and trace of T_α is the sum of all eigenvalues. One can then deduce[6] that
$$N_{K/F}(\alpha) = (-1)^n a_0^{n/d}$$

[6]by Questions 17 and 18 of Section 14.2 of DF, cite[DummitFoote]

and

$$\mathrm{Tr}_{K/F}(\alpha) = \frac{-n}{d} a_{d-1}.$$

As K/F is separable, therefore we may write $p(x) = m_{\alpha,F}(x) = (x-\lambda_1) \cdot \dots \cdot (x-\lambda_d)$ where λ_i are distinct eigenvalues of T_α or equivalently, F-conjugates of α. It is now sufficient to show that each eigenvalue λ_i has algebraic multiplicity n/d.

Let $\Phi(x) \in F[x]$ be the characteristic polynomial of T_α. Since $p(x)$ and $\Phi(x)$ have same irreducible factors and $p(x)$ is irreducible, it follows that $\Phi(x) = p(x)^k$ for some $k \geq 1$. As $\Phi(x)$ has degree n and $p(x)$ has degree d, therefore we conclude that $k = n/d$, as required. □

Norm & trace in general

We now define norm and trace for an arbitrary finite extension using the observation made in Lemma 16.6.6.12.

Definition 16.6.6.13 (Norm & trace). Let K/F be a finite extension and $\alpha \in K$. Let $T_\alpha : K \to K$ be the F-linear transformation obtained by multiplication by α. Then, we define

$$N_{K/F}(\alpha) = \det T_\alpha$$
$$\mathrm{Tr}_{K/F}(\alpha) = \mathrm{Tr} T_\alpha.$$

The main theorem here is the following characterization of separability of a finite extension.

Theorem 16.6.6.14 (Trace pairing & separability)**.** *Let K/F be a finite extension. Then the following are equivalent.*

1. *K/F is separable.*
2. *The trace pairing*

$$\begin{aligned} \langle -,- \rangle : K \times K &\longrightarrow F \\ (\alpha,\beta) &\longmapsto \mathrm{Tr}_{K/F}(\alpha\beta) \end{aligned}$$

is a non-degenerate bilinear map.

Recall that a bilinear map $T : V \times V \to k$ on a k-vector space V is non-degenerate if for any k-basis $\{v_i\}_{i=1}^n$ of V, the matrix $(T(v_i,v_j))_{1\leq i,j\leq n}$ is a non-singular matrix.

In order to prove the above theorem, we would require transitivity of trace. To this end, we first have the following basic results.

Lemma 16.6.6.15. *Let K/F be a finite extension of degree n. Then, for any $x, y \in K$ and $c \in F$, we have*

1. $\mathrm{Tr}_{K/F}(x+y) = \mathrm{Tr}_{K/F}(x) + \mathrm{Tr}_{K/F}(y)$.
2. $\mathrm{Tr}_{K/F}(cx) = c\mathrm{Tr}_{K/F}(x)$.
3. $N_{K/F}(xy) = N_{K/F}(x) N_{K/F}(y)$.
4. $N_{K/F}(cx) = c^n N_{K/F}(x)$.

Proof. Immediate. □

The following result can be used for inductive arguments.

Proposition 16.6.6.16. *Let K/F be a finite extension and $x \in K$. Then, for any intermediate extension $K/L/F$ such that $x \in L$, we have*

$$\begin{aligned}\mathrm{Tr}_{K/F}(x) &= [K:L]\cdot \mathrm{Tr}_{L/F}(x)\\ N_{K/F}(x) &= \left(N_{L/F}(x)\right)^{[K:L]}.\end{aligned}$$

Proof. Let $\{w_1, \dots, w_e\}$ be an L-basis of K. It then follows that the linear operator $T_x : K \to K$ obtained by multiplication by x is such that it restricts to an operator on each Lw_i, $i = 1, \dots, e$. Hence the matrix of T_x will be a diagonal block matrix where the block M_i will be the matrix of $T_x|_{Lw_i}$. Taking trace, we deduce that

$$\mathrm{Tr}_{K/F}(x) = \sum_{i=1}^{e} \mathrm{Tr}(M_i) = \sum_{i=1}^{e} \mathrm{Tr}_{L/F}(x) = \mathrm{Tr}_{L/F}(x)\cdot e = [K:L]\cdot \mathrm{Tr}_{L/F}(x),$$

as required. Similarly for determinant. □

The following states how to calculate trace of x in $F(x)/F$.

Lemma 16.6.6.17. *Let K/F be a finite extension and $x \in K$. Let $m_{x,F}(z) = z^d + a_{d-1}z^{d-1} + \dots + a_1 z + a_0$ where $d = [F(x):F]$. Then,*

$$\begin{aligned}\mathrm{Tr}_{F(x)/F}(x) &= -a_{d-1}\\ N_{F(x)/F} &= (-1)^d a_0.\end{aligned}$$

Proof. Omitted. □

We can now state an important formula for calculation of norm and trace in terms of conjugates and inseparability index (see §16.6.8).

Proposition 16.6.6.18. *Let K/F be a finite extension and $x \in K$. Then we have*

$$\begin{aligned}\mathrm{Tr}_{K/F}(x) &= \left(\sum_{\sigma \in \mathrm{Hom}_F(K,\bar{F})} \sigma(x)\right)\cdot [K:F]_i\\ N_{K/F}(x) &= \left(\prod_{\sigma \in \mathrm{Hom}_F(K,\bar{F})} \sigma(x)\right)^{[K:F]_i}.\end{aligned}$$

Proof. Note that in both the claims above, we need only show the above equality for K/F being inseparable. Indeed, for separable case, we can deduce this equality from Lemma 16.6.6.12.

Let K/F be inseparable and thus let F be of characteristic $p > 0$. By Lemma 16.6.8.15, we deduce that $[K:F]_i = p^n$ and thus RHS = 0 in the first equation above. We thus need only see that $\mathrm{Tr}_{K/F}(x) = 0$ as well. Indeed, by Lemma 16.6.6.17, we need only show that the sum of all conjugates is a multiple of p. Indeed, by Corollary 16.6.8.9, we have that each root of $m_{x,F}$ has common multiplicity p^n. Thus sum of roots of $m_{x,F}$ will be a multiple of p^n, thus 0, as required. One similarly proceeds for showing the same for norm. □

The following is an important result.

Theorem 16.6.6.19 (Transitivity of trace & norm)**.** *Let $L/K/F$ be finite extensions and $\alpha \in L$.*

$$\begin{aligned}\mathrm{Tr}_{K/F}(\mathrm{Tr}_{L/K}(\alpha)) &= \mathrm{Tr}_{L/F}(\alpha)\\ N_{K/F}(N_{L/K}(\alpha)) &= N_{L/F}(\alpha).\end{aligned}$$

Proof. Applying Proposition 16.6.6.18 in our case, we get

$$\mathrm{Tr}_{L/K}(\alpha) = [L:K]_i \cdot \sum_{\sigma \in \hom_K(L,\bar{K})} \sigma(\alpha).$$

Applying $\mathrm{Tr}_{K/F}$ onto above, we yield (note that $\bar{F} = \bar{K}$ as K/F is finite and Lemma 16.6.8.14):

$$\begin{aligned}\mathrm{Tr}_{K/F}\left(\mathrm{Tr}_{L/K}(\alpha)\right) &= \mathrm{Tr}_{K/F}\left([L:K]_i \cdot \sum_{\sigma \in \hom_K(L,\bar{K})} \sigma(\alpha)\right)\\ &= [L:K]_i[K:F]_i \cdot \sum_{\tau \in \hom_F(K,\bar{F})} \tau\left(\sum_{\sigma \in \hom_K(L,\bar{K})} \sigma(\alpha)\right)\\ &= [L:F]_i \cdot \sum_{\tau \in \hom_F(K,\bar{F})} \sum_{\sigma \in \hom_K(L,\bar{K})} \tilde{\tau}(\sigma(\alpha))\end{aligned}$$

where $\tilde{\tau}$ is an extension of $\tau : K \to \bar{F}$ to $\tilde{\tau} : \bar{K} \to \bar{F}$. We now define a bijection

$$\begin{aligned}\varphi : \hom_K(L,\bar{K}) \times \hom_F(K,\bar{F}) &\longrightarrow \hom_F(L,\bar{F})\\ (\sigma,\tau) &\longmapsto \tilde{\tau} \circ \sigma.\end{aligned}$$

Note that $\tilde{\tau} \circ \sigma$ is id on F and τ on k. This is injective as if $\tilde{\tau} \circ \sigma = \tilde{\tau}_1 \circ \sigma_1$, then restricting to K we get $\tau = \tau_1$ and thus, $\sigma = \sigma_1$. Moreover, this is surjective as the size of domain is $[L:K]_s \cdot [K:F]_s$ which is same as the size of codomain $[L:F]_s$. It follows that φ is a bijection.

We can now write the above equation as

$$\begin{aligned}\mathrm{Tr}_{K/F}(\mathrm{Tr}_{L/K}(\alpha)) &= [L:F]_i \cdot \sum_{\kappa \in \hom_F(L,\bar{F})} \kappa(\alpha)\\ &= \mathrm{Tr}_{L/F}(\alpha),\end{aligned}$$

as required. One can follow exact same procedure to show that

$$N_{K/F}(N_{L/K}(\alpha)) = N_{L/F}(\alpha),$$

as required. □

We may now prove the main theorem stated at the beginning of the section.

Proof of Theorem 16.6.6.14. (2. ⇒ 1.) Suppose K/F is inseparable such that $\mathrm{char}(F) = p > 0$. Then $[K : F]_i = p^n$ by Lemma 16.6.8.15. Hence, by Proposition 16.6.6.18, it follows that $\langle \alpha, \beta \rangle = 0$ for each $\alpha, \beta \in K$. Hence $\langle -, - \rangle$ is a degenerate bilinear map, a contradiction.
(1. ⇒ 2.) Suppose K/F is separable. Note it suffices to show that for each non-zero $\alpha \in K$, there exists $\beta \in K$ such that $\mathrm{Tr}_{K/F}(\alpha\beta) = \langle \alpha, \beta \rangle \neq 0$. Indeed, we first show this for $L/K/F$ the Galois closure of K/F. Observe that if for some $\alpha \in K$ non-zero we have that for all $\alpha' \in L$ we get $\mathrm{Tr}_{L/F}(\alpha\alpha') = 0$, then

$$\mathrm{Tr}_{L/F}(\alpha\alpha') = \sum_{\sigma \in \mathrm{Gal}(L/F)} \sigma(\alpha)\sigma(\alpha').$$

By linear independence of characters, we deduce that $\sigma(\alpha) = 0$ for all $\sigma \in \mathrm{Gal}\,(L/F)$, a contradiction as each σ is an automorphism and $\alpha \neq 0$ in K. It follows that there exists $\alpha' \in L$ such that $\mathrm{Tr}_{L/F}(\alpha\alpha') \neq 0$. By transitivity of trace (Theorem 16.6.6.19), we deduce

$$0 \neq \mathrm{Tr}_{L/F}(\alpha\alpha') = \mathrm{Tr}_{K/F}\left(\mathrm{Tr}_{L/K}(\alpha\alpha')\right) = \mathrm{Tr}_{K/F}(\alpha \cdot \mathrm{Tr}_{L/K}(\alpha')).$$

Letting $\beta = \mathrm{Tr}_{L/K}(\alpha')$, we conclude the proof. □

Galois groups of ≤ 4 degree polynomials

Recall that an *elementary symmetric function* s_i is the sum of the products of $\{x_1, \dots, x_n\}$ taken i at a time, that is, $s_1 = x_1 + \cdots + x_n$, $s_2 = x_1x_2 + \ldots x_{n-1}x_n$, $s_n = x_1 \ldots x_n$. Further recall that S_n acts on $F(x_1, \dots, x_n)$ by permuting x_i. A *symmetric function* is a rational function invariant under the action of S_n. We first have the fundamental theorem of symmetric functions.

Theorem 16.6.6.20. *Let F be a field. The fixed field of $F(x_1, \dots, x_n)$ under the action of S_n is $F(s_1, \dots, s_n)$. Thus every symmetric function is a rational function in $s_1, \dots, s_n$.*

This has major consequences.

Corollary 16.6.6.21. *Let F be a field. Then, $F(x_1, \dots, x_n)/F(s_1, \dots, s_n)$ is a Galois extension with Galois group S_n.*

Proof. Follows from Theorem 16.6.6.20 and 16.6.5.6. □

Next result tells us that if a polynomial has algebraically independent elements/indeterminates as roots, then that polynomial is special in that its Galois group has maximal symmetry. This is an important result as if we wish to find a closed form solution of roots in terms of the coefficients, then we ought to take coefficients as algebraically independent elements. In such a situation, the following result then tells us the Galois group of a "general" n-degree polynomial whose roots we assume to be indeterminates.

Theorem 16.6.6.22. *Let $x_1, \dots, x_n$ be indeterminates and F be a field. Then,*
1. *The polynomial $f(x) = (x - x_1) \dots (x - x_n)$ can be expressed as*

$$f(x) = x^n - s_1x^{n-1} + s_2x^{n-2} - \cdots + (-1)^{n-1}s_{n-1}x + (-1)^n s_n$$

where s_i are elementary symmetric polynomials in $x_1, \dots, x_n$.

2. *The polynomial $f(x)$ as above is separable and its splitting field over $F(s_1, \dots, s_n)$ is $F(x_1, \dots, x_n)$ with Galois group S_n.*
3. *If a polynomial $g(x)$ has indeterminates as coefficients, then its roots are also indeterminates.*

Remark 16.6.6.23. As Corollary 16.6.5.4 guarantees, the above theorem tells us exactly the polynomial whose splitting field is the Galois extension $F(x_1, \dots, x_n)/F(s_1, \dots, s_n)$ of Theorem 16.6.6.20.

We now use discriminants of a polynomial to get information about its Galois group.

Definition 16.6.6.24 (Discriminant). Let $f(x) \in F[x]$ be a polynomial with roots $x_1, \dots, x_n$. Then the discriminant of $f(x)$ is defined to be

$$D_f := \prod_{i<j} (x_i - x_j)^2.$$

Before beginning, we need some observations.

Lemma 16.6.6.25. *Let F be a field and $f(x) \in F[x]$ be a separable polynomial (so that $D_f \neq 0$). Let K/F be the splitting field of $f(x)$ over F. Then,*
1. *$D_f \in F$.*
2. *$\sqrt{D_f} \in K$.*

Proof. Let $\alpha_1, \dots, \alpha_n \in K$ be the distinct roots of $f(x)$. Then, for any $\sigma \in \mathrm{Gal}(K/F)$, $\sigma(D_f) = D_f$ as $\sigma(\alpha_i) = \alpha_j$ bijectively. This proves item 1. Item 2 is immediate. □

Remark 16.6.6.26. Let $f(x) \in F[x]$ be a separable polynomial and let K/F be the splitting field of $f(x)$. For any $\sigma \in \mathrm{Gal}(K/F)$, we get a permutation of $Z(f) \subseteq K$, the zero set of $f(x)$, that is, $Z(f)$ is a $\mathrm{Gal}(K/F)$-set. If there are n roots of $f(x)$, then we get a group homomorphism

$$\mathrm{Gal}(K/F) \hookrightarrow S_n$$

which is furthermore injective as if any $\sigma \in \mathrm{Gal}(K/F)$ gives the identity permutation of roots, then it is the identity map $K \to K$. We now always view Galois group of a separable polynomial $f(x)$ as a subgroup of S_n where n is the number of roots of $f(x)$, all of which are distinct as $f(x)$ is separable.

We make the most important statement about the discriminants now.

Proposition 16.6.6.27. *Let $f(x) \in F[x]$ be a separable polynomial with splitting field K/F. Then the following are equivalent.*
1. *$\mathrm{Gal}(K/F)$ is a subgroup of A_n.*
2. *$D_f \in F$ is a square of an element in F and that element is $\sqrt{D_f}$. That is, $\sqrt{D_f} \in F$.*

Proof. (1. ⇒ 2.) As any $\sigma \in A_n$ fixes $\prod_{i<j}(x_i - x_j) \in \mathbb{Z}[x_1, \dots, x_n]$ where $x_1, \dots, x_n$ are the roots of $f(x)$, therefore $\sigma \in \mathrm{Gal}(K/F)$ fixes $\sqrt{D_f}$. It follows by fundamental theorem (Theorem 16.6.5.7) that $\sqrt{D_f} \in F$.

(2. ⇒ 1.) Pick any element $\sigma \in \mathrm{Gal}(K/F)$. To show that $\sigma \in A_n$, we wish to show the criterion mentioned above. This critetion is equivalent to showing that $\sigma(\sqrt{D_f}) = \sqrt{D_f}$. This is equivalent by fundamental theorem to showing that $\sqrt{D_f} \in F$, which is what we are given. □

Solvability by radicals

We next discuss the various results surrounding solvability of a polynomials.

Definition 16.6.6.28 (Elements & polynomials solvable by radicals). Let K/F be an extension. An algebraic element $\alpha \in K$ over F is solvable by radicals if there exists simple radical extensions

$$F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots \subseteq K_n \ni \alpha$$

where $K_{i+1} = K_i(a_i^{1/n_i})$ where $a_i \in K_{i-1}$, $n_i \geq 1$. The field K_n are called roots extensions. A polynomial $f(x) \in F[x]$ is solvable by radicals if all its roots are solvable by radicals.

Remark 16.6.6.29. Note that if $f(x)$ is solvable, then its root extension contains the splitting field.

Definition 16.6.6.30 (Solvable extensions). An extension K/F is solvable if it is Galois and the Galois group Gal (K/F) is solvable[7].

We have the following main theorem.

Theorem 16.6.6.31 (Solvability by radicals)**.** *Let F be a characteristic 0 field and $f(x) \in F[x]$. Then the following are equivalent:*
1. *$f(x)$ is solvable by radicals.*
2. *If K/F is the splitting field of $f(x)$, then K/F is a solvable extension.*

Corollary 16.6.6.32 (Abel-Ruffini)**.** *Let F be a characteristic 0 field. For $n \geq 5$, the general polynomial $f(x) = x^n - s_{n-1}x^{n-1} + s_{n-2}x^{n-2} - \cdots + (-1)^n s_0$ where s_i are elementary symmetric functions of roots $x_1, \ldots, x_n$, is not solvable over $F(s_1, \ldots, s_n)$.*

Proof. By Theorem 16.6.6.22, we deduce that its splitting field is $K(x_1, \ldots, x_n)$ and its Galois group is S_n. For $n \geq 5$, we know that S_n is not solvable. It follows by Theorem 16.6.6.31 that $f(x)$ is not solvable by radicals, that is, there is no root extension of $f(x)$. This means that the roots of $f(x)$ are not obtained by radicals in coefficients. $\square$

Linearly disjoint extensions

We begin by following observation.

Lemma 16.6.6.33. *Let L/F and K/F be two finite extensions of F contained in some large field Ω. Then the following conditions are equivalent.*
1. *Any F-basis of L/F is a K-basis of LK/K.*
2. *Any F-basis of K/F is an L-basis of LK/L.*
3. *$[LK : K] = [L : F]$.*
4. *$[LK : F] = [L : F] \cdot [K : F]$.*

Proof. Fairly standard arguments, hence omitted. $\square$

This allows us to define the following.

[7] A group G is solvable if there exists a normal series with prime cyclic factors.

Definition 16.6.6.34 (**Linearly disjoint extensions**)**.** Let L/F and K/F be two finite extensions of F contained in some large field Ω. Then L/F and K/F are said to be linearly disjoint if they satisfy any of the equivalent conditions of Lemma 16.6.6.33.

The name is motivated by the following observation.

Lemma 16.6.6.35. *Let L/F and K/F be two finite extensions which are linearly disjoint. Then $L \cap K = F$.*

Proof. As we have an isomorphism $\mathrm{Gal}\,(K \cdot L/L) \cong \mathrm{Gal}\,(K/L \cap K)$ by Proposition 16.6.6.5, hence it follows that we have an equality in degree $[K \cdot L : L] = [K : L \cap K]$. By linear disjointness, $[K \cdot L : L] = [K : F]$. As $L \cap K \supseteq F$, thus by tower law we deduce that $[L \cap K : F] = 1$, as required. □

The following shows that above criterion is necessary, but not sufficient.

Example 16.6.6.36. Here is an example of extensions K/F and L/F such that $L \cap K = F$ but still they are not linearly disjoint. For $F = \mathbb{Q}$, take $K = \mathbb{Q}(2^{1/3})$ and $L = \mathbb{Q}(\omega 2^{1/3})$. Observe that $K \cap L = F$. However, as $[K \cdot L : F] = 6$ and $[K : F] = 3 = [L : F]$, we deduce that L/F and K/F are not linearly disjoint.

The following theorem shows a sufficient criterion which when satisfied together with $L \cap K = F$, makes L/F and K/F linearly disjoint.

Lemma 16.6.6.37. *Let L/F and K/F be two finite extensions. If K/F is Galois and $L \cap K = F$, then L/F and K/F are linear disjoint.*

Proof. By Proposition 16.6.6.5, we have that $K \cdot L/L$ is Galois and we have an isomorphism $\mathrm{Gal}\,(K \cdot L/L) \cong \mathrm{Gal}\,(K/L \cap K) = \mathrm{Gal}\,(K/F)$. Thus, we have an equality $[K \cdot L : L] = [K : F]$, hence K/F and L/F are linearly disjoint. □

Hence, we may summarize this discussion as follows.

Corollary 16.6.6.38. *Let K/F and L/F be two finite extensions. If K/F and L/F are linearly disjoint, then $K \cap L = F$. The converse holds if any of the K/F or L/F is a Galois extension.* □

16.6.7 Cyclotomic extensions

We discuss the extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ where ζ_n is an n^{th}-root of unity, that is, a solution of $x^n - 1$ in $\mathbb{C}$. We will see that n^{th}-roots of unity form a cyclic group $\mu_n \cong \mathbb{Z}/n\mathbb{Z}$, therefore we define a *primitive* n^{th} root of unity to be a generator of $\mathbb{Z}/n\mathbb{Z}$. Thus, there are $\varphi(n)$ many primitive n-th roots of unity, where φ is the Euler totient function. We also discuss the main theorems of abelian and cyclic extensions (Kronecker-Weber and Kummer).

We denote the group of n-th roots of unity as μ_n. Some basic facts about μ_n are as follows.

Lemma 16.6.7.1. *Let $n \in \mathbb{N}$. Then,*

1. *μ_n is a finite cyclic group isomorphic to $\mathbb{Z}/n\mathbb{Z}$.*
2. *If $d|n$, then $\mu_d \hookrightarrow \mu_n$.*

Proof. 1. μ_n is finite of size n since its the set of roots of $x^n - 1$ in $\mathbb{C}$. This is a group since product of any two n-th roots of unity is an n-th root of unity. Thus μ_n is a finite subgroup of the multiplicative group $\mathbb{C}^\times$. It follows that μ_n is cyclic.

2. Consider the map

$$\begin{aligned} \varphi : \mu_d &\longrightarrow \mu_n \\ \zeta &\longmapsto \zeta. \end{aligned}$$

This is well-defined since a d-th root of unity is also an n-th root of unity if $d|n$. Further, this is clearly a group homomorphism. □

Thus $\mu_d \leq \mu_n$ is precisely the subgroup of order d-elements of μ_n.

Definition 16.6.7.2 (n^{th}-cyclotomic polynomial). Let $n \in \mathbb{N}$. The n^{th}-cyclotomic polynomial is defined to be the polynomial $\Phi_n(x) = \prod_{\zeta \in \mu_n^\times}(x - \zeta)$, that is, the polynomial whose all roots are the primitive n^{th}-roots of unity.

We immediately have the following observations.

Lemma 16.6.7.3. *Let $\Phi_n(x)$ be the n^{th}-cyclotomic polynomial. Then,*
1. *$\Phi_n(x)|x^n - 1$.*
2. *$x^n - 1 = \prod_{d|n} \Phi_d(x)$.*

Proof. Follows from the observation that $x^n - 1 = \prod_{\zeta^n=1}(x - \zeta)$. □

Remark 16.6.7.4. Using Lemma 16.6.7.3, we see that we can calculate $\Phi_n(x)$ recursively by finding Φ_d for all $d|n$ and $d \neq n$. In particular,

$$\Phi_n(x) = \frac{x^n - 1}{\prod_{d|n, d\neq n} \Phi_d(x)}.$$

We now state and prove the following theorem, which in particular tells us that cyclotomic polynomial $\Phi_n(x)$ is monic irreducible of degree $\varphi(n)$. Once shown, we would be able to conclude that the the minimal polynomial of a primitive n^{th}-root of unity is $\Phi_n(x)$.

Theorem 16.6.7.5. *Let $n \in \mathbb{N}$. Then,*
1. *$\Phi_n(x)$ is a monic polynomial of degree $\varphi(n)$ in $\mathbb{Z}[x]$.*
2. *$\Phi_n(x)$ is an irreducible polynomial in $\mathbb{Z}[x]$.*
3. *$\Phi_n(x)$ is the minimal polynomial of any primitive n^{th}-root of unity $\zeta_n \in \mathbb{C}$.*
4. *If ζ_n is a primitive n^{th}-root of unity, then $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is a degree $\varphi(n)$ extension.*

Proof. 1. The fact that degree of $\Phi_n)(x)$ is $\varphi(n)$ follows from the fact that in $\mathbb{C}$ it is a product of $\varphi(n)$ many linear factors. This also shows that $\Phi_n(x)$ is a monic polynomial. We need only show that coefficients lie in $\mathbb{Z}$. To this end, we proceed by induction. For $n = 1$, $\Phi_n(x) = x - 1 \in \mathbb{Z}[x]$. For $n = 2$, $\Phi_2(x) = x + 1 \in \mathbb{Z}[x]$. Now suppose that for all $d < n$ $\Phi_d(x) \in \mathbb{Z}[x]$. Then we have

$$\Phi_n(x) = \frac{x^n - 1}{\prod_{d|n, d\neq n} \Phi_d(x)},$$

thus $f(x) := \prod_{d|n, d\neq n} \Phi_d(x) \in \mathbb{Z}[x]$ by inductive hypothesis. As $f(x)|x^n - 1$ in $\mathbb{Q}[x]$ and $f(x) \in \mathbb{Z}[x]$, therefore by results surrounding Gauss' lemma, we get $f(x)|x^n - 1$ in $\mathbb{Z}[x]$, that is, $\Phi_n(x) \in \mathbb{Z}[x]$.

2. Let $\Phi_n(x) = f(x)g(x)$ in $\mathbb{Z}[x]$ where we assume that $f(x)$ is an irreducible factor of $\Phi_n(x)$ (by $\mathbb{Z}[x]$ being an UFD). We claim that $f(x)$ has all primitive n^{th}-roots of unity as a root over $\mathbb{C}$, so that $f(x) = \Phi_n(x)$ over $\mathbb{Z}$. Indeed, let $\zeta^a \in \mu_n$ be any other primitive root, then $(a, n) = 1$ and so we may write $a = p_1 \ldots p_k$ where p_i are primes not dividing n. We wish to show that ζ^a is a root of $f(x)$. It suffices to show that if ζ is a root of $f(x)$, then ζ^p is a root of $f(x)$ as well for any prime p not dividing n. This is what we will show now.

Indeed, let $\zeta \in \mu_n$ a primitive n^{th}-root of unity which is a root of $f(x)$. As $f(x)$ is irreducible over $\mathbb{Z}[x]$, therefore irreducible over $\mathbb{Q}[x]$ as well, hence $f(x)$ is the minimal polynomial of ζ over $\mathbb{Q}$. Consider p a prime not dividing n. We wish to show that ζ^p is also a root of $f(x)$. Indeed, as $\Phi_n(x)$ has ζ^p as a root, therefore either $f(\zeta^p) = 0$ or $g(\zeta^p) = 0$ over $\mathbb{C}$. Suppose the latter is true. Thus $g(x^p)$ has ζ as a root. As $g(x^p) \in \mathbb{Q}[x]$, therefore $f(x)|g(x^p)$ in $\mathbb{Q}[x]$. As $f(x), g(x^p) \in \mathbb{Z}[x]$, therefore by results surrounding Gauss' lemma, we conclude that $f(x)|g(x^p)$ in $\mathbb{Z}[x]$. Let $g(x^p) = f(x) \cdot h(x)$ where $h(x) \in \mathbb{Z}[x]$. Going modulo p, we get that $\bar{g}(x^p) = (\bar{g}(x))^p$. Thus, $(\bar{g}(x))^p = \bar{f}(x)\bar{h}(x)$ in $\mathbb{F}_p[x]$. Thus, $\bar{g}$ and $\bar{f}$ have a common factor in $\mathbb{F}_p[x]$ as both have ζ as a root. Thus, $\bar{\Phi}_n(x) = \bar{f}(x)\bar{g}(x)$ has a repeated factor, thus, $\Phi_n(x)$ is not separable over over $\mathbb{F}_p$. But since $\Phi'_n(x) = nx^{n-1} \neq 0$ has only $x = 0$ as a root, therefore $\Phi_n(x)$ is separable. It follows that we have a contradiction to the separability of $x^n - 1$ as $\Phi_n(x)$ is a factor of $x^n - 1$, thus ζ^p cannot be a root of $g(x)$, as required.

3. As $\Phi_n(\zeta_n) = 0$ for any primitive n^{th}-root of unity, therefore we get that $m_{\zeta_n,\mathbb{Q}}|\Phi_n(x)$. As $m_{\zeta_n,\mathbb{Q}}$ is irreducible and so is $\Phi_n(x)$, thus $m_{\zeta_n,\mathbb{Q}} = \Phi_n$, as required.

4. As $\Phi_n(x)$ is the minimal polynomial of ζ_n which has degree $\varphi(n)$, the result follows. □

We now wish to study the Galois group of a cyclotomic extension.

Definition 16.6.7.6 (Cyclotomic extension). Let $\zeta_n \in \mathbb{C}$ be a primitive n^{th}-root of unity. The extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is called a cyclotomic extension.

It is easy to see that every cyclotomic extension is Galois.

Lemma 16.6.7.7. *Let $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ be a cyclotomic extension. Then $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is a Galois extension.*

Proof. By Theorem 16.6.7.5, 4, it follows that $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is finite. Observe that $m_{\zeta_n,\mathbb{Q}}(x) \in \mathbb{Q}[x]$ is $\Phi_n(x)$ by Theorem 16.6.7.5, 3 which is separable. As ζ_n is the primitive n^{th}-root of unity, therefore it generates all other roots of unity. Consequently, $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is normal as well, as required. □

Calculation of Galois group of $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is quite easy.

Theorem 16.6.7.8. *Let $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ be a cyclotomic extension where ζ_n is a primitive n^{th}-root. Then, the map*

$$\begin{aligned} (\mathbb{Z}/n\mathbb{Z})^\times &\longrightarrow \mathrm{Gal}\,(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \\ a &\longmapsto \sigma_a : \zeta_n \mapsto \zeta_n^a \end{aligned}$$

is an isomorphism.

Proof. Immediate. □

Cyclotomic extensions are a particular example of an abelian extension.

Definition 16.6.7.9 (**Abelian extension**)**.** Let K/F be a field extension. If K/F is Galois and $\mathrm{Gal}\,(K/F)$ is an abelian group, then K/F is called an abelian extension.

Remark 16.6.7.10. If $K_1, K_2/F$ are abelian extensions, then any subfield $K_1/L/F$ is an abelian extension by fundamental theorem (Theorem 16.6.5.7) and compositum $K_1 \cdot K_2/F$ is also abelian by Proposition 16.6.6.5.

An important result in the theory of finite abelian extensions is the fact that any extension of $\mathbb{Q}$ is abelian if and only if it is contained in a cyclotomic extension. Using this result, one can heuristically say that finite abelian groups are to groups what are cyclotomic extensions are to field extensions(!)

Theorem 16.6.7.11 (Kronecker-Weber)**.** *Let $K/\mathbb{Q}$ be an extension. Then the following are equivalent:*

1. *$K/\mathbb{Q}$ is a finite abelian.*
2. *$K \subseteq \mathbb{Q}(\zeta_n)$ for some $n \in \mathbb{N}$.*

Moreover, if G is any finite abelian group, then there exists $K/\mathbb{Q}$ finite abelian such that $\mathrm{Gal}\,(K/\mathbb{Q}) \cong G$.

Another important line of thought around cyclotomic extensions is the situation when Galois group is cyclic. We have seen that Galois groups of finite fields are cyclic (Proposition 16.6.6.2). Moreover, if p is a prime, then by Theorem 16.6.7.8, the cyclotomic extension $\mathbb{Q}(\zeta_p)/\mathbb{Q}$ also has cyclic Galois group for ζ_p a primitive p^{th}-root of unity. We now see that such Galois extensions are of a very simple type.

Definition 16.6.7.12 (**Cyclic extensions**)**.** An extension K/F is said to be cyclic if it is Galois and $\mathrm{Gal}\,(K/F)$ is cyclic.

Theorem 16.6.7.13 (Kummer-I)**.** *Let F be a characteristic $p > 0$ field and $\zeta_n \in F$ where ζ_n is a primitive n^{th}-root of unity for* $\gcd(n, p) = 1$.

1. *If $K = F(a^{1/n})$ for some non-zero $a \in F$, then K/F is a cyclic extension of degree d where $d|n$.*
2. *If K/F is a cyclic extension of degree n, then $K = F(a^{1/n})$ for some non-zero $a \in F$.*

Proof. 1. We first show that K/F is Galois. Let $\alpha = a^{1/n}$. Finiteness is clear as $m_{\alpha,F}(x)|x^n - a$. We wish to show that $m_{\alpha,F}(x)$ is separable. Indeed, since $x^n - a$ has derivative nx^{n-1} which is a non-zero polynomial (as $\gcd(n, p) = 1$) whose only root is 0, therefore $x^n - a$ is separable and thus so is $m_{\alpha,F}(x)$. Finally, as all roots of $x^n - a$ are $\{\zeta_n^k \alpha\}_{k=0,\dots,n-1}$, which are in K as $\zeta_n \in F$, therefore $x^n - a$ splits in K into linear factors, and hence so does $m_{\alpha,F}(x)$. Indeed, K is the splitting field of $x^n - a$ over F.

Next, we show that K/F is cyclic. Indeed, consider the map

$$\begin{aligned} \varphi : \mathrm{Gal}\,(K/F) &\longrightarrow \mu_n \\ \sigma &\longmapsto \frac{\sigma(\alpha)}{\alpha}. \end{aligned}$$

This is well defined as $\sigma(\alpha) = \zeta_n^{k_\sigma} \alpha$, some conjugate of α. Thus, $\varphi(\sigma) = \zeta_n^{k_\sigma}$. We claim that this is an injective group homomorphism, and thus $\mathrm{Gal}\,(K/F)$ is cyclic.

Indeed, this is a group homomorphism as $\varphi(\sigma \circ \tau) = \sigma(\tau(\alpha)) = \sigma(\zeta_n^{k_\tau} \alpha)/\alpha = \zeta_n^{k_\sigma} \zeta_n^{k_\tau}$. Hence, it is a group homomorphism. It is moreover injective as if $\zeta_n^{k_\sigma} = \zeta_n^{k_\tau}$, then $\sigma(\alpha) = \tau(\alpha)$. As σ, τ

are F-automorphisms of $K = F(\alpha)$ mapping α to the same element, therefore $\sigma = \tau$, as needed. Furthermore, as $|\mathrm{Gal}\,(K/F)| \,|\, |\mu_n|$, therefore $[K : F] = d$ where $d|n$.

2. We wish to find an n^{th}-root of some a in K and show that it generates K. As $\mathrm{Gal}\,(K/F) = \langle\sigma\rangle$ is cyclic, therefore consider the following element of K constructed out of any $\alpha \in K$:

$$\beta = \alpha + \zeta_n\sigma(\alpha) + \zeta_n^2\sigma^2(\alpha) + \cdots + \zeta_n^{n-1}\sigma^{n-1}(\alpha).$$

Observe that

$$\begin{aligned}\sigma(\beta) &= \sigma(\alpha) + \zeta_n\sigma^2(\alpha) + \zeta_n^2\sigma^3(\alpha) + \cdots + \zeta_n^{n-1}\alpha\\ &= \zeta_n^{n-1}\beta.\end{aligned}$$

Similarly, we get for each $0 \le k \le n-1$ the following relation:

$$\sigma^k(\beta) = \zeta_n^{n-k}\beta.$$

Hence, we see that $p(x) = x^n - \beta^n$ has all roots in K given by $\{\zeta_n^{n-k}\beta\}_{0\le k\le n-1}$.

We claim that $\beta^n \in F$. Indeed, we show that for $G = \mathrm{Gal}\,(K/F) = \langle\sigma\rangle$, the element β^n is in K^G and since $K^G = F$ by fundamental theorem (Theorem 16.6.5.7), hence we will be done. As $\sigma(\beta^n) = (\zeta_n^{n-1}\beta)^n = \beta^n$, therefore $\beta^n \in K^G = F$, as required. Hence, $\beta = a^{1/n}$ for $a = \beta^n \in F$.

We finally claim that $F(\beta) = K$. Indeed, as $K/F(\beta)/F$ is an intermediate extension and $\mathrm{Gal}\,(K/F)$ is cyclic hence abelian, therefore $F(\beta)/F$ is Galois by fundamental theorem (Theorem 16.6.5.7). As $\sigma \in \mathrm{Gal}\,(F(\beta)/F)$, therefore $|\mathrm{Gal}\,(F(\beta)/F)| \ge n$. But by fundamental theorem, $\mathrm{Gal}\,(F(\beta)/F) = \frac{\mathrm{Gal}(K/F)}{\mathrm{Gal}(K/F(\beta))}$, thus, $|\mathrm{Gal}\,(K/F(\beta))| = [K : F(\beta)] = 1$, thus, $[K : F] = [K : F(\beta)][F(\beta) : F] = [F(\beta) : F]$, thus showing that $F(\beta) = K$, as required. □

An important corollary strengthening the second statement of Kummer is as follows.

Corollary 16.6.7.14 (Kummer-II). *Let F be a field of characteristic $p > 0$ and $\zeta_n \in F$ be a primitive n^{th}-root of unity where $\gcd(n,p) = 1$. If K/F is a cyclic extension of degree d where $d|n$, then $K = F(a^{1/d})$ for some $a \in F$ non-zero*[8].

Proof. Note that as $\zeta_n \in F$, therefore $\mu_n \subseteq F^\times$. Recall from Lemma 16.6.7.1 that $\mu_d \hookrightarrow \mu_n$. It follows that F contains a primitive d^{th}-root of unity. As $\gcd(n,p) = 1$, it follows that $\gcd(d,p) = 1$. By Kummer-I (Theorem 16.6.7.13, 2), it follows that $K = F(a^{1/d}) \subseteq F(a^{1/n})$, for some non-zero $a \in F$, as required. □

Remark 16.6.7.15. Assuming the hypothesis of Corollary 16.6.7.14, we see that if K/F is cyclic of degree d where $d|n$, then $K = F(a^{1/d})$. Now note that we can write $K = F((a^{n/d})^{1/n}) = F(b^{1/n})$ where $b = a^{n/d}$.

[8]Note that as $d|n$, hence $F(a^{1/d}) \subseteq F(a^{1/n})$.

16.6.8 Inseparable & purely inseparable extensions

We now study a type of extension which is prevalent in the study of varieties of characteristic $p > 0$. Recall that an extension K/F is *inseparable* if it is not separable, that is, there is some element in K whose minimal polynomial over F is inseparable.

Some of our main results are characterizations of irreducible and minimal polynomials in characteristic $p > 0$ fields as stated in Proposition 16.6.8.2 and Corollary 16.6.8.9.

Definition 16.6.8.1 (Purely inseparable extension). Let F be a field of characteristic $p > 0$ and K/F be an extension. An element $\alpha \in K$ is said to be purely inseparable if for some $n \geq 0$, we have

$$\alpha^{p^n} \in F.$$

If every element of K is purely inseparable, then K/F is said to be purely inseparable.

Before beginning the study of purely inseparable fields, we need a fundamental result about irreducible polynomials in positive characteristic fields.

Proposition 16.6.8.2 ("Polynomial Frobenius")**.** *Let F be a field of characteristic $p > 0$. If $f(x) \in F[x]$ is an irreducible polynomial, then there exists an irreducible and separable polynomial $g(x) \in F[x]$ such that*

$$f(x) = g(x^{p^n})$$

for some $n \geq 0$.

Proof. Suppose that $f(x)$ is separable. Then $g = f$ and $n = 0$ would do. Hence we may assume that $f(x)$ is inseparable. Thus, by Lemma 16.6.4.14, it follows that $f'(x) = 0$. Writing

$$f(x) = \sum_{j=0}^{m} a_j x^j,$$

we deduce that $j = pk_j$. Thus, we may write

$$f(x) = \sum_{j=0}^{m} a_j x^{pk_j} = h(x^p)$$

where $h(x) = \sum_{j=0}^{m} a_j x^{k_j} \in F[x]$. As $f(x)$ is irreducible, therefore it follows that $h(x)$ is irreducible. Note that degree of h is $\frac{\deg f}{p}$. If h is separable, then we are done. If not, then we repeat the process, starting from $h(x)$, to yield $h_1(x)$ satisfying $h_1(x^p) = h(x)$ and thus $h_1(x^{p^2}) = f(x)$. As at each step the resulting polynomial has degree strictly smaller than that of previous, hence the process has to stop. As the process at a separable polynomial, we thus obtain $g(x)$ separable and irreducible such that $g(x^{p^n}) = f(x)$, as required. □

There are some other restatements of the definition, which are important to keep in mind. All of these uses the "Polynomial Frobenius" (Proposition 16.6.8.2) in a crucial manner.

Theorem 16.6.8.3. *Let F be a characteristic $p > 0$ field and K/F be an algebraic extension. The following are equivalent:*

1. *K/F is purely inseparable.*
2. *For every $\alpha \in K$ not in F, the minimal polynomial $m_{\alpha,F}(x)$ in $F[x]$ is an inseparable polynomial.*
3. *For every $\alpha \in K$, the minimal polynomial $m_{\alpha,F}(x)$ in $F[x]$ is of the form*

$$m_{\alpha,F}(x) = x^{p^n} - a$$

for some $a \in F$.

Proof. (1. $\Rightarrow$ 2.) For some $n \in \mathbb{N}$, we have $\alpha^{p^n} = a \in F$. Thus, $m_{\alpha,F}(x)|x^{p^n} - a$. As $f(x) = x^{p^n} - a$ and derivative $f'(x) = 0$ as $\mathrm{char}(F) = p$, therefore $x^{p^n} - a$ has repeated roots. Now suppose $x^{p^n} - a = f_1(x) \dots f_k(x)$ where each $f_i(x) \in F[x]$ is an irreducible factor of $x^{p^n} - a$. Since $x^{p^n} - a = (x - \alpha)^{p^n}$ in $K[x]$, it follows that each $f_i(x)$ divides $(x - \alpha)^{m_i}$ in $K[x]$. In particular, each $f_i(x)$ is inseparable. As $m_{\alpha,F}(x) = f_i(x)$ for some i as $m_{\alpha,F}(x)$ is irreducible dividing $x^{p^n} - a$ in $F[x]$, it follows that $m_{\alpha,F}(x)$ is inseparable.

(2. $\Rightarrow$ 1.) Pick any $\alpha \in K$. We wish to find $n \geq 0$ such that $\alpha^{p^n} \in F$. This is equivalent to showing that $m_{\alpha,F}(x)|x^{p^n} - a$ for some $a \in F$. If $\alpha \in F$, we are done. We may thus assume $\alpha \in K \setminus F$. Consider the minimal polynomial $m_{\alpha,F}(x) \in F[x]$. As it is irreducible, by Proposition 16.6.8.2 it follows that $m_{\alpha,F}(x) = f(x^{p^n})$ where $f(x) \in F[x]$ is irreducible and separable. As $f(\alpha^{p^n}) = 0$, it follows that $m_{\alpha^{p^n},F}(x)|f(x)$. As both are irreducible, it follows at once that $m_{\alpha^{p^n},F}(x) = f(x)$. We deduce that $m_{\alpha^{p^n},F}(x)$ is separable. By our hypothesis, it follows that $\alpha^{p^n} \in F$, as required.

(2. $\Rightarrow$ 3.) Pick any $\alpha \in K$. If $\alpha \in F$, there is nothing to do. We may thus assume $\alpha \in K \setminus F$. Consider $m_{\alpha,F}(x) \in F[x]$ which is irreducible and by hypothesis is inseparable. Observe by Polynomial Frobenius (Proposition 16.6.8.2) that there exists $g(x) \in F[x]$ irreducible and separable such that for some $n \geq 0$ we get

$$m_{\alpha,F}(x) = g(x^{p^n}).$$

It follows that $g(\alpha^{p^n}) = 0$ and thus $m_{\alpha^{p^n},F}(x) = g(x)$. We thus further deduce that $m_{\alpha^{p^n},F}(x)$ is irreducible and separable. By our hypothesis, we must have $\alpha^{p^n} = a \in F$ and thus $m_{\alpha^{p^n},F}(x) = g(x) = x - a$. As $m_{\alpha,F}(x) = g(x^{p^n}) = x^{p^n} - a$, hence we get the desired result.

(3. $\Rightarrow$ 1.) Pick any element $\alpha \in K$ not in F. As $m_{\alpha,F}(x) = x^{p^n} - a$ and $a \in F$, therefore $\alpha^{p^n} = a \in F$, as required. $\square$

It is clear from above that any non-trivial purely inseparable extension is inseparable. A simple corollary states that perfect fields don't have non-trivial inseparable extensions.

Corollary 16.6.8.4. *Let F be a field. Then, the following are equivalent:*

1. *An algebraic extension K/F is inseparable*[9].
2. *F is not a perfect field*[10].

[9] that is, there is an element whose minimal polynomial is inseparable.

[10] see Definition 16.6.4.4

Proof. This is just the contrapositive of Theorem 16.6.4.19. □

Corollary 16.6.8.5. *Let F be a perfect field. If K/F is purely inseparable, then $K = F$.*

Proof. Suppose K/F is non-trivial. By Corollary 16.6.8.4, it follows that F is not perfect, a contradiction. □

The following shows that the subfield generated by a purely inseparable element is purely inseparable.

Proposition 16.6.8.6. *Let F be a characteristic $p > 0$ field and K/F be a field extension and $\alpha \in K$ be an algebraic element which is a purely inseparable element over F. Then $F(\alpha)/F$ is purely inseparable.*

Proof. As $\alpha \in K$ is algebraic over F, therefore $F(\alpha) = F[\alpha]$. Pick any $\beta \in F[\alpha]$. We may write

$$\beta = a_m\alpha^m + \cdots + a_1\alpha + a_0.$$

As $\alpha^{p^n} \in F$, thus we get

$$\beta^{p^n} = a_m^{p^n}\alpha^{mp^n} + \cdots + a_1^{p^n}\alpha^{p^n} + a_0^{p^n} \in F,$$

as needed. □

The following result is important for it says that the separable closure of an algebraic extension completely divides the extension into separable and a purely inseparable part.

Proposition 16.6.8.7. *Let F be a field of characteristic $p > 0$ and K/F be an algebraic extension. Let L/F be the separable closure*[11] *of F in K. Then, K/L is purely inseparable.*

Proof. Pick any element $\alpha \in K$ not in L. We wish to show that $\alpha^{p^n} \in L$ for some $n \geq 0$. Consider $m_{\alpha,F}(x) \in F[x]$. Observe that $m_{\alpha,F}(x)$ is inseparable as $\alpha \notin L$. By Polynomial Frobenius (Proposition 16.6.8.2), it follows that $m_{\alpha,F}(x) = f(x^{p^n})$ for some irreducible separable $f(x) \in F[x]$. It follows that $m_{\alpha^{p^n},F}(x) = f(x)$ and thus α^{p^n} is a separable element, that is, $\alpha^{p^n} \in L$, as needed. □

Inseparability index

Our goal now is tom understand the deviation of an algebraic extension from separability. Recall that perfect fields have no deviation (Theorem 16.6.4.19). Hence, answering this question would shed light on characteristic $p > 0$ algebra.

We first begin by observing that separable degree always divides the degree in characteristic $p > 0$(!)

Proposition 16.6.8.8. *Let K/F be a finite extension where* $\text{char}(F) = p > 0$*. Then*

$$[K : F]_s \mid [K : F].$$

[11] see Definition 16.6.4.16.

Proof. By Proposition 16.6.4.10, it suffices to show the above statement for $K = F(\alpha)$ for some $\alpha \in K$. Now since

$$\begin{aligned}[F(\alpha) : F]_s &= \left|\mathrm{Hom}_F\left(F(\alpha), \bar{F}\right)\right| \\ &= \text{\# of distinct roots of } m_{\alpha,F}(x) \text{ in } \bar{F}.\end{aligned}$$

Further, since

$$\begin{aligned}[F(\alpha) : F] &= \deg m_{\alpha,F}(x) \\ &= \text{\# of total roots of } m_{\alpha,F}(x) \text{ in } \bar{F},\end{aligned}$$

therefore it suffices to show that each root $m_{\alpha,F}(x)$ is repeated same no. of times in $\bar{F}$, that is, multiplicity of each root of $m_{\alpha,F}(x)$ is same. Indeed, by Polynomial Frobenius (Proposition 16.6.8.2), we have an irreducible and separable $f(x) \in F[x]$ such that

$$m_{\alpha,F}(x) = f(x^{p^n})$$

for some $n \geq 0$. Let $\alpha_1, \dots, \alpha_m \in \bar{F}$ be the distinct roots of $m_{\alpha,F}(x)$. Observe that $\alpha_i^{p^n}$ is a root of $f(x)$ for each $i = 1, \dots, m$. It is clear that the function

$$\begin{aligned}\{\text{Roots of } m_{\alpha,F}(x)\} &\longrightarrow \{\text{Roots of } f(x)\} \\ \alpha_i &\longmapsto \alpha_i^{p^n}\end{aligned}$$

is surjective. Indeed, since $\deg m_{\alpha,F}(x) \geq \deg f(x)$. Thus, every root of $f(x)$ is of the form $\alpha_i^{p^n}$. Thus, we get

$$f(x) = (x - \alpha_1^{p^n}) \dots (x - \alpha_m^{p^n}).$$

Thus

$$\begin{aligned}m_{\alpha,F}(x) = f(x^{p^n}) &= (x^{p^n} - \alpha_1^{p^n}) \dots (x^{p^n} - \alpha_m^{p^n}) \\ &= (x - \alpha_1)^{p^n} \dots (x - \alpha_m)^{p^n},\end{aligned}$$

as needed. This completes the proof. ∎

We state one of the important consequences of the proof above.

Corollary 16.6.8.9 (Minimal polynomials in char p)**.** *Let K/F be a finite extension where* $\mathrm{char}(F) = p > 0$*. If $\alpha \in K$, then every root of $m_{\alpha,F}(x)$ has same multiplicity equal to p^n for some $n \geq 0$. In particular, $p | \deg m_{\alpha,F}(x)$.*

Proof. In the proof of Proposition 16.6.8.8, we deduced that if $\alpha_1, \dots, \alpha_m \in \bar{F}$ are roots of $m_{\alpha,F}(x)$, then

$$m_{\alpha,F}(x) = (x - \alpha_1)^{p^n} \dots (x - \alpha_m)^{p^n}$$

as required. ∎

Remark 16.6.8.10. The Corollary 16.6.8.9 generalizes the statement in Theorem 16.6.8.3, 3, in the sense that a purely inseparable extension is a finite extension of F with $\mathrm{char}(F) = p > 0$ such that every element has minimal polynomial with only one root with multiplicity p^n. In precise terms, we have the following result.

Corollary 16.6.8.11. *Let K/F be a finite extension where $\mathrm{char}(F) = p > 0$. Then the following are equivalent:*

1. *K/F is a purely inseparable extension.*
2. *Every element $\alpha \in K$ not in F has minimal polynomial which has only one distinct root.*

Proof. (1. $\Rightarrow$ 2.) This is clear from Theorem 16.6.8.3.

(2. $\Rightarrow$ 1.) As K/F is finite, therefore by Corollary 16.6.8.9, we have

$$m_{\alpha,F}(x) = (x - \alpha)^{p^n} = x^{p^n} - \alpha^{p^n}$$

in $K[x]$. However, comparing the equality above in $F[x]$, we deduce that $\alpha^{p^n} \in F$, as required. $\square$

Remark 16.6.8.12. Now consider $K = F(\alpha)$ over F where $\mathrm{char}(F) = p > 0$ and α algebraic over F. Then we saw in Corollary 16.6.8.9 that $m_{\alpha,F}(x)$ has every root repeated p^n many times for some $n \geq 0$. We can capture this common multiplicity of roots as $[K : F]/[K : F]_s$ since $[K : F]$ is the total number of roots of $m_{\alpha,F}(x)$ and $[K : F]_s$ is the number of distinct roots of $m_{\alpha,F}(x)$, so that the ratio will yield us the common multiplicity which is p^n. If $p^n = 1$ (i.e. $n = 0$), then we see that $m_{\alpha,F}(x)$ has no repeated roots. It follows that $F(\alpha)/F$ would then be separable. This fraction is thus storing information about separability of an extension. We now generalize this for not necessarily principal extensions.

Definition 16.6.8.13 (**Inseparability index**)**.** Let K/F be a finite extension where $\mathrm{char}(F) = p > 0$. Then the inseparability index of K/F is defined to be

$$[K : F]_i := \frac{[K : F]}{[K : F]_s}.$$

As both usual degree and separable degree satisfies tower law, therefore inseparability index also satisfies tower law.

Lemma 16.6.8.14. *Let $L/K/F$ be finite extensions where $\mathrm{char}(F) = p > 0$. Then,*

$$[L : F]_i = [L : K]_i \cdot [K : F]_i.$$

Proof. Immediate. $\square$

Using the tower law, we observe that inseparability index is always a power of characteristic.

Lemma 16.6.8.15. *Let K/F be a finite extension where $\mathrm{char}(F) = p > 0$. Then $[K : F]_i = p^k$ for some $k \geq 0$.*

Proof. As K/F is finite and inseparability index satisfies tower law (Lemma 16.6.8.14), we may reduce to showing that $[F(\alpha) : F]_i$ is a power of p. Indeed, observe that $[F(\alpha) : F] = \deg m_{\alpha,F}$ and $[F(\alpha) : F]_i = \#$ distinct roots of $m_{\alpha,F}$. By Corollary 16.6.8.9, we deduce that $\deg m_{\alpha,F} = (\#$ distinct roots of $m_{\alpha,F}\ ) \cdot (p^n)$ where p^n is the common multiplicity of each root of $m_{\alpha,F}$. Hence, $[F(\alpha) : F]_i$ is p^n, as required. $\square$

It should be clear that if K/F is purely inseparable, then $[K : F]_i = 1$. We now correctly prove it.

Lemma 16.6.8.16. *Let K/F be a finite and purely inseparable extension where* $\mathrm{char}(F) = p > 0$. *Then,*

$$[K : F]_s = 1.$$

Proof. By tower law for separable degree (Proposition 16.6.4.10), we may assume that $K = F(\alpha)$. As $[F(\alpha) : F]_s$ is the number of distinct zeroes of $m_{\alpha,F}(x)$, therefore by Corollary 16.6.8.11, we win. □

The following is a simple, yet enlightening observation.

Lemma 16.6.8.17. *Let F be a field of characteristic $p > 0$. If K/F a purely inseparable extension, then it is normal.*

Proof. Indeed, as $\alpha \in K$ is such that $m_{\alpha,F}(x)|x^{p^n} - a$ for some $a = \alpha^{p^n} \in F$, therefore all distinct roots of $m_{\alpha,F}(x)$ are distinct roots of $x^{p^n} - a$ as well. However, over K we have $x^{p^n} - a = x^{p^n} - \alpha^{p^n} = (x - \alpha)^{p^n}$. Thus, $x^{p^n} - a$ has only one distinct root, it follows that $m_{\alpha,F}(x)$ has only one distinct root, $\alpha \in K$. Since $\alpha \in K$ is arbitrary, hence K/F is normal, as required. □

16.6.9 Transcendence degree

Definition 16.6.9.1. (**Transcendence**) Let K/k be a field extension.

1. A collection of elements $\{\alpha_i\}_{i\in I}$ of K is said to be *algebraically independent* if the map
$$\begin{aligned} k[x_i \mid i \in I] &\longrightarrow K \\ x_i &\longmapsto \alpha_i \end{aligned}$$
is injective.
2. A *transcendence basis* of K/k is defined to be an algebraically independent set $\{\alpha_i \mid i \in I\}$ of K/k such that $K/k(\alpha_i \mid i \in I)$ is an algebraic extension.
3. The extension K/k is said to be *purely transcendental* if $K \cong k(x_i \mid i \in I)$ for some indexing set I.

Lemma 16.6.9.2. *Let K/k be a field extension. Then, $\{\alpha_i\}_{i\in I}$ is a transcendence basis of K/k if and only if $\{\alpha_i\}_{i\in I}$ is a maximal algebraically independent set of K/k.*

Proof. (L $\Rightarrow$ R) If $\{\alpha_i\}_{i\in I}$ is not maximal, then there exists $S \subset K$ containing $\{\alpha_i\}_{i\in I}$ such that S is algebraically independent. Let $\beta \in S \setminus \{\alpha_i\}_{i\in I}$. But since $K/k(\{\alpha_i\}_{i\in I})$ is an algebraic extension and $\beta \notin k(\{\alpha_i\}_{i\in I})$ by algebraic independence of S, therefore we have a contradiction to algebraic nature of the extension $K/k(\{\alpha_i\}_{i\in I})$.
(R $\Rightarrow$ L) Suppose $K/k(\{\alpha_i\}_{i\in I})$ is not algebraic. Then there exists $\beta \in K$ which is transcendental over $k(\{\alpha_i\}_{i\in I})$. Thus the set $\{\alpha_i\}_{i\in I} \cup \{\beta\}$ is a larger algebraically independent set, contradicting the maximality. □

Lemma 16.6.9.3. *Let K/k be a field extension. Then any two transcendence basis have the same cardinality.*

Proof. See Tag 030F of cite[Stacksproject]. □

Definition 16.6.9.4. (**Transcendence degree**) Let K/k be a field extension. The cardinality of any transcendence basis is said to be the transcendence degree, denoted trdeg K/k. Furthermore, if A is a domain containing k, then we define trdeg A/k to be the transcendence degree of A_0, the field of fractions of A, over k.

Remark 16.6.9.5. Let K/k be a field extension. If trdeg $K/k = 1$, then there exists $\alpha \in K$ such that α is not an algebraic element over k but $K/k(\alpha)$ is algebraic. In particular, for any transcendental element $\alpha \in K$ over k, the set $\{\alpha\}$ is algebraically independent over k. Precisely, there is a one-to-one bijection between the set of all singletons which are algebraically independent and all transcendental elements of K/k.

Example 16.6.9.6. There are some basic examples which reader might have encountered. For example, one knows that $\mathbb{Q}(\pi)/\mathbb{Q}$ is transcendental as $\pi \in \mathbb{Q}(\pi)$ is not algebraic over $\mathbb{Q}$. Consequently, trdeg $\mathbb{Q}(\pi)/\mathbb{Q}$ is 1, as $\mathbb{Q}(\pi)/\mathbb{Q}(\pi)$ is algebraic.

For another example, consider the next obvious situation of $\mathbb{Q}(e,\pi)/\mathbb{Q}$. Since $\{e\}$ and $\{\pi\}$ are algebraically independent sets over $\mathbb{Q}$, therefore trdeg in this case is ≥ 1. But it is an unknown problem whether $\{e,\pi\}$ forms an algebraically independent set over $\mathbb{Q}$(!) Consequently, if they do, then trdeg $\mathbb{Q}(e,\pi)/\mathbb{Q} = 2$ and if they don't, then the best we can say is trdeg $\mathbb{Q}(e,\pi)/\mathbb{Q} \geq 1$.

Example 16.6.9.7. We have trdeg $k(x_1, \dots, x_n)/k = n$ as $\{x_1, \dots, x_n\}$ forms a maximal algebraically independent set.

We observe some basic first properties of transcendence degree. First, transcendence degree satisfies additive tower law.

Lemma 16.6.9.8 (Additive tower law)**.** *Let $L/K/k$ be field extensions. Then*

$$\text{trdeg } L/k = \text{trdeg } L/K + \text{trdeg } K/k.$$

The following shows that that whatever transcendence degree of a k-algebra may be, there will be that many transcendental elements in it.

Lemma 16.6.9.9. *Let $A = k[\alpha_1, \dots, \alpha_n]$ be an integral domain where $\alpha_i \in K$ for some field extension K/k. If* trdeg $A/k = r > 0$*, then there exists $\alpha_{i_1}, \dots, \alpha_{i_r}$ which are transcendental over k.*

16.7 Integral dependence and normal domains

The main topic of interest of study in this section is the following question: "*let R be a ring and S be an R-algebra. How do all those elements of S behave like which satisfy a polynomial with coefficients in R?*".

16.7.1 Definitions and basic theory

In order to investigate this further, let us bring some definitions.

Definition 16.7.1.1. (**Integral elements and integral algebra**) Let R be a ring and S be an R-algebra. An element $s \in S$ for which there exists $p(x) \in R[x]$ such that $p(s) = 0$ in S is said to be an *integral element* over R. Further, S is said to be *integral over* R if every element of S is integral over R.

To begin deriving properties, we would need a fundamental result about endomorphisms of finitely generated modules.

Theorem 16.7.1.2 (Cayley-Hamilton)**.** *Let R be a ring, M be a finitely generated R-module generated by n elements and $I \leq R$ be an ideal. If $\varphi : M \to M$ is an R-linear map such that*

$$\varphi(M) \subseteq IM,$$

then there exists a monic polynomial

$$p(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

in $R[x]$ such that $p(\varphi) = 0$ in $\mathrm{Hom}_R(M, M)$ and $a_k \in I^k$ for $k = 1, \ldots, n$.

Proof. See Theorem 4.3, pp 120, [cite Eisenbud]. □

There are two immediate corollaries of Cayley-Hamilton which will remind the reader of finite-dimensional vector space case.

Corollary 16.7.1.3. *Let R be a ring and M be a finitely generated R-module. If $\phi : M \to M$ is a surjective R-module homomorphism, then ϕ is an isomorphism.*

Proof. Using ϕ, we may regard M as an $R[z]$-module. Note that M is a finitely generated $R[z]$-module. Let $I = \langle z \rangle \leq R[z]$. Since the action of z on M is by ϕ and ϕ is surjective, therefore $IM = M$. We may use Cayley-Hamilton with $\varphi = \mathrm{id}$ to deduce that there is a polynomial $p(x, z) \in R[x, z]$ such that $p(z, \mathrm{id}) = 0$ and $p(x, z)$ is a monic polynomial in $R[z][x]$. Consequently, we can write $0 = p(z, \mathrm{id}) = 1 + q(z)z$ for some $q(z) \in R[z]$. It follows that $-q(z)$ is the inverse of z in $R[z]$. Since $z \in R[z]$ denotes the endomorphism ϕ, so we have found an R-linear inverse of ϕ, namely the one corresponding to $-q(z)$, as required. □

Corollary 16.7.1.4. *Let R be a ring and M be a finitely generated R-module. If $M \cong R^n$, then any generating set of n elements of M is linearly independent. In particular, any generating set of n elements of M is a basis.*

Proof. Denote $f : M \to R^n$ to be the given isomorphism. Pick $S = \{s_1, \dots, s_n\}$ to be a generating set of M. This yields a surjection $g : R^n \to M$. We wish to show that g is an isomorphim. Observe that $gf : M \to M$ is surjective. It follows from Corollary 16.7.1.3 that gf is an isomorphism. Since f is an isomorphism, hence it follows that g is an isomorphism, as required. □

Here is an another application of Cayley-Hamilton.

Corollary 16.7.1.5. *Let $R \neq 0$ be a ring. If $\varphi : R^m \to R^n$ is an injective R-linear map, then $m \leq n$*[12].

Proof. Assuming to the contrary, assume that $m > n$ and φ is injective. Then we have the composite which is also injective:

$$\psi : R^m \xrightarrow{\varphi} R^n \hookrightarrow R^m,$$

where the latter is the inclusion into first n-coordinates. By Cayley-Hamilton, we get that ψ satisfies a monic polynomial in R:

$$\psi^k + r_{k-1}\psi^{k-1} + \cdots + r_1\psi + r_0 = 0.$$

We may assume that k is least possible. If $r_0 = 0$, then by injectivity of ψ, we would have an even smaller degree polynomial which annihilates ψ, not possible. Hence $r_0 \neq 0$ and thus applying the above polynomial at $e_m = (0, 0, \dots, 0, 1)$ gives that $r_0 = 0$, a contradiction. □

The fundamental result which drives the basic results about integral algebras is the following equivalence.

Proposition 16.7.1.6. *Let $R \to S$ be an R-algebra and $s \in S$. Then the following are equivalent.*

1. *$s \in S$ is integral over R.*
2. *$R[s] \subseteq S$ is a finite R-algebra.*
3. *$R[s] \subseteq S$ is contained in a finite R-algebra.*
4. *There is a faithful $R[s]$-module M which when restricted to R is finitely generated as an R-module.*

Proof. $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4$ follows at once. We do $4 \Rightarrow 1$. Indeed, let $I = \langle s \rangle \leq R[s]$ be the ideal generated by $s \in R[s]$. Consequently, s induces an endomorphism $m_s : M \to M$ by scalar multiplication. Observe that $m_s(M) = IM$. It follows by Cayley-Hamilton (Theorem 16.7.1.2) that there exists a monic $p(x) \in R[s][x]$ such that $p(m_s) = 0$ as an $R[s]$-linear map $M \to M$. Consequently, for any $a \in M$, we have $p(m_s)(a) = 0$, where upon expanding one sees that $p(m_s) = m_{q(s)}$ for some $q(s) \in R$, $q(x) \in R[x]$. But since M is faithful, therefore $q(s) = 0$, as required. □

Lemma 16.7.1.7. *Let $R \to S$ be an R-algebra and $s_1, \dots, s_n \in S$ be integral over R. Then $R[s_1, \dots, s_n]$ is a finite R-algebra.*

Proof. We proceed by induction over n. Base case follows from Proposition 16.7.1.6. Assume that $R_k = R[s_1, \dots, s_k]$ is a finite R-algebra. Since $s_{k+1} \in S$ is integral over R, therefore it is integral over R_k. It follows from Proposition 16.7.1.6 that $R_k[s_{k+1}]$ is a finite R_k-algebra. Since R_k is a finite R-algebra, therefore $R_k[s_{k+1}] = R[s_1, \dots, s_{k+1}]$ is a finite R-algebra, as required. □

[12]proof is taken from here.

One then obtains that finite generation of an algebra by integral elements as an algebra is equivalent to finite generation as an R-module.

Lemma 16.7.1.8. *Any finite R-algebra is integral over R.*

Proof. Let S be a finite R-algebra and let $s \in S$ be an element. Let $m_s : S \to S$ be the R-linear given by multiplication by s. As S is a finitely generated R-module, then by Cayley-Hamilton (Theorem 16.7.1.2), it follows that there is a monic $p(x) \in R[x]$ such that $p(m_s) = 0$ as an R-linear map. Applying $p(m_s)$ to $1 \in S$ yields $p(s) = 0$, as required. □

Proposition 16.7.1.9. *Let R be a ring and S be an R-algebra. Then the following are equivalent.*

1. *S is a finite R-algebra.*
2. *$S = R[s_1, \dots, s_n]$ where $s_1, \dots, s_n \in S$ are integral over R. In particular, S is integral over R.*

That is, an R-algebra is finite if and only if it is a finite type and integral R-algebra.

Proof. Observe that 2. ⇒ 1. is just Lemma 16.7.1.7. For 1. ⇒ 2. proceed as follows. By Lemma 16.7.1.8, it follows that S is integral over R. Let $s_1, \dots, s_n \in S$ be a generating set of S as an R-module. It is now clear that $R[s_1, \dots, s_n] = S$ as S is finitely generated. □

The following result show that all integral elements form a subring of S.

Proposition 16.7.1.10. *Let R be a ring and S be an R-algebra. The set of all elements of S integral over R forms a subalgebra of S, called the integral closure of R in S.*

Proof. Let $s, t \in S$ be integral over R. Then $R[s, t]$ is a subalgebra of S. It suffices to show that every element of $R[s, t]$ is integral over R. By Proposition 16.7.1.9, the algebra $R[s, t]$ is integral over R as it is finite by Lemma 16.7.1.7. □

With this, a natural situation is when every element of S is integral over R.

Definition 16.7.1.11. (**Normalization & integral extension**) Let R be a ring and S be an R-algebra. The subalgebra A of all integral elements of S over R is said to be the *integral closure of S over R*. One also calls A the *normalization of R in S*. If S is fraction field of R, then A is also denoted by $\tilde{R}$. Further, if $R \hookrightarrow S$ is a ring extension and every element of S is integral over R, then S is said to be an integral extension of R. If $f : R \to S$ is an integral R-algebra, then the map f is said to be *integral*.

Composition of integral maps is integral.

Lemma 16.7.1.12. *Let $R \to S$ and $S \to T$ be integral maps. Then the composite $R \to S \to T$ is integral.*

Proof. Pick any element $t \in T$. We wish to show that $R[t]$ is contained in a finite R-algebra by Proposition 16.7.1.6. As $S \to T$ is integral, there exists $p(x) \in S[x]$ monic such that $p(t) = 0$. So we have

$$t^n + s_{n-1}t^{n-1} + \cdots + s_1 t + s_0 = 0$$

in T where $s_i \in S$. Let $S' = R[s_0, \dots, s_{n-1}]$. As $R \to S$ is integral, therefore S' is a finite R-algebra by Lemma 16.7.1.7. Note that $R \subseteq S'$. By the above equation, it then follows that $S'[t]$ is a finite S'-algebra. As composition of finite maps is finite, therefore $S'[t]$ is a finite R-algebra containing $R[t]$, as required. □

Another trivial observation is that a map which factors an integral map becomes integral.

Lemma 16.7.1.13. *Let $A \to C$ be an integral map. If there is a map $A \to B$ such that*

$$\begin{tikzcd} A \arrow[r] \arrow[d] & C \\ B \arrow[ur] & \end{tikzcd}$$

commutes, then $B \to C$ is an integral map.

Proof. Pick any element $c \in C$. There exists non-zero monic $p(x) \in A[x]$ such that $p(x)$ is non-zero in $C[x]$ and $p(c) = 0$ in C. Observe that $p(x) \in B[x]$ is also a non-zero monic as if not then $p(x)$ would be zero in $C[x]$ because the above triangle commutes. The result then follows. □

The following observation is simple to see, but comes in very handy while handling intermediate rings that pop-up while subsequent localizations.

Lemma 16.7.1.14. *Let k be a field and A be an integral k-algebra. Then A is a field.*

Proof. Pick any element $a \in A$. By integrality, there exists $c_i \in k$ such that

$$a^n + c_{n-1}a^{n-1} + \cdots + c_1 a + c_0 = 0$$

in A. Consider this equation in the fraction field $Q(A)$ to multiply by a^{-1}, so that we may get

$$a^{n-1} + c_{n-1}a^{n-2} + \cdots + c_2 a + c_1 + c_0 a^{-1} = 0$$

in $Q(A)$. It thus follows that a^{-1} is a polynomial in A with coefficients in k, that is, $a^{-1} \in A$, as required. □

16.7.2 Normalization & normal domains

A special situation in Definition 16.7.1.11 is when R is a domain and S is its fraction field. These domains will play a crucial role later on, especially in arithmetic.

Definition 16.7.2.1. (**Normal domain**) Let R be a domain and S be its fraction field. If the normalization of R in S is R itself, then R is said to be a normal domain.

Example 16.7.2.2. Let R be a domain, K its fraction field and $\tilde{R}$ be the normalization of R in K. It follows that $\tilde{R} \hookrightarrow K$ is a normal domain. Indeed, let $\hat{R}$ be normalization of $\tilde{R}$ in K. Then, we have maps

$$R \hookrightarrow \tilde{R} \hookrightarrow \hat{R}$$

where both inclusions are integral maps by construction. It follows from Lemma 16.7.1.12 that the inclusion $R \hookrightarrow \hat{R}$ is integral, forcing $\hat{R} \subseteq \tilde{R}$ which further implies $\tilde{R} = \hat{R}$.

Further investigation into normal domains lets us identify all UFDs as normal domains.

Proposition 16.7.2.3. *All unique factorization domains are normal domains.*

Proof. Let R be a UFD and K be its fraction field. Let $\frac{a}{b} \in K$ with $\gcd(a, b) = 1$. Suppose $\frac{a}{b}$ is integral over R so that there exists $p(x) = x^n + c_{n-1}x^{n-1} + \dots c_1 x + c_0 \in R[x]$ such that $p(a/b) = 0$. It follows by rearrangement that

$$a^n + c_{n-1}ba^{n-1} + \dots c_1 b^{n-1} + c_0 b^n = 0.$$

Hence, $b|a^n$. As $\gcd(a, b) = 1$, hence we deduce thay $b|a$, a contradiction. □

Example 16.7.2.4. Consequently, $\mathbb{Z}$ and $\mathbb{Z}[x_1, \dots, x_n]$ are normal as well. Moreover, as Gauss' lemma states that R is UFD if and only if $R[x]$ is UFD, therefore we deduce that $R[x_1, \dots, x_n]$ is a normal domain if R is UFD.

We have something similar to Gauss' lemma for normal domains.

Proposition 16.7.2.5. *A ring R is normal if and only if $R[x]$ is normal.*

Proof. **TODO.** □

Further, we can obtain a generalization of the fact that a monic irreducible in $\mathbb{Z}[x]$ is irreducible in $\mathbb{Q}[x]$.

Proposition 16.7.2.6. *Let $R \hookrightarrow S$ be a ring extension and let $f \in R[x]$ be a monic polynomial. If $f = gh$ in $S[x]$ where g and h are monic, then the coefficients of g and h are integral over R.*

We also obtain that any monic irreducible in the polynomial ring in one variable over a normal domain is prime.

Lemma 16.7.2.7. *Let R be a ring and $f(x) \in R[x]$ be a monic irreducible. If R is a normal domain, then $f(x)$ is a prime element.*

Thus, for normal domains R, monic irreducible and monic prime polynomials are equivalent concepts.

We now show that normalization is a very hereditary process as it preserves many properties of the original ring. Indeed, we first show that normalization and localization commutes.

Proposition 16.7.2.8. *Let $f : R \to S$ be an R-algebra and $M \subseteq R$ be a multiplicative set. If $A \subseteq S$ is the integral closure of R in S, then $M^{-1}A$ is the integral closure of $M^{-1}R$ in $M^{-1}S$.*

Proof. We may assume that f is inclusion of a subring of S by replacing R by $f(R)$ and M by $f(M)$. Consequently, we have inclusions $R \hookrightarrow A \hookrightarrow S$ which induces inclusions $M^{-1}R \hookrightarrow M^{-1}A \hookrightarrow M^{-1}S$. We wish to show that $M^{-1}A$ is the integral closure of $M^{-1}R$ in $M^{-1}S$. Pick an element $s/m \in M^{-1}S$ where $m \in M$ which is integral over $M^{-1}R$. Consequently, there exists $r_i/m_i \in M^{-1}R$ for $0 \leq i \leq k-1$ such that

$$\left(\frac{s}{m}\right)^k + \frac{r_{k-1}}{m_{k-1}}\left(\frac{s}{m}\right)^{k-1} + \dots \frac{r_1}{s_1}\left(\frac{s}{m}\right) + \frac{r_0}{m_0} = 0$$

in $M^{-1}S$. Multiplying by product of denominators and absorbing coefficients into r_i, we get

$$m' s^k + r_{k-1}s^{k-1} + \dots + r_1 s + r_0 = 0$$

which we may multiply by $(m')^{k-1}$ to get

$$(m's)^k + r_{k-1}(m's)^{k-1} + \cdots + r_1(m')^{k-2}(m's) + r_0(m')^{k-1} = 0.$$

It follows that $m's \in A$, thus $s/1 \in M^{-1}A$ and thus $s/m \in M^{-1}A$.

Conversely, pick an element $a/m \in M^{-1}A$. We wish to show that it is integral over $M^{-1}R$. As $a \in A$, therefore we have

$$a^n + r_{n-1}a^{n-1} + \ldots a_1 r + a_0 = 0$$

for $r_i \in R$. This equation in $M^{-1}S$ can be divided by m^n to obtain

$$\left(\frac{a}{m}\right)^n + \frac{r_{n-1}}{m}\left(\frac{a}{m}\right)^{n-1} + \cdots + \frac{r_1}{m^{n-1}}\left(\frac{a}{m}\right) + \frac{r_0}{m^n} = 0.$$

It follows that a/m is integral over $M^{-1}R$, as required. □

An immediate, but important corollary of the above is the following.

Corollary 16.7.2.9. *Let A be a domain, K be its fraction field and $\tilde{A}$ be its normalization. Then, for all $g \in A$, we have $\tilde{A}_g = \widetilde{A_g}$ in K.* □

Another important corollary is that being a normal domain is a local property.

Proposition 16.7.2.10. *Let R be a domain. Then the following are equivalent:*

1. *R is a normal domain.*
2. *$R_{\mathfrak{p}}$ is a normal domain for each prime $\mathfrak{p} \in \operatorname{Spec}(R)$.*
3. *$R_{\mathfrak{m}}$ is a normal domain for each maximal $\mathfrak{m} \in \operatorname{Spec}(R)$.*

Proof. By Proposition 16.7.2.8, we immediately have that (1. ⇒ 2.) and (1. ⇒ 3.). The (2. ⇒ 3.) is immediate. We thus show (3. ⇒ 1.). Let K be the fraction field of R. Observe that each $R_{\mathfrak{m}}$ is a domain and have fraction field K again, where $\mathfrak{m} \in \operatorname{Spec}(R)$ is a maximal ideal. Thus we have $R \hookrightarrow R_{\mathfrak{m}} \hookrightarrow K$. Pick $x \in K$ which satisfies a monic polynomial over R. It follows that x satisifes a monic polynomial over $R_{\mathfrak{m}}$ for each maximal $\mathfrak{m} \in \operatorname{Spec}(R)$. Thus $x \in R_{\mathfrak{m}}$ for each $\mathfrak{m}$ as $R_{\mathfrak{m}}$ is a normal domain. We thus deduce from Lemma 16.1.2.12 that $x \in \bigcap_{\mathfrak{m} \neq R} R_{\mathfrak{m}} = R$, as required. □

Remark 16.7.2.11 (*Normalization is a strongly local construction*)**.** Let A be an arbitrary domain. Then we get an inclusion $\varphi_A : A \hookrightarrow \tilde{A}$ where $\tilde{A}$ is the normalization of A in its fraction field. We claim that the collection of maps $\{\varphi_A : A \hookrightarrow \tilde{A}\}$ one for each domain is a construction which is strongly local on domains (see Definitions 1.6.2.3 & 1.6.2.4).

Indeed, first $\{\varphi_A : A \hookrightarrow \tilde{A}\}$ is a construction on domains as if $\eta : A \to B$ is an isomorphism, then we have an isomorphism $\tilde{\eta} : \tilde{A} \to \tilde{B}$ given as follows: we have an isomorphism $\bar{\eta} : K_A \to K_B$ between their fraction fields, given by $a/a' \mapsto \eta(a)/\eta(b)$. Now $a/a' \in K_A$ is integral over A if and only if $\eta(a)/\eta(a') \in K_B$ is integral over B. This shows that $\bar{\eta} : K_A \to K_B$ restricts to an isomorphism $\tilde{\eta} : \tilde{A} \to \tilde{B}$. Moreover, if $\eta : A \to A$ is id, then so is $\tilde{\eta}$ and it satisfies the square and cocycle condition as well of Definition 1.6.2.3. We now claim that normalization is strongly local.

Indeed, pick $g \in A$ non-zero. Then, the localization of the inclusion $\varphi_A : A \hookrightarrow \tilde{A}$ at element g yields $(\varphi_A)_g : A_g \hookrightarrow \tilde{A}_g = \widetilde{A_g}$ which is equal to the normalization of the domain $\varphi_{A_g} : A_g \hookrightarrow \widetilde{A_g}$. It follows that any integral scheme X admits a *normalization* in light of Theorem 1.6.2.10. Indeed, this is what is the content of Theorem 1.6.6.3.

We have a universal property for normalization of domains.

Proposition 16.7.2.12. *Let A be a domain and $\tilde{A}$ be the normalization of A in its fraction field. Then for any normal domain B and an injective map $A \hookrightarrow B$, there exists a unique map $\tilde{A} \to B$ such that following commutes:*

$$\begin{tikzcd} \tilde{A} \arrow[r, dashed] & B \\ A \arrow[u, hook] \arrow[ur, hook] & \end{tikzcd}.$$

Proof. Let $f : A \hookrightarrow B$. This, by universal property of fraction fields, induces a unique injective map $\varphi : K \hookrightarrow L$ from fraction field of A to that of B such that $\varphi|_A = f$. Let $x \in \tilde{A}$. Then

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

holds in K where $a_i \in A$. Applying φ on the above equation yields

$$\varphi(x)^n + f(a_{n-1})\varphi(x)^{n-1} + \cdots + f(a_1)\varphi(x) + f(a_0) = 0$$

in L. It follows that $\varphi(x)$ is an integral element of L over B. As B is normal it follows that $\varphi(x) \in B$. Consequently, we have a unique map

$$\varphi|_{\tilde{A}} : \tilde{A} \to B$$

such that the triangle commutes, as required. □

In certain situations (especially those arising in geometry and arithmetic), normalization preserves noetherian property. **TODO.**

16.7.3 Noether normalization lemma

Finally, as a big use of normalization in geometry, we obtain the following famous result.

Theorem 16.7.3.1. [13] *Let k be a field and A be a finite type k-algebra. Then, there exists elements $y_1, \dots, y_r \in A$ algebraically independent over k such that the inclusion $k[y_1, \dots, y_r] \hookrightarrow A$ is an integral map.*

Proof. Let us assume that k is infinite. Let $x_1, \dots, x_n \in A$ be generators of A as a k-algebra. Suppose there is no algebraically independent subset of $\{x_1, \dots, x_n\}$. Thus, each $x_1, \dots, x_n$ is integral over k. As $A = k[x_1, \dots, x_n]$, therefore by Proposition 16.7.1.9 it follows that A is integral over k, so there is nothing to show here.

Consequently, we may assume that there is a largest algebraically independent subset of $\{x_1, \dots, x_n\}$, denoted $\{x_1, \dots, x_r\}$. It follows that each $x_{r+1}, \dots x_n$ is integral/algebraic over k. If $r = n$, then A is the affine n-ring over k, so there is nothing to show. Consequently, we may assume that $n > r$. We now proceed by induction over n.

In the base case, we have $n = 1$, and thus $r < 1$. It follows that $A = k[x]$ where $x \in A$ is algebraically dependent over k, that is, x is integral over k. Consequently, A is integral over k by

[13]Exercise 5.16 of AMD.

Lemma 16.7.1.7 and there is nothing to show. We now do the inductive case.

Assume that every finite type k-algebra $B \subseteq A$ with $n-1$ generators have elements $\{y_1, \dots, y_m\} \subseteq B$ algebraically independent over k such that B is integral over $k[y_1, \dots, y_m]$. Denote $A_{n-1} = k[x_1, \dots, x_{n-1}] \subseteq A$. It now suffices to find a finite type k-algebra $B \subseteq A$ generated by $n-1$ elements not containing x_n such that the following two statements hold about B:

1. $x_n \in A$ is integral over B,
2. $B[x_n] = A$.

For if such a B exists, then we have integral maps $k[y_1, \dots, y_m] \hookrightarrow B$ and $B \hookrightarrow B[x_n] = A$ (Proposition 16.7.1.6). Then, by Lemma 16.7.1.12, it follows that $k[y_1, \dots, y_m] \hookrightarrow A$ is integral, as needed.

Indeed, first observe that since x_n is algebraic over k and $k \subseteq A_{n-1}$, therefore x_n is algebraic over A_{n-1}. Consequently, there is a polynomial $f(z_1, \dots, z_{n-1}, z_n) \in k[z_1, \dots, z_n]$ of total degree N such that $f(x_1, \dots, x_{n-1}, x_n) = 0$. Using this, we now construct the required algebra B as follows. Let F be the highest degree homogeneous part of f and denote it by

$$F(z_1, \dots, z_n) = \sum_{i_1 + \dots + i_n = N} c_{i_1 \dots i_n} z_1^{i_1} \dots z_n^{i_n}$$

where $c_{i_1 \dots i_n}$ and is 0 for those indices which are not present in F and is 1 for those which are present. Let $(\lambda_1, \dots, \lambda_{n-1}) \in k^{n-1}$ be a tuple such that $F(\lambda_1, \dots, \lambda_{n-1}, 1) \neq 0$. Such a tuple exists because the field is infinite (n might be arbitrarily large). Consequently, for each $0 \leq i \leq n-1$, consider the following elements of A:

$$x_i' = x_i - \lambda_i x_n.$$

Let $B = k[x_1', \dots, x_{n-1}'] \subseteq A$. We now show that above two hypotheses are satisfied by B. This will conclude the proof. First, we immediately have the second hypothesis as $B[x_n] = k[x_1', \dots, x_{n-1}', x_n] = k[x_1, \dots, x_n] = A$. We thus need only show that x_n is integral over B. This also follows by the way of construction of B; consider the polynomial

$$g(z_1, \dots, z_{n-1}, z_n) := f(z_1 + \lambda_1 z_n, \dots, z_{n-1} + \lambda_{n-1} z_n, z_n)$$

in $k[z_1, \dots, z_{n-1}, z_n]$. We wish to show the following two items

1. $g(z_1, \dots, z_{n-1}, z_n)$ is monic in z_n,
2. $g(x_1', \dots, x_{n-1}', x_n) = 0$.

This would suffice as a polynomial in $B[z_n]$ is just a polynomial in $k[x_1', \dots, x_{n-1}', z_n]$. Indeed, we see that

$$\begin{aligned}
g(z_1, \dots, z_{n-1}, z_n) &= f(z_1 + \lambda_1 z_n, \dots, z_{n-1} + \lambda_{n-1} z_n, z_n) \\
&= F(z_1 + \lambda_1 z_n, \dots, z_{n-1} + \lambda_{n-1} z_n, z_n) + \cdots \\
&= \sum_{i_1 + \dots + i_n = N} c_{i_1 \dots i_n} (z_1 + \lambda_1 z_n)^{i_1} \dots (z_{n-1} + \lambda_{n-1} z_n)^{i_{n-1}} z_n^{i_n} + \cdots \\
&= \left(\sum_{i_1 + \dots + i_n = N} c_{i_1 \dots i_n} \lambda_1^{i_1} z_n^{i_1} \dots \lambda_{n-1}^{i_{n-1}} z_n^{i_{n-1}} z_n^{i_n} \right) + \dots \\
&= z_n^N \left(\sum_{i_1 + \dots + i_n = N} c_{i_1 \dots i_n} \lambda_1^{i_1} \dots \lambda_{n-1}^{i_{n-1}} \right) + \cdots \\
&= z_n^N F(\lambda_1, \dots, \lambda_{n-1}, 1) + \cdots .
\end{aligned}$$

It follows that g is monic in z_n and $g(x_1', \ldots, x_{n-1}', x_n) = f(x_1, \ldots, x_{n-1}, x_n) = 0$. This completes the proof. $\square$

16.7.4 Dimension of integral algebras

We will cover Cohen-Seidenberg theorems about primes in an integral extension. The main theorem will allow us to deduce that, apart from other things, dimension of an integral R-algebra is equal to that of R.

16.8 Dimension theory

We will discuss the notion of dimension of rings and how that notion corresponds to dimension of the corresponding affine scheme. Further, the notion of dimension applied to algebraic geometry will garnish us with a concrete geometric intuition to situations which otherwise may feel completely sterile.

16.8.1 Dimension, height & coheight

As usual, all rings are commutative with 1.

Definition 16.8.1.1. (**Dimension of a ring**) Let R be a ring. Then $\dim R$ is defined as follows

$$\dim R := \sup_r \{\mathfrak{p}_0 \supsetneq \mathfrak{p}_1 \supsetneq \cdots \supsetneq \mathfrak{p}_r \mid \mathfrak{p}_i \text{ are prime ideals of } R\}.$$

Definition 16.8.1.2. (**Height/coheight of a prime ideal**) Let R be a ring and $\mathfrak{p} \lneq R$ be a prime ideal. Then height of $\mathfrak{p}$ is defined as follows:

$$\operatorname{ht} \mathfrak{p} := \sup_r \{\mathfrak{p} = \mathfrak{p}_0 \supsetneq \mathfrak{p}_1 \supsetneq \cdots \supsetneq \mathfrak{p}_r \mid \mathfrak{p}_i \text{ are prime ideals of } R\}.$$

Similarly, the coheight of $\mathfrak{p}$ is defined by

$$\operatorname{coht} \mathfrak{p} := \sup_r \{\mathfrak{p} = \mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_r \mid \mathfrak{p}_i \text{ are prime ideals of } R\}$$

Remark 16.8.1.3. Note that the dimension of a prime ideal $\mathfrak{p}$ as a ring may not be same as its height in R, as there might be many more primes in $\mathfrak{p}$ which may fail to be primes in the ring R. But clearly, $\dim \mathfrak{p} \geq \operatorname{ht} \mathfrak{p}$.

Recall that the *dimension of a topological space* X is defined as

$$\dim X = \sup_r \{Z_0 \supsetneq Z_1 \supsetneq \cdots \supsetneq Z_r \mid Z_i \text{ are irreducible closed subsets of } X\}.$$

We now have some immediate observations about height, coheight and dimension.

Lemma 16.8.1.4. *Let R be a ring. Then,*

1. $\operatorname{ht} \mathfrak{p} = \dim R_{\mathfrak{p}}$,
2. $\operatorname{coht} \mathfrak{p} = \dim R/\mathfrak{p}$,
3. $\operatorname{ht} \mathfrak{p} + \operatorname{coht} \mathfrak{p} \leq \dim R$.

Proof. Prime ideals of $R/\mathfrak{p}$ are in one-to-one order preserving bijection with prime ideals of R containing $\mathfrak{p}$. Prime ideals of $R_{\mathfrak{p}}$ are in one-to-one order preserving bijection with prime ideals of R contained in $\mathfrak{p}$. Let Y denote the length of all chains of prime ideals of R passing through $\mathfrak{p}$. Consequently, $\sup Y \leq \dim X$. But $\sup Y = \operatorname{ht} \mathfrak{p} + \operatorname{coht} \mathfrak{p}$. □

Lemma 16.8.1.5. *Let R be a PID. Then,* $\dim R = 1$. *Consequently, $\mathbb{Z}$ and $k[x]$ are one dimensional rings for any field k*[14].

[14] as the intuition agrees!

Proof. Any chain is either of the form $\langle 0\rangle$ or $\langle x\rangle \supsetneq \langle 0\rangle$. □

Further, by Theorem 16.1.5.3 we see the following.

Lemma 16.8.1.6. *If R is a PID which is not a field, then* $\dim R[x] = 2$.

Proof. Indeed, by Theorem 16.1.5.3, the longest chain of prime ideals of the form $o \lneq \langle f(x)\rangle \lneq \langle p, h(x)\rangle$ where $f(x)$ is irreducible and $h(x)$ is irreducible modulo prime $p \in R$, as one can see immediately. □

The following is also a simple assertion, which basically is why one introduces dimension of a ring.

Lemma 16.8.1.7. *Let R be a ring. Then,*

$$\dim \operatorname{Spec}(A) = \dim A.$$

Proof. Immediate from definitions and Lemma 1.2.1.1. □

Let us now give some more helpful notions, especially the dimension of an R-module.

Definition 16.8.1.8. (**Dimension of a module and height of ideals**) Let M be an R-module. Then the dimension of M is defined as

$$\dim M := \dim R/\operatorname{Ann}(M).$$

Further, for an ideal $I \leq R$, we define the height of I as the infimum of heights of all prime ideals above I:

$$\operatorname{ht} I := \inf\{\operatorname{ht} \mathfrak{p} \mid \mathfrak{p} \supseteq I,\ \mathfrak{p} \in \operatorname{Spec}(R)\}.$$

We have the corresponding topological result.

Lemma 16.8.1.9. *Let R be a ring and M be a finitely generated R-module. Then,*

$$\dim M = \dim \operatorname{Supp}(M)$$

where $\operatorname{Supp}(M) \subseteq \operatorname{Spec}(R)$ *is the support of the module M.*

Proof. The result follows as $\operatorname{Supp}(M)$ is the closed subset $V(\operatorname{Ann} M)$ so that any irreducible closed set in $\operatorname{Supp}(M)$ will be irreducible closed in $\operatorname{Spec}(R)$ and then we can use Lemma 1.2.1.1. □

16.8.2 Dimension of finite type k-algebras

In algebraic geometry, one is principally interested in finite type algebras over a field. Thus it is natural to engage in the study of their dimensions. We discuss some elementary results in this direction in this section. See Section 16.1.6 for basics of finite type k-algebras.

The main results are as follows.

Theorem 16.8.2.1. *Let k be a field and A be a finite type k-algebra which is a domain*[15]*. Then,*

$$\dim A = \operatorname{trdeg} A/k.$$

Theorem 16.8.2.2. *Let k be a field and A be a finite type k-algebra which is a domain and let $\mathfrak{p} \lneq A$ be a prime ideal. Then,*

$$\operatorname{ht} \mathfrak{p} + \dim A/\mathfrak{p} = \dim A.$$

16.8.3 Fundamental results

We begin with the fundamental theorem of dimension theory.

Theorem 16.8.3.1 (Fundamental theorem)**.**

The following is the famous principal ideal theorem.

Theorem 16.8.3.2 (Krull's Hauptidealsatz)**.** *Let R be a noetherian ring. If $I \leq R$ is a principal ideal, then any minimal prime $\mathfrak{p}$ containing I is such that*

$$\operatorname{ht}(\mathfrak{p}) \leq 1.$$

In particular, if $I \neq 0$ and R a domain, then any minimal prime containing I has height 1.

[15]note that such algebras are exactly the ones which correspond to affine algebraic varieties.

16.9 Completions

Do from Chapter 7 of Eisenbud

16.9.1 Hensel's lemma

Theorem 16.9.1.1 (Hensel)**.** *Let $(A, \mathfrak{m})$ be a noetherian local ring and $(\hat{A}, \hat{\mathfrak{m}})$ be the $\mathfrak{m}$-adic completion of A. Let $f \in \hat{A}[x_1, \dots, x_n]$. If $\bar{f}$ has a solution $(\bar{a}_1, \dots, \bar{a}_n)$ in $\hat{A}/\hat{\mathfrak{m}} = A/\mathfrak{m}$, then $(\bar{a}_1, \dots, \bar{a}_n)$ can be extended to $(a_1, \dots, a_n)$ in $\mathbb{A}_k^n$ on which f vanishes.*

16.10 Valuation rings

We begin with the basic theory of valuation rings.

16.10.1 Valuations & discrete valuations

Definition 16.10.1.1. (**Valuation on a field**) Let K be a field and G be an abelian group. A function $v : K \to G \cup \{\infty\}$ is said to be a valuation of K with values in G if v satisfies

1. $v(xy) = v(x) + v(y)$,
2. $v(x + y) \geq \min\{v(x), v(y)\}$,
3. $v(x) = \infty$ if and only if $x = 0$.

Let $\mathrm{Val}(K, G)$ denote the set of all valuations over K with values in G.

Few immediate observations are in order.

Lemma 16.10.1.2. *Let K be a field, G be an abelian group and $v \in \mathrm{Val}(K, G)$ be a valuation. Then,*

1. *$R = \{x \in K \mid v(x) \geq 0\} \cup \{0\}$ is a subring of K,*
2. *$\mathfrak{m} = \{x \in K \mid v(x) > 0\} \cup \{0\}$ is a maximal ideal of R,*
3. *$(R, \mathfrak{m})$ is a local ring,*
4. *R is an integral domain,*
5. *$R_{\langle 0 \rangle} = K$,*
6. *$\forall x \in K$, $x \in R$ or $x^{-1} \in R$.*

Proof. Items 1 and 4 are immediate from the axioms of valuations. Items 2 and 3 are immediate from the observation that $\{x \in K \mid v(x) = 0\} \cup \{0\}$ is a field in R. For items 5 and 6, we need to observe that $v(1) = 0$ and for any $x \in K^{\times}$, $v(x^{-1}) = -v(x)$. □

Remark 16.10.1.3. We call the subring $R \subset K$ above corresponding to a valuation v over K to be the *value ring of v.*

Definition 16.10.1.4. (**Valuation rings**) Let R be an integral domain. Then R is said to be a valuation ring if it is the value ring of some valuation over $K = R_{\langle 0 \rangle}$.

Definition 16.10.1.5. (**Domination**) Let K be a field and $A, B \subset K$ be two local rings in K. Then B is said to dominate A if $B \supseteq A$ and $\mathfrak{m}_B \cap A = \mathfrak{m}_A$.

There is an important characterization of valuation rings inside a field K with respect to all local rings in K.

Theorem 16.10.1.6. *Let K be a field and $R \subset K$ be a local ring. Denote $\mathrm{Loc}(K)$ to be the set of all local rings in K together with the partial order of domination. Then, the following are equivalent,*

1. *R is a valuation ring.*
2. *R is a maximal element of the poset $\mathrm{Loc}(K)$.*

Furthermore, for every local ring $S \in \mathrm{Loc}(K)$, there exists a valuation ring $R \in \mathrm{Loc}(K)$ which dominates S.

Proof. See Tag 00I8 of cite[Stacksproject]. □

An important type of valuation rings are where the value group is the integers.

Definition 16.10.1.7. (**Discrete valuation rings**) Let R be a valuation ring. Then R is said to be a discrete valuation ring (DVR) if the value group of R is the integers $\mathbb{Z}$.

It turns out that noetherian local domains of dimension 1 have some important characterizations, one of them being that they are exactly local Dedekind domains.

Theorem 16.10.1.8. *Let A be a noetherian local domain of dimension 1. Then the following are equivalent*

1. *A is a DVR.*
2. *A is a normal domain (that is, a local Dedekind domain).*
3. *A is a regular local ring.*
4. *The maximal ideal of A is principal and the generator t is called the "local parameter" of A.*

Proof. *Do it from Atiyah-Macdonald page 94.* □

It is a simple fact to see the following.

Proposition 16.10.1.9. *Let R be a DVR with local parameter $t \in R$ and $F = Q(R)$. Then,*

1. *Every element of R is of the form ut^n for $u \in R^\times = R \setminus tR$ and $n \in \mathbb{N} \cup \{0\}$.*
2. *We have that R is a PID. In particular, every ideal is generated by some power of the local parameter.*
3. *The discrete valuation of R is given by (note that $F = \{ut^n \mid n \in \mathbb{Z}\}$)*

$$\begin{aligned} \nu : F &\longrightarrow \mathbb{Z} \\ ut^n &\longmapsto n. \end{aligned}$$

Proof. 1. Let $a \in R$ which is not a unit, hence $a \in tR$, thus $a = rt$ where $r \in R$. As $r \in tR$, then $r = r_1 t$ and thus $a = r_1 t^2$. Doing the same on r_1 and continuing, we get an ascending chain, which terminates by noetherian condition, yielding us the factorization $a = ut^n$ where $u \in R$ is a unit and $n \in \mathbb{N}$, as required.

2. By Theorem 16.10.1.8, (R, tR) is a local ring. Let I be a proper ideal. We wish to show that it is generated by some t^n. To this end, we first show that I is a free R-module. Indeed, as R is a Dedekind domain (Theorem 16.10.1.8), we deduce that any ideal is a line bundle (Theorem 16.11.0.4, 5). As projective modules over local rings are free, it follows that $I \cong R$. Consequently, $I = aR$ and we conclude by 1.

3. We need only check that ν is a valuation and its value ring is R. Indeed the latter is immediate by item 1. The former is immediate by definition. □

Example 16.10.1.10. ($\mathbb{Z}_{\langle p \rangle}$ and $k[x]_{\langle p(x) \rangle}$) Let $p \in \mathbb{Z}$ be a prime and $p(x) \in k[x]$ be irreducible. Then both $\mathbb{Z}_{\langle p \rangle}$ and $k[x]_{\langle p(x) \rangle}$ are DVRs as they are local rings of PIDs (see Theorem 16.10.1.8). Moreover, their local parameters are $p \in \mathbb{Z}_{\langle p \rangle}$ and $p(x) \in k[x]_{\langle p(x) \rangle}$.

Example 16.10.1.11 (p-adic integers, $\hat{\mathbb{Z}}_p$)**.** Let p be a prime and consider the p-adic integer ring $\hat{\mathbb{Z}}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$. An element x of $\hat{\mathbb{Z}}_p$ can be written as

$$x = (x_1, \dots, x_n, \dots)$$

such that for all $k < l$, $x_k = x_l \mod p^k$. This defines $\hat{\mathbb{Z}}_p$ as a quotient of $\prod_{n \geq 1} \mathbb{Z}/p^n\mathbb{Z}$. We will follow the above characterization of elements of $\hat{\mathbb{Z}}_p$.

Then we claim that $\hat{\mathbb{Z}}_p$ is a DVR. Indeed, we first show that $\hat{\mathbb{Z}}_p$ is a domain. Let $x = (x_1, \dots, x_n, \dots)$ and $y = (y_1, \dots, y_n, \dots)$ be two p-adic integers. If none of x or y is zero, we claim that $xy = (x_1y_1, \dots, x_ny_n, \dots) \neq 0$ as well. Indeed, let $k = \nu(x)$, that is, the largest k such that $x_k = 0 \mod p^k$ and similarly let $l = \nu(y)$. Then, it is easy to see that $x_{k+l}y_{k+l}$ is the largest term of xy which is non-zero. Thus, $xy \neq 0$. This shows that $\hat{\mathbb{Z}}_p$ is a domain.

We also denote by $\hat{\mathbb{Q}}_p$ the fraction field of $\hat{\mathbb{Z}}_p$, the field of p-adic rationals. We construct a discrete valuation on $\hat{\mathbb{Q}}_p$ with value ring being $\hat{\mathbb{Z}}_p$. Indeed, consider

$$\begin{aligned} \nu_p : \hat{\mathbb{Q}}_p &\longrightarrow \mathbb{Z} \\ \frac{x}{y} &\longmapsto \nu_p(x) - \nu_p(y) \end{aligned}$$

where $\nu_p(x)$ for $x \in \hat{\mathbb{Z}}_p$ is the largest n such that $x_n = 0 \mod p^n$. It can easily be seen that this defines a discrete valuation whose value ring is $\hat{\mathbb{Z}}_p$, thus showing that $\hat{\mathbb{Z}}_p$ is a DVR. As $\nu(p) = 1$ where $p = (0, p, p, \dots, p, \dots) \in \hat{\mathbb{Z}}_p$, hence the local parameter of $\hat{\mathbb{Z}}_p$ is p.

16.10.2 Absolute values

We discuss the basics of absolute values and places, which will be used to state Ostrowski's theorem which classifies the places of $\mathbb{Q}$.

16.11 Dedekind domains

We will now discuss a class of rings which forms the right context for doing number theory in more abstract setting. We give here the barebones, rest will be developed as needed elsewhere.

Definition 16.11.0.1 (**Dedekind domain**)**.** A noetherian normal domain of dimension 1 is defined to be a Dedekind domain.

The following are some of the many equivalent characterizations of a Dedekind domain.

Theorem 16.11.0.2. *Let R be a noetherian domain of dimension 1. Then the following are equivalent:*

1. *R is normal (equivalently, Dedekind).*
2. *Every primary ideal $\mathfrak{q}$ of R is of the form $\mathfrak{q} = \mathfrak{p}^n$ for some prime ideal $\mathfrak{p}$ and $n \geq 0$.*
3. *$R_{\mathfrak{p}}$ is a DVR for each non-zero prime $\mathfrak{p}$.*

Theorem 16.11.0.3. *Let R be a domain. Then the following are equivalent:*

1. *R is a Dedekind domain.*
2. *Every fractional ideal of R is invertible.*

The following are some of the striking consequences of Dedekind condition.

Theorem 16.11.0.4. *Let R be a Dedekind domain.*

1. *Any finitely generated torsion-free R-module is projective.*
2. *Any ideal $I \leq R$ is a unique product of positive prime powers upto permutation, that is,*
$$I = \mathfrak{p}_1^{n_1} \dots \mathfrak{p}_r^{n_r},\ n_i \geq 1.$$
3. *Any invertible ideal $I \in \mathrm{Cart}(R)$ is a unique product of prime powers upto permutation, that is,*
$$I = \mathfrak{p}_1^{n_1} \dots \mathfrak{p}_r^{n_r},\ n_i \in \mathbb{Z} \setminus \{0\}$$
where a negative power of $\mathfrak{p}_i$ has the obvious meaning.
4. $\mathrm{Cart}(R)$ *is the free abelian group generated by* $\mathrm{Spec}\,(R) \setminus \{\mathfrak{o}\}$*:*
$$\mathrm{Cart}(R) \cong \mathbb{Z}(\mathrm{Spec}\,(R) \setminus \{\mathfrak{o}\}).$$
5. $\mathrm{Pic}(R)$ *is the group of isomorphism classes of ideals of R under multiplication:*
$$\mathrm{Pic}(R) \cong \{0 \neq I \leq R \text{ upto } R\text{-linear isomorphism}\}.$$

Remark 16.11.0.5 (The Dedekind philosophy)**.** Let R be a Dedekind domain. Then, the ideals are "generalized numbers of R with multiplication" and they are upto isomorphism given by the Picard group $\mathrm{Pic}(R)$, which are the line bundles upto isomorphism. Hence the analogy

"Generalized numbers of R upto isomorphism" $\rightsquigarrow$ Line bundles on R upto isomorphism.

Similarly, the invertible ideals of R, that is, Cartier divisors of R[16] are "generalized fractions of R with multiplication" and they are given by the Cartier group $\mathrm{Cart}(R)$. Hence the analogy

"Generalized fractions of R" $\rightsquigarrow$ Cartier divisors on R.

Proof of Theorem 16.11.0.4, 1. Let $F = Q(R)$ and M a finitely generated torsion-free module over R. We proceed by induction on $\dim_F M \otimes_R F$. If $\dim_F M \otimes_R F = 0$, then $M \otimes_R F = M_{\mathfrak{o}} = 0$, thus any non-zero element of M is torsion, which is not possible and thus $M = 0$, which is free so projective[17]. Now suppose all finitely generated torsion-free modules with $\dim_F M \otimes_R F \leq n$

[16] which, we would like to remind, are codimension-1 cycles on R(!)

[17] Essentially this is where we will be using the torsion-free hypothesis, the rest can be done without it, as can be seen.

are projective. Let M be torsion-free finitely generated with $\dim_F M \otimes_R F = n+1$. Hence, $M \otimes_R F \cong F^{n+1}$ as F-vector spaces. Now observe that as M is torsion-free, therefore the map

$$\begin{aligned} M &\longrightarrow M \otimes_R F \cong M_\circ \\ m &\longmapsto m \otimes 1 \mapsto \frac{m}{1} \end{aligned}$$

is an injection. Consequently, we may consider $M \subseteq F^{n+1}$. Consider any projection map $F^{n+1} \to F$. As any finitely generated submodule of F is a fractional ideal, therefore $I = \operatorname{Im}\left(M \hookrightarrow F^{n+1} \to F\right)$ is a fractional ideal. As R is Dedekind, so I is invertible (Theorem 16.11.0.3). As we have a surjection $M \twoheadrightarrow I$ and I is projective (see Cart-Pic sequence, Theorem 1.10.4.5), thus, the surjection is split and we have $M \cong N \oplus I$. As I is rank 1 projective, therefore by additivity of dimension, we have $\dim_F N \otimes_R F = n$. As M is torsion-free, so is N. By inductive hypothesis, N is projective, hence $M \cong N \oplus I$ is projective as well. $\square$

The following are some basic examples of Dedekind domains.

Example 16.11.0.6 (PIDs are Dedekind)**.** Let R be a PID. Then R is Dedekind as PIDs are noetherian, normal (since UFD) and of dimension 1 as every finite prime chain has length 1.

If $R = \mathbb{Z}$, then by Theorem 16.11.0.4, 4 & 5, we deduce that

$$\begin{aligned} \operatorname{Pic}(\mathbb{Z}) &= 0 \\ \operatorname{Cart}(\mathbb{Z}) &\cong \mathbb{Q}^\times. \end{aligned}$$

Similarly, if $R = k[x]$ for some field k, then,

$$\begin{aligned} \operatorname{Pic}(k[x]) &= 0 \\ \operatorname{Cart}(k[x]) &\cong k(x)^\times. \end{aligned}$$

Example 16.11.0.7 (Local Dedekind domains (i.e. DVRs))**.** By Theorem 16.10.1.8, local Dedekind domains are equivalent to DVRs, so DVRs forms another class of important Dedekind domains. Indeed, by Theorem 16.11.0.2, DVRs are exactly the local rings of Dedekind domains.

Hence for $\mathbb{Z}$, the local rings $\mathbb{Z}_{\langle p \rangle}$ for each prime $p \in \mathbb{Z}$ give local Dedekind domains and so does $k[x]_{\langle p(x) \rangle}$ for each irreducible polynomial $p(x) \in k[x]$.

16.12 Tor and Ext functors

We discuss two important functors in this section.

16.12.1 Some computations

Do exercises from Bruzzo.

16.13 Projective and injective modules

In this section we define an important object in the study of algebraic K-theory, projective modules. These generalize finitely generated free R-modules. This notion is further used in a very important geometric concept called depth and Cohen-Macaulay condition. In order to reach there, we would need a concept called projective dimension, which we cover here.

16.13.1 Projective modules

All rings will be associative with 1, but may not be commutative, unless stated otherwise. We denote $\mathbf{Proj}(R)$ to be the category of finitely generated projective left R-modules. Below are some easy to prove equivalent characterizations of projective modules and some of their properties.

Proposition 16.13.1.1. *Let R be a ring and P be a left R-module. Then the following are equivalent:*

1. *P is finitely generated projective.*
2. *Any short exact sequence $0 \to M \to N \to P \to 0$ is split exact.*
3. *There exists a module Q such that $P \oplus Q \cong R^n$.*
4. *There exists a surjection $\pi : R^n \twoheadrightarrow P$ which splits.*
5. *The functor $\mathrm{Hom}_R(P, -) : \mathbf{Mod}(R) \to \mathbf{Ab}$ is an exact functor, where $\mathbf{Mod}(R)$ is the category of left R-modules.*

Proposition 16.13.1.2. *Let $P, Q \in \mathbf{Proj}(R)$ be two finitely generated projective modules. Then*[18]*,*

1. *$P \oplus Q$ is a finitely generated projective module,*
2. *Any direct summand of P is a finitely generated projective module.*
3. *If R is commutative, then $P \otimes_R Q$ is a finitely generated projective R-module.*
4. *If R is commutative, then P is flat.*
5. *♣We have that $\check{P} = \mathrm{Hom}_R(P, R)$ is a projective R^{op}-module. If R is commutative, then $\check{P}$ is a projective R-module.*
6. *♣If R is commutative, then $\mathrm{rank}(\check{P}) = \mathrm{rank}(P)$.*
7. *♣If R is commutative, then trace of P, that is $\tau_P := \mathrm{Im}\left(\mathrm{ev} : \check{P} \otimes_R P \to R\right)$, is an idempotent ideal of R.*

Proof. † Item 1. and 2. are immediate from Proposition 16.13.1.1. For item 3, observe that if $P \oplus P' = R^{\oplus n}$, then $(P \otimes_R Q) \oplus (P' \otimes_R Q) = (P \oplus P') \otimes_R Q = R^{\oplus n} \otimes_R Q = Q^{\oplus n}$. As Q is projective, therefore $Q^{\oplus n}$ is projective by item 1. We conclude by item 2.

For item 4, we need only show that for an injective map $f : M' \to M$, the map $f \otimes \mathrm{id} : M' \otimes_R P \to M \otimes_R P$ is also injective. As P is projective, so there exists Q f.g. projective module such that $P \oplus Q = R^n$. Consequently, we get the commutative diagram as below:

$$\begin{tikzcd}
(M' \otimes_R P) \oplus (M' \otimes_R Q) \arrow[r, phantom, "\cong"] \arrow[d, "(f\otimes\mathrm{id})\oplus(f\otimes\mathrm{id})"] & M' \otimes_R (P \oplus Q) \arrow[r, phantom, "\cong"] \arrow[d, "f\otimes\mathrm{id}"] & (M')^{\oplus n} \arrow[d, "f^{\oplus n}"] \\
(M \otimes_R P) \oplus (M \otimes_R Q) \arrow[r, phantom, "\cong"] & M \otimes_R (P \oplus Q) \arrow[r, phantom, "\cong"] & M^{\oplus n}
\end{tikzcd}.$$

The right vertical map is injective by hypothesis. By commutativity of the diagram above, the rest of the two vertical maps are also injective. Hence, $f \otimes \mathrm{id} : M' \otimes_R P \to M \otimes_R P$ is injective as well,

[18] We put ♣ wherever finite generation of P and Q are not needed, i.e. if only projectivity of P and Q are needed.

as required.

Item 5 follows from existence of Q such that $P \oplus Q \cong R^n$ and that direct sum in first variable commutes with hom.

For item 6, first observe that $P \otimes_R \kappa(\mathfrak{p}) \cong P_\mathfrak{p}/\mathfrak{p}P_\mathfrak{p}$. Since $\mathrm{Hom}_R(P,R)_\mathfrak{p} \cong \mathrm{Hom}_{R_\mathfrak{p}}(P_\mathfrak{p}, R_\mathfrak{p}) = \check{P}_\mathfrak{p}$ as one of the modules in the hom is finitely presented (see Proposition 16.1.2.13), therefore we need only show that $P_\mathfrak{p} \cong \mathrm{Hom}_{R_\mathfrak{p}}(P_\mathfrak{p}, R_\mathfrak{p})$. To this end, as localization of projective modules is projective since localization is exact, we deduce that $P_\mathfrak{p}$ is projective $R_\mathfrak{p}$-module. Consequently, $P_\mathfrak{p}$ is free as $R_\mathfrak{p}$ is local (see Theorem 16.23.0.9). Hence the required isomorphism $P_\mathfrak{p} \cong \mathrm{Hom}_{R_\mathfrak{p}}(P_\mathfrak{p}, R_\mathfrak{p})$ is immediate.

For item 7, the fact that τ_P is an ideal is immediate from definition of ev as $\varphi \otimes x \mapsto \varphi(x)$. We now show that $\tau_P^2 = \tau_P$. To this end, we need only show that $\tau_P \subseteq \tau_P^2$. It can be seen that it is sufficient to show that any element $x \in P$ can be written as $x = \sum_{i=1}^n \psi_i(x)x_i$ for $x_i \in P$ and $\psi_i \in \check{P}$. Indeed, as there exists Q such that $P \oplus Q = R^F$, therefore for any $x \in P$, we may write $x = \sum_{i=1}^n r_i x_i$ where $r_i = f_i(x)$ where $\{f_i\}_{i\in F}$ is the dual basis of $(R^F)\check{}$. This completes the proof. □

Recall that an R-module M is locally free if for all $\mathfrak{p} \in \mathrm{Spec}(R)$, there exists a basic open $\mathfrak{p} \in D(f) \subseteq \mathrm{Spec}(R)$ such that M_f is a free R_f-module[19]. An important local characterization of projective modules is the following.

Theorem 16.13.1.3. *Let R be a commutative ring and M be an R-module. Then the following are equivalent:*

1. *M is finitely generated projective.*
2. *M is locally free of finite rank.*

Proof. (1. ⇒ 2.) Pick $\mathfrak{p} \in \mathrm{Spec}(R)$. Then, $M_\mathfrak{p}$ is a finitely generated $R_\mathfrak{p}$-module which is also projective as localization is exact. It follows from Theorem 17.1.2.3 that $M_\mathfrak{p} = (R_\mathfrak{p})^{\oplus n}$. Let $\{m_i/s_i\}_{i=1,\dots,n}$ be an $R_\mathfrak{p}$-basis of $M_\mathfrak{p}$. It follows by multipliying by $s_1 \dots s_n$ that we have a map $f : R^n \to M$ which may not be surjective, however, $f_\mathfrak{p} : R_\mathfrak{p}^n \to M_\mathfrak{p}$ is surjective. Denoting $N = \mathrm{CoKer}(f)$, we deduce that $N_\mathfrak{p} = 0$. As N is finitely generated, it follows that there exists $s \in R$ such that $N_s = 0$. But since $N_s = \mathrm{CoKer}(f_s)$, where $f_s : R_s^n \to M_s$, thus, we deduce that f_s is surjective. Since M_s is a projective R_s-module, therefore $M_s \oplus P = R_s^n$ where P is a finitely generated projective R_s-module. Localizing at $\mathfrak{p}$ again, we see that $M_\mathfrak{p} \oplus P_\mathfrak{p} = R_\mathfrak{p}^n$, but since $M_\mathfrak{p} = R_\mathfrak{p}^n$, thus, $P_\mathfrak{p} = 0$. It follows by finite generation that there exists $t \in R$ such that $t \cdot P = 0$ and thus $P_t = 0$. It follows that $M_{st} \oplus P_t = R_{st}^n$ and thus $M_{st} = R_{st}^n$ so that $f = st$ will do the job.

(2. ⇒ 1.) The proof is in two steps. In step 1, one shows that a locally free module of finite rank is finitely presented with free stalks. This follows from faithfully flat descent. In step 2, one shows that finitely presented modules with free stalks are projective. Indeed, let M be such a module. Then, we have an exact sequence

$$R^m \to R^n \xrightarrow{\pi} M \to 0.$$

By Proposition 16.13.1.1, it suffices to show that π splits. To this end, it is sufficient to show that $\pi_* : \mathrm{Hom}_R(M, R^n) \to \mathrm{Hom}_R(M, M)$ is surjective, as then id_M will have a section, as required.

[19] That is, $\tilde{M}$ is locally free, i.e. a vector bundle over $\mathrm{Spec}(R)$.

Indeed, as surjectivity of maps of modules is a local property ($f_\mathfrak{p} : M_\mathfrak{p} \to N_\mathfrak{p}$ is surjective for all $\mathfrak{p} \in \operatorname{Spec}(R)$ if and only if $f : M \to N$ is surjective), thus we reduce to showing that $(\pi_*)_\mathfrak{p}$ is surjective. As Hom and localization commutes if one of the modules is finitely presented (see Proposition 23.1.2.13 of [**FoG**]), therefore we wish to show that $\pi_{\mathfrak{p}*} : \operatorname{Hom}_{R_\mathfrak{p}}\left(M_\mathfrak{p}, R_\mathfrak{p}^n\right) \to \operatorname{Hom}_{R_\mathfrak{p}}(M_\mathfrak{p}, M_\mathfrak{p})$ is surjective. This is true as the map $\pi_\mathfrak{p} : R_\mathfrak{p}^n \to M_\mathfrak{p}$ is surjective by exactness of localization and since $M_\mathfrak{p}$ is a projective $R_\mathfrak{p}$-module as it is free by our hypothesis. This concludes the proof. □

Remark 16.13.1.4. By Theorem 16.13.1.3, it follows that vector bundles over $\operatorname{Spec}(R)$ are in one-to-one bijection with projective modules over R.

Using the above result, we can show that rank of a projective module is a continuous function from $\operatorname{Spec}(R)$ to $\mathbb{Z}$.

Proposition 16.13.1.5. *Let R be a commutative ring and M be a projective R-module. Then* $\operatorname{rank} : \operatorname{Spec}(R) \to \mathbb{Z}$ *is a continuous map.*

Proof. † By discreteness of $\mathbb{Z}$, it suffices to show that each fibre of rank is an open set. Indeed,

$$\begin{aligned}\operatorname{rank}^{-1}(n) &= \{\mathfrak{p} \in \operatorname{Spec}(R) \mid \dim_{\kappa(\mathfrak{p})} M \otimes_R \kappa(\mathfrak{p}) = n\}\\ &= \{\mathfrak{p} \in \operatorname{Spec}(R) \mid \dim_{\kappa(\mathfrak{p})} M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p} = n\}.\end{aligned}$$

By Theorem 16.13.1.3, M is locally free, hence $M_\mathfrak{p} \cong R_\mathfrak{p}^k$ for all $\mathfrak{p}$ in some largest open set $U \subseteq \operatorname{Spec}(R)$. Consequently, $\dim_{\kappa(\mathfrak{p})} M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p} = \dim_{\kappa(\mathfrak{p})}(R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p})^k = \dim_{\kappa(\mathfrak{p})} \kappa(\mathfrak{p})^k = k$ for all $\mathfrak{p} \in U$. Thus the above fibre is either empty or non-empty open set, as required. □

A simple example shows that $\mathbf{Proj}(R)$ cannot be abelian.

Example 16.13.1.6. Let $\mathbb{Z}$ be free $\mathbb{Z}$-module of rank 1. Observe that $2\mathbb{Z} \subseteq \mathbb{Z}$ is also a free module of rank 1. Hence both $\mathbb{Z}$ and $2\mathbb{Z}$ are projective $\mathbb{Z}$-modules. However, $\mathbb{Z}/2\mathbb{Z}$ is not a projective $\mathbb{Z}$-module as it cannot be a direct summand of $\mathbb{Z}^{\oplus n}$ for any $n \in \mathbb{N}$ since $\mathbb{Z}^{\oplus n}$ doesn't have any 2-torsion element. Consequently, $\mathbf{Proj}(R)$ is not abelian.

One observes that rank of a constant rank projective module remains same under extension of scalars.

Proposition 16.13.1.7. *Let $f : R \to S$ be a ring homomorphism between commutative rings. If P is a finitely generated projective R-module, then*

$$\operatorname{rank}(P \otimes_R S) = \operatorname{rank}(P) \circ f^*.$$

Hence, if P is constant rank n, then so is $P \otimes_R S$.

Proof. Let $\mathfrak{q} \in \operatorname{Spec}(S)$ and $f^*(\mathfrak{q}) = f^{-1}(\mathfrak{q}) = \mathfrak{p} \in \operatorname{Spec}(R)$. We need only show that if $P \otimes_R \kappa(\mathfrak{p}) \cong$

$\kappa(\mathfrak{p})^n$, then $(P \otimes_R S) \otimes_S \kappa(\mathfrak{q}) \cong \kappa(\mathfrak{q})^n$. Indeed, as

$$\begin{aligned}(P \otimes_R S) \otimes_S \kappa(\mathfrak{q}) &\cong P \otimes_R \kappa(\mathfrak{q}) \\ &\cong P \otimes_R S_\mathfrak{q} \otimes_S S/\mathfrak{q} \\ &\cong P \otimes_R R_\mathfrak{p} \otimes_{R_\mathfrak{p}} S_\mathfrak{q} \otimes_S S/\mathfrak{q} \\ &\cong P_\mathfrak{p} \otimes_{R_\mathfrak{p}} S_\mathfrak{q} \otimes_S S/\mathfrak{q} \\ &\cong R_\mathfrak{p}^n \otimes_{R_\mathfrak{p}} S_\mathfrak{q} \otimes_S S/\mathfrak{q} \\ &\cong S_\mathfrak{q}^n \otimes_S S/\mathfrak{q} \\ &\cong (S_\mathfrak{q} \otimes_S S/\mathfrak{q})^n \\ &\cong \kappa(\mathfrak{q})^n,\end{aligned}$$

as required. □

It is quite intuitive to claim that finite rank projective modules ought to be finitely generated. Indeed it is true.

Proposition 16.13.1.8. *Let R be a commutative ring and M be a finite rank projective module. Then M is finitely generated.*

Proof. † A result of Kaplansky states that a module over commutative ring R is projective if and only if it is locally free (we have done the finite case above in Theorem 16.13.1.3). Since by Theorem 16.13.1.3, it is sufficient to show that M is locally free of finite rank, where by above we already know it is locally free, we need only show that M is also finitely locally free. Let $f \in R$ be such that $M_f \cong R_f^F$. We wish to show that $|F| < \infty$. As M is finite rank, therefore for each $\mathfrak{p} \in \operatorname{Spec}(R)$, $\dim_{\kappa(\mathfrak{p})} M \otimes_R \kappa(\mathfrak{p}) < \infty$. If $f \notin \mathfrak{p}$, then since $M_\mathfrak{p} = (M_f)_\mathfrak{p} \cong (R_f^F)_\mathfrak{p} \cong R_\mathfrak{p}^F$, we deduce that $M \otimes_R \kappa(\mathfrak{p}) \cong M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p} \cong (R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p})^F = \kappa(\mathfrak{p})^F$. Thus $|F| < \infty$, as required. □

An important conceptual result which will guide us in defining higher K-groups is the cofinality of free modules in projective modules.

Lemma 16.13.1.9. *Let R be a ring and let $\mathbf{Free}(R)^{\cong}$ be the isomorphism classes of finitely generated free R-modules. This is a monoid under direct sum with identity 0. Then $\mathbf{Free}(R)^{\cong}$ is cofinal in $\mathbf{Proj}(R)^{\cong}$.* □

There is also a characterization of finitely generated projective modules in terms of flatness.

Proposition 16.13.1.10. *Let R be a commutative ring and M be an R-module. Then the following are equivalent:*

1. *M is a finitely presented flat R-module.*
2. *M is a finitely generated projective R-module.*

Proof. (1. ⇒ 2.) As M is finitely generated, thus to show that it is flat, it suffices to show that $M_\mathfrak{p}$ is a free $R_\mathfrak{p}$-module for each $\mathfrak{p} \in \operatorname{Spec}(R)$. As localization is exact, we reduce to assuming that R is a local ring and M is a finitely presented flat R-module. By Corollary 6.6 of cite[Eisenbud], it follows that M is projective R-module. As projective modules over local rings are free (Theorem 16.23.0.9), thus M is free, as required.

(2. ⇒ 1.) As M is finitely generated projective, then it is finitely presented as if $M \oplus N \cong R^n$ where N is thus also finitely generated projective, then we get a presentation $N \to R^n \to M \to 0$, as required. Clearly, M is flat by Proposition 16.13.1.2, 4. □

16.13.2 Divisible modules and Baer's criterion

Baer's criterion gives a characterization of injective R-modules. It consequently helps to show that divisible modules are injective in $\mathbf{Mod}(R)$ and thus that $\mathbf{Mod}(R)$ has enough injectives.

Definition 16.13.2.1 (Divisible modules). An R-module M is said to be divisible if for every $r \in R$, the multiplication by r map $\mu_r : M \to M$ is surjective.

Theorem 16.13.2.2. *(Baer's criterion) Let R be a ring and M be an R-module. The following are equivalent:*

1. *M is an injective R-module.*
2. *For any ideal $\mathfrak{a} \leq R$ and any map $f : \mathfrak{a} \to M$, there exists an extension $g : R \to M$ such that the following commutes:*

$$\begin{tikzcd} 0 \arrow[r] & \mathfrak{a} \arrow[r] \arrow[d, "f"'] & R \arrow[dl, dashed, "g"] \\ & M & \end{tikzcd} .$$

That is, one needs to check injectivity condition along inclusions of submodules of R.

Proof. 1. $\Rightarrow$ 2. is immediate from definition. For 2. $\Rightarrow$ 1. we proceed as follows. Pick $i : A \to B$ an injection of submodule $A \leq B$ and a map $f : A \to M$. We wish to extend this to $g : B \to M$. Indeed, consider the poset $\mathcal{P}$ of tuples (A', f'), $f' : A' \to M$ an extension of f with $(A', f') \leq (A'', f'')$ such that $A' \subseteq A''$ and f'' extends f'. By Zorn's lemma, we have a maximal extension $\bar{f} : \bar{A} \to M$. We reduce to showing that $\bar{A} = B$. If not, then there is $b \in B \setminus \bar{A}$. Consider $\tilde{A} = Rb + \bar{A}$. We claim that there is a map $\tilde{f} : \tilde{A} \to M$ extending f. Indeed, consider the ideal $\mathfrak{a} = \{r \in R \mid rb \in \bar{A}\}$. The map $\bar{f}$ defines a map $\mathfrak{a} \to M$ given by $r \mapsto \bar{f}(rm)$. By hypothesis, this has an extension, say $\kappa : R \to M$. Thus, we may define $g : \tilde{A} \to M$ as $rb + \bar{a} \mapsto \kappa(r) + \bar{f}(\bar{a})$. This extends f as if $rb + \bar{a} \in A$, then $rb \in \bar{A}$. Consequently, $\kappa(r) + \bar{f}(\bar{a}) = \bar{f}(rb) + \bar{f}(\bar{a}) = \bar{f}(rb + \bar{a}) = f(rb + \bar{a})$, as needed. $\square$

As a corollary, we see that injective R-modules are divisible.

Corollary 16.13.2.3. *Let R be a ring and M be an R-module. If M is injective, then M is divisible.*

Proof. Pick any $m \in M$ and $r \in R$. Then, we have an R-linear map $\mu_r : \langle r \rangle \to M$ given by $r \mapsto m$. By Theorem 16.13.2.2, 2, this extends to an R-linear homomorphism $g : R \to M$ where $\mu_r(r) = g(r) = rg(1) = m$, Thus $g(1) \in M$ is such that $rg(1) = m$, as needed. $\square$

16.14 Multiplicities

We study Hilbert polynomial and multiplicity of a graded module at a prime. This is useful to do intersection theory in projective spaces. In the general setting, we will assign a Hilbert polynomial to each projective variety, which yields invariants of the variety in question.

16.14.1 Length

We begin by studying length of modules.

Definition 16.14.1.1 (**Length of a module**). Let R be a ring and M be an R-module. Then the length of M is given by the length of the longest ascending chain of submodules of M:

$$\mathrm{len}_R(M) := \sup\{r \in \mathbb{N} \mid M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_r \text{ is a chain of submodules of } M\}.$$

A finite chain $M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_r$ is called a maximal length chain if it cannot be extended, that is, each factor M_i/M_{i-1} is a simple module. A maximal length chain is also called a composition series. Consequently, length of a module M is defined to be the length of the longest composition series.

An important result about length of modules is the fact that over a local ring R, any two composition series have the same length and composition factors.

Theorem 16.14.1.2 (Jordan-Hölder). *Let R be a local ring and M be an R-module which contains a composition series. Then any other composition series has the same length and composition factors. That is, length of M is equal to length of any composition series.*

The following are essential properties of length which one uses while dealing with maps.

Lemma 16.14.1.3. *Let $f : R \to S$ be a map of rings and M be an S-module. Then $\mathrm{len}_R(M) \geq \mathrm{len}_S(M)$ and equality holds if f is surjective.*

Proof. Follows from correspondence of submodules via a quotient map. □

We wish to characterize finite length modules over a noetherian ring. We begin with a lemma.

Lemma 16.14.1.4. *Any finite length R-module is finitely generated.*

Proof. If M is not finitely generated, then let $\{f_\alpha\}_{\alpha \in I}$ be a generating set of M and let $\{f_n\}_n$ be a subsequence. Then, the chain

$$0 \subsetneq \langle f_1 \rangle \subsetneq \langle f_1, f_2 \rangle \subsetneq \dots$$

is a chain of submodules of M which doesn't stabilizes, a contradiction to finite length. □

Using results on artinian rings (§16.3.1), we see an important characterization of artinian rings and finite length rings.

Theorem 16.14.1.5. *Let R be a ring. The following are equivalent:*

1. *R is artinian.*

2. *R has finite length.*

Proof. (1. $\Rightarrow$ 2.) By Theorem 16.3.1.3, 3, we reduce to assuming R is local artinian, $(R, \mathfrak{m})$. By Proposition 16.3.1.2, 3, Jacobson radical of R is nilpotent, which is just $\mathfrak{m}$. We construct a chain of ideals of R, where each subquotient has finite length. Indeed, consider the chain

$$0 = \mathfrak{m}^n \subsetneq \mathfrak{m}^{n-1} \subsetneq \cdots \subsetneq \mathfrak{m}^2 \subsetneq \mathfrak{m} \subsetneq R.$$

Note that $\mathfrak{m}^{i-1}/\mathfrak{m}^i$ is an $\kappa = R/\mathfrak{m}$-module. If any one of $\mathfrak{m}^i/\mathfrak{m}^{i-1}$ is infinite dimensional as an κ-vector space, then the above chain of ideals can be refined to an infinite chain of strictly decreasing ideals, a contradiction to artinian condition. Hence each subquotient is a finite dimensional κ-module and hence its length as an R-module is equal to its dimension as a κ-module (Lemma 16.14.1.3).

(2. $\Rightarrow$ 1.) Take any descending chain of ideals $I_0 \supsetneq I_1 \supsetneq I_2 \supsetneq \dots$. If it doesn't stabilize, then we have an infinite length chain, so that $\mathrm{len}(R)$ is not finite, a contradiction. □

The following is an essential result which we'll use later.

Proposition 16.14.1.6. *Let R be a noetherian ring and M be a finitely generated R-module. If $\mathfrak{p} \in \mathrm{Supp}\,(M)$ is a minimal prime of M, then $M_\mathfrak{p}$ is a finite length $R_\mathfrak{p}$-module.*

Proof. As $\mathrm{Supp}\,(M) = V(\mathrm{Ann}(M))$, therefore a minimal prime $\mathfrak{p} \in \mathrm{Supp}\,(M)$ is an isolated/minimal prime of $\mathrm{Ann}(M)$. As $M_\mathfrak{p}$ is an $R_\mathfrak{p}$-module, therefore it suffices to construct a composition series of $M_\mathfrak{p}$. Let M be generated by $f_1, \dots, f_n \in M$, so that $M_\mathfrak{p}$ is also generated by their respective images. We thus get the following chain:

$$0 \subseteq \langle f_1 \rangle \subseteq \langle f_1, f_2 \rangle \subseteq \cdots \subseteq \langle f_1, \dots, f_n \rangle = M_\mathfrak{p}.$$

It suffices to show that $\frac{\langle f_1, \dots, f_i \rangle}{\langle f_1, \dots, f_{i-1} \rangle}$ is a finite length $R_\mathfrak{p}$-module. Indeed, we have a surjection

$$\langle f_i \rangle \twoheadrightarrow \frac{\langle f_1, \dots, f_i \rangle}{\langle f_1, \dots, f_{i-1} \rangle},$$

hence it suffices to show that $\langle f_i \rangle$ is a finite length $R_\mathfrak{p}$-module. To this end, pick any $x \in M$. We'll show that $xR_\mathfrak{p}$ is a finite length $R_\mathfrak{p}$-module. Observe that $\langle x \rangle = xR_\mathfrak{p}$ is isomorphic to $R_\mathfrak{p}/I$ where I is the annihilator of x in $R_\mathfrak{p}$. We may write $I = \mathfrak{a}R_\mathfrak{p}$ where $\mathfrak{a} \leq R$ is contained in $\mathfrak{p}$. Hence, we wish to show that $S = R_\mathfrak{p}/\mathfrak{a}R_\mathfrak{p}$ is a finite length $R_\mathfrak{p}$-module, that is S is a finite length ring. Indeed, as $S = (R/\mathfrak{a})_\mathfrak{p}$ and $\mathfrak{p}$ is a minimal prime in $\mathrm{Supp}\,(M)$, that is, minimal prime containing $\mathrm{Ann}(M)$, and since $\mathrm{Ann}(M) \subseteq \mathfrak{a} \subseteq \mathfrak{p}$, therefore $\mathfrak{p}$ is a minimal prime of $\mathfrak{a}$ as well. It follows that $S = (R/\mathfrak{a})_\mathfrak{p}$ is a dimension 0 ring. Since R is noetherian and noetherian property is inherited by quotients and localizations, therefore S is a noetherian ring of dimension 0, hence artinian. From Theorem 16.14.1.5, it follows that S is of finite length, as required. □

Proposition 16.14.1.7. *Let R be a noetherian ring and M be an R-module. Then the following are equivalent:*

1. *M has finite length.*
2. *M is finitely generated and* $\dim R/\mathrm{Ann}(M) = 0$*, i.e. $R/\mathrm{Ann}(M)$ is an artinian ring.*

Proof. (1. $\Rightarrow$ 2.) By Lemma 16.14.1.4, M is finitely generated. Let $0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_r = M$ be a composition series of M, which exists as $\mathrm{len}(M) < \infty$. We thus get that $M_i/M_{i-1} \cong R/\mathfrak{m}_i$ for some maximal ideals $\mathfrak{m}_i$ by Lemma 16.1.1.1. Note that $\dim R/\mathrm{Ann}(M) = 0$ if and only if $\mathrm{Supp}\,(M)$ consists only of maximal ideals. So let $\mathfrak{p} \in \mathrm{Supp}\,(M)$. Thus $M_\mathfrak{p} \neq 0$. It follows that for some i, $(M_i/M_{i-1}) \neq 0$. As $(M_i/M_{i-1}) = (R/\mathfrak{m}_i)_\mathfrak{p}$, therefore this can only happen if $\mathfrak{m}_i \subseteq \mathfrak{p}$, i.e. $\mathfrak{m}_i = \mathfrak{p}$, as required. This also shows that $\mathrm{Supp}\,(M) = \{\mathfrak{m}_1, \dots, \mathfrak{m}_r\}$.

(2. $\Rightarrow$ 1.) We need only construct a composition series of M. We have $\mathrm{Supp}\,(M)$ consists only of maximal ideals. Consider $\mathrm{Supp}\,(M) \subseteq \mathrm{Spec}\,(R)$. As M is finitely generated, say by $f_1, \dots, f_n$. Then we get a chain of submodules

$$0 \subsetneq \langle f_1 \rangle \subsetneq \langle f_1, f_2 \rangle \subsetneq \cdots \subsetneq \langle f_1, \dots, f_n \rangle = M.$$

We need only show that each subquotient is a finite length R-module. Indeed, as we have a surjection

$$\langle f_i \rangle \twoheadrightarrow \frac{\langle f_1, \dots, f_i \rangle}{\langle f_1, \dots, f_{i-1} \rangle},$$

so it suffices to show that $\langle f_i \rangle$ is a finite length R-module. To this end, it suffices to show that for each $x \in M$, the submodule Rx is of finite length. Indeed, we have $Rx \cong R/I$ where $I = \mathrm{Ann}(x)$. As $I \supseteq \mathrm{Ann}(M)$, therefore

$$R/I \cong \frac{R/\mathrm{Ann}(M)}{I/\mathrm{Ann}(M)}.$$

As $R/\mathrm{Ann}(M)$ is an artinian ring and any quotient of artinian ring is an artinian ring, it follows at once that R/I is an artinian ring. By Theorem 16.14.1.5, $R/I \cong Rx$ is of finite length, as required. $\square$

16.14.2 Degree of a graded module

We begin by studying multplicity at a prime.

Definition 16.14.2.1 (Multiplicity at a prime). Let R be a ring and M be an R-module. The multiplicity of M at prime $\mathfrak{p} \in \mathrm{Spec}\,(R)$ is given by

$$\mu_\mathfrak{p}(M) := \mathrm{len}_{R_\mathfrak{p}} M_\mathfrak{p}.$$

Definition 16.14.2.2 (Hilbert function). Let R be a ring and M be a graded $k[x_0, \dots, x_n]$-module. The Hilbert function of M is defined to be the following

$$\begin{aligned} \varphi_M : \mathbb{Z} &\longrightarrow \mathbb{Z} \\ d &\longmapsto \dim_k M_d. \end{aligned}$$

The main theorem on Hilbert functions is that it is actually a numerical polynomial (a rational polynomial which on large integers give integers), and that this polynomial is unique.

Theorem 16.14.2.3 (Hilbert-Serre). *Let $S = k[x_0, \dots, x_n]$ and M be a finitely generated graded S-module. Then, there exists a polynomial $P_M(x) \in \mathbb{Q}[x]$ such that it is unique with respecto the following properties:*

1. *there exists $D \in \mathbb{N}$ such that for all $d \geq D$, we have*

$$\varphi_M(d) = P_M(d),$$

   that is, P_M is a numerical polynomial,

2. *the degree of $P_M(x)$ is equal to* $\dim V(\mathrm{Ann}_R(M))$[20].

The $P_M(x)$ is called the Hilbert polynomial of M.

Proof. See Theorem 7.5 of cite[Hartshorne]. □

We will now define the degree of a graded $S = k[x_0, \dots, x_n]$-module. This will allow us to do define the notion of degree of projective schemes over k.

Definition 16.14.2.4 (Degree of a graded S-module). Let $S = k[x_0, \dots, x_n]$ and M be a graded S-module. Then, we define

$$\deg_S M := \deg(P_M)! \cdot c_{\uparrow}(P_M)$$

where $c_{\uparrow}(P_M)$ denotes the leading coefficient of the Hilbert polynomial P_M.

Remark 16.14.2.5. Let $r = \deg(P_M)$. We may alternatively view the degree of M as

$$\deg_S M = P_M^{(r)}(x),$$

that is, the r^{th}-derivative of P_M.

[20] It is a simple exercise to see that the annihilator ideal of a graded S-module is homogeneous.

16.15 Kähler differentials

We study analogues of tangent and cotangent bundles of topology in commutative algebra.

Definition 16.15.0.1 (Derivations & Kähler differentials). Let S be an R-algebra and M be an S-module. An R-linear derivation $d : S \to M$ is a group homomorphism such that it satisfies Leibnitz's rule:

$$d(fg) = fd(g) + gd(f).$$

The set of all R-linear derivations $S \to M$ forms an S-module denoted $\mathrm{Der}_R(S, M)$. We define the S-module of Kähler differentials of S/R as the follows. Define $X = \{d(f) \mid f \in S\}$ be a set of free symbols one for each $f \in S$. Then define $\mathrm{Käh}(S)$ to be the S-submodule of $S^{\oplus X}$ generated by

$$d(fg) - fd(g) - gd(f),\ d(af + bg) - ad(f) - bd(g)$$

for all $a, b \in R$ and $f, g \in S$. We then define $\Omega_{S/R}$ to be the following quotient:

$$0 \to \mathrm{Käh}(S) \to S^{\oplus X} \to \Omega_{S/R} \to 0.$$

Observe that $d(a) = 0$ in $\Omega_{S/R}$ for all $a \in R$, thus if $R \twoheadrightarrow S$ is surjective, then $\Omega_{S/R} = 0$. The canonical map

$$\begin{aligned} d : S &\longrightarrow \Omega_{S/R} \\ f &\longmapsto d(f) \end{aligned}$$

is an R-linear derivation of S in $\Omega_{S/R}$ called the universal R-linear derivation.

We immediately have the following helpful characterization.

Proposition 16.15.0.2 (Universal property of $\Omega_{S/R}$). *Let S be an R-algebra. The for any S-module M and any R-linear derivation $e : S \to M$, there exists a unique S-linear homomorphism $\tilde{d} : \Omega_{S/R} \to M$ such that the following commutes:*

$$\begin{tikzcd} \Omega_{S/R} \arrow[r, dashed, "\tilde{e}"] & M \\ S \arrow[u, "d"] \arrow[ur, "e"'] & \end{tikzcd}.$$

Proof. Consider the S-linear map

$$\begin{aligned} e^{\oplus X} : S^{\oplus X} &\longrightarrow M \\ \sum_{i=1}^{n} f_i dg_i &\longmapsto \sum_{i=1}^{n} f_i eg_i. \end{aligned}$$

It follows at once that $\mathrm{Ker}\left(e^{\oplus X}\right) \supseteq \mathrm{Käh}(S)$. By universal property of cokernels, we thus obtain a unique S-linear map

$$\tilde{e} : \Omega_{S/R} \longrightarrow M$$

such that the required triangle commutes. □

Corollary 16.15.0.3. *Let S be an R-algebra and M be an S-module. Then there is an S-linear isomorphism*

$$\mathrm{Der}_R(S, M) \cong \mathrm{Hom}_S\left(\Omega_{S/R}, M\right).$$

Proof. The S-linear isomorphism is given by $e \mapsto \tilde{e}$, which is injective by universal property and surjective by composition with universal R-linear derivation d. $\square$

Remark 16.15.0.4. Just as tensor product is the representing object of bilinear maps from $M \times N$, similarly $\Omega_{S/R}$ is the representing object of R-linear derivations from S.

Example 16.15.0.5. Let R be a ring and $S = R[x_1, \dots, x_n]$. Then we claim that $\Omega_{S/R}$ is free S-module of rank n given by

$$\Omega_{S/R} = Sdx_1 \oplus \cdots \oplus Sdx_n.$$

Indeed, as $\Omega_{S/R}$ is a finitely generated S-module by $dx_1, \dots, dx_n$ via Leibnitz's rule, therefore we have an S-linear surjection

$$\begin{aligned} S^{\oplus n} &\longrightarrow \Omega_{S/R} \\ (p_1, \dots, p_n) &\longmapsto \sum_{i=1}^{n} p_i dx_i. \end{aligned}$$

This has an inverse given by the unique maps $\partial_i : \Omega_{S/R} \to S$ induced by the R-linear derivations $\partial_i : S \to S$ mapping $p \mapsto \frac{\partial}{\partial x_i} p$. This completes the proof.

Remark 16.15.0.6 (Relative cotangent functor)**.** The assignment of Kähler differentials is functorial. Indeed, by universal properties, we have

$$\begin{tikzcd} S \arrow[r, "\varphi"] & S' \arrow[rr, phantom, "\rightsquigarrow"] & & \Omega_{S/R} \arrow[r, "\tilde{\varphi}"] & \Omega_{S'/R'} \\ R \arrow[u] \arrow[r] & R' \arrow[u] & & S \arrow[u, "d"'] \arrow[r, "\varphi"'] & S' \arrow[u, "d"'] \end{tikzcd}.$$

Moreover, the S-linear map

$$\tilde{\varphi} : \Omega_{S/R} \to \Omega_{S'/R'}$$

is equivalent to the S'-linear map

$$\begin{aligned} S' \otimes_S \Omega_{S/R} &\longrightarrow \Omega_{S'/R'} \\ f' \otimes f dg &\longmapsto f' \varphi(f) d\varphi(g). \end{aligned}$$

We have two fundamental exact sequences aiding computations.

Proposition 16.15.0.7 (Cotangent sequence/First sequence)**.** *Let R be a ring and $R \to S \to T$ be ring homomorphisms. Then the following is an exact sequence of T-modules where the maps are the obvious ones:*

$$T \otimes_S \Omega_{S/R} \longrightarrow \Omega_{T/R} \longrightarrow \Omega_{T/S} \longrightarrow 0.$$

Proof. Kernel on the right is exactly the T-submodule generated by ds for $s \in S$. This is exactly the image of the left as well. □

Proposition 16.15.0.8 (Conormal sequence/Second sequence)**.** *Let S be an R-algebra and $I \leq S$ be an ideal. Denote $T = S/I$ to be the quotient S-algebra. Then the following is an exact sequence*

$$I/I^2 \xrightarrow{d} T \otimes_S \Omega_{S/R} \longrightarrow \Omega_{T/R} \longrightarrow 0.$$

The map $d : x + I^2 \mapsto 1 \otimes dx$ and the other is the natural map corresponding to $\pi : S \to S/I$.

It is wise to discuss the following result immediately, so that one can see how the geometric discussion of differentials might be carried.

Theorem 16.15.0.9 (Diagonal criterion)**.** *Let S be an R-algebra and $\varphi : S \otimes_R S \to S$ be the structure morphism. Let $I = \mathrm{Ker}(\varphi)$ which is an S-module as it is a submodule of S-module $S \otimes_R S$. Then, for the R-linear derivation $e : S \to I/I^2$ mapping $s \mapsto 1 \otimes s - s \otimes 1$, the pair $(I/I^2, e : S \to I/I^2)$ is isomorphic to $(\Omega_{S/R}, d : S \to \Omega_{S/R})$:*

$$(I/I^2, e) \cong (\Omega_{S/R}, d).$$

Proof. We need only prove that both the pairs satisfy the same universal property as stated in Proposition 16.15.0.2. **TODO** □

Kähler differentials behaves nicely with tensor products and localizations. The main idea behind both proofs is to use functoriality of Kähler differentials and the resulting maps and then form their inverses (see Remark 16.15.0.6).

Proposition 16.15.0.10 (Base change)**.** *Let R be a ring and R' and S be R-algebras. Consider the pushout square*

$$\begin{tikzcd} S \arrow[r] \arrow[dr, phantom, "\lrcorner" very near end] & S \otimes_R R' \\ R \arrow[u] \arrow[r] & R' \arrow[u] \end{tikzcd}.$$

Then,

$$\Omega_{S \otimes_R R'/R'} \cong \Omega_{S/R} \otimes_S (S \otimes_R R').$$

Proposition 16.15.0.11 (Localization)**.** *Let S be an R-algebra and $M \subseteq S$ be a multiplicative set. Consider the following commutative square*

$$\begin{tikzcd} S \arrow[r] & M^{-1}S \\ R \arrow[u] \arrow[r, "\mathrm{id}"'] & R \arrow[u] \end{tikzcd}.$$

Then

$$\Omega_{M^{-1}S/R} \cong M^{-1}\Omega_{S/R}.$$

16.16 Depth, Cohen-Macaulay & regularity

We now study some homological properties of commutative rings with 1.

16.16.1 Regular rings, projective & global dimension

Definition 16.16.1.1 (**Regular ring, projective and global dimension**). A noetherian ring R is said to be regular if every R-module M has a finite length projective resolution. That is, if for every R-module M, there exists an exact sequence

$$0 \to P_n \to P_{n-1} \to \cdots \to P_0 \to M \to 0$$

such that P_i are projective R-modules where n is the length of the projective resolution. The projective dimension of an R-module M is defined as

$$\mathrm{pd}(M) := \inf\{\text{length of projective resolution of } M\}.$$

Further we define global dimension of R as

$$\mathrm{gl}\dim(R) := \sup\{\mathrm{pd}(M) \mid M \in \mathbf{Mod}(R)\}.$$

By far the most important class for us is the regular local rings. We first establish the following to resolve the tension made in Definition 16.1.2.16, amongst other goals.

Theorem 16.16.1.2. *Let $(R, \mathfrak{m})$ be a local ring with $k = R/\mathfrak{m}$. Then the following are equivalent:*

1. *R is a regular local ring*[21].
2. $\dim_k \mathfrak{m}/\mathfrak{m}^2 = \dim R$.
3. *If $\mathfrak{m}$ has minimal generating set as $\{a_1, \ldots, a_n\}$, then $\dim A = n$.*
4. $\mathrm{gl}\dim(A) = \dim A < \infty$.

Some more properties of regular local rings are as follows.

Proposition 16.16.1.3. *Let $(R, \mathfrak{m})$ be a regular local ring.*

1. *R is a noetherian normal domain, in particular, a Krull domain (see Definition 1.10.2.1).*
2. *If $x \in \mathfrak{m} \setminus \mathfrak{m}^2$, then xR is a prime ideal.*

Localization of a regular ring at a prime is a regular local ring.

Lemma 16.16.1.4. *Let A be a regular ring. For any $\mathfrak{p} \in \mathrm{Spec}(A)$, the local ring $A_{\mathfrak{p}}$ is regular.*

Proof. Take any $A_{\mathfrak{p}}$-module M. We wish to show that M has a finite length projective resolution by $A_{\mathfrak{p}}$-modules. To this end, consider the localization map $A \to A_{\mathfrak{p}}$. By restriction, we have an A-module M_A. By Lemma 16.1.2.19, we have $M_A \otimes_A A_{\mathfrak{p}} \cong M$. Consequently, as M_A has a finite length resolution by projective A-modules, localizing at $\mathfrak{p}$, we get a finite length resolution of M by projective $A_{\mathfrak{p}}$-modules, as localization of projective modules is projective. □

Our first goal is to show that regular local rings are UFD. This will help us in showing that on a locally factorial domain (more generally locally factorial noetherian integral separated scheme), Weil and Cartier divisor groups agree. We will do this using the theory of Weil and Cartier divisors themselves.

[21] in the sense of Definition 16.16.1.1.

Theorem 16.16.1.5. *Let R be a regular local ring. Then R is a UFD.*

First observe the following important reduction.

Proposition 16.16.1.6. *Let R be a noetherian domain. Then the following are equivalent:*
1. *R is a UFD.*
2. *All height 1 primes of R are principal.*

Proof. (1. $\Rightarrow$ 2.) Let $\mathfrak{p}$ be a non-zero prime ideal. Pick any non-zero $a \in \mathfrak{p}$. As R is a UFD, we may write $a = p_1^{n_1} \dots p_k^{n_k}$ where $p_i \in R$ are primes. Assume $k \geq 2$. As $\mathfrak{p}$ is a prime and $a \in \mathfrak{p}$, it follows that there exists $p_i \in \mathfrak{p}$. Thus, $p_i R \subsetneq \mathfrak{p}$, which is a contradiction to height 1 of $\mathfrak{p}$. It follows that $k = 1$, and thus $\mathfrak{p} = p_i R$, as required.

(2. $\Rightarrow$ 1.) Observe that a noetherian domain is in particular a factorization domain. Consequently, we need only show that any irreducible element is prime. Let $f \in R$ be irreducible. We wish to show that fR is a prime ideal. By Krull's Hauptidealsatz (Theorem 16.8.3.2), if $\mathfrak{p}$ is a minimal prime containing fR, then since R is a domain, we deduce that $\mathfrak{p}$ is of height 1. By our hypothesis, $\mathfrak{p} = pR$ is principal where $p \in R$ is a prime element. As $fR \subseteq pR$, we deduce that $p|f$, i.e. $f = pr$ for some $r \in R$. But f is irreducible, therefore either p or r is a unit. As p is prime, so r is a unit and thus $fR = pR$ is a prime ideal, as required. $\square$

Proof of Theorem 16.16.1.5. By Proposition 16.16.1.6, we need only show that height 1 primes of R (prime divisors of R) are principal. We do this by induction on $\dim(R)$. If $\dim(R) = 1$, then by Theorem 16.10.1.8, we deduce that R is a DVR and thus is PID, so a UFD. Now assume that $\dim(R) = n$ and any regular local ring of dimension $< n$ is UFD. Let $f \in \mathfrak{m} \setminus \mathfrak{m}^2$. By relative Weil divisors (Proposition 1.10.2.22), as fR is principal (Proposition 16.16.1.3, 2), we get that $\mathrm{Cl}(R) \cong \mathrm{Cl}(R_f)$. By R UFD iff $\mathrm{Cl}(R) = 0$, we reduce to showing that $S = R_f$ is a UFD. By Proposition 16.16.1.6, it suffices to show that all height 1 primes of S are principal, which is same as showing that all height 1 primes are free of rank 1.

Let $\mathfrak{p}$ be a height 1 prime of S. As R is regular, $\mathfrak{p}$ is obtained by localizing a prime of R at f and localization being exact, we deduce that we have a free resolution of $\mathfrak{p}$ (finitely generated projective modules over local ring R) as

$$0 \to S^{k_n} \to \dots \to S^{k_0} \to \mathfrak{p} \to 0.$$

For any prime $\mathfrak{q} \in S$, $\mathfrak{p}_\mathfrak{q}$ is a prime ideal of $S_\mathfrak{q}$ of height 1 where $S_\mathfrak{q}$ is a regular local ring of dim $< n$, so that by inductive hypothesis, it is UFD and thus by Proposition 16.16.1.6, it follows that $\mathfrak{p}_\mathfrak{q}$ is principal and thus free. Hence $\mathfrak{p}_\mathfrak{q}$ is free at each prime of S, hence $\mathfrak{p}$ is projective module of rank 1 i.e. a line bundle.

By above resolution, we deduce that $\mathfrak{p}$ is a stably free line bundle over S. As stably free line bundles are free[22], we get that $\mathfrak{p}$ is free, as required. $\square$

[22]if M is a line bundle such that $M \oplus R^n = R^{n+1}$, then taking $\wedge^{n+1}$ both sides, we deduce that $\wedge^{n+1}(M \oplus R^n) \cong R$. Now $\wedge^{n+1}(M \oplus R^n) \cong \bigoplus_{i=0}^{n+1} \wedge^i M \oplus \wedge^{n+1-i} R^n = \bigoplus_{i=1}^{n+1} \wedge^i M \oplus \wedge^{n+1-i} R^n = \bigoplus_{i=1}^{n+1} (\wedge^i M)^{{}^nC_{n+1-i}}$. Localizing at $\mathfrak{p}$, we deduce that $R_\mathfrak{p} \cong \bigoplus_{i=1}^{n+1} (\wedge^i R_\mathfrak{p})^{{}^nC_{n+1-i}}$, from which we deduce that $\bigoplus_{i=2}^{n+1} (\wedge^i M)^{{}^nC_{n+1-i}}$ is zero at each prime $\mathfrak{p}$ and is thus 0 module. It follows that $R \cong \wedge^1 M \cong M$, as required.

16.16.2 Depth

We begin with the notion of depth of an R-module.

Definition 16.16.2.1 (**I-depth**)**.** Let R be a ring, $I \leq R$ be an ideal and M a finitely generated R-module such that $IM \subsetneq M$. Define the I-depth of M as

$$\mathrm{depth}_I(M) := \min\{i \geq 0 \mid \mathrm{Ext}^i_R(R/I, M) \neq 0\}.$$

For a local ring $(R, \mathfrak{m}, \kappa)$ and a finitely generated R-module M, we define

$$\mathrm{depth}(M) := \mathrm{depth}_{\mathfrak{m}}(M) = \min\{i \geq 0 \mid \mathrm{Ext}^i_R(\kappa, M) \neq 0\},$$

that is, depth of M is just the $\mathfrak{m}$-depth of M as an R-module[23].

The main theorem about depth is the famous Auslander-Buchsbaum theorem.

Theorem 16.16.2.2 (Auslander-Buchsbaum)**.** *Let R be a noetherian local ring. For a non-zero finitely generated R-module M with $\mathrm{pd}_R(M) < \infty$, we have*

$$\mathrm{pd}_R(M) + \mathrm{depth}(M) = \mathrm{depth}(R).$$

[23]One might as well call the depth of a local ring as its Ext-dimension.

16.17 Filtrations

Do from Chapter 5 of Eisenbud

16.18 Flatness

This is one of the important parts of commutative algebra, as this notion corresponds to the idea of a continuous family of schemes, in some sense, as is discussed in the respective part above.

Definition 16.18.0.1. (**Flat modules and flat map of rings**) Let R be a ring. An R-module M is said to be flat if for any short exact sequence of R-modules

$$0 \longrightarrow N_1 \longrightarrow N_2 \longrightarrow N_3 \longrightarrow 0$$

the following sequence is exact

$$0 \longrightarrow M \otimes_R N_1 \longrightarrow M \otimes_R N_2 \longrightarrow M \otimes_R N_3 \longrightarrow 0 .$$

A map $\varphi : A \to B$ is a flat map if B is a flat A-module. In this case one also calls B to be a flat A-algebra.

Remark 16.18.0.2.
1. By right exactness of tensor products, it is sufficient to check that the s.e.s. $0 \to N_1 \to N_2$ is taken to s.e.s $0 \to M \otimes_R N_1 \to M \otimes_R N_2$.
2. Since localisation is an exact functor (Lemma 16.1.2.2), thus the natural map $A \to S^{-1}A$ is a flat map for any multiplicative set $S \subseteq A$.

Recall that $\mathrm{Tor}_i^R(M, -)$ is the i^{th}-left derived functor of $N \mapsto M \otimes_R N$. A module M is said to be flat if the tensor functor $N \mapsto M \otimes_R N$ is exact. Here are equivalent notions of flatness:

Theorem 16.18.0.3. *Let R be a ring and M be an R-module. Then the following are equivalent:*
1. *M is a flat R-module.*
2. *$\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$ and R-modules N.*
3. *$\mathrm{Tor}_1^R(M, N) = 0$ for all R-modules N.*
4. *$\mathrm{Tor}_1^R(M, N) = 0$ for all finitely generated R-modules N.*
5. *$\mathrm{Tor}_1^R(M, R/I) = 0$ for every ideal $I \leq R$.*
6. *$I \otimes_R M \to M$ is injective for every ideal $I \leq R$.*
7. *$I \otimes_R M \to IM$ is an isomorphism for every ideal $I \leq R$.*
8. *$M_\mathfrak{p}$ is a flat $R_\mathfrak{p}$-module for every $\mathfrak{p} \in \mathrm{Spec}\,(R)$.*

Proof. (1. $\Leftrightarrow$ 2.) Pick any projective resolution of N as $P_\bullet \stackrel{\epsilon}{\to} N \to 0$ and consider the complex $M \otimes P_\bullet$. As $\mathrm{Tor}_i^R(M, N) = \mathrm{Ker}\,(d_i \otimes \mathrm{id})/\mathrm{Im}\,(d_{i+1} \otimes \mathrm{id})$, therefore to show $(1 \Rightarrow 2)$, it suffices to show that $\mathrm{Im}\,(d_{i+1} \otimes \mathrm{id}) = \mathrm{Ker}\,(d_i \otimes \mathrm{id})$. By applying $- \otimes M$ on $0 \to \mathrm{Im}\,(d_{i+1}) = \mathrm{Ker}\,(d_i) \stackrel{\iota_{i+1}}{\to} P_i \stackrel{d_i}{\to} \mathrm{Im}\,(d_i) \to 0$, we get that $\mathrm{Ker}\,(d_i \otimes \mathrm{id}) = \mathrm{Im}\,(\iota_i \otimes \mathrm{id})$, so that we reduce to showing $\mathrm{Im}\,(\iota_{i+1} \otimes \mathrm{id}) = \mathrm{Im}\,(d_{i+1} \otimes \mathrm{id})$. This is provided by tensoring the following diagram with M and using flatness of M:

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Im}\,(d_{i+1}) \arrow[r, "\iota_{i+1}"] & P_i \\ & & P_{i+1} \arrow[ul, two heads, "d_{i+1}"] \arrow[u, "d_{i+1}"'] \end{tikzcd} .$$

The converse is immediate by the long exact sequence of a derived functor associated to a short exact sequence. Note that (2. $\Rightarrow$ 3.) is easy.

(3. $\Leftrightarrow$ 1.) If $\mathrm{Tor}_1^R(M,N) = 0$ for any N, then M is flat as for any short exact sequence $0 \to N_1 \to N_2 \to N_3 \to 0$, the corresponding l.e.s. of Tor gives exactness of $0 \to M \otimes N_1 \to M \otimes N_2 \to M \otimes N_3 \to 0$.

(4. $\Leftrightarrow$ 3.) If N is any R-module, then we have a $\square$

Some more properties of flat modules are as follows.

Theorem 16.18.0.4. *Let R be a ring and M be an R-module.*

1. *If M is projective, then M is flat.*
2. *If R is local and M is flat, then M is free.*
3. *If M is finitely generated, then M is projective if and only if M is flat.*
4. *If M, N are flat R-modules, then so is $M \otimes_R N$.*
5. *If $M = \oplus_i M_i$, then M is flat if and only if M_i are flat.*
6. *If $S \subseteq R$ is a multiplicative set, then $S^{-1}A$ is flat.*
7. *If $0 \to M' \to M \to M'' \to 0$ is exact and M'' is flat, then M is flat if and only if M' is flat.*
8. *(Extension of scalars) If $f : R \to S$ is a ring homomorphism and M is flat, then $M \otimes_R S$ is a flat S-module.*
9. *(Restriction of scalars for flat maps) If $f : R \to S$ is a flat ring homomorphism and N is a flat S-module, then N is a flat R-module.*
10. *Rings $R[x_1, \dots, x_n]$ is a flat R-module.*
11. *If R is a PID, then M is flat if and only if M is torsion free.*

Proof. 3. Follows from associativity of tensor products at once.

4. ($\Rightarrow$) Take any s.e.s. $0 \to P \to Q \to P' \to 0$. Tensoring with M gives $0 \to \oplus_i P \otimes M_i \to \oplus_i Q \otimes M_i \to \oplus_i P' \otimes M_i \to 0$. It is clear that $0 \to P \otimes M_i \to Q \otimes M_i \to P' \otimes M_i \to 0$ is exact as $\mathrm{Ker}\,(\oplus_i P \otimes M_i \to \oplus_i Q \otimes M_i) = \oplus_i \mathrm{Ker}\,(P \otimes M_i \to Q \otimes M_i)$. The other side ($\Leftarrow$) is easy.

5.

6.

7. Let $0 \to P \to Q$ be injective map of R-modules. By flatness of S as an R-module, we have $0 \to P \otimes_R S \to Q \otimes_R S$ is an exact sequence of S-modules. By flatness of N as an S-module, we have $0 \to (P \otimes_R S) \otimes_S N \to (Q \otimes_R S)_S N$ is exact. By associativity and $S \otimes_S N \cong N$, the result follows.

8. Note that $R[x_1, \dots, x_n]$ is isomorphic as R-module to $\bigoplus_{i \in \mathbb{N}} R$. The proof then follows from item 4. $\square$

16.19 Lifting properties : Étale maps

16.20 Lifting properties : Unramified maps

16.21 Lifting properties : Smooth maps

16.22 Simple, semisimple and separable algebras

These algebras are at the heart of the Galois phenomenology, i.e. all things related to polynomials splitting in a bigger field or not. Our study of these objects will thus motivate the study of the corresponding geometrical picture.

write more
se algebras,
.6.

16.22.1 Semisimple algebras

Definition 16.22.1.1. (**Semisimple algebras over a field** k) Let A be a k-algebra. Then A is a semisimple k-algebra if the Jacobson radical of A is 0.

16.22.2 Separable algebras

We will first study a rather special type of separable algebras, which are finitely generated and free as modules. Let us first give an example of such an algebra which is motivating our definition given later.

Example 16.22.2.1. Consider a ring A and the A-algebra A^n. There is something special about A^n; it is "separated" into finitely pieces which looks like A. This can be formalized. Indeed, we have the most obvious fact about such algebras that the obvious map

$$\begin{aligned} \varphi : A^n &\longrightarrow \operatorname{Hom}_A(A^n, A) \\ (a_1, \dots, a_n) &\longmapsto e_i \mapsto a_i \end{aligned}$$

is an isomorphism of A-algebras. More specifically, the map φ takes $(a_i) = (a_1, \dots, a_n)$ to the following mapping

$$\begin{aligned} \varphi((a_i)) : A^n &\longrightarrow A \\ (b_1, \dots, b_n) &\longmapsto a_1 b_1 + \cdots + a_n b_n. \end{aligned}$$

We now wish to generalize this. That is to say, taking above phenomenon as a definition we want to generalize when an A-algebra B "separates" into simple pieces. For this to work, we need to find an alternate characterization of the above phenomenon. For this, a little bit of thought shows that the above map is obtained as the dual map of the $\phi \in \operatorname{Hom}_A(A^n, \operatorname{Hom}_A(A^n, A))$ under the $\otimes$-Hom adjunction

$$\operatorname{Hom}_A(A^n \times A^n, A) \cong \operatorname{Hom}_A(A^n, \operatorname{Hom}_A(A^n, A))$$

where the isomorphism is given by

$$(A^n \times A^n \xrightarrow{f} A) \longmapsto ((a_i) \mapsto ((b_i) \mapsto f((a_i), (b_i)))) .$$

Now, consider the map

$$\begin{aligned} \tilde{\phi} : A^n \times A^n &\longrightarrow A \\ ((a_i), (b_i)) &\longmapsto \sum_{i=1}^{n} a_i b_i. \end{aligned}$$

The tensor-hom isomorphism tells us that $\tilde{\phi}$ is the dual map of ϕ above. Now notice that this dual map $\tilde{\phi}$ has a very simple description; it is given by the following commutative diagram:

$$\begin{tikzcd} A^n \times A^n \arrow[r, "\tilde{\phi}"] \arrow[d] & A \\ \mathrm{Hom}_A\left(A^n, A^n\right) \arrow[ur, "\mathrm{Tr}"'] & \end{tikzcd}.$$

It is this dual map that we shall generalize to the setting of arbitrary A-algebra B which is finitely generated and free of rank n. Indeed, for any A-algebra B and chose any generating set of B as an A-module, so that for any element $b \in B$, we can write $b = (b_1, \ldots, b_n) \in A^n$. We thus get a natural map $\tilde{\phi}$ as in the diagram below

$$\begin{tikzcd} (b,c) \arrow[d, mapsto] & B \times B \arrow[r, "\tilde{\phi}"] \arrow[d, "\kappa"'] & A \\ (b_i c_j)_{1\leq i,j\leq n} & \mathrm{Hom}_A\left(B, B\right) \arrow[ur, "\mathrm{Tr}"'] & \end{tikzcd}.$$

Now, consider the tensor-hom dual of $\tilde{\phi}$ to obtain

$$\begin{aligned} \phi : B &\longrightarrow \mathrm{Hom}_A\left(B, A\right) \\ b &\longmapsto \left(c \mapsto \tilde{\phi}(b,c)\right). \end{aligned}$$

In order to mimic the case of A^n, we would require the map ϕ to be an isomorphism. Indeed, this is what we do in the definition given below.

Before defining a nice class of separable algebras, let us define an A-algebra B to be *finitely free* if B is finitely generated and free as an A-module.

Definition 16.22.2.2. (**Free separable algebras**) Let A be a ring and B be a finitely free A-algebra of rank n and chose a generating set of B, so for $b \in B$, we can write $b = (b_1, \ldots, b_n)$ for $b_i \in A$. Define $\tilde{\varphi}$ to be the following map

$$\begin{tikzcd} (b,c) \arrow[d, mapsto] & B \times B \arrow[r, "\tilde{\varphi}"] \arrow[d, "\kappa"'] & A \\ (b_i c_j)_{1\leq i,j\leq n} & \mathrm{Hom}_A\left(B, B\right) \arrow[ur, "\mathrm{Tr}"'] & \end{tikzcd}.$$

Then B is said to be a separable A-algebra if the tensor-hom dual map $\varphi : B \to \mathrm{Hom}_A\left(B, A\right)$ is an isomorphism of A-algebras.

We would now like to show how separable algebras become familiar in the case of algebras over a field.

Proposition 16.22.2.3. *Let k be a field and A be an k-algebra. Then, the following are equivalent*

1. *A is a free separable k-algebra.*

2. *$A = \prod_{i=1}^{n} K_i$ where K_i are finite separable extensions of field k.*

Proof. ☐

Important e
Chapter 16.

Another characterization of separable algebras is as follows.

Lemma 16.22.2.4. *Let A be a ring and B be a finitely free A-algebra. Then the following are equivalent.*

1. *B is a separable A-algebra.*
2. *For all $\{w_1, \dots, w_n\}$ in B which is a generating set of free A-module B, we have*

$$\det\left(\mathrm{Tr}(w_i w_j)_{1\le i,j\le n}\right) \in A^{\times}.$$

t exercise,
.6.

Proof. ☐

16.23 Miscellaneous

We collect in this section results which so far doesn't fit in any other prior section. Perhaps this means our arrangement of material is not optimal.

The following result is a generalization of Lagrange interpolation formula.

Lemma 16.23.0.1. *Let K/F be an algebraic field extension. Then for any $\alpha_1, \dots, \alpha_n \in K$, such that α_i is not equal to any α_j nor any of its conjugate, and for any choice $\beta_1, \dots, \beta_n \in K$, there exists a polynomial $f(x) \in F[x]$ such that $f(\alpha_i) = \beta_i$ for all $i = 1, \dots, n$.*

Proof. Let $\alpha_1, \dots, \alpha_n \in K$ be such that α_j is not equal to α_i nor any of its conjugates for any $j \neq i$. Let $\beta_1, \dots, \beta_n \in K[\alpha_i]$. We wish to find a polynomial $f(x) \in F[x]$ such that $f(\alpha_i) = \beta_i$ for each $i = 1, \dots, n$.

We first observe that as K is an algebraic extension of F, therefore there exists $p_i(x) \in F[x]$ which is the minimal polynomial of $\alpha_i \in K$. This polynomial is obtained by looking at the kernel of evaluation at α_i, $\varphi_i : F[x] \to K$ where $x \mapsto \alpha_i$. Consequently, $p_i(x)$ is a monic irreducible polynomial of least degree in $F[x]$ such that $p_i(\alpha_i) = 0$, for each $i = 1, \dots, n$.

As $\mathfrak{m}_i := \langle p_i(x) \rangle \le F[x]$ are maximal ideals and $p_i(x) \neq p_j(x)$ because $\alpha_i \neq \alpha_j, \overline{\alpha_j}$[24], therefore $\mathfrak{m}_i + \mathfrak{m}_j = F[x]$ for all $i \neq j$. Hence $\mathfrak{m}_i$ are comaximal. Consequently, we obtain by Chinese remainder theorem that

$$\begin{array}{ccccccc} F[x] & \twoheadrightarrow & \frac{F[x]}{\mathfrak{m}_1 \dots \mathfrak{m}_n} & \xrightarrow{\cong} & \frac{F[x]}{\mathfrak{m}_1} \times \dots \times \frac{F[x]}{\mathfrak{m}_n} & \xrightarrow{\cong} & F[\alpha_1] \times \dots \times F[\alpha_n] \\ f(x) & \longmapsto & f(x) + \mathfrak{m}_1 \dots \mathfrak{m}_n & \longmapsto & (f(x) + \mathfrak{m}_i)_i & \longmapsto & (f(\alpha_1), \dots, f(\alpha_n)) \end{array}.$$

Consequently, by above diagram, for the elements $(\beta_,1 \dots, \beta_n) \in F[\alpha_1] \times \dots \times F[\alpha_n]$, there exists a polynomial $f(x) \in F[x]$ such that $(f(\alpha_1), \dots, f(\alpha_n)) = (\beta_1, \dots, \beta_n)$. Hence $f(\alpha_i) = \beta_i$ for each $i = 1, \dots, n$. This completes the proof. ☐

The following is a general exercise in basic ideal theory.

[24]because conjugates have same minimal polynomials.

Lemma 16.23.0.2. *Let R be a commutative ring with unity. Let $\mathfrak{p} \lneq R$ be a prime ideal and $I, J \leq R$ be ideals. Then,*

1. *$I^k \subseteq \mathfrak{p}$ for some $k \geq 0$ implies $I \subseteq \mathfrak{p}$,*
2. *the following are equivalent:*
   (a) *$\sqrt{I} + \sqrt{J} = R$,*
   (b) *$I + J = R$,*
   (c) *$I^k + J^l = R$ for all $k, l > 0$.*

Proof. 1. Let $I \leq R$ be an ideal and $\mathfrak{p} \lneq \mathbb{R}$ be a prime ideal. Then, we wish to show that $I^k \subseteq \mathfrak{p} \implies I \subseteq \mathfrak{p}$ for any $k \in \mathbb{N}$.

Indeed, pick any $x \in I$. As $x^k \in I$, therefore $x^k \in \mathfrak{p}$. As $x^k = x \cdot x^{k-1} \in \mathfrak{p}$, therefore either $x \in \mathfrak{p}$ of $x^{k-1} \in \mathfrak{p}$. If the former, then we are done. If the latter, then we have $x^{k-1} = x \cdot x^{k-2} \in \mathfrak{p}$. Continuing in this manner, we eventually reach to the conclusion that $x \in \mathfrak{p}$.
2. ((a) $\Rightarrow$ (b)) : As we have $x \in \sqrt{I}$ and $y \in \sqrt{J}$ such that $x + y = 1$, therefore for some $n, m \in \mathbb{N}$ we have $x^n \in I$ and $y^m \in J$. Now, observe that

$$\begin{aligned} 1 = 1^{n+m} = (x+y)^{n+m} &= \sum_{r=0}^{n+m} {}^{n+m}C_r x^r y^{n+m-r} \\ &= \sum_{r=0}^{n} {}^{n+m}C_r x^r y^{n+m-r} + \sum_{r=n+1}^{n+m} {}^{n+m}C_r x^r y^{n+m-r}. \end{aligned}$$

If $0 \leq r \leq n$, then $y^{n+m-r} \in J$ and if $n+1 \leq r \leq n+m$, then $x^r \in I$. Hence $\sum_{r=0}^{n} {}^{n+m}C_r x^r y^{n+m-r} \in J$ and $\sum_{r=n+1}^{n+m} {}^{n+m}C_r x^r y^{n+m-r} \in I$. This shows that there exists $a \in I$ and $b \in J$ then $a + b = 1$.

((b) $\Rightarrow$ (c)) : As we have $x \in I$ and $y \in J$ such that $x + y = 1$, thus writing $1 = 1^{k+l}$ again, we see

$$\begin{aligned} 1 &= 1^{k+l} = (x+y)^{k+l} \\ &= \sum_{r=0}^{k+l} {}^{k+l}C_r x^r y^{k+l-r} \\ &= \sum_{r=0}^{k} {}^{k+l}C_r x^r y^{k+l-r} + \sum_{r=k+1}^{k+l} {}^{k+l}C_r x^r y^{k+l-r}. \end{aligned}$$

If $0 \leq r \leq k$, then $y^{k+l-r} \in J^l$ and if $k+1 \leq r \leq k+l$, then $x^r \in I^k$. Consequently, we have $\sum_{r=0}^{k} {}^{k+l}C_r x^r y^{k+l-r} \in J^l$ and $\sum_{r=k+1}^{k+l} {}^{k+l}C_r x^r y^{k+l-r} \in I^k$. Hence there exists $a \in I^k$ and $b \in J^l$ such that $a + b = 1$.
((c) $\Rightarrow$ (a)) : Setting $k = l = 1$, we have that there exists $x \in I$ and $y \in J$ such that $x + y = 1$. As $\sqrt{I} \supseteq I$ and $\sqrt{J} \supseteq J$, therefore $x \in \sqrt{I}$ and $y \in \sqrt{J}$ such that $x + y = 1$. Hence $\sqrt{I} + \sqrt{J} = R$. This completes the proof. □

The following is a counterexample to the claim that a sub-algebra of a finite type algebra is a finite type algebra.

Lemma 16.23.0.3. *Let R be a ring. The ring $R[t, tx, tx^2, \dots, tx^i, \dots]$ is neither a finite type R-algebra nor a finite type $R[t]$-algebra.*

Proof. Let $S = R[t, tx, tx^2, tx^3, \dots]$. We wish to show that S is not a finitely generated R or $R[t]$ algebra.

a) We first show that S is not finitely generated R-algebra. Indeed, let $p_1, \dots, p_n \in S$ be generators of S as an R-algebra. Then, we have that $p_i \in R[t, tx, \dots, tx^{m_i}]$ as a polynomial can atmost be in finitely many indeterminates. Hence, letting $M = \max_i m_i$, we obtain that $p_1, \dots, p_n \in R[t, tx, \dots, tx^M]$. It then follows that the R-algebra generated by $p_1, \dots, p_n$ will only be inside $R[t, tx, \dots, tx^M]$. We consequently reduce to showing that $R[t, tx, \dots, tx^M] \neq S$.

Let $tx^{M+1} \in S$. We claim that $tx^{M+1} \notin R[t, tx, \dots, tx^M]$. Assuming to the contrary, we have that for some $a_{k_0,\dots,k_M} \in R$

$$\begin{aligned} tx^{M+1} &= \sum_{k_0,\dots,k_M} a_{k_0,\dots,k_M} t^{k_0} \dots (tx^M)^{k_M} \\ &= \sum_{k_0,\dots,k_M} a_{k_0,\dots,k_M} t^{k_0+\dots+k_M} \cdot x^{k_1+2k_2+\dots+Mk_M}. \end{aligned}$$

We thus deduce that $a_{k_0,\dots,k_M} \neq 0$ if and only if $k_0 + \dots + k_M = 1$. As $k_i \in \mathbb{Z}_{\geq 0}$, we further deduce that the only non-zero coefficients are $a_{1,0,\dots,0}, a_{0,1,\dots,0}, \dots, a_{0,0,\dots,1}$. Hence, the above equation reduces to

$$tx^{M+1} = a_{1,0,\dots,0}t + a_{0,1,\dots,0}tx + \dots + a_{0,0,\dots,1}tx^M.$$

Clearly, for no choice of coefficients $a_{1,0,\dots,0}, a_{0,1,\dots,0}, \dots, a_{0,0,\dots,1}$ in R can we make both sides equal in $R[t, x]$. This is a contradiction.

b) We now wish to show that S is not finitely generated as an $R[t]$-algebra. Assuming to the contrary, there exists $p_1, \dots, p_n \in S$ such that S is generated by them as an $R[t]$-algebra. Again for the same reason as in a), we see that $p_1, \dots, p_n \in R[t, tx, \dots, tx^M]$ for some $M \in \mathbb{Z}_{>0}$. Now, as $R[t, tx, \dots, tx^M] = R[t][tx, tx^2, \dots, tx^M]$, therefore the $R[t]$-algebra generated by $p_1, \dots, p_n$ will only be inside $R[t][tx, tx^2, \dots, tx^M]$. Hence, we reduce to showing that $R[t][tx, tx^2, \dots, tx^M] \neq S$. To this end, the exact same technique as in part a) works verbatim, as we need only show that $tx^{M+1} \notin R[t][tx, tx^2, \dots, tx^M] = R[t, tx, \dots, tx^M]$.

This completes the proof. $\square$

The following result characterizes all ideals of $F[[x]]$, yielding that $F[[x]]$ is a local PID, i.e. a DVR, and tells us that localization of $F[[x]]$ at the local parameter x yields the Laurent series ring, i.e. the fraction field of $F[[x]]$.

Proposition 16.23.0.4. *Let F be a field and $R = F[[x]]$.*

1. *An element in $a = a_0 + a_1x + \dots \in R$ is a unit if and only if $a_0 \neq 0$.*
2. *Every non-zero ideal of R is of the form x^kR.*
3. *$R[x^{-1}] = Q(R) = F((x))$.*

Proof. 1. ($\Rightarrow$) Since $\sum_{i\geq 0} a_ix^i$ is a unit in $F[[x]]$, therefore there exists $\sum_{i\geq 0} b_ix^i$ which is an inverse of $\sum_{i\geq 0} a_ix^i$. Consequently, we have

$$\begin{aligned} (a_0 + a_1x + \dots) \cdot (b_0 + b_1x + \dots) &= 1 \\ (a_0b_0 + (a_1b_0 + a_0b_1)x + \dots) &= 1. \end{aligned}$$

Comparing the degree 0 term both sides, we obtain $a_0b_0 = 1$. Therefore, if $a_0 = 0$, then $a_0b_0 = 0$ and we would thus obtain a contradiction.
($\Leftarrow$) Suppose $a_0 \neq 0$. We wish to find $\sum_{i\geq 0} b_i x^i$ such that $\left(\sum_{i\geq 0} a_i x^i\right) \cdot \left(\sum_{j\geq 0} b_j x^j\right) = 1$. We can calculate what b_is should be by observing the following:

$$\left(\sum_{i\geq 0} a_i x^i\right) \cdot \left(\sum_{j\geq 0} b_j x^j\right) = \sum_{k\geq 0} c_k x^k$$

where $c_k = \sum_{i+j=k} a_i b_j$. We now claim that there exists a unique solution for each b_i in the equations given by setting $c_0 = 1$ and $c_k = 0$ for all $k \geq 1$. We show this by strong induction. Indeed, for $c_0 = a_0 b_0 = 1$ yields that $b_0 = a_0^{-1}$. For $k = 1$, we have $c_1 = a_1 b_0 + a_0 b_1 = 0$ which thus yields $b_1 = -a_0^{-1} a_1 b_0$. We now wish to show that if b_l has a unique solution for all $l = 0, \dots k-1$, then b_k has a unique solution as well. Indeed, b_k satisfies the following equation coming from $c_k = 0$:

$$\begin{aligned} 0 &= \sum_{i+j=k} a_i b_j \\ &= a_0 b_k + \sum_{i+j=k, j<k} a_i b_j. \end{aligned}$$

By inductive hypothesis, for all $0 \leq j < k$, b_j has a unique solution. Consequently by the above, b_k has a unique solution as well. This completes the induction which yields the required formal power series.

$\sum_{j\geq 0} b_j x^j$ which acts as the inverse of $\sum_{i\geq 0} a_i x^i$. 2. We wish to show that any non-zero ideal $I \leq R$ is of the form $I = x^k R$ where $k \in \mathbb{N}$. Pick any ideal $I \leq R$. For any power series $p(x) = c_n x^n + c_{n+1} x^{n+1} + \dots$ where $c_n \neq 0$, we define n to be the ***co-degree*** of $p(x)$. Then, let $p(x) = c_k x^k + c_{k+1} x^{k+1} + \dots$ be the element of I with least co-degree (such an element exists by virtue of well-ordering of $\mathbb{N}$). Consequently, we obtain $p(x) = x^k(c_k + c_{k+1}x + \dots)$.

We thus claim that $I = x^k R$. Indeed, pick any $f(x) \in I$. Then, $f(x) = d_n x^n + d_{n+1} x^{n+1} + \dots$ where $d_n \neq 0$. Hence, we may write $f(x) = x^n(d_n + d_{n+1}x + \dots)$. By item 1, we know that $d_n + d_{n+1}x + \dots$ is a unit in R, so that we may write $f(x) = x^n u$, $u \in R$ is a unit. Now, as $f(x) \in I$, thus co-degree of f is atleast k as $p(x) \in I$ with co-degree k is the least co-degree element. Consequently, we may write $f(x) = x^k x^{n-k} u$. Hence $f(x) \in x^k R$. Conversely, pick any $x^k g(x) \in x^k R$. Since $p(x) = x^k(c_k + c_{k+1}x + \dots)$ where $c_k \neq 0$, therefore $c_k + c_{k+1}x + \dots$ is a unit, hence $p(x) = x^k v$ for some unit $v \in R$. Thus, $x^k \in I$ and hence $x^k g(x) \in I$. This completes the proof.

3. We wish to show that $R[\frac{1}{x}] = Q(R)$, the fraction field of R, i.e. $F((x))$. Indeed, as $x \in R$ is a non-zero element, therefore $1/x \in Q(R)$ and consequently, $R[\frac{1}{x}] \subseteq Q(R)$. We now wish to show that converse also holds.

Pick any $\frac{f(x)}{g(x)} \in Q(R)$ where $f(x), g(x) \in R$ are power series. Let $f(x)$ have co-degree n and $g(x)$ have co-degree m. We may then write

$$\frac{f(x)}{g(x)} = \frac{c_n x^n + c_{n+1} x^{n+1} + \dots}{d_m x^m + d_{m+1} x^{m+1} + \dots}$$

where $c_n, d_m \neq 0$. We may further write above as

$$\frac{f(x)}{g(x)} = \frac{x^n u}{x^m v}$$

for units $u = c_n + c_{n+1}x + \dots, v = d_m + d_{m+1}x + \dots \in R$ (by item 1).

If $n > m$, then $\frac{f(x)}{g(x)} = \frac{x^{n-m}w}{1}$ for some unit $w \in R$ and we know that $\frac{x^{n-m}}{1} \in R[\frac{1}{x}]$. If $n < m$, then $\frac{f(x)}{g(x)} = \frac{w}{x^{m-n}}$ for some unit $w \in R$ and we know that $\frac{1}{x^{m-n}} \in R[\frac{1}{x}]$. Finally if $n = m$, then $\frac{f(x)}{g(x)}$ is a unit of R and hence of $R[\frac{1}{x}]$.

Hence in all cases, $\frac{f(x)}{g(x)} \in R[\frac{1}{x}]$. We thus conclude $Q(R) \subseteq R[\frac{1}{x}]$, completing the proof. □

In the following theorem, we show some important properties of the ring $\mathbb{Z}[\omega]$, where ω is a third root of unity.

Theorem 16.23.0.5. *Let $R = \mathbb{Z}[\omega]$ where $\omega = e^{\frac{2\pi i}{3}}$ is a cube root of unity.*

1. *R is a Euclidean domain.*
2. *The function given by*

$$f : \operatorname{Spec}(\mathbb{Z}[\omega]) \longrightarrow \operatorname{Spec}(\mathbb{Z})$$
$$\pi \longmapsto \begin{cases} p & \text{if } \pi = p \text{ upto associates,} \\ \pi\bar{\pi} & \text{else.} \end{cases}$$

   is surjective such that $f^{-1}(p)$ is either $\{\pi, \bar{\pi}\}$ or $\{p\}$ (upto associates) for any prime $p \in \operatorname{Spec}(\mathbb{Z})$.
3. *Let $p \in \mathbb{Z}$ be a prime. The following are equivalent:*
   (i) *p splits in $\mathbb{Z}[\omega]$, that is $p = \alpha\bar{\alpha}$ for some $\alpha \in \mathbb{Z}[\omega]$,*
   (ii) *$x^2 \pm x + 1$ has a root in $\mathbb{F}_p$, that is, $\exists a \in \mathbb{F}_p$ such that $a \neq 1$ and $a^3 = \pm 1$,*
   (iii) *either $p = 3$ or $p = 1 \mod 3$.*
4. *Take any $n \in \mathbb{Z}$. The following are equivalent:*
   (i) *$n = a^2 \pm ab + b^2$ for some $a, b \in \mathbb{Z}$,*
   (ii) *primes $2 \mod 3$ occurs evenly many times in the prime factorization of n.*

Proof. 1. We first wish to show that R is a Euclidean domain. We claim that the following function

$$d : R \setminus \{0\} \longrightarrow \mathbb{N} \cup \{0\}$$
$$\alpha = a + b\omega \longmapsto \alpha\bar{\alpha} = a^2 + b^2 - ab$$

satisfies the axiom of size function for R. Indeed, pick any $\alpha, \beta \in R$ where $\beta \neq 0$. We may then write

$$\frac{\alpha}{\beta} = \frac{\alpha\bar{\beta}}{\beta\bar{\beta}} = \frac{\alpha\bar{\beta}}{c} = a + ib$$

where $a, b \in \mathbb{Q}$. As any rational $x \in \mathbb{Q}$ can be written as $x = n + q$ where $n \in \mathbb{Z}$ and $0 \leq q \leq 1/2$, therefore we may write

$$\frac{\alpha}{\beta} = a + ib = (n_1 + r_1) + \omega(n_2 + r_2)$$

where $n_1, n_2 \in \mathbb{Z}$ and $0 \leq r_1, r_2 \leq 1/2$. Thus,

$$\alpha = \beta(n_1 + \omega n_2) + \beta(r_1 + \omega r_2) \tag{1.1}$$

As $\alpha, \beta(n_1 + \omega n_2) \in R$, therefore by (1.1) we deduce that $\beta(r_1 + \omega r_2) \in R$.
Note that since the size function d is the norm map, which is actually a multiplicative map defined on whole of $\mathbb{C}$ as

$$\begin{aligned} \mathbb{C} &\longrightarrow \mathbb{R} \\ z &\longmapsto z\bar{z}, \end{aligned}$$

hence, we see that

$$\begin{aligned} d(\beta(r_1 + \omega r_2)) &= \beta\bar{\beta}(r_1^2 + r_2^2 - r_1 r_2) \\ &\leq \beta\bar{\beta}\left(\frac{1}{2^2} + \frac{1}{2^2}\right) \\ &= \frac{\beta\bar{\beta}}{2} \\ &< \beta\bar{\beta} \\ &= d(\beta). \end{aligned}$$

Thus, Eq. (1.1) is the required division of α by β. This proves that R is a Euclidean domain.

2. Let R be an arbitrary Euclidean domain and let Spec (R) denote the set of all prime ideals of R. As R is a Euclidean domain, therefore it is a PID. Consequently, Spec (R) is in one-to-one bijection with prime/irreducible elements of R together with 0. Hence, we write $p \in$ Spec (R) to mean a prime element of R. We know that $\mathbb{Z}[\omega]$ and $\mathbb{Z}$ are Euclidean domains. We wish to show that there is a surjective map

$$\begin{aligned} f : \operatorname{Spec}(\mathbb{Z}[\omega]) &\longrightarrow \operatorname{Spec}(\mathbb{Z}) \\ \pi &\longmapsto \begin{cases} p & \text{if } \pi = p \text{ upto associates,} \\ \pi\bar{\pi} & \text{else.} \end{cases} \end{aligned}$$

such that $f^{-1}(p)$ is either $\{\pi, \bar{\pi}\}$ or $\{p\}$ (upto associates) for any prime $p \in$ Spec $(\mathbb{Z})$ where $\pi \in$ Spec $(\mathbb{Z}[\omega])$ is a prime element.
We first observe that $\mathbb{Z}[\omega]$ has a non-trivial automorphism given by $\alpha = a + b\omega \mapsto \bar{\alpha} = a + b\omega^2$. Pick $\pi \in$ Spec $(\mathbb{Z}[\omega])$ a non-zero prime element. Observe that automorphisms takes a prime element to a prime element. As $\mathbb{Z}$ is a UFD, therefore for $p_1, \ldots, p_l \in$ Spec $(\mathbb{Z})$ non-zero primes, and $\pi_1, \ldots, \pi_k \in$ Spec $(\mathbb{Z}[\omega])$ non-zero primes, we may write

$$\begin{aligned} \pi\bar{\pi} &= a^2 + b^2 - ab \\ &= p_1 \ldots p_l \\ &= \pi_1 \ldots \pi_k \end{aligned}$$

where the last equality comes from writing prime factorization of each p_i in $\mathbb{Z}[\omega]$.
Now, as $\mathbb{Z}[\omega]$ is a UFD, therefore $k = 2$ and hence $l \leq 2$. We now have two cases

(i) If $l = 2$, then $\pi\bar{\pi} = p_1 p_2$. Expanding each p_i into product of primes in $\mathbb{Z}[\omega]$, we immediately deduce by unique factorization in $\mathbb{Z}[\omega]$ that $p_1 = \pi$ and $p_2 = \bar{\pi}$ upto associates (wlog). Hence, $\bar{\pi} = p_2 = p_1$. That is,

$$\pi\bar{\pi} = p^2.$$

(ii) If $l = 1$, then

$$\pi\bar{\pi} = p$$

for some non-zero prime $p \in \mathrm{Spec}\,(\mathbb{Z})$.

This defines the function $f : \mathrm{Spec}\,(\mathbb{Z}[\omega]) \to \mathrm{Spec}\,(\mathbb{Z})$. Next, we wish to show that this is surjective. Indeed, pick any non-zero $p \in \mathrm{Spec}\,(\mathbb{Z})$. Using prime factorization in $\mathbb{Z}[\omega]$, we obtain primes $\pi_1, \dots, \pi_k$ in $\mathbb{Z}[\omega]$ such that

$$p = \pi_1 \dots \pi_k.$$

Again using the conjugation automorphism yields us

$$p^2 = (\pi_1\bar{\pi_1}) \dots (\pi_k\bar{\pi_k}).$$

Note $\pi_i\bar{\pi_i} \in \mathbb{Z}$. Hence, by unique factorization of $\mathbb{Z}$, we obtain $k \leq 2$. We now have two cases

(i) If $k = 2$, then $p^2 = (\pi_1\bar{\pi_1})(\pi_2\bar{\pi_2})$. As π_i are not units, we deduce that $p = \pi_1\bar{\pi_1}$ and $p = \pi_2\bar{\pi_2}$. Consequently, we have $\pi_1\bar{\pi_1} = \pi_2\bar{\pi_2}$. Thus, by unique factorization of $\mathbb{Z}[\omega]$, we further deduce that $\pi_1 = \pi_2$ or $\bar{\pi_2}$. Hence, $p = \pi\bar{\pi}$ for a unique $\pi \in \mathrm{Spec}\,(\mathbb{Z}[\omega])$.

(ii) If $k = 1$, then

$$p^2 = \pi\bar{\pi}$$

for some $\pi \in \mathrm{Spec}\,(\mathbb{Z}[\omega])$. Writing p as a product of primes in $\mathbb{Z}[\omega]$, we immediately deduce of unique factorization of $\mathbb{Z}[\omega]$ that $p = \pi'$ upto units for some non-zero prime $\pi' \in \mathrm{Spec}\,(\mathbb{Z}[\omega])$. Consequently, $p^2 = \pi'\bar{\pi}' = \pi\bar{\pi}$. Again by unique factorization of $\mathbb{Z}[\omega]$, we immediately deduce that $\pi = \pi'$ upto units.

This shows the surjectivity of the map f.

3. (i) $\iff$ (ii) : By part b), p splits in $\mathbb{Z}[\omega]$ iff p is not prime in $\mathbb{Z}[\omega]$. This happens iff $\mathbb{Z}[\omega]/p$ is not a domain. We now observe

$$\begin{aligned}\frac{\mathbb{Z}[\omega]}{p\mathbb{Z}[\omega]} &\cong \frac{\frac{\mathbb{Z}[x]}{\langle x^2+x+1\rangle}}{\frac{\langle p, x^2+x+1\rangle}{\langle x^2+x+1\rangle}} \\ &\cong \frac{\mathbb{Z}[x]}{\langle p, x^2+x+1\rangle} \\ &\cong \frac{\frac{\mathbb{Z}[x]}{p\mathbb{Z}[x]}}{\frac{\langle p, x^2+x+1\rangle}{p\mathbb{Z}[x]}} \\ &\cong \frac{\mathbb{F}_p[x]}{\langle x^2+x+1\rangle}.\end{aligned}$$

Hence, p is not prime in $\mathbb{Z}[\omega]$ iff x^2+x+1 is reducible in $\mathbb{F}_p[x]$. As a polynomial of degree 2 or 3 over a field is reducible iff it has a root in the field, therefore p is not prime in $\mathbb{Z}[\omega]$ iff x^2+x+1 has a root in $\mathbb{F}_p$. Similarly, since ω^2 has minimal polynomial x^2-x+1 and $\mathbb{Z}[\omega]=\mathbb{Z}[\omega^2]$, hence repeating the above yields p is not prime in $\mathbb{Z}[\omega]$ iff x^2-x+1 has a root in $\mathbb{F}_p[x]$.

(ii) $\Rightarrow$ (iii) : If $p=2$, then $x^2\pm x+1$ has no roots in $\mathbb{F}_2$. Consequently, let $p\neq 2,3$. We then wish to show that $p=1 \mod 3$. Let $a\in\mathbb{F}_p$ be the root of $f(x)=x^2\pm x+1$. Thus, $a^3=\pm 1$. Observe that $a\neq\pm 1$ as if $a=1$, then $f(1)$ and $f(-1)$ are either 1 or 3 and since $p\neq 3$, therefore $f(1),f(-1)\neq 0$, a contradiction.
As $a^3=\pm 1$ and $a\neq\pm 1$, therefore the order of $a\in\mathbb{F}_p^*$ is either 3 or 6. In either case, as $\left|\mathbb{F}_p^*\right|=p-1$, therefore by Lagrange's theorem, $3|p-1$ or $6|p-1$. But in both cases, we have $p=1 \mod 3$.

(iii) $\Rightarrow$ (ii) : If $p=3$, then $1\in\mathbb{F}_3$ is root of x^2+x+1 and 2 is the root of x^2-x+1. If $p=1 \mod 3$, then we proceed as follows. As $\mathbb{F}_p^*$ is a cyclic group of order $p-1$ and since $p-1=3k$ for some $k\in\mathbb{Z}$, hence there exists an element $a\in\mathbb{F}_p$ of order 3. Consequently, we have $a^3=1$ and thus x^3-1 in $\mathbb{F}_p[x]$ has a root. As $x^3-1=(x-1)(x^2+x+1)$ and $a\neq 1$, hence a is a root of x^2+x+1.
Now since

$$\frac{\mathbb{F}_p[x]}{\langle x^2+x+1\rangle}\cong\frac{\mathbb{F}_p[x-1]}{\langle (x-1)^2+(x-1)+1\rangle}=\frac{\mathbb{F}_p[x]}{\langle x^2-x+1\rangle}$$

therefore if x^2+x+1 has a root in $\mathbb{F}_p$, then so does x^2-x+1.

4. (i) $\Rightarrow$ (ii) : Write the prime factorization of n in $\mathbb{Z}[\omega]$ as follows

$$\begin{aligned} n &= (a+b\omega)(a+b\omega^2)\\ &= (\pi_1\ldots\pi_k)(\bar{\pi}_1\ldots\bar{\pi}_k)\\ &= (\pi_1\bar{\pi}_1)\ldots(\pi_k\bar{\pi}_k).\end{aligned}$$

From parts b) and c), we know that for any prime element $\pi\in\mathbb{Z}[\omega]$, we have $\pi\bar{\pi}=p$ iff $p=3$ or $1 \mod 3$ and $\pi\bar{\pi}=p^2$ iff $p=2 \mod 3$. Consequently, we have

$$n=(p_1\ldots p_m)(p_{m+1}^2\ldots p_k^2)$$

where we call primes $p_1,\ldots,p_m$ which are either 3 or $1 \mod 3$ of ***split type***. Similarly, we call the primes $p_{m+1},\ldots,p_k$ which are $2 \mod 3$ of ***unsplit type***. From above it is clear that unsplit type primes appear evenly many times (they appear in squares) in the prime factorization of n.

(ii) $\Rightarrow$ (i) : Let $n\in\mathbb{Z}$ be such that its prime factorization in $\mathbb{Z}$ is as follows

$$n=(p_1\ldots p_m)(q_1^{2k_1}\ldots q_n^{2k_n})$$

where q_i are primes of unsplit type, that is, $q_i=2 \mod 3$ and p_i are of split type, that is, 3 or $1 \mod 3$. Now, by part b), we may write $p_i=\pi_i\bar{\pi}_i$ as they split in $\mathbb{Z}[\omega]$ and $q_i=\xi_i$, where

ξ_i, π_i are primes in $\mathbb{Z}[\omega]$.
It follows that we may write

$$\begin{aligned} n &= (\pi_1\bar{\pi_1}\dots\pi_m\bar{\pi_m})\left(\xi_1^{2k_1}\dots\xi_n^{2k_n}\right) \\ &= (\xi_1^{k_1}\dots\xi_n^{k_n})(\pi_1\dots\pi_m)\cdot(\xi_1^{k_1}\dots\xi_n^{k_n})(\bar{\pi_1}\dots\bar{\pi_m}) \\ &= \alpha\bar{\alpha} \end{aligned}$$

where $\alpha = (\xi_1^{k_1}\dots\xi_n^{k_n})(\pi_1\dots\pi_m) = a + b\omega$, as required.
This completes the proof. □

Example 16.23.0.6. As an example use of above we may now find all ordered tuples $(a, b) \in \mathbb{Z}^2$ such that $2100 = a^2 - ab + b^2$.

Observe that

$$\begin{aligned} 2100 &= 2^2 \cdot 3 \cdot 5^2 \cdot 7 \\ &= 2^2 \cdot 5^2 \cdot (2+\omega)(2+\omega^2)(3+\omega)(3+\omega^2). \end{aligned}$$

We now wish to find the distinct $\alpha \in \mathbb{Z}[\omega]$ such that $2100 = \alpha\bar{\alpha}$. For this, we first need to find all units of $\mathbb{Z}[\omega]$.

Indeed, we claim that the units of $\mathbb{Z}[\omega]$ are $1, -1, \omega, -\omega, 1+\omega, -1-\omega$. We give a terse proof of this fact as follows. Let $a+b\omega \in \mathbb{Z}[\omega]$ be a unit, so that there exists $c+d\omega$ such that $(a+b\omega)(c+d\omega) = 1$. Then, the multiplicative map

$$\begin{aligned} \mathbb{Z}[\omega] &\to \mathbb{Z} \\ \alpha &\mapsto \alpha\bar{\alpha} \end{aligned}$$

yields in $\mathbb{Z}$ that $(a^2+b^2-ab)(c^2+d^2-cd) = 1$. This forces $a^2+b^2-ab = 1 = c^2+d^2-cd$. From these equations one can deduce that $c+d\omega = (a-b)-b\omega$. Hence, $a+b\omega$ is a unit iff $a^2+b^2-ab = 1$. It follows by AM-GM inequality on a^2 and b^2 that $ab \le 1$. Hence, we deduce that $a = 1, b = 1$ or $a = -1, b = -1$ or $a = 0$ or $b = 0$. Correspondingly, we get the six units of $\mathbb{Z}[\omega]$ as mentioned above.

In order to count the number of distinct pairs $(a, b) \in \mathbb{Z}^2$ such that $n = a^2+b^2-ab = (a+b\omega)(a+b\omega^2)$ properly, let us bring some notations. Let $X_n = \{(a+b\omega) \mid (a+b\omega)(a+b\omega^2) = n\} \subseteq \mathbb{Z}[\omega]$. Denote $f : \mathbb{Z}[\omega] \to \mathbb{Z}$ to be the multiplicative map $\alpha \mapsto \alpha\bar{\alpha}$. We thus have $X_n = f^{-1}(n)$. Now observe that

1. for each $a+b\omega \in X_n$, we have $b+a\omega \in X_n$,
2. for each $a+b\omega \in X_n$, we have $a+b\omega^2 \in X_n$,
3. for each $a+b\omega \in X_n$ and $u \in \mathbb{Z}[\omega]$ a unit, we have $u(a+b\omega) \in X_n$. This is because in $\mathbb{Z}[\omega]$, inverse of a unit is its conjugate.

Our goal is to count ordered tuples $(a, b) \in \mathbb{Z}^2$ such that $n = a^2+b^2-ab$. Immediately, we see that such ordered tuples are in bijection with X_n. Hence, we reduce to counting X_n.

From the above discussion, we see the elements in X_n obtained by multiplying by units are

- $2\cdot 5\cdot 1\cdot(2+\omega)(3+\omega) = 50+40\omega$,
- $2\cdot 5\cdot -1\cdot(2+\omega)(3+\omega) = -50-40\omega$,
- $2\cdot 5\cdot \omega\cdot(2+\omega)(3+\omega) = -40+10\omega$,

- $2 \cdot 5 \cdot (-\omega) \cdot (2+\omega)(3+\omega) = 40 - 10\omega$,
- $2 \cdot 5 \cdot (1+\omega) \cdot (2+\omega)(3+\omega) = 10 + 50\omega$,
- $2 \cdot 5 \cdot (-1-\omega) \cdot (2+\omega)(3+\omega) = -10 - 50\omega$,
- $2 \cdot 5 \cdot 1 \cdot (2+\omega^2)(3+\omega) = 40 - 10\omega$.
- $2 \cdot 5 \cdot -1 \cdot (2+\omega^2)(3+\omega) = -40 + 10\omega$,
- $2 \cdot 5 \cdot \omega \cdot (2+\omega^2)(3+\omega) = 10 + 50\omega$,
- $2 \cdot 5 \cdot (-\omega) \cdot (2+\omega^2)(3+\omega) = -10 - 50\omega$,
- $2 \cdot 5 \cdot (1+\omega) \cdot (2+\omega^2)(3+\omega) = 50 + 40\omega$,
- $2 \cdot 5 \cdot (-1-\omega) \cdot (2+\omega^2(3+\omega) = -50 - 40\omega$.

Similarly, those obtained by swapping are

- $40 + 50\omega$,
- $-40 - 50\omega$,
- $10 - 40\omega$,
- $50 + 10\omega$,
- $-50 - 10\omega$.

Hence, there are 12 such ordered tuples $(a, b) \in \mathbb{Z}^2$ given by $(40, 50), (-40, -50), (10, -40), (50, 10), (-50, 10), (50, 4$

The following is a simple but powerful lemma about certain type of k-algebras.

Lemma 16.23.0.7. *Let k be a field and A be a k-algebra such that there is a maximal ideal $\mathfrak{m} \lneq A$ for which $A/\mathfrak{m} \cong k$. Then,*

$$A \cong k \oplus \mathfrak{m}$$

where $k \oplus \mathfrak{m}$ obtains the k-algebra structure from A.

Proof. Consider the triangle

$$\begin{tikzcd} A \arrow[d, two heads, "\pi"'] & k \arrow[l, hook'] \arrow[dl, "\cong"] \\ A/\mathfrak{m} & \end{tikzcd} .$$

Pick any $a \in A$. We have $\pi(a) \in A/\mathfrak{m} \cong k$, so let $\pi(a) \in k$ by identifying under that isomorphism. Consequently, we may write $a = \pi(a) + (a - \pi(a))$. Note since $\pi(a - \pi(a)) = \pi(a) - \pi(\pi(a)) = \pi(a) - \pi(a) = 0$ by the commutativity of the above, therefore $a \in \mathfrak{m}$. Furthermore $\mathfrak{m} \cap k = 0$ is immediate as $\mathfrak{m}$ is a proper ideal. It follows that $A = k \oplus \mathfrak{m}$ as k-linear subspaces, and thus $k \oplus \mathfrak{m}$ is a k-algebra as well, isomorphic to A, where, since $(k_1 + m_1) \cdot (k_2 + m_2) = k_1k_2 + k_1m_2 + k_2m_1 + m_1m_2$ inside of A, hence we may define the k-algebra structure on $k \oplus \mathfrak{m}$ as

$$(k_1, m_1) \cdot (k_2, m_2) = (k_1k_2, k_1m_2 + k_2m_1 + m_1m_2)$$

for $(k_i, m_i) \in k \oplus \mathfrak{m}$. $\square$

The following proposition shows that any submodule of a free module over a PID is free (which is not true in general). This is also a main ingredient in computation of K_0 of a PID (that it is $\mathbb{Z}$).

Proposition 16.23.0.8. *Let R be a PID and X an indexing set. Then any submodule of $R^{\oplus X}$ is free.*

Proof. Let $M \leq R^{\oplus X}$ be a submodule. For each $Y \subseteq X$, consider the submodule

$$M_Y := M \cap R^{\oplus Y}.$$

Denote by $\mathbb{T}$ the following partially ordered set

$$\mathbb{T} = \left\{ (B, Y) \mid Y \subseteq X,\ B \subseteq M \text{ s.t. } M_Y = \bigoplus_{b \in B} Rb \right\}$$

where $(B_1, Y_1) \leq (B_2, Y_2)$ if and only if $B_1 \subseteq B_2$ and $Y_1 \subseteq Y_2$.

We first claim that $\mathbb{T}$ is non-empty. Indeed, consider any finite subset $Y \subseteq X$. We claim that $M \cap R^{\oplus Y}$ is free. To this end, first observe that $M \cap R^{\oplus Y} \leq R^{\oplus Y}$. As finite direct sum of noetherian modules is noetherian, therefore $R^{\oplus Y}$ is noetherian. As a module is noetherian if and only if every submodule is finitely generated, therefore $M \cap R^{\oplus Y}$ is finitely generated.

By structure theorem of finitely generated modules over a PID, we deduce that

$$M \cap R^{\oplus Y} \cong \frac{R}{d_1 R} \oplus \cdots \oplus \frac{R}{d_k R} \oplus R^n. \tag{5.1}$$

As R is a PID, so in particular a domain, therefore $R^{\oplus Y}$ has no R-torsion element. Consequently, in Eq. (5.1), we conclude that $d_i = 1$ for each $i = 1, \ldots, k$, that is, $M \cap R^{\oplus Y} \cong R^n$. Hence, $M \cap R^{\oplus Y}$ is free, as required. More generally this argument shows that any submodule of R^X where X is finite is free. This shows that $\mathbb{T}$ is non-empty.

We next wish to show that $\mathbb{T}$ has a maximal element. We will use Zorn's lemma on $\mathbb{T}$ for this. Pick any totally ordered subset $\mathcal{T} \subseteq \mathbb{T}$. We wish to show that $\mathcal{T}$ has an upper bound. Indeed, denote

$$C = \bigcup_{(B,Y) \in \mathcal{T}} B \;\&\; Z = \bigcup_{(B,Y) \in \mathcal{T}} Y.$$

We claim that

$$M_Z := M \cap R^{\oplus Z} = \bigoplus_{c \in C} Rc.$$

For $(\subseteq)$, pick an element $m \in M_Z$. We may write

$$m = (m_\alpha)_{\alpha \in Z}$$

where $m_\alpha \in R$ for each $\alpha \in Z$ and $m_{\alpha_i} \neq 0$ only for $i = 1, \ldots, k$. As $\alpha_i \in Z$ and $\mathcal{T}$ is totally ordered, therefore for some $(B, Y) \in \mathcal{T}$, we have $\alpha_i \in Y$ for each $i = 1, \ldots, k$. Thus, $m \in M \cap R^Y = \bigoplus_{b \in B} Rb$. In particular, $m \in \bigoplus_{b \in B} Rb \subseteq \bigoplus_{c \in C} Rc$ as $B \subseteq C$. This shows $(\subseteq)$. For $(\supseteq)$, pick any $(m_c)_{c \in C} \in \bigoplus_{c \in C} Rc$. Then $m_c = 0$ for all but finitely many $c_1, \ldots, c_k$. As $\mathcal{T}$ is totally ordered and $m_{c_i} \in Rc_i$, therefore there exists $(B, Y) \in \mathcal{T}$ such that all $c_i \in B$ for $i = 1, \ldots, k$. We then conclude that $m \in \bigoplus_{b \in B} Rb = M \cap R^{\oplus Y} \subseteq M \cap R^{\oplus Z}$, as needed. This shows that $(C, Z) \in \mathbb{T}$.

It is clear that for any $(B, Y) \in \mathcal{T}$, we have $(B, Y) \leq (C, Z)$ by construction. Hence we have produced an upper bound for any toset of $\mathbb{T}$. It follows by Zorn's lemma that $\mathbb{T}$ has a maximal element. Let it be denoted by $(\tilde{B}, \tilde{Y})$.

It now suffices to show that $\tilde{Y} = X$ as it would imply $M = M \cap R^{\oplus X} \in \mathbb{T}$, and hence is free. To this end, suppose $\tilde{Y} \subsetneq X$. Then there exists $\tilde{Y} \subsetneq Y'$ such that $Y' \setminus \tilde{Y}$ is finite. We shall now construct an element $(B', Y') \in \mathbb{T}$ such that $(\tilde{B}, \tilde{Y}) \leq (B', Y')$ and $(\tilde{B}, \tilde{Y}) \neq (B', Y')$, thus contradicting the maximality of $(\tilde{B}, \tilde{Y})$.

We first have the following exact sequence

$$0 \longrightarrow M \cap R^{\oplus \tilde{Y}} \xhookrightarrow{i} M \cap R^{\oplus Y'} \xrightarrow{\pi} \mathrm{CoKer}\,(()i) \longrightarrow 0 \tag{5.2}$$

We claim that CoKer $(()i)$ is a free module. To this end, we first claim that

$$\mathrm{CoKer}\,(i) = \frac{M \cap R^{\oplus Y'}}{M \cap R^{\oplus \tilde{Y}}} \cong K$$

where $K \leq R^{\oplus Y' \setminus \tilde{Y}}$ is a submodule. Indeed, consider the map $\tilde{\varphi}$ obtained by the universal property of quotients

$$\begin{tikzcd} M \cap R^{\oplus Y'} \arrow[r, "\varphi"] \arrow[d, two heads] & R^{\oplus Y' \setminus \tilde{Y}} \\ \frac{M \cap R^{\oplus Y'}}{M \cap R^{\oplus \tilde{Y}}} \arrow[ur, dashed, "\tilde{\varphi}"'] & \end{tikzcd}$$

where φ is the R-linear map which takes $(m_\alpha)_{\alpha \in Y'} \mapsto (m_\alpha)_{\alpha \in Y' \setminus \tilde{Y}}$. It is clear that $\mathrm{Ker}\,(\varphi) = M \cap R^{\oplus \tilde{Y}}$. Consequently, $\tilde{\varphi}$ is an inclusion and let $K \leq R^{\oplus Y' \setminus \tilde{Y}}$ be its image.

As $Y' \setminus \tilde{Y}$ is finite and we showed above that every submodule of a finitely generated free module is free, therefore

$$K = \bigoplus_{z \in Z} Rz \cong R^{\oplus Z}.$$

where $Z \subseteq R^{\oplus Y' \setminus Y}$. This shows that CoKer $(()i) \cong R^{\oplus Z}$ is a free R-module. In particular, it is projective. Consequently, the exact sequence of (5.2) is split exact so that there exists $j : \mathrm{CoKer}\,(()i) \hookrightarrow M \cap R^{\oplus Y'}$ such that $\pi j = \mathrm{id}_{\mathrm{CoKer}(()i)}$. It now follows immediately that

$$\begin{aligned} M \cap R^{\oplus Y'} &= \mathrm{Ker}\,(\pi) \oplus j\,(\mathrm{CoKer}\,(i)) \\ &= \left(M \cap R^{\oplus \tilde{Y}}\right) \oplus j\,(\mathrm{CoKer}\,(i)) \end{aligned}$$

where $j\,(\mathrm{CoKer}\,(i)) \cong R^{\oplus Z}$ so it is free. Hence, we see that $B' \supsetneq \tilde{B}$. This shows that $(B', Y') \geq (\tilde{B}, \tilde{Y})$, completing the proof. $\square$

A similar result to the above yields that projective modules over a local ring are free.

Theorem 16.23.0.9. *Let $(R, \mathfrak{m})$ be a local ring*[25]*. If P is a finitely generated projective R-module, then P is free. Moreover,* $\mathrm{rank} P = \dim_{R/\mathfrak{m}} P/\mathfrak{m}P$.

Let us digress for a moment and first show a crucial property of local rings which is the technical heart of the proof.

[25] the argument works also for non-commutative local rings.

Proposition 16.23.0.10. *Let $(R, \mathfrak{m})$ be a local ring. If $\{\bar{x}_1, \dots, \bar{x}_n\}$ is an $R/\mathfrak{m}$-basis of $(R/\mathfrak{m})^{\oplus n}$ for $x_i \in R$, then $\{x_1, \dots, x_n\}$ is an R-basis of $R^{\oplus n}$.*

Proof. Let $x_i = (a_{i1}, \dots, a_{in}) \in R^n$. Consequently, we get a matrix $A = (a_{ij}) \in M_n(R)$ whose rows are x_i. Note that it is sufficient to show that A is invertible, that is $A \in \mathrm{GL}_n(R)$. Denote $\bar{A} \in M_n(R/\mathfrak{m})$ to be the matrix reduced mod $\mathfrak{m}$. Note that $\bar{A}$ is invertible, that is, $\bar{A} \in \mathrm{GL}_n(R/\mathfrak{m})$, as it is a basis of $(R/\mathfrak{m})^n$. Consequently, there exists $B \in M_n(R)$ such that $\bar{A} \cdot \bar{B} = I_n = \bar{B} \cdot \bar{A}$. We now construct an inverse of A in $\mathrm{GL}_n(R)$. Note that we have $A \cdot B = (c_{ij})$ where $c_{ii} \in R^\times$ and $c_{ij} \in \mathfrak{m}$ for $i \neq j$. Doing an elemeantry column operations on $A \cdot B$, we deduce that there exists $E \in GL_n(R)$ such that $(A \cdot B) \cdot E$ is a diagonal matrix with diagonal entries being units of R, as required. $\square$

The proof is now immediate.

Proof of Theorem 16.23.0.9. Let P be a finitely generated projective R-module. Denote $\kappa = R/\mathfrak{m}$ be the residue field of $(R, \mathfrak{m})$. Let $\dim_\kappa P/\mathfrak{m}P = n$ and $\{\bar{x}_i\}_{i=1,\dots,n} \in P/\mathfrak{m}P$ be a κ-basis of $P/\mathfrak{m}P$ for $x_i \in P$. We claim that $\{x_i\}_{i=1,\dots,n}$ is an R-basis of P. Indeed, as P is projective, there exists a projective module Q such that $P \oplus Q = R^{m+n}$. Going modulo $\mathfrak{m}$, we get that $\dim_\kappa Q/\mathfrak{m}Q = m$. Let $\{\bar{x}_{n+i}\}_{i=1,\dots,m}$ be a κ-basis of $Q/\mathfrak{m}Q$ for $x_{n+i} \in Q$. Consequently, $\{x_i\}_{i=1,\dots,n+m} \subseteq R^{n+m}$ is such that $\{\bar{x}_i\}_{i=1,\dots,n+m}$ forms a κ-basis of $(R/\mathfrak{m})^{n+m}$. By Proposition 16.23.0.10, it follows that $\{x_i\}_{i=1,\dots,n+m}$ is an R-basis of $R^{n+m} = P \oplus Q$. It is clear from $R^{m+n} = P \oplus Q$ that $\{x_1, \dots, x_n\} \subseteq R^{n+m}$ spans P and are linearly independent, as required. $\square$

Chapter 17

K-Theory of Rings

Contents

17.1 K_0

We study the first K-group, which is also the easiest to construct and understand. We begin by studying this for rings, before looking at it more geometrically via schemes.

17.1.1 K_0 of a ring & basic properties

Definition 17.1.1.1 ($K_0(R)$)**.** Let R be a ring and consider $F(R)$ to be the free abelian group generated by objects of skeleton of $\mathbf{Proj}(R)$, denoted $\mathbf{Proj}(R)^{\cong}$. Consider the subgroup of $F(R)$

$$E = \langle\{[P \oplus Q] - [P] - [Q] \mid P, Q \in \mathbf{Proj}(R)\}\rangle$$

Then we define

$$K_0(R) := F(R)/E.$$

That is,

$$\mathbf{Proj}(R)^{\cong} \hookrightarrow F(R) \twoheadrightarrow K_0(R).$$

Consequently, $K_0(R)$ is an abelian group as it is quotient of the free abelian group $F(R)$. In particular, the addition in $K_0(R)$ of $P, Q \in \mathbf{Proj}(R)$ is $[P] + [Q] = [P \oplus Q]$.

Some observations from the definitions are as follows.

Lemma 17.1.1.2. *Let R be a ring and $P, Q \in \mathbf{Proj}(R)$ be two finitely generated projective left R-modules. Then, the following are equivalent:*

1. *$[P] = [Q]$ in $K_0(R)$.*
2. *There exists $P' \in \mathbf{Proj}(R)$ such that $P \oplus P' \cong Q \oplus P'$*[1]*.*
3. *There exists $n \geq 0$ such that $P \oplus R^n \cong Q \oplus R^n$.*

Proof. (1. $\Rightarrow$ 2.) Unravelling the definition, we deduce that there exists $P_i, Q_i, P'_j, Q'_j \in \mathbf{Proj}(R)$ such that

$$P - Q = \sum_{i=1}^{n} P_i \oplus Q_i - P_i - Q_i - \left(\sum_{j=1}^{m} P'_j \oplus Q'_j - P'_j - Q'_j\right)$$

in $F(R)$. By rearrangement, we deduce that P is either isomorphic to $P_i \oplus Q_i$ or P_i or Q_i and Q is either isomorphic to $P'_j \oplus Q'_j$ or P'_j or Q'_j. Conseuqently, the summand may be taken as P' can be taken to be the direct sum of all P_i, Q_i, P'_j, Q'_j which will be projective.

(2. $\Rightarrow$ 3.) As $P' \in \mathbf{Proj}(R)$, thus, there exists $Q' \in \mathbf{Proj}(R)$ such that $P' \oplus Q' = R^n$ for some $n \in \mathbb{N}$. Hence taking direct sum with Q' in the given isomorphism will give us the required isomorphism.

(3. $\Rightarrow$ 1.) As $[P] = [P \oplus R^n] - [R^n] = [Q \oplus R^n] - [R^n] = [Q]$ in $K_0(R)$, hence we get the desired result. □

A simple corollary yields the precise meaning of $[P] = [R^n]$ in $K_0(R)$.

Corollary 17.1.1.3. *Let R be a ring and P be a finitely generated projective module. If $[P] = [R^n]$, then P is stably free.*

Proof. By Lemma 17.1.1.2, we deduce that $P \oplus R^k \cong R^{n+k}$, as required. □

[1] This is at times also called stable isomorphism of two modules.

We now establish that K_0 is a functor on **Ring** to **Ab**.

Construction 17.1.1.4 (Functor K_0)**.** Let $f : A \to B$ be a map of rings. We define $f_* : K_0(A) \to K_0(B)$ by extension of scalars:

$$\begin{aligned} f_* : K_0(A) &\longrightarrow K_0(B) \\ [P] &\longmapsto [P \otimes_A B]. \end{aligned}$$

As $(f \circ g)_* = f_* \circ g_*$, therefore K_0 is a functor.

The following shows that if R is a commutative ring then $K_0(R)$ is a commutative ring.

Lemma 17.1.1.5. *Let R be a commutative ring. Then the operation $([P],[Q]) \mapsto [P \otimes_R Q]$ for $P, Q \in$ $\mathbf{Proj}(R)$ defines a commuative ring structure on $K_0(R)$.*

Proof. Indeed, this is immediate by commutativity of $\otimes$ upto isomorphism for commutative rings and distributivity of $\otimes$ over $\oplus$. □

The following states that K_0 preserves products.

Lemma 17.1.1.6. *Let R, R_1, R_2 be rings. If $R = R_1 \times R_2$, then*

$$K_0(R) \cong K_0(R_1) \times K_0(R_2).$$

Proof. We need only show a bijection $\mathbf{Proj}(R)^{\cong} \cong \mathbf{Proj}(R_1)^{\cong} \times \mathbf{Proj}(R_2)^{\cong}$. Indeed, consider the function $P \mapsto (P \otimes_R R_1, P \otimes_R R_2)$. We claim that the map $P_1 \oplus P_2 \mapsfrom (P_1, P_2)$ is an inverse of above. Indeed, we see $(P \otimes_R R_1) \oplus (P \otimes_R R_2) = P \otimes_R (R_1 \oplus R_2) = P \otimes_R R \cong P$. Similarly, $(P_1 \oplus P_2) \otimes_R R_1 = (P_1 \oplus P_2) \otimes_R \frac{R}{0 \times R_2} \cong \frac{P_1 \oplus P_2}{(0 \times R_2) \cdot (P_1 \oplus P_2)} \cong \frac{P_1 \oplus P_2}{P_2} \cong P_1$, as required. □

The following result says that K_0 is invariant of reducing the structure.

Proposition 17.1.1.7. *Let R be a ring and $I \leq R$ be a nilpotent ideal. Then $K_0(R) \cong K_0(R/I)$. In particular, $K_0(R) \cong K_0(R_{\mathrm{red}})$.*

Proof. It is sufficient to show that $\mathbf{Proj}(R)^{\cong} \cong \mathbf{Proj}(R/I)^{\cong}$. Indeed, this is true by idempotent lifting (see Exercise I.2.2 of **[WeibK]**). □

The following is a simple characterization of units of the commutative ring $K_0(R)$.

Proposition 17.1.1.8. *Let R be a commutative ring. Then*

$$\mathrm{Pic}(R) \hookrightarrow K_0(R)^{\times}$$

and every element of the form $[P] \in K_0(R)^{\times}$ is in $\mathrm{Pic}(R)$.

Proof. † Consider the map

$$\begin{aligned} \varphi : \mathrm{Pic}(R) &\longrightarrow K_0(R)^{\times} \\ [P] &\longmapsto [P] \end{aligned}$$

which takes the isomorphism class of a line bundle to its K_0-class in $K_0(R)$. This is a group homomorphism as $\varphi([P]\cdot[Q]) = \varphi([P\otimes_R Q]) = [P\otimes_R Q] = [P]\cdot[Q]$. Now take any $[P]\in K_0(R)^\times$, then there is $[Q]\in K_0(R)^\times$ such that $[P\otimes_R Q] = [R]$. It follows by Lemma 17.1.1.2 that $P\otimes_R Q$ is stably free, i.e. $(P\otimes_R Q)\oplus R^n \cong R^{n+1}$. Comparing the rank, we see that $P\otimes_R Q$ is constant rank 1. Thus, $P\otimes_R Q$ is a line bundle which is stably free, so by an argument involving top exterior power, we deduce that $P\otimes_R Q$ is free of rank 1. By another rank argument, we deduce that P and Q are line bundles. Thus $\varphi([P]) = [P]$.

This is injective as if P is a line bundle such that $[P] = [R]$ in $K_0(R)$, then P is stably free by Lemma 17.1.1.2, and thus is free of rank 1, i.e. $P\cong R$ and is thus the identity of $\mathrm{Pic}(R)$, as required. □

17.1.2 Computations

There are few main computations for K_0 of a ring; PIDs, local rings and more generally, Dedekind domains. We recall that rings may not be commutative.

Theorem 17.1.2.1 (K_0 of a PID)**.** *Let R be a PID. Then the map*

$$\begin{aligned}\varphi : \mathbb{Z} &\longmapsto K_0(R)\\ 1 &\longmapsto [R]\end{aligned}$$

is an isomorphism.

Proof. Observe that $\varphi(n) = [R^n] = [R]+\cdots+[R]$ n-times. To see injectivity, observe that if $\varphi(n) = [0]$, then $[R^n] = [0]$. It follows by Lemma 17.1.1.2, 3, that $R^{n+m}\cong R^m$. Going modulo any maximal ideal of R, we deduce that we have an $R/\mathfrak{m}$-vector space isomorphism $(R/\mathfrak{m})^{\oplus m+n}\cong (R/\mathfrak{m})^{\oplus m}$. It follows at once that $n=0$, as required.

To see surjectivity, we need only show that $\mathrm{Im}\,(\varphi)$ contains the image of $\mathbf{Proj}(R)^{\cong}$ in $K_0(R)$. To this end, take any $P\in\mathbf{Proj}(R)^{\cong}$. We need only show that for some $n\in\mathbb{N}$, $[P] = [R^n]$ in $K_0(R)$. It suffices to show that every projective module over a PID is free. Indeed, this is true (see Proposition 16.23.0.8). □

As a field is a PID, we deduce the following.

Corollary 17.1.2.2 (K_0 of a field)**.** *Let F be a field. Then $K_0(F)\cong\mathbb{Z}$.* □

Theorem 17.1.2.3 (K_0 of a local ring)**.** *Let $(R,\mathfrak{m})$ be a local ring. Then the group homomorphism*

$$\begin{aligned}\varphi : \mathbb{Z} &\longrightarrow K_0(R)\\ 1 &\longmapsto [R]\end{aligned}$$

is an isomorphism.

Proof. Observe that the map $\rho' : F(R)\to F(R/\mathfrak{m})$ given by $P\mapsto P\otimes_R R/\mathfrak{m} = P/\mathfrak{m}P$ defines a group homomorphism $\rho : K_0(R)\to K_0(R/\mathfrak{m}) = \mathbb{Z}$. Note that $\rho([R^n]) = [R^n/\mathfrak{m}R^n] = [(R/\mathfrak{m}R)^{\oplus n}]$ which yields $n\in\mathbb{Z}$ under ρ. Consequently, ρ is surjective. We need only show that ρ is injective. To this end, observe that it is sufficient to show that any finitely generated projective R-module is free, as ρ is injective over this subset of $K_0(R)$. Indeed, this is true (see Theorem 16.23.0.9). □

The next computation is for Dedekind domains.

Theorem 17.1.2.4 (K_0 of a Dedekind domain)**.** *Let R be a Dedekind domain. Then*

$$K_0(R) \cong \mathbb{Z} \oplus \mathrm{Cl}(R)$$

is an isomorphism of groups where $\mathrm{Cl}(R)$ *is the ideal class group of* R.

However, we will show something general, which will showcase to us the use of geometric viewpoint. Recall that $H_0(X) = \mathbb{Z}^{\oplus r}$ where r is the number of path-components of X. Note that we may interpret $H_0(X) = C(X, \mathbb{Z})$, the set of all continuous functions from X to the discrete space $\mathbb{Z}$. We denote $H_0(R) = H_0(\mathrm{Spec}\,(R))$. Let R be a commutative ring. Then, any finitely generated projective module P gives a continuous map $\mathrm{rank}(P) : \mathrm{Spec}\,(R) \to \mathbb{Z}$ given by $\mathfrak{p} \mapsto \dim_{\kappa(\mathfrak{p})} P \otimes_R \kappa(\mathfrak{p})$ (see Exercise I.2.11 of [**WeibK**]). Thus, we get the following result.

Lemma 17.1.2.5 (The rank map)**.** *Let R be a commutative ring. Then, the map*

$$\begin{aligned} \mathrm{rank} : K_0(R) &\longrightarrow H_0(R) \\ [P] &\longmapsto \mathrm{rank}(P) \end{aligned}$$

is a ring homomorphism.

Proof. We need only show that $\mathrm{rank}(P \oplus Q) = \mathrm{rank}(P) + \mathrm{rank}(Q)$ and $\mathrm{rank}(P \otimes_R Q) = \mathrm{rank}(P) \cdot \mathrm{rank}(Q)$. The former is immediate. For the latter, we need only observe that $P \otimes_R Q \otimes_R \kappa(\mathfrak{p}) \cong (P \otimes_R \kappa(\mathfrak{p})) \otimes_{\kappa(\mathfrak{p})} (\kappa(\mathfrak{p}) \otimes_R Q)$, then the result follows from the basic fact of dimension of tensor product of vector spaces. □

Lemma 17.1.2.6. *Let R be a commutative ring. If R is noetherian, then $H_0(R)$ is a direct summand of $K_0(R)$. We write*

$$K_0(R) \cong H_0(R) \oplus \tilde{K}_0(R),$$

where $\tilde{K}_0(R) = \mathrm{Ker}\,(\mathrm{rank})$.

Proof. By Exercise I.2.4 of [**WeibK**], we have that for any continuous map $f : \mathrm{Spec}\,(R) \to \mathbb{Z}$, there is a decomposition $R = R_1 \times \dots R_k$ where on each $\mathrm{Spec}\,(R_i)$, f is constant n_i, say. Thus, for each such f, we construct the R-module

$$R^f = R_1^{n_1} \times \cdots \times R_k^{n_k}$$

which we claim is finitely generated projective R-module. Indeed this can be seen by observing that if P_1 is projective R_1 and P_2 is projective R_2 modules, then $P_1 \oplus P_2$ is projective $R_1 \times R_2$-module, by making $P_1 \oplus P_2$ a direct summand of an $R_1 \times R_2$-free module. Thus, we define a map

$$\begin{aligned} H_0(R) &\longrightarrow K_0(R) \\ f &\longmapsto [R^f]. \end{aligned}$$

Now observe that the composite

$$H_0(R) \longrightarrow K_0(R) \xrightarrow{\mathrm{rank}} H_0(R)$$

is such that $f \mapsto [R^f] \mapsto \mathrm{rank}(R^f)$. Since $\mathrm{rank}(R^f) = f$ by definition of R^f, thus the composite is id. It follows that we have a decomposition $K_0(R) \cong H_0(R) \oplus \mathrm{Ker}\,(\mathrm{rank})$, as required. □

As a Dedekind domain is noetherian, this gives us a hint towards the above result (Theorem 17.1.2.4).

Construction 17.1.2.7 (Picard group and determinant bundle)**.** Next map that we wish to discuss is the determinant map for a commutative ring, which will be a map from $K_0(R)$ to $\mathrm{Pic}(R)$, the Picard group of scheme $\mathrm{Spec}\,(R)$, which may be described as the commutative group of isomorphism classes of all finitely generated projective modules[2] of constant rank 1, where the group operation is $\otimes$ and inverse is taking dual module.

For a projective module P, we may define a rank 1 projective module given as follows. If P has constant rank, then $\det(P) = \wedge^{\mathrm{rank}(P)}P$, the top exterior power of P, which has rank 1 and is projective as it is locally free (since exterior powers commute with tensoring). Whereas if P doesn't have constant rank, then writing $R = R_1 \times \cdots \times R_n$ such that $\mathrm{rank}(P)$ is constant on each $\mathrm{Spec}\,(R_i)$. We thus get a decomposition $P = P_1 \times \cdots \times P_n$ where each P_i is a projective R_i-module. We then define $\det(P) = \wedge^{\mathrm{rank}(P_1)}P_1 \times \cdots \times \wedge^{\mathrm{rank}(P_n)}P_n$. As an R-module, this has rank 1 by a simple tensor calculation.

The main observation here is the following.

Lemma 17.1.2.8 (The det map)**.** *Let R be a commutative ring. Then the map*

$$\det : K_0(R) \to \mathrm{Pic}(R)$$

is a surjective group homomorphism.

Proof. We need only show that det is a group homomorphism on the generators of $F(R)$. That is, we wish to show that $\det([P \oplus Q]) = \det([P]) \otimes_R \det([Q])$. To this end, by above discussion, we may reduce to assuming P has constant rank n and Q has constant rank m. Now, by binomial sum formula, we get, $\det(P \oplus Q) = \wedge^{n+m}(P \oplus Q) = \bigoplus_{i=0}^{n+m} \wedge^i P \otimes \wedge^{n+m-i}Q$. All of the terms except $\wedge^n P \otimes_R \wedge^m Q$ are zero above. Consequently, we get the required result. To see surjectivity, observe that any rank 1 projective module P is such that $\det([P]) = \wedge^1 P \cong P$. □

We come to the main theorem.

Theorem 17.1.2.9. *Let R be a commutative ring. Then the map*

$$\mathrm{rank} \oplus \det : K_0(R) \longrightarrow H_0(R) \oplus \mathrm{Pic}(R)$$

is surjective with kernel being the ideal

$$SK_0(R) = \langle [P] - [R^m] \mid P \in \mathbf{Proj}(R)^{\cong} \text{ of constant rank } m \ \&\ \wedge^m P \cong R \rangle.$$

Proof. Indeed, $SK_0(R)$ is in the $\mathrm{Ker}\,(\mathrm{rank} \oplus \det)$. Conversely, as $\mathrm{Ker}\,(\mathrm{rank} \oplus \det) \subseteq \tilde{K}_0(R) = \mathrm{Ker}\,(\mathrm{rank})$ and $\tilde{K}_0(R)$ is the filtered limit of the set $F_n(R) = \{[P] \mid P \in \mathbf{Proj}(R)^{\cong} \text{ of constant rank } n\}$ via the map $F_n(R) \to \tilde{K}_0(R)$ mapping as $[Q] \mapsto [Q] - [R^n]$ (see Lemma 2.3.1 of **[WeibK]**), thus we deduce that any $[P] \in \mathrm{Ker}\,(\mathrm{rank} \oplus \det)$ is of form $[P] = [Q] - [R^m]$ where Q is projective of constant rank m.

As $[P] \in \mathrm{Ker}\,(\mathrm{rank} \oplus \det)$, therefore $\wedge^m Q = R$, as can be seen easily. □

[2] also called algebraic line bundles over $\mathrm{Spec}\,(R)$.

Corollary 17.1.2.10. *Let R be a commutative noetherian ring of dimension one. Then*

$$\text{rank} \oplus \det : K_0(R) \longrightarrow H_0(R) \oplus \text{Pic}(R)$$

is an isomorphism.

The proof of Theorem 17.1.2.4 is immediate as $\text{Pic}(R)$ of a Dedekind domain is the ideal class group.

Proof. By classification of finitely generated projective modules over a commutative noetherian ring of dimension one, it follows that if $P \in \mathbf{Proj}(R)^{\cong}$ is of constant rank, then it is isomorphic to $\det(P) \oplus R^{\text{rank}(P)-1}$. Consequently, the ideal $SK_0(R) = 0$ and we conclude by Theorem 17.1.2.9. □

Another simple calculation is of Artin rings.

Proposition 17.1.2.11. *Let R be an Artinian ring with $|\text{mSpec}\,(R)| = n$. Then*

$$K_0(R) \cong \mathbb{Z}^{\oplus n}.$$

Proof. † By structure theorem for Artinian rings, we know that $A \cong \prod_{i=1}^n A/\mathfrak{m}_i^k$ for some $k > 0$ where $\text{mSpec}\,(R) = \{\mathfrak{m}_1, \dots, \mathfrak{m}_n\}$ and each $A/\mathfrak{m}_i^k$ is an Artin local ring (see Theorem 8.7, pp 90, [**AMD**]). By Lemma 17.1.1.6, we deduce that

$$K_0(R) \cong \prod_{i=1}^n K_0(A/\mathfrak{m}_i^k).$$

By Theorem 17.1.2.3, we have that $K_0(A/\mathfrak{m}_i^k) \cong \mathbb{Z}$, giving the required result. □

There is a famous important calculation of K_0 done by Quillen and Suslin around the same time.

Theorem 17.1.2.12 (Quillen-Suslin)**.** *Let R be a PID (for example, a field). Then all projective modules over $R[x_1, \dots, x_n]$ is free. In particular,*

$$K_0(R[x_1, \dots, x_n]) \cong \mathbb{Z}.$$

17.1.3 Homological properties of K_0

We have already seen the reduced K_0, denoted $\tilde{K}_0$ in Lemma 17.1.2.6. We thus discuss some other homology-type results for K_0.

Relative exact sequence for K_0

To discuss excision and Mayer-Vietoris, we first need to understand what will be the analogue of union of two subspaces in this context. It is perhaps not that surprising that the right answer is geometrically motivated.

Lemma 17.1.3.1 (Milnor squares)**.** *Let $f : R \to S$ be a ring homomorphism and $I \leq R$ be an ideal such that $f|_I : I \to f(I)$ is a bijection onto an ideal $f(I) \leq S$ which we also denote by I. Then the following is a pullback square of rings:*

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, two heads] \arrow[dr, phantom, "\lrcorner", very near start] & S \arrow[d, two heads] \\ R/I \arrow[r, "\bar{f}"'] & S/I \end{tikzcd}.$$

We call them Milnor squares.

Proof. Recall that

$$R/I \times_{\bar{f}} S := \frac{R/I \times S}{\langle (r+I, s) \in R/I \times S \mid f(r) - s \in I \rangle}.$$

Consider the map

$$\begin{aligned} R &\longrightarrow R/I \times_{\bar{f}} S \\ r &\longmapsto (r + I, f(r)). \end{aligned}$$

This is injective as if $r \in I$ and $f(r) = 0 \in I$, then since $f|_I$ is bijective, then $r = 0$. This is surjective as for any $(r+I, s) \in R/I \times_{\bar{f}} S$, we have that $f(r) - s = f(i), i \in I$ and thus $s = f(r-i)$. Consequently, $r - i \mapsto (r - i + I, f(r-i)) = (r + I, s)$, as required. $\square$

Remark 17.1.3.2. A Milnor square equivalently yields the following pushout diagram of affine schemes:

$$\begin{tikzcd} \mathrm{Spec}\,(R) & \mathrm{Spec}\,(S) \arrow[l, "f^*"'] \\ \mathrm{Spec}\,(R/I) \arrow[u, hook] \arrow[ur, phantom, "\ulcorner", very near end] & \mathrm{Spec}\,(S/I) \arrow[l, "\bar{f}^*"] \arrow[u, hook] \end{tikzcd}.$$

Consequently, we get that Spec (R) is obtained by gluing Spec (R/I) to Spec (S) along the closed subspace Spec $(S/I) \hookrightarrow$ Spec (S) via the map $\bar{f}^*$, that is,

$$\mathrm{Spec}\,(R) = \mathrm{Spec}\,(S) \amalg_{\bar{f}^*} \mathrm{Spec}\,(R/I).$$

Remark 17.1.3.3 (Milnor squares and excisive triples)**.** By the above remark, it is clear that Milnor squares behave as excisive triples as seen in topology. That is a Milnor square

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, two heads] \arrow[dr, phantom, "\lrcorner", very near start] & S \arrow[d, two heads] \\ R/I \arrow[r, "\bar{f}"'] & S/I \end{tikzcd}$$

can be seen as the excisive triple (Spec (R), Spec (S), Spec (R/I)) for the space Spec (R).

Using the above idea, we can now define K-groups relative to an ideal I as follows.

Definition 17.1.3.4 (Relative K_0). Let R be a commutative ring and $I \leq R$ be an ideal. Consider the commutative ring $R \oplus I$ with 1 whose operation is

$$(r,x)\cdot(s,y) := (rs, ry + sx + xy),$$

where identity is $(1,0)$. We call this the augmented ring. Consider the projection homomorphism $p : R \oplus I \to R$. This induces the map $p_* : K_0(R \oplus I) \to K_0(R)$. We thus define

$$K_0(R,I) := \operatorname{Ker}(p_* : K_0(R \oplus I) \to K_0(R)).$$

We call $K_0(R,I)$ the relative K_0-group w.r.t ideal I. As the composite of the ring homomorphisms $R \to R \oplus I \to R$ is identity, therefore after applying K_0, we get $K_0(R) \to K_0(R \oplus I) \to K_0(R)$ is identity. It follows by splitting lemma that we have

$$K_0(R \oplus I) \cong K_0(R) \oplus K_0(R,I).$$

We define now a group which we will meet consistently.

Definition 17.1.3.5 ($\mathrm{GL}(R)$). Let R be a ring and $\mathrm{GL}_n(R) = \operatorname{Aut}(R^n)$, the group of R-linear automorphisms of the free module R^n where the group operation is composition. We may think of $\mathrm{GL}_n(R)$ as $n \times n$ invertible matrix over R. Observe that we have injective maps

$$\begin{aligned}\mathrm{GL}_n(R) &\overset{\iota_n}{\hookrightarrow} \mathrm{GL}_{n+1}(R)\\ g_n &\longmapsto \begin{bmatrix} g_n & 0 \\ 0 & 1\end{bmatrix}.\end{aligned}$$

Consequently we have a directed system $\{\mathrm{GL}_n(R), \iota_n\}_n$. We define

$$\mathrm{GL}(R) := \varinjlim_n \mathrm{GL}_n(R)$$

An element $[g_n] \in \mathrm{GL}(R)$ for some $g_n \in \mathrm{GL}_n(R)$ is an equivalence class where $g_n \sim h_m$ for some $h_m \in \mathrm{GL}_m(R)$ if and only if $\iota_{n,k}(g_n) = \iota_{m,k}(h_m)$ in $\mathrm{GL}_k(R)$ for some $k \geq n, m$. Observe that the group operation of $\mathrm{GL}(R)$ is as follows: for $g_n, h_m \in \mathrm{GL}(R)$, $[g_n]\cdot[h_m] = [\iota_{n,p}(g_n)\cdot\iota_{m,p}(h_m)]$ for some $p \geq n, m$. It is clear that $[g_n]$ only consists of all the elements $\iota_{n,k}(g_n)$ for all $k \geq n$, i.e. $[g_n]$ is the infinite invertible matrix obtained by padding by 1 on diagonals.

The main theorem here is the following:

Theorem 17.1.3.6. *Let R, S be commutative rings and let $I \leq R$ be an ideal. Then we have a natural exact sequence*

$$\mathrm{GL}(R) \to \mathrm{GL}(R/I) \xrightarrow{\partial} K_0(R,I) \to K_0(R) \to K_0(R/I).$$

Excision for K_0

A main observation here yields the independence of $K_0(R,I)$ on R.

Theorem 17.1.3.7. *Let R be a ring and $I \leq R$ be an ideal. If $f : R \to S$ is a ring homomorphism which maps ideal I isomorphically to an ideal $f(I)$ of S (which we denote by I again), then*

$$K_0(R,I) \cong K_0(S,I).$$

Proof. See Exercise II.2.3 of **[WeibK]**. □

Mayer-Vietoris for K_0

We begin by constructing a Mayer-Vietoris sequence for K_0, which will be later extended to a long-exact sequence while discussing higher K-groups, just like in homology theory. We do this essentially by using excision, as is usually done in singular homology.

Theorem 17.1.3.8 (Mayer-Vietoris)**.** *Consider a Milnor square*

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, two heads] \arrow[dr, phantom, "\lrcorner", very near start] & S \arrow[d, two heads] \\ R/I \arrow[r, "\bar{f}"'] & S/I \end{tikzcd}.$$

Then there is a long exact sequence

$$\mathrm{GL}(S/I) \to K_0(R) \to K_0(S) \oplus K_0(R/I) \to K_0(S/I).$$

Proof. Observe that we have maps relative sequences of Theorem 17.1.3.6 as follows:

$$\begin{tikzcd} \mathrm{GL}(R) \arrow[r] \arrow[d] & \mathrm{GL}(R/I) \arrow[r, "\partial"] \arrow[d] & K_0(R,I) \arrow[r] \arrow[d, "\cong"'] & K_0(R) \arrow[r] \arrow[d] & K_0(R/I) \arrow[d] \\ \mathrm{GL}(S) \arrow[r] & \mathrm{GL}(S/I) \arrow[r, "\partial"'] & K_0(S,I) \arrow[r] & K_0(S) \arrow[r] & K_0(S/I) \end{tikzcd}$$

where the middle vertical map is the excision isomorphism of Theorem 17.1.3.7. By Barrat-Whitehead lemma (Lemma 14.6, **[HarpAT]**), we get an exact sequence

$$\mathrm{GL}(S/I) \to K_0(R) \to K_0(S) \oplus K_0(R/I) \to K_0(S/I),$$

as required. □

17.1.4 K_0 of a scheme

Let X be a scheme. We will define the 0^{th} K-group of X in essentially the same manner as we have done for rings, by keeping in mind the fact that the finite rank algebraic vector bundles over X are analogues of finitely generated projective modules (in the affine case, both are indeed same). Using this we will define and prove some basic properties of $K_0(X)$.

Definition 17.1.4.1 (K_0 **of a scheme** X)**. TODO.**

We begin by observing that K_0 of a ring R is same as K_0 of the affine scheme Spec (R).

17.1.5 Applications

We present some applications of the K_0 and its calculations.

Wall's finiteness obstruction

Construction 17.1.5.1 (0^{th}-Whitehead group of a group G)**.** Let G be a group and $\mathbb{Z}[G]$ be the group ring of G[3]. Then we have the following commutative diagram of rings:

$$\begin{tikzcd} \mathbb{Z} \arrow[r, "\iota"] \arrow[rr, bend right, "\mathrm{id}"'] & \mathbb{Z}[G] \arrow[r, "\epsilon"] & \mathbb{Z} \end{tikzcd} .$$

Applying K_0, we get the following commutative diagram of groups:

$$\begin{tikzcd} K_0(\mathbb{Z}) \arrow[r, "\iota_*"] \arrow[rr, bend right, "\mathrm{id}"'] & K_0(\mathbb{Z}[G]) \arrow[r, "\epsilon_*"] & K_0(\mathbb{Z}) \end{tikzcd} .$$

It hence follows that $\iota_* : K_0(\mathbb{Z}) \hookrightarrow K_0(\mathbb{Z}[G])$ is an injective map. We define the 0^{th}-Whitehead group of G to be the cokernel of ι_*:

$$\mathrm{Wh}_0(G) := \mathrm{CoKer}\,(\iota_*) = K_0(\mathbb{Z}[G])/K_0(\mathbb{Z}).$$

Moreover, as $\epsilon_* \circ \iota_* = \mathrm{id}_*$, therefore the following s.e.s. is split on the left:

$$\begin{tikzcd} 0 \arrow[r] & K_0(\mathbb{Z}) \arrow[r, "\iota_*"] & K_0(\mathbb{Z}[G]) \arrow[r] \arrow[l, bend left, "\epsilon_*"] & \mathrm{Wh}_0(G) \arrow[r] & 0 \end{tikzcd} .$$

Consequently, we get the following decomposition of $K_0(\mathbb{Z}[G])$:

$$K_0(\mathbb{Z}[G]) \cong K_0(\mathbb{Z}) \oplus \mathrm{Wh}_0(G).$$

Remark 17.1.5.2 (Information in $\mathrm{Wh}_0(G)$.)**.** Observe that the map $\iota_* : K_0(\mathbb{Z}) \to K_0(\mathbb{Z}[G])$ maps the generator $[\mathbb{Z}]$ to $[\mathbb{Z}[G]]$. As $\mathrm{Wh}_0(G) := \mathrm{CoKer}\,(\iota_*)$, thus we deduce that $\mathrm{Wh}_0(G)$ is that part of $K_0(\mathbb{Z}[G])$ which stores the information of non-stably free projective f.g. $\mathbb{Z}[G]$-modules (see Corollary 17.1.1.3). Thus, we deduce that

$$\mathrm{Wh}_0(G) = 0 \implies K_0(\mathbb{Z}[G]) \cong \mathbb{Z}.$$

Moreover, if $[P] = [Q]$ in $\mathrm{Wh}_0(G)$, then $[P] - [Q] \in K_0(\mathbb{Z})$ and thus $[P] - [Q] = [\mathbb{Z}[G]^n]$ in $K_0(\mathbb{Z}[G])$. It follows that $[P] = [Q \oplus \mathbb{Z}[G]^n]$ and hence by Lemma 17.1.1.2, we deduce that $P \oplus \mathbb{Z}[G]^k \cong Q \oplus \mathbb{Z}[G]^{n+k}$. We thus deduce that

$$[P] = [Q] \text{ in } \mathrm{Wh}_0(G) \iff \exists n, k \in \mathbb{N} \text{ s.t. } P \oplus \mathbb{Z}[G]^k \cong Q \oplus \mathbb{Z}[G]^{n+k}.$$

The information stored in $\mathrm{Wh}_0(G)$ is thus quite non-trivial (K_0-classes of non-stably free f.g. projective $\mathbb{Z}[G]$-modules).

The following theorem explains the interest in $\mathrm{Wh}_0(G)$. Recall that a CW-complex X is *dominated* by a complex K if there is a map $f : K \to X$ which exhibits X as a homotopy retract of K.

[3]Note that if G is infinite cyclic group with generator x, then $\mathbb{Z}[G] = \mathbb{Z}[x]$, the polynomial ring.

Theorem 17.1.5.3 (Wall's finiteness obstruction)**.** *Let X be a CW-complex which is dominated by a finite CW-complex K. Denote $G = \pi_1(X)$.*

1. *The tuple (X, K, G) determines an element $w(X) \in \mathrm{Wh}_0(G)$ which is independent of dominating complex K.*
2. *The following are equivalent:*
   (a) *X is homotopy equivalent to a finite CW-complex.*
   (b) *$w(X) = 0$ in $\mathrm{Wh}_0(X)$.*

17.2 K_1

It is said that the Grothendieck group $K_0(R)$ is orthogonal to all higher K-groups in the sense that the former looks at the spread of projective modules while the rest only look at the eventual behaviour as the size of those modules grows.

Hyman Bass's group $K_1(R)$ is the intended "value group for the determinant" of an invertible matrix over R and can be defined as the abelianization of the direct limit of automorphism groups of finitely generated projective modules. Gaussian elimination and Dieudonne's determinant map will help us to compute $K_1(R)$ in familiar examples.

17.2.1 K_1 of a ring & basic properties

We begin by defining $K_1(R)$.

Definition 17.2.1.1 ($K_1(R)$)**.** Let R be a ring. We define $K_1(R)$ to be the abelianisation of $\mathrm{GL}(R)$:

$$K_1(R) := \frac{\mathrm{GL}(R)}{[\mathrm{GL}(R), \mathrm{GL}(R)]}.$$

We immediately have a decomposition of $K_1(R)$ as follows.

Construction 17.2.1.2 (Dieudonné's det for K_1)**.** Let R be a commutative ring. Consider the map

$$\det : \mathrm{GL}_n(R) \to R^\times$$

which is the determinant map. As $\det \circ \iota_n = \det$ by the determinant of block diagonals, therefore we obtain a map

$$\begin{aligned} \det : \mathrm{GL}(R) &\to R^\times \\ [g_n] &\mapsto \det(g_n). \end{aligned}$$

It is easy to see that this is a group homomorphism. Moreover, this map is surjective. Observe that since det of an element in the commutator is 1, thus by universal property of quotients, we get a surjective group homomorphism

$$\begin{aligned} \det : K_1(R) &\longrightarrow R^\times \\ \overline{[g_n]} &\longmapsto \det(g_n) \end{aligned}$$

which is called Dieudonné's determinant[4]. We further denote its kernel as

$$SK_1(R) := \mathrm{Ker}\left(\det : K_1(R) \twoheadrightarrow R^\times\right).$$

As the composite $R^\times \to K_1(R) \to R^\times$ given by $u \mapsto \overline{[u]} \mapsto \det(u) = u$ is identity, it follows that we have a splitting:

$$K_1(R) \cong R^\times \oplus SK_1(R).$$

Note that the kernel of $\det : \mathrm{GL}_n(R) \to R^\times$ is exactly $\mathrm{SL}_n(R)$.

[4]We sometimes call $\det : \mathrm{GL}(R) \to R^\times$ as Dieudonné's determinant as well.

The K_1 of a ring commutes with product.

Lemma 17.2.1.3. *Let $R = R_1 \times R_2$ be a ring where R_i are rings. Then*

$$K_1(R) \cong K_1(R_1) \times K_1(R_2).$$

Proof. Since $\mathrm{GL}_n(R) \cong \mathrm{GL}_n(R_1) \times \mathrm{GL}_n(R_2)$ via the map $(m^1_{ij}, m^2_{ij}) \mapsto ((m^1_{ij}), (m^2_{ij}))$, therefore this yields an isomorphism $\mathrm{GL}(R) \cong \mathrm{GL}(R_1) \times \mathrm{GL}(R_2)$. Moreover, under the same isomorphism, it can be checked that $[\mathrm{GL}(R) : \mathrm{GL}(R)] \cong [\mathrm{GL}(R_1) : \mathrm{GL}(R_1)] \times [\mathrm{GL}(R_2) : \mathrm{GL}(R_2)]$. This completes the proof. $\square$

We next see that K_1 of a ring and matrix ring are equivalent.

Proposition 17.2.1.4. *Let R be a ring and $n \in \mathbb{N}$. Then*

$$K_1(R) \cong K_1(M_n(R)).$$

Proof. To this end, it suffices to show that $\mathrm{GL}(R) \cong \mathrm{GL}(M_n(R))$. We do this via basic Morita theory. We know that R and $M_n(R)$ are Morita equivalent, where the Morita functors are

$$\begin{aligned} \mathbf{Mod}(R) &\leftrightarrow \mathbf{Mod}(M_n(R)) \\ M &\mapsto M \otimes_R R^n \\ N \otimes_{M_n(R)} R^n &\leftarrow\!\shortmid N \end{aligned}$$

where $M \otimes_R R^n \cong M^n$ is an $M_n(R)$-module by matrix multiplication: $(r_{ij}) \cdot (m_1, \dots, m_n) = \left(\sum_j r_{1j} m_j, \dots, \sum_j r_{nj} m_j\right)$. Similarly, $N \otimes_{M_n(R)} R^n$ is an R-module. The equivalence thus takes the $M_n(R)$-module to the R-module $M_n(R)^m$ to $M_n(R)^m \otimes_{M_n(R)} R^n \cong R^{nm} \cong R^n \otimes_R R^m$. Consequently, the equivalence of categories gives an isomorphism of the endomorphism groups $M_{mn}(R) \cong \mathrm{End}_R(R^n \otimes_R R^m) \cong \mathrm{End}_{M_n(R)}(M_n(R)^m) = M_m(M_n(R))$. Thus, the group of units are also isomorphic, yielding

$$\mathrm{GL}_{mn}(R) \cong \mathrm{GL}_m(M_n(R)).$$

As the multiples of n are cofinal in $\mathbb{N}$, thus we get an isomorphism in the direct limit

$$\mathrm{GL}(R) \cong \mathrm{GL}(M_n(R)).$$

Thus the commutators are also isomorphic and the above isomorphism thus extends to the isomorphism $K_1(R) \cong K_1(M_n(R))$. $\square$

We will later see that actually K_1 is a Morita invariant.

Remark 17.2.1.5 (Information in $K_1(R)$)**.** While $K_0(R)$ tells us about the spread of f.g. projective R-modules, $K_1(R)$ tells us about the eventual spread of automorphisms of f.g. projective modules. The latter may not be evident as of now, but the next few results will showcase exactly this, as will be seen by Whitehead's lemma and Bass' theorem.

Our first task is to show that the commutator of $\mathrm{GL}(R)$ is a familiar object.

Definition 17.2.1.6 ($E(R)$)**.** Consider the subgroup of $\mathrm{GL}_n(R)$ denoted $E_n(R)$ generated by elementary n-matrices of type 1, that is, the invertible matrices $e_{ij}(r) \in \mathrm{GL}_n(R)$ where $i \neq j$ such that their diagonal is all 1, $e_{ij} = r$ at (i, j) entry and 0 in the rest. Thus we have an injection

$$E_n(R) \hookrightarrow \mathrm{GL}_n(R).$$

Taking direct limits, we get an injection

$$E(R) \hookrightarrow \mathrm{GL}(R).$$

A simple exercise shows that $E_n(F) = \mathrm{SL}_n(F)$ for F a field.

Lemma 17.2.1.7. *Let R be an Euclidean domain. Then*

$$E_n(R) = \mathrm{SL}_n(R).$$

Consequently, $E(R) = \mathrm{SL}(R)$ in $\mathrm{GL}(R)$.

Proof. We proceed by induction on $n \in \mathbb{N}$. Indeed, for $n = 1$, we have $E_1(R) = \mathrm{SL}_1(R)$. For the inductive step, fix $n \in \mathbb{N}$. We wish to show that $\mathrm{SL}_n(R) \subseteq E_n(R)$. Pick $g \in \mathrm{SL}_n(R)$ and write

$$g = \begin{bmatrix} g_{11} & g_{12} & \cdots \\ g_{21} & g_{22} & \cdots \\ \vdots & & \end{bmatrix}.$$

As R is a Euclidean domain, therefore by Euclidean division, we get

$$\begin{aligned} g_{21} &= g_{11}q_1 + r_1 \\ g_{11} &= r_1 q_2 + r_2 \\ r_1 &= r_2 q_3 + r_3 \\ \vdots &= \vdots \\ r_{k-1} &= r_k q_{k+1} \end{aligned}$$

Note all the above operations can be done by type 1 row operations. Consequently, by doing these row and column operations we can reduce g to

$$ege' = \begin{bmatrix} g'_{11} & 0 & \cdots \\ 0 & g'_{22} & \cdots \\ \vdots & & \end{bmatrix} = \begin{bmatrix} g'_{11} & 0 \\ 0 & M_{n-1} \end{bmatrix}$$

Now, by inductive hypothesis, there exists $e_{n-1}, e'_{n-1} \in E_{n-1}(R) \hookrightarrow E_n(R)$ such that $e_{n-1}M_{n-1}e'_{n-1} = 1_{n-1}$. Taking the image of e_{n-1}, e'_{n-1} in $E_n(R)$, we get $e_n, e'_n \in E_n(R)$ such that

$$e_n ege' e'_n = \begin{bmatrix} g'_{11} & 0 \\ 0 & I_{n-1} \end{bmatrix}.$$

As $\det(e_n ege' e'_n) = 1$, therefore we deduce that $g'_{11} = 1$, and thus

$$e_n ege' e'_n = I_n,$$

hence $g = e^{-1}e_n^{-1}e_n'^{-1}e'^{-1} \in E_n(R)$, as required. This completes the proof. $\square$

Theorem 17.2.1.8 (Whitehead's lemma)**.** *Let R be a ring. Then $E(R) = [\mathrm{GL}(R), \mathrm{GL}(R)]$. Consequently,*

$$K_1(R) = \frac{\mathrm{GL}(R)}{E(R)}.$$

Remark 17.2.1.9. It follows that $K_1(R)$ consists of all classes of all infinite invertible matrices which are not similar to each other by type 1 elementary matrices. Recall that two matrices $g, h \in M_n(R)$ are similar to each other if there exists e, e' an elementary matrix (not necessarily only of type 1) such that $g = e'he$. Thus, it is in this sense does $K_1(R)$ measures the eventual spread of automorphisms of free modules. In the next section, we will see that more is true.

Remark 17.2.1.10 ($E(R)$ is perfect in $\mathrm{GL}(R)$)**.** Observe the following three basic relations for elements in $E_n(R)$ for $n \geq 3$:

$$\begin{aligned}
&e_{ij}^{(n)}(\lambda) \cdot e_{ij}^{(n)}(\mu) = e_{ij}^{(n)}(\lambda + \mu) \\
&[e_{ij}^{(n)}(\lambda), e_{kl}^{(n)}(\mu)] = 1, \quad i \neq l,\ j \neq k \\
&[e_{ij}^{(n)}(\lambda), e_{jk}^{(n)}(\mu)] = e_{ik}^{(n)}(\lambda\mu), \quad i \neq k.
\end{aligned}$$

The last relation combined with Theorem 17.2.1.8 immediately tells us that $E_n(R)$ is perfect and hence so is $E(R)$.

Basic computations

Using Whitehead's lemma and Dieudonné's determinant, we can do some basic computations.

Lemma 17.2.1.11. *Let R be an Euclidean domain. Then,*

$$K_1(R) \cong R^{\times}.$$

Proof. Using Dieudonné's determinant, we have a surjective map

$$\det : \mathrm{GL}(R) \twoheadrightarrow R^{\times}.$$

As $\mathrm{Ker}\,(\det) = \mathrm{SL}(R)$ and by Lemma 17.2.1.7, $\mathrm{SL}(R) = E(R)$, thus by Whitehead's lemma (Theorem 17.2.1.8), we conclude that $K_1(R) \cong \mathrm{GL}(R)/\mathrm{Ker}\,(\det) \cong R^{\times}$. □

Corollary 17.2.1.12. *Let F be a field. Then $K_1(F) = F^{\times}$.* □

Proposition 17.2.1.13. *Let D be a division ring (a non-commutative field). Then,*

$$K_1(D) \cong \frac{D^{\times}}{[D^{\times}, D^{\times}]}.$$

Proof. This follows from Dieudonné's theorem that for a division ring,

$$\frac{\mathrm{GL}_n(D)}{E_n(D)} \cong \frac{D^{\times}}{[D^{\times}, D^{\times}]}$$

for $n \geq 3$. This induces the required isomorphism. □

Proposition 17.2.1.14. *Let R be a semilocal ring. Then $SK_1(R) = 1$ and thus*

$$K_1(R) \cong R^\times.$$

Proof. See Lemma 1.4, pp 202 of **[WeibK]**. □

Corollary 17.2.1.15. *Let R be an Artinian ring. Then, $K_1(R) \cong R^\times$.* □

17.2.2 More properties of K_1

We study some properties of $K_1(R)$ which depends on R-modules. We begin by relating group homology and $K_1(R)$.

Characteristic map & Bass' result

We now see how $K_1(R)$ is actually the eventual spread of automorphisms of f.g. projective R-modules. We first show that every $P \in \mathbf{Proj}(R)$ yields a group homomorphism $\mathrm{Aut}\,(P) \to K_1(R)$.

Construction 17.2.2.1 ($\chi_P : \mathrm{Aut}\,(P) \to K_1(R)$)**.** Let P be a f.g. projective R-module. We construct a characteristic map

$$\chi_P : \mathrm{Aut}\,(P) \longrightarrow K_1(R).$$

For each $Q \in \mathbf{Proj}(R)$ and $n \in \mathbb{N}$ such that $P \oplus Q \cong R^n$ via a map $\theta : P \oplus Q \to R^n$, we get a map

$$\begin{aligned} \chi_{\theta,Q,n} : \mathrm{Aut}\,(P) &\longrightarrow K_1(R) \\ \varphi &\longmapsto \overline{[\theta \circ (\varphi \oplus \mathrm{id}_Q) \circ \theta^{-1}]} \end{aligned}$$

We first show that $\chi_{\theta,Q,n}$ is independent of θ if n and Q are fixed. Indeed, it is immediate to see that if we take $\theta' : P \oplus Q \to R^n$ a different isomorphism, then $\chi_{\theta,Q,n}$ and $\chi_{\theta',Q,n}$ are conjugates by an $\theta \circ \theta'^{-1} \in \mathrm{GL}_n(R)$. Thus, in $K_1(R)$, $\chi_{\theta,Q,n}$ determines a unique class for each $\varphi \in \mathrm{Aut}\,(P)$, independent of θ. We may thus write

$$\chi_{Q,n} : \mathrm{Aut}\,(P) \longrightarrow K_1(R).$$

Next we show that $\chi_{Q\oplus R^k,n+k} = \chi_{Q,n}$, that is, χ is stable. Indeed, this is immediate as the image of the above maps factor through $\mathrm{GL}(R)$.

Now suppose we have an isomorphism

$$P \oplus Q' \cong R^m$$

where $m > n$. As $P \oplus Q \cong R^n$, so we get $P \oplus (Q \oplus R^m) \cong R^{m+n} \cong P \oplus (Q' \oplus R^n)$. Thus we get

$$Q \oplus R^{m+k} \cong Q' \oplus R^{n+k}.$$

Hence by previous, we have $\chi_{Q,n} = \chi_{Q\oplus R^{m+k},m+n+k} = \chi_{Q'\oplus R^{n+k},m+n+k} = \chi_{Q',m}$, as required.

Definition 17.2.2.2 (Characteristic map). Let R be a ring and $P \in \mathbf{Proj}(R)$. Then the map χ_P : $\mathrm{Aut}\,(P) \to K_1(R)$ constructed in Construction 17.2.2.1 will be called the characteristic map of P.

We now state a classical theorem of Bass, which shows the real intent behind defining $K_1(R)$.

Theorem 17.2.2.3 (Bass)**.** *Let R be a ring and denote by $t\mathbf{Proj}(R)^{\cong}$ to be the filtered category whose objects are isomorphism classes of f.g. projective R-modules and an arrow $P \to P'$ is an isomorphism class of Q such that $P \oplus Q \cong P'$. Then,*

$$K_1(R) \cong \varinjlim_{P \in t\mathbf{Proj}(R)^{\cong}} \frac{\mathrm{Aut}\,(P)}{[\mathrm{Aut}\,(P), \mathrm{Aut}\,(P)]}.$$

Proof. Note that the mapping $P \mapsto \frac{\mathrm{Aut}(P)}{[\mathrm{Aut}(P),\mathrm{Aut}(P)]}$ is a functor $t\mathbf{Proj}(R)^{\cong} \to \mathbf{Ab}$ where for $P \to P'$ given by $\theta : P \oplus Q \stackrel{\cong}{\to} P'$, the functor maps it to the mapping

$$\begin{aligned} \frac{\mathrm{Aut}\,(P)}{[\mathrm{Aut}\,(P), \mathrm{Aut}\,(P)]} &\longrightarrow \frac{\mathrm{Aut}\,(P')}{[\mathrm{Aut}\,(P'), \mathrm{Aut}\,(P')]} \\ \bar{\varphi} &\longmapsto \overline{\theta \circ (\varphi \oplus \mathrm{id}_Q) \circ \theta^{-1}}. \end{aligned}$$

It can be shown that $t\mathbf{Proj}(R)^{\cong}$ is a filtered category[5]. As we have

$$K_1(R) = \frac{\mathrm{GL}(R)}{[\mathrm{GL}(R), \mathrm{GL}(R)]} \cong \varinjlim_{n} \frac{\mathrm{GL}_n(R)}{[\mathrm{GL}_n(R), \mathrm{GL}_n(R)]}$$

and free modules are cofinal in the filtered category $t\mathbf{Proj}(R)^{\cong}$, thus we get that

$$\varinjlim_{n} \frac{\mathrm{GL}_n(R)}{[\mathrm{GL}_n(R), \mathrm{GL}_n(R)]} \cong \varinjlim_{P \in t\mathbf{Proj}(R)^{\cong}} \frac{\mathrm{Aut}\,(P)}{[\mathrm{Aut}\,(P), \mathrm{Aut}\,(P)]},$$

as required. $\square$

Remark 17.2.2.4. The proof of Theorem 17.2.2.3 shows that for any f.g. projective module P, the characteristic map

$$\chi_P : \mathrm{Aut}\,(P) \to K_1(R)$$

are the maps into the filtered limit $\varinjlim_P \frac{\mathrm{Aut}(P)}{[\mathrm{Aut}(P),\mathrm{Aut}(P)]}$. Indeed, this is evident from the functor $P \mapsto \frac{\mathrm{Aut}(P)}{[\mathrm{Aut}(P),\mathrm{Aut}(P)]}$ and Construction 17.2.2.1. In particular the following diagram commutes:

$$\begin{tikzcd} K_1(R) & \varinjlim_{Q \in t\mathbf{Proj}(R)^{\cong}} \frac{\mathrm{Aut}(Q)}{[\mathrm{Aut}(Q),\mathrm{Aut}(Q)]} \arrow[l, "\cong"'] \\ \mathrm{Aut}\,(P) \arrow[u, "\chi_P"] \arrow[r, two heads, "\pi"'] & \frac{\mathrm{Aut}(P)}{[\mathrm{Aut}(P),\mathrm{Aut}(P)]} \arrow[u, "\iota_P"'] \end{tikzcd}.$$

[5]For two objects P, P', we have $P \oplus Q \cong R^n$ and $P' \oplus Q' \cong R^m$. Consequently, we have $P \oplus Q \oplus R^m \cong R^{n+m}$ and $P' \oplus Q' \oplus R^n \cong R^{n+m}$. For existence of coequalizers, observe that for $P \oplus Q_1 \cong P' \cong P \oplus Q_2$, we get $[Q_1] = [Q_2]$ in $K_0(R)$. By Lemma 17.1.1.2, we deduce that $Q_1 \oplus R^n \cong Q_2 \oplus R^n$. Thus, the map

$$P' \xrightarrow{R^n} P' \oplus R^n$$

coequalizes both Q_1 and Q_2, as required.

Generators for $SK_1(R)$

Recall by Construction 17.2.1.2 that we have a decomposition

$$K_1(R) \cong R^\times \oplus SK_1(R)$$

where $SK_1(R) = \mathrm{Ker}\,(\det : K_1(R) \to R^\times)$. Hence, to understand $K_1(R)$, it is sufficient to understand the subgroup $SK_1(R)$. Indeed, in certain cases on R, there is a class of elements of $SK_1(R)$ which is known to be its generating set.

Definition 17.2.2.5 (**Mennicke symbols**)**.** Let R be a commutative ring. A Mennicke symbol is an element of $SK_1(R)$ which is determined by following. Consider $a, b \in R$ such that $aR + bR = R$. Thus there exists $c, d \in R$ such that $ad - bc = 1$. The matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(R)$$

is such that its class in $K_1(R)$ is in $\mathrm{Ker}\,(\det : K_1(R) \to R^\times) = SK_1(R)$. Thus we denote the class of the above matrix in $SK_1(R)$ as $[a, b]$, which we call the Mennicke symbol correspoding to a, b.

The obvious observation here is that Mennicke symbol $[a, b]$ doesn't depend on c, d.

Lemma 17.2.2.6. *Let R be a commutative ring and let $aR + bR = R$ such that $ad - bc = 1$. Then the Mennicke symbol $[a, b]$ is independent of $c, d \in R$.*

Proof. Indeed, we need only observe that if $ad - bc = 1 = ad' - bc'$, then we have

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c' & d' \end{bmatrix}^{-1} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d' & -b \\ -c' & a \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ cd' - dc' & 1 \end{bmatrix} \in E_2(R) \end{aligned}$$

and thus the class of

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c' & d' \end{bmatrix}^{-1}$$

in $K_1(R)$ is contained in $E(R) = [\mathrm{GL}(R), \mathrm{GL}(R)]$ (Theorem 17.2.1.8). Hence in $K_1(R)$ both the classes are same and is in $SK_1(R)$. $\square$

The main point of Mennicke symbols is the following result which clarifies the dimension 1 case.

Theorem 17.2.2.7. *Let R be a commutative noetherian ring of dimension 1.*

1. *Mennicke symbols generates $SK_1(R)$.*
2. *If $\kappa(\mathfrak{m})$ is a finite field for all $\mathfrak{m} \in \mathrm{mSpec}\,(R)$, then every element of $SK_1(R)$ has torsion.*

A famous result calculates the $SK_1(\mathcal{O}_F)$ for a number field $F/\mathbb{Q}$.

Theorem 17.2.2.8 (Serre, Milnor, Bass)**.** *Let $K/\mathbb{Q}$ be a number field. Then,*

$$SK_1(\mathcal{O}_F) = 0.$$

17.3 K_2

We discuss some basic results about K_2, relegating the more homological discussions about it to higher K-theory.

Definition 17.3.0.1 (Steinberg group & K_2). Recall that the n^{th}-Steinberg group $\mathrm{St}_n(R)$ is the quotient of the free group generated by the symbols $x_{ij}^{(n)}(\lambda)$, $\lambda \in R$ and $1 \leq i \neq j \leq n$ by the subgroup generated by the known relations which elementary $n \times n$-matrices of type 1 satisfies:

$$\begin{aligned}
&x_{ij}^{(n)}(\lambda) \cdot x_{ij}^{(n)}(\mu) \cdot x_{ij}^{(n)}(\lambda+\mu)^{-1} \\
&[x_{ij}^{(n)}(\lambda), x_{kl}^{(n)}(\mu)], \quad i \neq l,\ j \neq k \\
&[x_{ij}^{(n)}(\lambda), x_{jk}^{(n)}(\mu)] \cdot x_{ik}^{(n)}(\lambda\mu)^{-1}, \quad i \neq k.
\end{aligned}$$

We call these the Steinberg relations. Consider the group homomorphism for each $n \in \mathbb{N}$

$$\begin{aligned}
\mathrm{St}_n(R) &\longrightarrow \mathrm{St}_{n+1}(R) \\
\overline{x_{ij}^{(n)}(\lambda)} &\longmapsto \overline{x_{ij}^{(n+1)}(\lambda)}.
\end{aligned}$$

The Steinberg group is defined to be the direct limit

$$\mathrm{St}(R) = \varinjlim_n \mathrm{St}_n(R)$$

where we denote the class of $x_{ij}^{(n)}(\lambda)$ as $x_{ij}(\lambda)$. As for each $n \in \mathbb{N}$, we have a surjective group homomorphism

$$\begin{aligned}
\phi_n : \mathrm{St}_n(R) &\longrightarrow E_n(R) \\
\overline{x_{ij}^{(n)}(\lambda)} &\longmapsto e_{ij}^{(n)}(\lambda),
\end{aligned}$$

thus we get a unique surjective map

$$\phi : \mathrm{St}(R) \longrightarrow E(R).$$

We thus define

$$K_2(R) := \mathrm{Ker}\,(\phi : \mathrm{St}(R) \to E(R)).$$

Thus, we have

$$E(R) \cong \frac{\mathrm{St}(R)}{K_2(R)}.$$

17.3.1 Central extensions & $K_2(R)$

Our goal is to show the following theorem, which, amongst other things, says that $K_2(R)$ is an abelian group.

Theorem 17.3.1.1. *Let R be a ring. Then the extension*

$$1 \to K_2(R) \to \mathrm{St}(R) \xrightarrow{\phi} E(R) \to 1$$

exhibits $\mathrm{St}(R)$ *as a universal central extension of $E(R)$. Moreover, $K_2(R)$ is the center of* $\mathrm{St}(R)$*.*

Proof. In order to show that this extension is central, we first need to show that $K_2(R)$ is in the center of $\mathrm{St}(R)$. In-fact we see that $K_2(R)$ is the center of $\mathrm{St}(R)$. Indeed, as the center of $E(R)$ is 1 and the center of $\mathrm{St}(R)$ is contained in the inverse image of the center of $E(R)$, thus center of $\mathrm{St}(R) \subseteq K_2(R)$. Conversely, we need to show that $K_2(R)$ is in center of $\mathrm{St}(R)$.

Pick an element $\alpha \in K_2(R)$. We wish to show that every element of $\mathrm{St}(R)$ commutes with α. As every element of $\mathrm{St}(R)$ is a word in elements $x_{ij}(\lambda)$, it suffices to show that $[\alpha, x_{ij}(\lambda)] = 1$. Write $x_{ij}(\lambda)$ as the class of some element $x_{ij}^{(n)}(\lambda) \in \mathrm{St}_n(R)$. By the third relation of $\mathrm{St}_n(R)$, we can write

$$x_{ij}^{(n)}(\lambda) = [x_{in}^{(n)}(\lambda), x_{nj}^{(n)}(1)]$$

in $\mathrm{St}_n(R)$. A little algebra makes it clear that it is sufficient to show that $[\alpha, x_{in}(\lambda)] = 1 = [\alpha, x_{nj}(\lambda)]$. We hence reduce to showing that

$$[\alpha, x_{in}(\lambda)] = 1$$

for any $1 \leq i \leq n-1$, $n \in \mathbb{N}$ and $\lambda \in R$.

We now exploit the fact that $\alpha \in K_2(R) = \mathrm{Ker}\,(\phi)$ by using the map $\phi : \mathrm{St}(R) \to E(R)$. Consider the subgroup

$$G_n = \langle x_{in}(\lambda) \mid 1 \leq i \leq n-1,\ \lambda \in R \rangle \leq \mathrm{St}(R).$$

We wish to understand the structure of G_n. Pick two elements $x_{in}(\lambda), x_{jn}(\mu) \in G_n$. Observe that as $i \neq n$ and $n \neq j$, we deduce by second relation that

$$x_{in}(\lambda) \cdot x_{jn}(\mu) = x_{jn}(\mu) \cdot x_{in}(\lambda),$$

that is, G_n is commutative. Hence, any arbitrary element of G_n is of the form $\beta = x_{1n}(\lambda_1) \cdot \dots \cdot x_{n-1n}(\lambda_{n-1})$.

We thus see that $\phi|_{G_n} : G_n \to E(R)$ is injective as if $\beta \in G_n$ as above goes to 1 in $E(R)$, then

$$e_{1n}(\lambda_1) \cdot \dots \cdot e_{n-1n}(\lambda_{n-1}) = \begin{bmatrix} 1 & 0 & \cdots & \lambda_1 \\ 0 & 1 & \cdots & \lambda_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = 1.$$

It follows that $\lambda_i = 0$ and thus $x_{in}(\lambda_i) = 1$, which further implies that $\beta = 1$, as required. Now as shown above, the group G_n is commutative and $\phi|_{G_n}$ is injective. We first claim that α is a normalizer of G_n, that is $\alpha G_n = G_n \alpha$. Indeed, α can be expressed as a product $x_{ij}(\mu)$. Using the Steinberg relations, we first easily see that (recall that $x_{ij}(\mu)^{-1} = x_{ij}(-\mu)$)

$$x_{ij}(\mu) \cdot x_{kn}(\lambda) = \begin{cases} x_{kn}(\lambda) x_{ij}(\mu) & \text{if } j \neq k \\ x_{in}(\lambda\mu) x_{kn}(\lambda) x_{ij}(\mu) & \text{if } j = k. \end{cases}$$

Then, for any $\beta \in G_n$, we have $\alpha \cdot \beta = \beta' \cdot \alpha$ for some $\beta' \in G_n$. Applying the injective map $\varphi|_{G_n}$, we see that $1 \cdot \varphi(\beta) = \varphi(\beta') \cdot 1$, thus $\beta = \beta'$. This shows that α is actually a centralizer of G_n, thus, every element of G_n commutes with α, as required. This shows that the extension is central and $K_2(R)$ is the center of $\mathrm{St}(R)$.

We now show that the extension is moreover universal central. By Theorem 6.9.7 of **[WeibHA]** (Recognition principle), it suffices to prove that $\mathrm{St}(R)$ is perfect and every central extension of $\mathrm{St}(R)$ splits. The fact that $\mathrm{St}(R)$ is perfect follows immediately from the third Steinberg relation. Thus we reduce to showing that every central extension of $\mathrm{St}(R)$ splits. This is the major part of the proof of Theorem 5.10 of **[MilnKTh]**. □

Corollary 17.3.1.2. *Let R be a ring. Then*

$$K_2(R) \cong H_2(E(R); \mathbb{Z}).$$

Proof. Follows from Theorem 17.3.1.1 and Theorem 6.9.5 of **[WeibHA]**. □

17.3.2 G-representations of $K_2(F)$

Observe that $K_2(R)$ is an abelian group since by Theorem 17.3.1.1 it is the center of $\mathrm{St}(R)$. Hence, one may try to find a presentation of $K_2(R)$ for any ring R. Matsumoto gives us one such presentation for K_2 of a field. This has intricate connections to algebra.

Theorem 17.3.2.1 (Matsumoto)**.** *Let F be a field. There is a presentation of $K_2(F)$ (multiplicatively written) as in*

$$1 \to \mathrm{Mats}(F) \to \mathbb{Z}^{\oplus F^\times \times F^\times} \to K_2(F) \to 1$$

where $\mathrm{Mats}(F)$ *is the subgroup of* $\mathbb{Z}^{\oplus F^\times \times F^\times}$ *generated by the following relations for* $a_i, b_i, a, b \in F^\times$

$$\begin{aligned}(a_1a_2, b) &= (a_1, b) \cdot (a_1, b)\\ (a, b) &= (b, a)^{-1}\\ (a, 1-a) &= 1,\ a \neq 1.\end{aligned}$$

We denote the class of (a, b) in $K_2(F)$ by $\{a, b\}$. We call the above relations Matsumoto's relations. Thus $\{a, b\}$ satisfies Matsumoto's relations in $K_2(F)$.

The important technique here is that Matsumoto's theorem allows us to give an equivalence between G-representations of $K_2(F)$ and functions $F^\times \times F^\times \to G$ of certain type, for any abelian group G.

Definition 17.3.2.2 (Symbols on a field). Let F be a field and G be an abelian group. A G-valued symbol over F is a G-valued function $F^\times \times F^\times$, denoted

$$(,) : F^\times \times F^\times \to G$$

which satisfies the Matsumoto relations verbatim as stated in Theorem 17.3.2.1. We denote by $\mathrm{Symb}_G(F)$ the set of all symbols of F over G.

Here's the equivalence.

Theorem 17.3.2.3. *Let F be a field and G be an abelian group. Then there is a bijection*

$$\mathrm{Hom}_{\mathbf{Ab}}(K_2(F), G) \cong \mathrm{Symb}_G(F).$$

Proof. Define the bijection as

$$\begin{aligned}\mathrm{Hom}_{\mathbf{Ab}}(K_2(F), G) &\longleftrightarrow \mathrm{Symb}_G(F)\\ \varphi &\longmapsto (a,b) := \varphi(\{a,b\})\\ \tilde{\varphi} &\longleftarrow (,)_s\end{aligned}$$

where $\tilde{\varphi} : K_2(F) \to G$ is obtained by extending the group homomorphism $\mathbb{Z}^{\oplus F^\times \times F^\times} \to G$ given by $(a,b) \mapsto (a,b)_s$ to the unique map $\tilde{\varphi} : K_2(F) \to G$ via universal property of cokernels ($(,)_s$ satsifies Matsumoto's relations) in the presentation of Theorem 17.3.2.1. $\square$

Using this we can immediately find K_2 of a finite field.

Proposition 17.3.2.4. *Let $\mathbb{F}_q$ be a finite field of characteristic $p > 0$. Then*

$$K_2(\mathbb{F}_q) = 1.$$

Proof. By Theorem 17.3.2.3 and Yoneda's lemma, it suffices to show that for any abelian group G, $\mathrm{Symb}_G(\mathbb{F}_q) = 1$. Let $(,)_s \in \mathrm{Symb}_G(\mathbb{F}_q)$. As $\mathbb{F}_q^\times$ is a finite cyclic group, therefore let $g \in \mathbb{F}_q^\times$ be its generator. Then, any element in $\mathbb{F}_q^\times \times \mathbb{F}_q^\times$ is of the form (g^n, g^m). By Matsumoto's relations we deduce that $(g^n, g^m)_s = (g, g^m)_s^n = (g^m, g)_s^{-n} = (g,g)_s^{-nm}$. Moreover, as $(g,g) = (g,g)^{-1}$, therefore (g,g) is an element of order atmost 2 in G. Let us assume that order of (g,g) is precisely 2.

Note that $(-,g) : \mathbb{F}_q^\times \to G$ is a group homomorphism by first Matsumoto's relation. Thus $1 = (1,g) = (g^{|g|}, g) = (g,g)^{|g|}$, so we conclude that $2 \mid |g| = q-1$. If $p = 2$, then $q-1$ is odd and thus we have a contradiction. Thus, we may assume $p > 2$.

We wish to show that $(g,g) = 1$. To this end, we use the third Matsumoto's relation. Note that it suffices to show that for some $k \in \mathbb{N}$ odd, we have $(g,g)^k = 1$. Writing $k = nm$, we see that we need $(g^n, g^m) = 1$. Thus we reduce to showing that there exists $h \in \mathbb{F}_q^\times - 1$ such that $h = g^n$, $1 - h = g^m$ and n, m are odd. To this end, observe that

$$\mathbb{F}_q^\times - 1 = \{g, g^2, g^3, \dots, g^{q-2}\}.$$

Thus there are $(q-3)/2$ squares of g and $(q-1)/2$ non-squares of g in $\mathbb{F}_q^\times - 1$. Next, observe that the map

$$\begin{aligned}\mathbb{F}_q^\times - 1 &\longrightarrow \mathbb{F}_q^\times - 1\\ h &\longmapsto 1 - h\end{aligned}$$

is an injective map and since $\mathbb{F}_q^\times$ is finite, thus is a bijection. Thus, if for all $h \in \mathbb{F}_q^\times - 1$ which is non-square, $1-h$ is a square, then by the above bijection, there are atleast $(q-1)/2$ squares, which is more than $(q-3)/2$, a contradiction. Hence there is h such that it is a non-square and $1-h$ is also a non-square, as required. This completes the proof. $\square$

17.3.3 p-divisibility of $K_2(F)$

Recall that a multiplicatively written abelian group G is a *p-divisible* group if for all $g \in G$, there is $h \in G$ such that $h^p = g$. It is *uniquely p-divisible* if G moreover has no p-torsion element, i.e. no non-trivial element of G has order p. It is *divisible* if G is p-divisible for each prime p. Finally G is *uniquely divisible* if it is uniquely p-divisible for each prime p.

Observe that a uniquely p-divisible group G has a unique p^{th}-root for every element in G. Thus, a uniquely divisible group G has unique p^{th}-root for every element and for every prime p. We would like to show the following result.

Theorem 17.3.3.1 (Bass-Tate)**.** *Let F be a field and p be a prime such that every polynomial $x^p - a$, $a \in F$ splits in $F[x]$ into linear factors. Then $K_2(F)$ is uniquely p-divisible.*

Proof. We first have to show that $K_2(F)$ is a p-divisible group. As the image of a p-divisible group is p-divisible, it suffices to find a p-divisible group which surjects on $K_2(F)$. Observe that $F^\times$ is a p-divisible group by the given hypothesis. It can be checked easily that tensor product of two uniquely p-divisible groups is unqiuely p-divisible[6]. By Matsumoto's theorem (Theorem 17.3.2.1), we have a surjective homomorphism

$$\varphi : F^\times \otimes_{\mathbb{Z}} F^\times \to K_2(F).$$

This shows that $K_2(F)$ is p-divisible. We need only show that $K_2(F)$ has no p-torsion. To this end, we first relate the unique p-divisibility of $F^\times \otimes_{\mathbb{Z}} F^\times$ to $K_2(F)$. Note that we have the following commutative diagram whose rows are exact:

$$\begin{tikzcd}
 & 0 \arrow[r] \arrow[d] & F^\times \otimes_{\mathbb{Z}} F^\times \arrow[r, "m_p"] \arrow[d, "\varphi"] & F^\times \otimes_{\mathbb{Z}} F^\times \arrow[r] \arrow[d, "\varphi"] & 0 \\
0 \arrow[r] & T_p(K_2(F)) \arrow[r] & K_2(F) \arrow[r, "m_p"] & K_2(F) &
\end{tikzcd}.$$

By Snake lemma, we deduce the following exact sequence

$$0 \to \mathrm{Ker}\,(\varphi) \stackrel{m_p}{\to} \mathrm{Ker}\,(\varphi) \to T_p(K_2(F)) \to 0 \to 0.$$

Thus, we have

$$T_p(K_2(F)) \cong \mathrm{Ker}\,(\varphi)/\mathrm{Im}\,(m_p).$$

We wish to show that $T_p(K_2(F)) = 1$. To this end, it suffices to thus show that $\mathrm{Ker}\,(\varphi)$ is p-divisible.

Note that $\mathrm{Ker}\,(\varphi)$ is generated by elements of form $a \otimes 1 - a$ for $a \in F^\times$ by Matsumoto's theorem (the other relations are trivially satisfied by $\mathrm{Ker}\,(\varphi)$ as it is a subgroup of $F^\times \otimes_{\mathbb{Z}} F^\times$). Thus, it suffices to show that for any $a \in F^\times$, the element $a \otimes (1 - a)$ is a multiple of p of some other element in $\mathrm{Ker}\,(\varphi)$. By hypothesis, we have

$$x^p - a = \prod_i (x - b_i)$$

---

[6]Let G, H be two uniquely p-divisible groups. Consider the isomorphism $m_p : G \to G$. Tensoring $0 \to 0 \to G \stackrel{m_p}{\to} G \to 0$ by H, we get $0 \otimes H \to G \otimes H \stackrel{m_p \otimes \mathrm{id}}{\to} G \otimes H \to 0$ is exact. As $0 \otimes H = 0$, thus $m_p \otimes \mathrm{id}$ is an isomorphism.

for $b_i \in F^\times$. It follows that $1 - a = \prod_i (1 - b_i)$ and since the notation for tensor product is additive, so we get

$$a \otimes (1-a) = \sum_i a \otimes (1-b_i) = \sum_i b_j^p \otimes (1-b_i) = p \sum_i b_j \otimes (1-b_i).$$

As $\sum_i b_j \otimes (1-b_i) \in \mathrm{Ker}\,(\varphi)$ as it is a sum of generators of it, hence this completes the proof. □

We have two immediate corollaries.

Corollary 17.3.3.2. *Let F be a field.*

1. *If F is algebraically closed, then $K_2(F)$ is uniquely divisible.*
2. *If F is perfect of characteristic $p > 0$, then $K_2(F)$ is uniquely p-divisible.*

Proof. Both follows immediately from Theorem 17.3.3.1. For the latter, observe that by perfection, for each $a \in F^\times$, there is a $b \in F^\times$ such that $b^p = a$ and thus $x^p - a = (x-b)^p$ by algebra in characteristic $p > 0$. □

17.3.4 Brauer group & Galois symbol

Brauer group is another subtle invariant of a field. Its main uses are in a) algebraic geometry, where it classifies certain type of projective varieties over a field, and in b) algebraic number theory where it is used to construct the Hasse invariant.

We will here construct a representation of $K_2(F)$ for certain fields in the Brauer group of F, denoted $\mathrm{Br}(F)$. For that, we first study a generalization of matrix algebras, *central simple algebras* (CSA) over F. Recall A is a CSA over F if A is a finite dimensional associative unital algebra over F which is simple and whose center is exactly F. The following basic observations is all we require of them.

Proposition 17.3.4.1. *Let F be a field.*

1. *$M_n(F)$ is a central simple algebra over F.*
2. *If A, B are two central simple algebras over F, then so is $A \otimes_F B$.*

Proof. 1. The center of $M_n(F)$ is the diagonal matrices $\mathrm{diag}(\lambda, \lambda, \dots, \lambda)$ for $\lambda \in F^\times$. This is immediate by considering elementary matrices. The fact that $M_n(F)$ is simple can be seen by Smith normal form and performing elementary row operations together amongst others to complete the diagonal with non-zero entries.

2. Observe that the map $K \to A \otimes_K B$ mapping $k \mapsto k \otimes 1$ is injective. We first show centrality of $A \otimes_F B$. Let Z be the center of $A \otimes_F B$ and $x = \sum_i a_i \otimes b_i \in Z$. We wish to show that $x \in K$. Indeed, consider $a \in A$. We have $(a \otimes 1) \cdot x = x \cdot (a \otimes 1)$. This yields

$$\sum_i (a_i a - a a_i) \otimes b_i = 0.$$

Write $a_i a - a a_i = \sum_k c_k m_k$ where $\{m_k\}$ forms a basis of A. We may assume that $\{b_i\}$ is a basis of B. It follows that $\sum_{i,k} c_k (m_k \otimes b_i) = 0$. As $\{m_k \otimes b_i\}_{k,i}$ forms a K-basis of $A \otimes_K B$, therefore $c_k = 0$ for all k. It follows at once that $a_i a = a a_i$. Hence $a_i \in K$ for all i. Similarly, $b_i \in K$ for all i. This shows that $x \in K$.

Next we show simplicity of $A \otimes_F B$. Suppose $0 \neq I \leq A \otimes_F B$ is a two-sided non-zero ideal. Let $x = \sum_i a_i \otimes b_i \in I$. As the two-sided ideal generated by b_1 is B by simplicity of B, therefore we have $b', b'' \in B$ such that $b'b_1b'' = 1$. Similarly, for $a_2 \in A$ we have $a'a_2a'' = 1$. We thus have that $x \cdot (1 \otimes b) - (1 \otimes b) \cdot x$ is an element in I which has length strictly smaller than n. We may thus put minimality hypothesis on n and hence deduce that $x \cdot (1 \otimes b) = (1 \otimes b) \cdot x$ for each $b \in B$. Similarly, we may get that $x \cdot (a \otimes 1) - (a \otimes 1) \cdot x$ for each $a \in A$. Now, any element of $A \otimes_F B$ is of form $\sum_i (a_i \otimes 1) \cdot (1 \otimes b_i)$. It follows that x is in the center of $A \otimes_F B$ which by above is F. This shows that $x \in I$ is a unit of $A \otimes_F B$, hence I is trivial. $\square$

Definition 17.3.4.2 (**Brauer group of a field**). Let F be a field. The Brauer group of F is defined to be quotient of the free abelian group generated by isomorphism classes of central simple algebras by the subgroup generated by the relations

$$[A \otimes_F B] = [A] \cdot [B]$$
$$[M_n(F)] = 1.$$

We denote the abelian group by $\mathrm{Br}(F)$. We denote the subgroup of n-torsion elements of $\mathrm{Br}(F)$ by ${}_n\mathrm{Br}(F) = \{[A] \in \mathrm{Br}(F) \mid [A]^n = 1\}$.

We will construct a group homomorphism $K_2(F) \to \mathrm{Br}(F)$. The map will take a Matsumoto symbol $\{a, b\}$ to a *cyclic algebra*, which will be a central simple algebra over F.

Definition 17.3.4.3 (Cyclic algebra). Let F be a field containing a primitive n^{th}-root of unity ζ. Fix two $a, b \in F^\times$. The cyclic algebra generated by a, b is the F-algebra $A_\zeta(a, b)$ generated by two elements x, y subject to the following relations:

$$x^n = a \cdot 1$$
$$y^n = b \cdot 1$$
$$yx = \zeta xy.$$

The following are few important properties of cyclic algebras.

Proposition 17.3.4.4. *Let F be a field containing a primitive n^{th}-root of unity ζ and $a, b \in F^\times$. Denote $A = A_\zeta(a, b)$ to be the cyclic algebra generated by a and b.*
1. *A is an n^2-dimensional F-algebra.*
2. *A is a central F-algebra.*
3. *A is a simple F-algebra.*

Proof. 1. Take any element $z \in A$. Then z can be written as a sum of monomials $x^i y^j$ where $1 \leq i, j \leq n - 1$. As each $x^i y^j$ is independent, therefore A has dimension n^2 over F, as required.

2. Let Z be the center of A. It contains F. Let $z \in Z$. We wish to show that $z \in F$. Indeed, we may write

$$z = \sum_{0 \leq i, j \leq n-1} c_{ij} x^i y^j$$

where note that the terms x^iy^j are independent. As z is in the center, we must have $z\cdot x = x\cdot z$ in particular. Expanding this, one yields,

$$\sum_{1\le i,j} c_{ij}(\zeta^j-1)x^{i+1}y^j = 0.$$

We hence deduce that $c_{ij}(\zeta^j-1)=0$ for all $1\le i,j\le n-1$. As ζ is primitive and $j\le n-1$, it follows at once that $\zeta^j-1\neq 0$ and thus $c_{ij}=0$ for all $1\le i,j\le n-1$. We hence deduce that $z=c_{00}\in F$, as required.

3. Pick any non-zero ideal $I\le A$ and let $z\in I$ be of form $z=\sum_{k=0}^m c_kx^{i_k}y^{j_k}$ of shortest length. We will show that $c_k=0$ for all k possibly except if $i_k,j_k=0$. Indeed, we may multiply z by x^{n-i_0} on left and y^{n-j_0} on right to get the first term of $x^{n-i_0}zy^{n-j_0}$ as c_0ab. We may multiply by $a^{-1}b^{-1}$ to further get the first term to be c_0. Now, the difference $\tilde{z}=a^{-1}b^{-1}x^{n-i_0}zy^{n-j_0}-z$ is zero as it is in I and has length strictly smaller than m. By repeating the same on $\tilde{z}$, one can show coefficients of all terms of $\tilde{z}$ are 0, from which we yield that $c_k=0$ for all k. Since in $\tilde{z}$, the term of z corresponding to $i_k,j_k=0$ is absent, hence $c_k=0$ for all k possibly except if $i_k,j_k=0$, as required. □

An important property of cyclic algebra is that they are *Brauer-torsion*.

Proposition 17.3.4.5. *Let F be a field containing a primitive n^{th}-root of unity ζ and $a,b\in F^\times$. Denote $A=A_\zeta(a,b)$ to be the cyclic algebra generated by a and b. Then, $A^{\otimes n}$ is isomorphic to a matrix algebra over F.*

We now show the existence of a symbol for fields with enough roots of unity.

Theorem 17.3.4.6 (Galois symbol-1)**.** *Let F be a field with a primitive n^{th}-root of unity ζ. Then the following map*

$$\begin{aligned}\varphi : K_2(F) &\longrightarrow \mathrm{Br}(F)\\ \{a,b\} &\longmapsto A_\zeta(a,b)\end{aligned}$$

is a homomorphism whose image is in ${}_n\mathrm{Br}(F)$. We call this map the n^{th} power norm residue symbol for F.

We will later do Merkurjev-Suslin theorem, which will tell us that the above map $\tilde{\varphi}$ is an isomorphism. Thus, all n-torsion elements of $\mathrm{Br}(F)$ are precisely the classes of cyclic algebras. We also construct Galois symbols for more general fields in §17.5.1.

Proof. Consider the function $F^\times\times F^\times\to \mathrm{Br}(F)$ given by $(a,b)\mapsto [A_\zeta(a,b)]$. To get a map from $K_2(F)$, by Matsumoto's theorem (Theorem 17.3.2.1), it suffices to show that the above map vanishes for Matsumoto's relations. To see that $[A_\zeta(\alpha,\beta)]\cdot[A_\zeta(\alpha,\gamma)]=[A_\zeta(\alpha,\beta\gamma)]$, we first observe the isomorphism $A_\zeta(\alpha,\beta)\otimes A_\zeta(\alpha,\gamma)\cong M_n(A_\zeta(\alpha,\beta\gamma))$ (Ex. 6.12, pp 266 of **[WeibHA]**) and observe that since $M_n(A)\cong A\otimes M_n(F)$, thus, $[M_n(A)]=[A]$ in $\mathrm{Br}(F)$. This shows the first relation. We next wish to show that $A_\zeta(a,1-a)\cong M_n(F)$ for all $a\neq 1$ in $F^\times$. By the Lemma 17.3.4.7 below, we reduce to showing that $A_\zeta(a,1-a)$ contains an n-torsion element. Indeed, as $x^n=a$, $y^n=1-a$ and since $(x+y)^n=x^n+y^n=a+(1-a)=1$[7], we thus we win. □

[7]for any cyclic algebra $A_\zeta(a,b)$, the equation $(x+y)^n=a+b$ holds. This can be checked by a calculation involving binomal theorem and then showing that the inner term sums to 0 by multiplying it on left by y and on right by x and observing the resultant pattern.

Lemma 17.3.4.7. *Let F be a field containing a primitive n^{th}-root of unity ζ. If A is a center simple algebra over F of dimension n^2 such that there exists $z \in A$ not in the center for which $z^n = 1$, then $A \cong M_n(F)$.*

Proof. Consider the subalgebra of A generated by z, denoted $F[z]$. As $z^n = 1$, therefore we have

$$F[z] \cong \frac{F[x]}{\langle x^n - 1\rangle}.$$

As $x^n - 1 \prod_{i=0}^{n-1}(x - \zeta^i)$ in $F[x]$ as F contains a primitive n^{th}-root of unity ζ, hence it follows that the quotient

$$F[z] \cong \frac{F[x]}{\langle x^n - 1\rangle} \cong \prod_{i=0}^{n-1} \frac{F[x]}{\langle x - \zeta^i\rangle}$$

by Chinese remainder theorem as the ideals $\langle x - \zeta^i\rangle$ are comaximal for $i \neq j$. Consequently, we have a splitting of the algebra $F[z]$ as:

$$F[z] \cong \underbrace{F \times \cdots \times F}_{n-\text{times}}$$

As $A \supseteq F[z]$, thus A has n-idempotents, which further induces a splitting of A into n-subalgebra (e_i are the idempotents of A):

$$A = e_1 A \times \cdots \times e_n A.$$

By Wedderburn-Artin theorem, every finite dimensional simple algebra B is isomorphic to $M_k(S)$ for some division ring S and B can atmost split into k-many right ideals. Consequently, $A \cong M_k(S)$ for some division algebra S over F and $n \leq k$. As $\dim_F A = n^2 = (\dim_F S)\cdot(\dim_S M_k(S)) = k^2 \cdot \dim_F S$, therefore we deduce that $k = n$ and $\dim_F S = 1$, yielding $A \cong M_n(F)$, as needed. $\square$

17.4 Higher K-theory of rings-I

We now wish to define higher K-groups for an associative ring R with 1. To this end, we will construct a space whose lower homotopy groups will agree with the K_1 and K_2 as defined above, and will then define higher K-groups to be the higher homotopy groups of this space. The K_0 case will need some special attention. We will require a lot of topological information to thoroughly discuss the definitions and their motivations, for which we may often refer to Foundational Homotopy Theory, Chapter 5.

We now make a series of observations to motivate the definition of what we need.

- Let R be a ring. The classifying space $\mathrm{BGL}(R)$ of $\mathrm{GL}(R)$ is such that
$$\pi_1(\mathrm{BGL}(R)) \cong \mathrm{GL}(R)$$
and thus it has a perfect subgroup $E(R)$ (Remark 17.2.1.10).
- Moreover, the homology of $\mathrm{BGL}(R)$ satisfies the following for any $\mathrm{GL}(R)$-module M:
$$H_\bullet(\mathrm{BGL}(R); M) \cong H_\bullet(\mathrm{GL}(R); M)$$
where on the right we have group homology of $\mathrm{GL}(R)$ with coefficients in M. We also know that $K_1(R) \cong H_1(\mathrm{GL}(R); \mathbb{Z})$ and $K_2(R) \cong H_2(E(R); \mathbb{Z})$ (Corollary 17.3.1.2). In particular, we have
$$K_1(R) \cong H_1(\mathrm{BGL}(R); \mathbb{Z}).$$
- As we have a natural quotient map
$$\pi : \mathrm{GL}(R) \to \frac{\mathrm{GL}(R)}{E(R)} = K_1(R),$$
hence, we may ask whether there is a space X and a map
$$i : \mathrm{BGL}(R) \to X$$
such that $\pi_1(X) \cong K_1(R)$ and the map
$$i_* : \pi_1(\mathrm{BGL}(R)) \to \pi_1(X)$$
is exactly the quotient map π?
In such a scenario, $\pi_1(X)$ is abelian and hence by Hurewicz, we must have
$$K_1(R) \cong \pi_1(X) \cong H_1(X; \mathbb{Z}).$$
Since we also have a map
$$i_* : H_1(\mathrm{BGL}(R); \mathbb{Z}) \to H_1(X; \mathbb{Z})$$
and by the previous observation $H_1(\mathrm{BGL}(R); \mathbb{Z}) \cong K_1(R)$, hence we have a map
$$i_* : K_1(R) \to K_1(R).$$
We would naturally like this to be an isomorphism. Hence we may wonder whether X and $i : \mathrm{BGL}(R) \to X$ can also be made such that
$$i_* : H_\bullet(\mathrm{BGL}(R); M) \to H_\bullet(X; M)$$
for any $\mathrm{GL}(R)$-module M is an isomorphism?

In conclusion, we want the following:

Construct a space X and map $i : \mathrm{BGL}(R) \to X$ such that (Q)
1. $\pi_1(X) \cong K_1(R)$,
2. $i_* : \pi_1(\mathrm{BGL}(R)) \to \pi_1(X)$ is the map $\pi : \mathrm{GL}(R) \twoheadrightarrow K_1(R)$,
3. For all $\mathrm{GL}(R)$-modules M, $i_* : H_\bullet(\mathrm{BGL}(R); M) \to H_\bullet(X; M)$ is an isomorphism.

We call the above three requirements to be the (Q)-criterion. What Quillen found that the space X can be constructed, but it will be well defined only upto homotopy equivalence (as the conditions we want is only about the homotopy groups of X). Consequently, there are many ways to construct X from $\mathrm{BGL}(R)$, all yielding homotopy equivalent spaces. We first discuss the definition of $\mathrm{BGL}(R)^+$.

17.4.1 The homotopy type $\mathrm{BGL}(R)^+$

The definition of $\mathrm{BGL}(R)^+$ is not enlightening, in-fact the criterion (Q) is what we will call a $\mathrm{BGL}(R)^+$.

Definition 17.4.1.1 ($\mathrm{BGL}(R)^+$)**.** Let R be an associative ring with 1. A pair (X, i) of a CW-complex X and map $i : \mathrm{BGL}(R) \to X$ satisfying the (Q)-criterion above is called a model of $\mathrm{BGL}(R)^+$. For a model X of $\mathrm{BGL}(R)^+$, we define the homotopy type of X to be the $\mathrm{BGL}(R)^+$. We will abuse the notation and sometimes write $\mathrm{BGL}(R)^+$ as a model of $\mathrm{BGL}(R)^+$ as well!

Remark 17.4.1.2 ($\mathrm{BGL}(R)^+$ is defined upto homotopy)**.** As the definition shows, the space $\mathrm{BGL}(R)^+$ is a homotopy type, not really a space. Of-course, we need to show that any two spaces X and Y satisfying (Q)-criterion are homotopy equivalent. This will require some work.

From now on, as was the case earlier, we will assume all the rings are associative with 1.

Theorem 17.4.1.3 (Quillen)**.** *Let R be a ring. If X and Y are two models for $\mathrm{BGL}(R)^+$, then $X \simeq Y$, that is, they are homotopy equivalent.*

This is proved later in Corollary 17.4.1.17. Before proving this, we first need to establish that our definition of homotopy type $\mathrm{BGL}(R)^+$ actually does serve the purpose. That is, we wish to show that $\pi_2(\mathrm{BGL}(R)^+) \cong K_2(R)$.

Theorem 17.4.1.4. *Let R be a ring. Then*

$$\pi_2(\mathrm{BGL}(R)^+) \cong K_2(R).$$

This is proved in Theorem 17.4.1.8. We now embark on a small homotopical study of $\mathrm{BGL}(R)^+$, to extract information about $\mathrm{BGL}(R)^+$ which will yield the results stated above.

We fix a model of $\mathrm{BGL}(R)^+$ and also call it $\mathrm{BGL}(R)^+$ in the following. Of-course, we have not proved existence of a model yet, but we will do so soon, after realizing that indeed the homotopy type $\mathrm{BGL}(R)^+$ works for our need.

Acyclic fiber, acyclic maps

As the map $i : \mathrm{BGL}(R) \to \mathrm{BGL}(R)^+$ is a quasi-isomorphism, thus it becomes a special map by the following important theorem.

Theorem 17.4.1.5 (Acyclic fiber theorem)**.** *Let $f : X \to Y$ be a based map of connected CW-complexes. Then the following are equivalent:*

1. *For all $k \geq 0$, we have*
$$f_* : H_k(X; M) \xrightarrow{\cong} H_k(Y; M)$$
for every $\pi_1(Y)$-module M[8]*.*
2. *The homotopy fiber Ff of f is acyclic*[9]*.*

Proof. (1. $\Rightarrow$ 2.) By replacing X by the fibration replacement of f (see Construction 5.3.1.11), we may assume that we have a fibration $Ff \xrightarrow{i} X \xrightarrow{f} Y$. Assume that $\pi_1(Y) = 0$, so that we have a Serre spectral sequence $E^2_{pq} = H_p(Y; H_q(Ff)) \Rightarrow H_{p+q}(X)$ and for the trivial fibration pt. $\to Y \xrightarrow{\mathrm{id}} Y$ which gives another Serre spectral sequence $'E^2_{pq} = H_p(Y; H_q(\mathrm{pt.})) \Rightarrow H_{p+q}(Y)$. We have a commutative diagram:

$$\begin{tikzcd}
Ff \arrow[r, "i"] \arrow[d] & X \arrow[r, "f"] \arrow[d, "f"'] & Y \arrow[d, "\mathrm{id}"'] \\
\mathrm{pt.} \arrow[r] & Y \arrow[r, "\mathrm{id}"'] & Y
\end{tikzcd}.$$

By comparison theorem (Proposition 5.13 of [**HatchSSeq**]), we deduce that Ff is acyclic. It follows that if Y is simply connected and f induces isomorphism on integral homology, then homotopy fiber of f is acyclic.

Now suppose $\pi_1(Y) \neq 0$. The main idea is to reduce to the simply connected case by going to universal cover of Y. Indeed, if $\tilde{Y}$ is the universal cover of Y, then we have the following pullback diagrams (by Lemma 5.2.1.2, we have that $\tilde{f}$ is a fibration):

$$\begin{tikzcd}
F\tilde{f} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & X \times_Y \tilde{Y} \arrow[r, "\tilde{f}"] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \tilde{Y} \arrow[d, "p"] \\
Ff \arrow[r] & X \arrow[r, "f"'] & Y
\end{tikzcd}.$$

Denote $\tilde{X} = X \times_Y \tilde{Y}$. It then follows by maps constructed by unique path lifting that $F\tilde{f} \cong Ff$. It thus suffices to show that $F\tilde{f}$ is acyclic. To this end, by above, we reduce to showing that we have an isomorphism $\tilde{f}_* : H_k(\tilde{X}; \mathbb{Z}) \to H_k(\tilde{Y}; \mathbb{Z})$ for all $k \geq 0$. This follows from the following comutative square with vertical maps being isomorphisms:

$$\begin{tikzcd}
H_k(\tilde{X}; \mathbb{Z}) \arrow[r, "\tilde{f}_*"] \arrow[d, "\cong"'] & H_k(\tilde{Y}; \mathbb{Z}) \arrow[d, "\cong"] \\
H_k(X; \mathbb{Z}[\pi_1(Y)]) \arrow[r, "f_*"'] & H_k(X; \mathbb{Z}[\pi_1(Y)])
\end{tikzcd}.$$

[8]That is, M is a left $\mathbb{Z}[\pi_1(Y)]$-module.
[9]that is, Ff has homology of a point.

As f_* is an isomorphism by hypothesis, we win.

(2. $\Rightarrow$ 1.) As before, we may assume that $Ff \xrightarrow{i} X \xrightarrow{f} Y$ is a fibration. Fix a $\pi_1(Y)$-module M. Observe that the E^2-page of Serre spectral sequence $E^2_{pq} = H_p(Y; H_q(Ff; M)) \Rightarrow H_{p+q}(X; M)$ is all 0 except possibly the bottom row (which consists of $H_q(Y; M)$) since $H_q(Ff; M) = 0$ forall $q \geq 1$ and $H_0(Ff; M) = M$ by a simple use of universal coefficients theorem. It follows that E collapses on the E^2-page, so that $H_n(X; M) \cong H_n(Y; M)$. In particular, this isomprhism comes from f_* as the above isomorphim is by the edge homomorphism which we know in Serre spectral sequence is via the map $f : X \to Y$ (see Addendum 2, Theorem 5.3.2 of **[WeibHA]**). □

Corollary 17.4.1.6. *Let R be a ring. Then the homotopy fiber of $i : \mathrm{BGL}(R) \to \mathrm{BGL}(R)^+$ is acyclic.*

Proof. Follows from definition and Theorem 17.4.1.5. □

Acyclicity is both homological and cohomological.

Lemma 17.4.1.7. *Let $f : X \to Y$ be a map of connected spaces and π be an abelian group. If $f_* : H_q(X; \pi) \to H_q(Y; \pi)$ is an isomorphism for all $q \geq 0$, then $f^* : H^q(Y; \pi) \to H^q(X; \pi)$ is an isomorphism for all $q \geq 0$.*

Proof. By universal coefficient theorem for cohomology, we have the following commutative diagram where rows are exact:

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Ext}^1_{\mathbb{Z}}(H_{n-1}(X), \pi) \arrow[r] & H^n(X; \pi) \arrow[r] & \mathrm{Hom}\,(H_n(X), \pi) \arrow[r] & 0 \\ 0 \arrow[r] & \mathrm{Ext}^1_{\mathbb{Z}}(H_{n-1}(Y), \pi) \arrow[r] \arrow[u] & H^n(Y; \pi) \arrow[r] \arrow[u, "f^*"'] & \mathrm{Hom}\,(H_n(Y), \pi) \arrow[r] \arrow[u] & 0 \end{tikzcd} .$$

As the vertical arrows on left and right are induced by $f_* : H_q(X) \to H_q(Y)$ which is an isomorphism, therefore they are also isomorphisms. By 5-lemma, we conclude that f^* is an isomorphism.

$K_2(R)$ **&** $\pi_2(\mathrm{BGL}(R)^+)$

We may now see that indeed $K_2(R) \cong \pi_2(\mathrm{BGL}(R)^+)$!

Theorem 17.4.1.8. *Let R be a ring. Then*

$$\pi_2(\mathrm{BGL}(R)^+) \cong H_2(E(R); \mathbb{Z}).$$

Proof. Main idea is to exhibit $\pi_2(\mathrm{BGL}(R)^+)$ in a universal central extension of $E(R)$ as follows:

$$0 \to \pi_2(\mathrm{BGL}(R)^+) \to ?? \to E(R) \to 1,$$

so that uniqueness of universal central extension would yield the proof together with Theorem 6.9.5 of **[WeibHA]**.

We may employ the long exact sequence of homotopy groups associated to a map (see Corollary 5.3.3.7). Indeed, we have the following part of a l.e.s:

$$\pi_2(\mathrm{BGL}(R)) \to \pi_2(\mathrm{BGL}(R)^+) \to \pi_1(Fi) \to \pi_1(\mathrm{BGL}(R)) \xrightarrow{i_*} \pi_1(\mathrm{BGL}(R)^+) \to \pi_0(Fi).$$

As Fi is acyclic, therefore $\pi_0(Fi) = 0$. As $\mathrm{BGL}(R)$ is $K(\mathrm{GL}(R), 1)$, therefore $\pi_2(\mathrm{BGL}(R)) = 0$. Moreover, $\pi_1(\mathrm{BGL}(R)) = \mathrm{GL}(R)$, $\pi_1(\mathrm{BGL}(R)^+) = K_1(R)$ and $i_* = \pi$ where $\pi : \mathrm{GL}(R) \to \frac{\mathrm{GL}(R)}{E(R)}$, by definition.

It follows that we have the following s.e.s.:

$$0 \to \pi_2(\mathrm{BGL}(R)^+) \xrightarrow{\rho} \pi_1(Fi) \to \mathrm{GL}(R) \xrightarrow{\pi} K_1(R) \to 0.$$

We then further deduce the following s.e.s.:

$$0 \to \pi_2(\mathrm{BGL}(R)^+) \xrightarrow{\rho} \pi_1(Fi) \to E(R) \to 1.$$

To complete the proof, we need only show that the above is a universal central extension of $E(R)$. To this end, we first have to show that image of ρ is in the center of $\pi_1(Fi)$. This follows from Corollary IV.3.5 of [**WhElem**]. Finally we wish to see that the above is universal central. To this end, by Recognition Criterion 6.9.7 of [**WeibHA**], it suffices to show that $\pi_1(Fi)$ is perfect and every central extension of $\pi_1(Fi)$ splits. The former is true as Fi is acyclic, so $H_1(Fi) = 0$ and thus $\pi_1(Fi)$ is perfect. For the latter, by Corollary 6.9.9 of [**WeibHA**], it suffices to show that $H_2(\pi_1(Fi); \mathbb{Z}) = 0$. This follows from the Propsotioin 17.4.1.9 mentioned below. □

Proposition 17.4.1.9. *Let E be an acyclic space with fundamental group G. Then $H_2(G; \mathbb{Z}) = 0$.*

Proof. Note that BG is a space with fundamental group G as well and moreover $H_2(BG; \mathbb{Z}) \cong H_2(G; \mathbb{Z})$. It hence suffices to show that $H_2(BG; \mathbb{Z}) = 0$. As BG is obtained by attaching cells to X, therefore we have a natural map $f : E \to BG$ which induces isomorphism on π_1. By considering the fibration replacement, we may assume f is a fibration and

$$Ff \to E \xrightarrow{f} BG$$

a fibration sequence. Considering the Serre spectral sequence of this fibration, we obtain $E^2_{pq} = H_p(BG; H_q(Ff)) \Rightarrow H_{p+q}(E) = 0$. Consequently, we have $E^\infty_{pq} = 0$ for all p, q except $p, q = 0$. An immediate observation of relevant differentials yield that $E^\infty_{00} = E^2_{00}$, from which it follows that $E^2_{00} = H_0(BG; H_0(Ff)) \cong H_0(E) = \mathbb{Z}$. As BG is path-connected, therefore $H_0(Ff) \cong \mathbb{Z}$, showing that Ff is path-connected. Another simple analysis of differentials at E^2_{20} yields that if $E^2_{01} = 0$, then $E^\infty_{20} = E^2_{20} = H_2(BG; H_0(Ff)) = H_2(BG; \mathbb{Z})$, so that it will follow that $H_2(BG; \mathbb{Z}) = 0$, as required. We thus reduce to showing that $E^2_{01} = H_0(BG; H_1(Ff)) \cong H_1(Ff)$ is 0.

From the fibration long exact sequence obtained from f and the fact that $\pi_k(BG) = 0$ for all $k \neq 1$, and $\pi_1(BG) = G$, we deduce that that map $i : Ff \to E$ induces an isomorphism on π_k for all $k \geq 2$ and moreover $\pi_0(Ff) = \pi_1(Ff) = 0$ as $f_* : \pi_1(E) \to G$ is an isomorphism. Thus, Ff and $\tilde{E}$, the universal cover of E has same homotopy groups via the map $Ff \to E$. One can then show by unique lifting criterion that there is a map $\tilde{i} : Ff \to \tilde{E}$ which is a weak equivalence. By Whitehead's theorem (Ff is a CW-complex as well), it follows that $\tilde{i}$ is a homotopy equivalence. Hence $H_1(Ff) \cong H_1(\tilde{E})$ and since $\pi_1(\tilde{E}) = 0$, by Hurewicz, we have $H_1(\tilde{E}) = 0$, as required. This completes the proof. □

The $+$-construction & uniqueness

We next wish to prove Theorem 17.4.1.3. That is, any two models of $\mathrm{BGL}(R)^+$ are homotopy equivalent. To this end, we first abstract out the necessary conditions from the definition of $\mathrm{BGL}(R)^+$.

Definition 17.4.1.10 ($+$**-construction).** Let X be a based connected CW-complex and G be a perfect normal subgroup of $\pi_1(X)$. Then a map of CW-complexes $f : X \to Y$ is called a $+$-construction on X w.r.t. G if f is acyclic and $\mathrm{Ker}\,(f_* : \pi_1(X) \to \pi_1(Y)) = G$.

Remark 17.4.1.11. Let $f : X \to Y$ be a $+$-construction w.r.t. $P \leq \pi_1(X)$ perfect normal subgroup. By homotopy long exact sequence corresponding to map $Ff \to X \xrightarrow{f} Y$, we can immediately get following exact sequence:

$$\pi_1(Ff) \to \pi_1(X) \xrightarrow{f_*} \pi_1(Y) \to \pi_0(Ff).$$

By Theorem 17.4.1.5, Ff is acyclic and thus $\pi_0(Ff) = 0$. Thus we have the exact sequence:

$$0 \to G \to \pi_1(X) \xrightarrow{f_*} \pi_1(Y) \to 0.$$

Construction 17.4.1.12 (The construction of X^+)**.** Let X be a based connected CW-complex and $G \leq \pi_1(X)$ a perfect normal subgroup. We construct an inclusion $i : X \to X^+$ which is a $+$-construction of X w.r.t. G. To this end, the main strategy is as follows:

1. First attach 2-cells to X to kill G in $\pi_1(X)$.
2. Then attach 3-cells to remove the extra homology classes added by step 1.

Let us denote G in generators as follows:

$$G = \langle g_\alpha \mid \alpha \in I \rangle.$$

As $g_\alpha \in \pi_1(X)$, therefore we may interpret them as loops

$$g_\alpha : S^1 \to X.$$

Now attach 2-cells to X along each of the g_α:

$$\begin{tikzcd} X' \arrow[dr, phantom, "\ulcorner", very near start] & \amalg_\alpha D^2 \arrow[l] \\ X \arrow[u, hook, "j_0"] & \amalg_\alpha S^1 \arrow[l, "\amalg_\alpha g_\alpha"] \arrow[u, hook, "i_0"'] \end{tikzcd} \quad . \tag{A1}$$

We first claim that $\pi_1(X')$ is $\pi_1(X)/G$ via j_0. Indeed, the map

$$j_{0*} : \pi_1(X) \longrightarrow \pi_1(X')$$

is surjective since any element $h : S^1 \to X'$ in $\pi_1(X')$ by cellular approximation theorem factors through the inclusion j_0. In particular, the 1-skeleton of X' is same as that of X. Consequently to prove our claim, we need only show that $\mathrm{Ker}\,(j_{0*}) = G$. Clearly, $\mathrm{Ker}\,(j_{0*}) \supseteq G$ by construction. Furthermore, if $k : S^1 \to X$ is null-homotopic in X', then k extends to $k' : D^2 \to X'$. By cellular approximation, we may assume that k' is a cellular map, so that k' is mapping in the 2-skeleton of

X'. It follows at once that if k is not in G, then k (which we assume, by cellular approximation, that it is in 1-skeleton of X) on composition with j_0 gives a non-contractible loop as X' only trivializes all loops in G, a contradiction.

This shows that

$$\pi_1(X') = \pi_1(X)/G.$$

To complete the proof, we have to now kill all "new" homology classes of X' with an arbitrary choice of coefficient system $\mathcal{L}$ whose groups are isomorphic to L. To this end, we will attach 3-cells to X' to obtain the space X^+.

To illustrate the idea, suppose we have constructed X^+ by attaching 3-cells to X'. Our goal is then to show that $H_k(X^+;\mathcal{L}) \cong H_k(X;\mathcal{L})$. We thus have a triplet (X^+, X', X). By homology l.e.s. for the pair (X^+, X) , it suffices to show that

$$H_k(X^+, X;\mathcal{L}) = 0$$

for all $k \geq 0$. Recall that the homology of pair (X^+, X') with coefficient $\mathcal{L}$ is given by the homology of complex $L \otimes_{\mathbb{Z}[\pi_1(X)/G]} C_\bullet(\widetilde{X^+}, \hat{X})$ where $\hat{X}$ is the pullback of $\widetilde{X^+}$ along $X \to X^+$. It is thus sufficient to show that $C_\bullet(\widetilde{X^+}, \hat{X})$ is an acyclic complex (whose homology in every degree is 0). As $\widetilde{X^+}/\hat{X}$ will be a 3-dimensional CW-complex with no 1-cells, it is thus sufficient to show that the differential

$$d : C_3(\widetilde{X^+}, \hat{X}) \to C_2(\widetilde{X^+}, \hat{X})$$

is an isomorphism.

Now since we have isomorphisms $C_3(\widetilde{X^+}, \hat{X}) \cong C_3(\widetilde{X^+}, \widetilde{X'}) \cong H_3(\widetilde{X^+}, \widetilde{X'})$ and $C_2(\widetilde{X^+}, \hat{X}) \cong C_2(\widetilde{X'}, \hat{X}) \cong H_2(\widetilde{X'}, \hat{X})$ by the fact that cells of universal cover are obtained by lifting, therefore we have to show that the boundary map obtained by the triplet l.e.s. for $(\widetilde{X^+}, \widetilde{X'}, \hat{X})$ is an isomorphism. This is how we construct X^+ and then show that for this construction the above actually holds.

In order to construct X^+, we need maps $S^2 \to X'$ through which we can attach 3-cells. In particular, these are elements of $\pi_2(X')$. Consider the following pullback square

$$\begin{tikzcd} \hat{X} \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \tilde{X}' \arrow[d, "\pi"] \\ X \arrow[r, hook, "j_0"'] & X' \end{tikzcd}$$

where $\tilde{X}' \to X'$ is the universal cover. As pullback of covering is a covering, thus the map $\hat{X} \to X$ is a covering. Now, it is clear that $\hat{X} = \pi^{-1}(X)$, thus the inclusion $\hat{X} \hookrightarrow \tilde{X}'$ is also induced by attaching 2-cells to $\hat{X}$. It follows that $\pi_1(\hat{X}) \cong G$.

Next, observe that in the homology l.e.s. of $(\tilde{X}', \hat{X})$, we get the following isomorphism by Hurewicz (as $\tilde{X}'$ is 1-connected)

$$\pi_2(\tilde{X}') \xrightarrow{\cong} H_2(\tilde{X}') \xrightarrow{j_*} H_2(\tilde{X}', \hat{X}) \longrightarrow H_1(\hat{X}).$$

Again, by Hurewicz, we have

$$H_1(\hat{X}) \cong \pi_1(\hat{X})^{ab} = G^{ab} = 0$$

as G is perfect. Hence the above sequence becomes

$$\pi_2(\tilde{X}') \xrightarrow{\cong} H_2(\tilde{X}') \overset{j_*}{\twoheadrightarrow} H_2(\tilde{X}', \hat{X}).$$

Using the above, we have a surjection $\pi_2(\tilde{X}') \twoheadrightarrow H_2(\tilde{X}', \hat{X})$. For each homology class $[c_\beta] \in H_2(\tilde{X}', \hat{X})$ in a fixed generating set, choose one and only element in the fiber $[\tilde{h}_\beta] \in \pi_2(\tilde{X}')$. We thus have a collection of maps $\{\tilde{h}_\beta : S^2 \to \tilde{X}'\}_\beta$. Composing them with $\pi : \tilde{X}' \to X'$ yields maps $\{h_\beta : S^2 \to X'\}_\beta$. We use these maps to attach 3-cells to X'. Indeed, consider the pushout space:

$$\begin{tikzcd} X^+ \arrow[dr, phantom, "\ulcorner", very near start] & \amalg_\beta D^3 \arrow[l, "\amalg_\beta d_\beta"'] \\ X' \arrow[u, hook, "k_0"] & \amalg_\beta S^2 \arrow[l, "\amalg_\beta h_\beta"] \arrow[u, hook] \end{tikzcd} . \tag{A2}$$

We thus have the following inclusions of subcomplexes of X^+:

$$X \overset{j_0}{\hookrightarrow} X' \overset{k_0}{\hookrightarrow} X^+.$$

We again pass to universal cover of X^+ in order and take pullback along $X' \hookrightarrow X^+$ to have better algebraic control via Hurewicz:

$$\begin{tikzcd} \hat{X}' \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \widetilde{X^+} \arrow[d, "\pi"] \\ X' \arrow[r, hook, "k_0"'] & X^+ \end{tikzcd} .$$

But $\pi_1(\hat{X}') = 0$ since k_{0*} is an isomorphism on π_1 and $\pi_1(\widetilde{X^+}) = 0$. Hence, we deduce that

$$\hat{X}' \cong \tilde{X}',$$

that is, $\hat{X}'$ is the universal cover of X'.

By naturality of Hurewicz, we have a map between the long exact sequences of homotopy groups induced by the map $\hat{X}' \hookrightarrow \widetilde{X^+}$ to that of homology groups

$$\begin{tikzcd} \cdots & \pi_{n+1}(\widetilde{X^+}, \hat{X}') \arrow[r] \arrow[d] & \pi_n(\hat{X}') \arrow[r] \arrow[d] & \pi_n(\widetilde{X^+}) \arrow[r] \arrow[d] & \pi_n(\widetilde{X^+}, \hat{X}') \arrow[d] & \cdots \\ \cdots & H_{n+1}(\widetilde{X^+}, \hat{X}') \arrow[r] & H_n(\hat{X}') \arrow[r] & H_n(\widetilde{X^+}) \arrow[r] & H_n(\widetilde{X^+}, \hat{X}') & \cdots \end{tikzcd} .$$

For $n = 3$, we get the following sequence from the above

$$\begin{tikzcd} \pi_3(\widetilde{X^+}, \widetilde{X'}) \arrow[r] \arrow[d] & \pi_2(\widetilde{X'}) \arrow[d, "\cong"] & \\ H_3(\widetilde{X^+}, \widetilde{X'}) \arrow[r, "\tilde{\partial}"'] & H_2(\widetilde{X'}) \arrow[r, two heads, "j_*"'] & H_2(\widetilde{X'}, \hat{X}) \end{tikzcd} .$$

We claim that $j_* \circ \tilde{\partial}$ is an isomorphism. Note that this is the boundary map of cellular complex. Indeed, observe that $H_3(\widetilde{X^+}, \widetilde{X'})$ is a free abelian group generated by the lift of 3-cells attached by $\tilde{h}_\beta$. We thus need only show that $j_* \circ \tilde{\partial}$ maps this bijectively onto the generators of $H_2(\widetilde{X'}, \hat{X})$ which we know are $[c_\beta]$. We know that the lifted map $\tilde{h}_\beta : S^2 \to \widetilde{X'}$ determines an element in $\pi_3(\widetilde{X^+}, \widetilde{X'})$ by definition of relative homotopy, whose image in $\pi_2(\widetilde{X'})$ is exactly $[\tilde{h}_\beta]$. Moreover, the class determined by $\tilde{h}_\beta$ in $\pi_3(\widetilde{X^+}, \widetilde{X'})$, under the Hurewicz map, determines a class $[l_\beta] \in H_3(\widetilde{X^+}, \widetilde{X'})$. By commutativity of above, it follows that $j_* \circ \tilde{\partial}$ maps $[l_\beta] \mapsto [c_\beta]$. As for each generator $[c_\beta] \in H_2(\widetilde{X'}, \hat{X})$, the element $[l_\beta]$ is unique by construction, we get that $j_* \circ \tilde{\partial}$ is an isomorphism, as required.

Remark 17.4.1.13. While it is rarely that we will use the explicit construction above, it is still good to keep in mind the precise way in which we found the 3-cells to attach to X' to get X^+. In particular, the attaching steps (A1) and (A2) are good to keep in mind.

Example 17.4.1.14 (+-construction of homology spheres)**.** Let X be a based connected CW-complex which is a homology n-sphere for $n > 1$ so that $\pi_1(X)$ is perfect. For $P = \pi_1(X)$, we claim that any $+$-construction of X w.r.t. P, $f : X \to X^+$, is such that $S^n \simeq X^+$.

Indeed, observe that $\pi_1(X)$ is perfect as X is a homology n-sphere. As f is a $+$-construction, therefore $\pi_1(X^+)$ is $\pi_1(X)/\pi_1(X) = 0$ by Remark 17.4.1.11. Moreover, X^+ itself is a homology n-sphere as $f : X \to X^+$ is acyclic. We now find a map $g : S^n \to X$ such that g is a weak equivalence, so that by Whitehead's theorem we will conclude that g is a homotopy equivalence, as required.

Indeed, observe that since X^+ is 1-connected, therefore by Hurewicz's theorem, we have $\pi_2(X^+) \cong H_2(X^+)$. If $n \neq 2$, then $\pi_2(X^+) = 0$ as X^+ is also a homology n-sphere. By induction and using Hurewicz repeatedly, we get that $\pi_k(X^+) = 0$ for all $0 \leq k \leq n-1$, so that X^+ is $n-1$-connected and thus by another application of Hurewicz, we have $\pi_n(X^+) \cong H_n(X^+) = \mathbb{Z}$. We thus have a non-trivial map $g : S^n \to X^+$ whose homology class is the generator. We finally claim that g induces an isomorphism in integral homology, which will complete the proof by Theorem 7.5.9 of **[SpaAT]** (Whitehead's theorem). To this end, as X^+ is a also a homology n-sphere, thus we need only show that $g_* : H_n(S^n) = \mathbb{Z} \to H_n(X^+) = \mathbb{Z}$ takes $[\mathrm{id}] \mapsto [g]$. Indeed, we have $g_*([\mathrm{id}]) = [g \circ \mathrm{id}] = [g] \in H_n(X^+)$, as needed.

Proposition 17.4.1.15. *Let $i : X \to X^+$ and $j : Y \to Y^+$ be +-constructions w.r.t. perfect normal subgroups $G \leq \pi_1(X)$ and $H \leq \pi_1(Y)$. Then*

$$i \times j : X \times Y \to X^+ \times Y^+$$

is a $+$-construction of $X \times Y$ w.r.t. the perfect normal subgroup $G \times H \leq \pi_1(X \times Y)$.

Proof. We first show acyclicity of $i \times j$. By unravelling definitions, one reduces to showing that $F(i \times j) \cong F(i) \times F(j)$ is acyclic. To this end, use Künneth formula to deduce that if X, Y are acyclic, then so is $X \times Y$. The fact that kernel of $(i \times j)_*$ is $G \times H$ follows from $(i \times j)_* = i_* \times j_* : \pi_1(X) \times \pi_1(Y) \to \pi_1(X^+) \times \pi_1(Y^+)$, as required. □

The following universal property of Quillen tells us what we need, and then some more.

Theorem 17.4.1.16 (Quillen)**.** *Let X be a CW-complex and P be a perfect normal subgroup of $\pi_1(X)$. Let $f : X \to Y$ be a +-construction on X w.r.t. P. If $g : X \to Z$ is a map such that*

$$P \subseteq \operatorname{Ker}(g_* : \pi_1(X) \to \pi_1(Z)),$$

then there exists a map $h : Y \to Z$ such that the following diagram of spaces commutes

$$\begin{tikzcd} Y \arrow[r, dashed, "h"] & Z \\ X \arrow[u, "f"] \arrow[ur, "g"'] & \end{tikzcd}$$

and h is unique upto homotopy.

An immediate corollary is what we seek.

Corollary 17.4.1.17 (Uniqueness of +-construction)**.** *Let X be a CW-complex and P be a perfect normal subgroup of $\pi_1(X)$. If $f : X \to Y$ and $g : X \to Z$ are two +-constructions, then there is a homotopy equivalence $h : Y \xrightarrow{\simeq} Z$.* □

Another important consequence is that we have maps in +-construction.

Lemma 17.4.1.18. *Let X, Y be two connected CW-complexes and $i : X \to X^+$ and $j : Y \to Y^+$ be +-constructions w.r.t. perfect normal subgroups $G \leq \pi_1(X)$ and $H \leq \pi_1(Y)$ respectively. If $f : X \to Y$ is a map such that $f_* : \pi_1(X) \to \pi_1(Y)$ maps G into H, then there exists a map $\tilde{f} : X^+ \to Y^+$ unique upto homotopy w.r.t. the commutativity of the following square of spaces:*

$$\begin{tikzcd} X^+ \arrow[r, dashed, "\tilde{f}"] & Y^+ \\ X \arrow[u, "i"] \arrow[r, "f"'] & Y \arrow[u, "j"'] \end{tikzcd} .$$

Proof. The map $j \circ f$ on π_1 takes G to 0, so by Theorem 17.4.1.16 gives the required map unique upto homotopy. □

We shall prove Theorem 17.4.1.16 by using obstruction theory as developed in **[WhElem]**, Chapter VI.

Proof of Theorem 17.4.1.16. Consider the based connected CW-complex X^+ obtained by Construction 17.4.1.12. Let $g : X \to Z$ be a map such that

$$P \subseteq \operatorname{Ker}(g_* : \pi_1(X) \to \pi_1(Z)).$$

We wish to extend g to $\tilde{g} : X^+ \to Z$. Consider the map $\theta : \pi_1(X)/P \to \pi_1(Z)$ as in the triangle below which exists by hypothesis on g_*:

$$\begin{tikzcd} \pi_1(X)/P \arrow[r, dashed, "\theta"] & \pi_1(Z) \\ \pi_1(X) \arrow[u, "i_*"] \arrow[ur, "g_*"'] & \end{tikzcd} .$$

We wish to show that g extends to $\tilde{g} : X^+ \to Z$ such that $\tilde{g}_* = \theta$. To this end, by obstruction theory, it is sufficient to show that

$$H^q(X^+, X; \mathcal{L}) = 0$$

for all $q \geq 3$ and all local coefficient systems $\mathcal{L}$ on X^+. Fix a local coefficient system $\mathcal{L}$ with group G. Note that we have

$$H^q(X^+, X; \mathcal{L}) \cong H^q\left(\mathrm{Hom}_{\mathbb{Z}[\pi_1(X^+)]}\left(C_\bullet(\widetilde{X^+}, \hat{X}), G\right)\right)$$

where we have the following pullback of the universal cover of X^+:

$$\begin{tikzcd} \hat{X} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \widetilde{X^+} \arrow[d] \\ X \arrow[r, hook] & X^+ \end{tikzcd}.$$

Now note from the Construction 17.4.1.12 that

$$C_k(\widetilde{X^+}, \hat{X}) = 0$$

for all $k \neq 2, 3$ and $d : C_3(\widetilde{X^+}, \hat{X}) \to C_2(\widetilde{X^+}, \hat{X})$ is an isomorphism. It follows at once that $H^q(X^+, X; \mathcal{L}) = 0$ for all $q \geq 0$, as required.

For uniqueness upto homotopy, obstruction theory further gives us a sufficient criterion that $H^2(X^+, X; \mathcal{L}) = 0$. Hence we are done. Moreover, by the long exact sequence of pairs for cohomology with local coefficients, we deduce that the map $i : X \hookrightarrow X^+$ induces isomorphism

$$i^* : H^q(X^+; \mathcal{L}) \to H^q(X; i^*\mathcal{L}),$$

that is, $i : X \to X^+$ is cohomologically acyclic as well. This shows the universal property for the explicit construction. We now show that any +-construction on X w.r.t. P is homotopy equivalent to the explicit one. This will then complete the proof.

Let $f : X \to Y$ be a +-construction w.r.t. P. Then by above there exists a map $\tilde{f} : X^+ \to Y$ as in the following triangle

$$\begin{tikzcd} X^+ \arrow[r, dashed, "\tilde{f}"] & Y \\ X \arrow[u, "i"] \arrow[ur, "f"'] & \end{tikzcd}.$$

We claim that the map $\tilde{f}$ is a homotopy equivalence. By Whitehead's theorem, it is sufficient to show that $\tilde{f}$ is a weak-equivalence. Observe that as i and f are homologically acyclic, it follows at once that $\tilde{f}$ is also acyclic. Moreover, $\tilde{f}$ induces isomorphism in fundamental groups. By acyclic fiber theorem (Theorem 17.4.1.5), it follows that the homotopy fiber $F\tilde{f}$ is acyclic. We further claim that $F\tilde{f}$ is 1-connected. Indeed, from the long exact sequence for homotopy groups for $\tilde{f}$ and that $\tilde{f}_* : \pi_1(X^+) \to \pi_1(Y)$ is an isomorphism, it follows that the map $\pi_1(F\tilde{f}) \to \pi_1(X^+)$ is the zero map. It suffices to show that the transgression $\pi_2(Y) \to \pi_1(F\tilde{f})$, which is surjective by exactness, is the zero map as well. As $F\tilde{f}$ is acyclic, therefore $\pi_1(F\tilde{f})$ is a perfect group. By above, it is also abelian, and thus the zero group, as required.

Hence $F\tilde{f}$ is a 1-connected acyclic space, so that by Hurewicz's theorem, all homotopy groups of $F\tilde{f}$ are 0. By homotopy long exact sequence of $\tilde{f}$, it follows that $\tilde{f}$ is a weak-equivalence, as required. This also proves Corollary 17.4.1.17. □

K-theory space

We now finally define higher K-groups as follows:

Definition 17.4.1.19 (**K-theory space & K-groups**)**.** Let R be a ring and let $\mathrm{BGL}(R)^+$ be as in Definition 17.4.1.1. Then for all $n \geq 1$, define

$$K_n(R) = \pi_n(\mathcal{K}(R))$$

where $\mathcal{K}(R) = K_0(R) \times \mathrm{BGL}(R)^+$ and $K_0(R)$ has discrete topology. The space $\mathcal{K}(R)$ is called the K-theory space of R.

We first have to show that $\pi_n(\mathcal{K}(R))$ is really independent of basepoints and that we are really computing homotopy groups of $\mathrm{BGL}(R)^+$ only, not of something else.

Lemma 17.4.1.20. *Let R be a ring and $\mathcal{K}(R)$ be the K-theory space of R.*
1. *$\pi_0(\mathcal{K}(R)) = K_0(R)$ in the usual sense.*
2. *$\pi_n(\mathcal{K}(R)) = \pi_n(\mathrm{BGL}(R)^+)$.*

Proof. Picking any base point of $\mathcal{K}(R)$, the path component of that point is homeomorphic to $\mathrm{BGL}(R)^+$, so the homotopy groups for different base points are all isomorphic. □

We immediately have maps in K-theory space.

Proposition 17.4.1.21 (Maps in K-theory)**.** *Let $f : R \to S$ be a ring homomorphism of commutative rings. Then there is an induced map*

$$\mathcal{K}(f) : \mathcal{K}(R) \longrightarrow \mathcal{K}(S)$$

unique upto homotopy w.r.t. commutativity of the following square:

$$\begin{tikzcd}
\mathcal{K}(R) \arrow[r, dashed, "\mathcal{K}(f)"] & \mathcal{K}(S) \\
K_0(R) \times \mathrm{BGL}(R) \arrow[u, "\mathrm{id}\times i"] \arrow[r, "\mathrm{id}\times Bf"'] & K_0(S) \times \mathrm{BGL}(S) \arrow[u, "\mathrm{id}\times j"']
\end{tikzcd}.$$

We denote the maps in K_n-groups by

$$f_* : K_n(R) \to K_n(S).$$

Proof. As f induces a group homomorphism $f : \mathrm{GL}_p(R) \to \mathrm{GL}_p(S)$ by taking a matrix A to $f(A)$ where f is applied on each entry and $f(A)$ is invertible as product of matrices is a polynomial in each entry. Taking direct limits both side yields map

$$f : \mathrm{GL}(R) \to \mathrm{GL}(S).$$

Applying $B(-)$ yields a map

$$Bf : \mathrm{BGL}(R) \to \mathrm{BGL}(S)$$

such that $(Bf)_* = f$. Note that f maps $E(R)$ into $E(S)$. It follows by Lemma 17.4.1.18 that we have a map

$$\tilde{Bf} : \mathrm{BGL}(R)^+ \to \mathrm{BGL}(S)^+$$

unique upto homotopy. As we already have a map $f_* : K_0(R) \to K_0(S)$ by Construction 17.1.1.4, it follows we have a continuous map

$$\mathcal{K}(f) : \mathcal{K}(R) \longrightarrow \mathcal{K}(S)$$

unique upto homotopy, as required. □

An important property of $\mathrm{BGL}(R)^+$ which will be exploited later is that it has a homotopy associative, unital and commutative operation.

Theorem 17.4.1.22 (H-spaces & $\mathrm{BGL}(R)^+$)**.** *Let R be a ring. Then $\mathrm{BGL}(R)^+$ is a homotopy commutative H-group.*

17.4.2 Cup product in K-theory

Construction 17.4.2.1 (Loday's product)**.** Let R, S be two rings. We wish to define a product ($\otimes$ over $\mathbb{Z}$)

$$K_p(R) \otimes K_q(S) \longrightarrow K_{p+q}(R \otimes S).$$

Indeed, recall that we have a product on homotopy groups induced by smash product

$$\pi_p(X) \otimes \pi_q(Y) \longrightarrow \pi_{p+q}(X \wedge Y)$$

for any spaces X and Y. Hence it is sufficient to define a map

$$\psi : \mathrm{BGL}(R)^+ \wedge \mathrm{BGL}(S)^+ \longrightarrow \mathrm{BGL}(R \otimes S)^+.$$

This is the main content of Loday's construction.

To construct ψ, we first construct a map

$$\tilde{\varphi}_{pq} : \mathrm{BGL}_p(R) \times \mathrm{BGL}_q(S) \longrightarrow \mathrm{BGL}_{pq}(R \otimes S)$$

as follows. Note we have a homomorphism

$$\begin{aligned} \theta_{pq} : \mathrm{GL}_p(R) \times \mathrm{GL}_q(S) &\longrightarrow \mathrm{GL}_{pq}(R \otimes S) \\ (A, B) &\longmapsto A \otimes B \end{aligned}$$

where $A \otimes B$ is the usual tensor product of matrices (A and B represent a class in $\mathrm{GL}(R)$ and $\mathrm{GL}(S)$ respectively). In more details, the map is given by

$$\left(\begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} & \cdots \\ b_{21} & b_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \right) \mapsto \begin{bmatrix} a_{11} \otimes b_{11} & a_{11} \otimes b_{12} & \cdots \\ a_{11} \otimes b_{21} & a_{11} \otimes b_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

Now observe that θ maps $E(R) \times E(S)$ into $E(R \otimes S)$, and thus we get maps φ_{pq} and $\tilde{\varphi}_{pq}$ by Theorem 17.4.1.16, unique upto homotopy, so that the following square commutes (see Propsition 17.4.1.15):

$$\begin{tikzcd}
\mathrm{BGL}_p(R)^+ \times \mathrm{BGL}_q(S)^+ \arrow[r, dashed, "\tilde{\varphi}_{pq}"] & \mathrm{BGL}_{pq}(R \otimes S)^+ \\
\mathrm{BGL}_p(R) \times \mathrm{BGL}_q(S) \arrow[r, "\varphi_{pq}"'] \arrow[u, "i \times j"] & \mathrm{BGL}_{pq}(R \otimes S) \arrow[u]
\end{tikzcd}.$$

To get a map in $\mathrm{BGL}_p(R)^+ \wedge \mathrm{BGL}_q(S)^+ \to \mathrm{BGL}_{pq}(R \otimes S)^+$, we need to show that the subspace $\{e\} \times \mathrm{BGL}_q(S)^+ \cup \mathrm{BGL}_p(R)^+ \times \{e\}$ is mapped to the basepoint of $\mathrm{BGL}_{pq}(R \otimes S)^+$ under the map $\tilde{\varphi}_{pq}$. However, the map $\tilde{\varphi}_{pq}$ may not satisfy this condition. So we exploit the H-space structure of $\mathrm{BGL}(R \otimes S)^+$. First, consider the map $\tilde{\varphi}$ with domain and codomain stabilized, which exists by another application of Theorem 17.4.1.16:

$$\begin{tikzcd}
\mathrm{BGL}(R)^+ \times \mathrm{BGL}(S)^+ \arrow[r, dashed, "\tilde{\varphi}"] & \mathrm{BGL}(R \otimes S)^+ \\
\mathrm{BGL}_p(R)^+ \times \mathrm{BGL}_q(S)^+ \arrow[r, "\tilde{\varphi}_{pq}"'] \arrow[u] & \mathrm{BGL}_{pq}(R \otimes S)^+ \arrow[u]
\end{tikzcd}.$$

Now construct the following map

$$\begin{aligned}
\psi : \mathrm{BGL}(R)^+ \times \mathrm{BGL}(S)^+ &\longrightarrow \mathrm{BGL}(R \otimes S)^+ \\
(x, y) &\longmapsto \tilde{\varphi}(x, y) - \tilde{\varphi}(x, \mathrm{pt.}) - \tilde{\varphi}(\mathrm{pt.}, y)
\end{aligned}$$

where pt. denotes the relevant basepoint and the addition is the homotopy commutative group operation on $\mathrm{BGL}(R \otimes S)^+$ by Theorem 17.4.1.22. Note that this map ψ is constant on the subspace $\mathrm{BGL}(R)^+ \wedge \mathrm{BGL}(S)^+$, so that it finally induces a map (which we again call ψ)

$$\psi : \mathrm{BGL}(R)^+ \wedge \mathrm{BGL}(S)^+ \longrightarrow \mathrm{BGL}(R \otimes S)^+.$$

This is the required map ψ. Thus, we get a map

$$\pi_p(\mathrm{BGL}(R)^+) \otimes \pi_q(\mathrm{BGL}(S)^+) \longrightarrow \pi_{p+q}(\mathrm{BGL}(R)^+ \wedge \mathrm{BGL}(S)^+) \xrightarrow{\psi_*} \pi_{p+q}(\mathrm{BGL}(R \otimes S)^+),$$

which in other words is

$$K_p(R) \otimes K_q(S) \longrightarrow K_{p+q}(R \otimes S),$$

as required.

The following is then the main theorem, which is suggestive of the heuristic that Loday's product behaves like cup product in cohomology. That is, the K-theory of schemes will come equipped with a cup product[10]!

Theorem 17.4.2.2 (Loday's theorem)**.** *Let R, S be rings.*

1. *The Loday's product $\psi : K_p(R) \otimes K_q(S) \to K_{p+q}(R \otimes S)$ is an associative and bilinear map.*

[10] and hence suggests how to do intersection theory on non-smooth schemes.

2. *If R is commutative, then the following Loday's product:*

$$K_p(R) \otimes K_q(R) \xrightarrow{\psi} K_{p+q}(R \otimes_{\mathbb{Z}} R) \longrightarrow K_{p+q}(R)$$

where $R \otimes_{\mathbb{Z}} R \to R$ is the structure map of ring R, makes $K(R) = \bigoplus_{i=0}^{\infty} K_i(R)$ into a graded-commutative ring. That is, if $x \in K_p(R)$ and $y \in K_q(R)$, then

$$x \cdot y = (-1)^{pq} y \cdot x.$$

17.4.3 Relative exact sequence for K_n

One of the main benefits of defining higher K-groups as homotopy groups of the K-theory space is that we get the familiar relative exact sequence for free.

Definition 17.4.3.1 (Relative K_n). Let R, S be commutative rings and $f : R \to S$ be a ring homomorphism. Then we get a map $\mathcal{K}f : \mathcal{K}(R) \to \mathcal{K}(S)$. Consider the homotopy fiber

$$F(\mathcal{K}f) \longrightarrow \mathcal{K}(R) \xrightarrow{\mathcal{K}f} \mathcal{K}(S).$$

Denote

$$K_i(f) := \pi_i(F(\mathcal{K}f))$$

which we call the relative K-group of map f. Note that then by homotopy l.e.s. associated to a map, we get the following:

$$\begin{tikzcd}
K_{i-1}(f) \arrow[r] & K_{i-1}(R) \arrow[r, "f_*"] & K_{i-1}(S) \arrow[dll] \\
K_i(f) \arrow[r] & K_i(R) \arrow[r, "f_*"] & K_i(S)
\end{tikzcd}.$$

Lemma 17.4.3.2. *Let R be a commutative ring, $I \leq R$ be an ideal and $\pi : R \twoheadrightarrow R/I$ be the corresponding quotient map. Then,*

$$K_0(R, I) \cong K_0(\pi)$$

where $K_0(R, I)$ is as defined in Definition 17.1.3.4.

Proof. Note $K_0(\pi) = \pi_0(F\mathcal{K}\pi)$, where we have the following fiber sequence

$$F\mathcal{K}\pi \to K_0(R) \times \mathrm{BGL}(R)^+ \overset{\mathcal{K}\pi}{\to} K_0(R/I) \times \mathrm{BGL}(R/I)^+.$$

By definition, we have

$$F\mathcal{K}\pi = \mathrm{Ker}\left(\pi_* : K_0(R) \to K_0(R/I)\right) \times F\widetilde{B\pi}$$

where $\widetilde{B\pi} : \mathrm{BGL}(R)^+ \to \mathrm{BGL}(R/I)^+$. Thus, we have

$$K_0(\pi) = \mathrm{Ker}\,(\pi_*) \times \pi_0(F\widetilde{B\pi}).$$

On the other hand, consider $p : R \oplus I \to R$ as in Definition 17.1.3.4. We then get the fiber sequence

$$F\mathcal{K}p \to K_0(R \oplus I) \times \mathrm{BGL}(R \oplus I)^+ \stackrel{\mathcal{K}p}{\to} K_0(R) \times \mathrm{BGL}(R)^+.$$

Again, we have

$$\begin{aligned} F\mathcal{K}p &= \mathrm{Ker}\,(p_* : K_0(R \oplus I) \to K_0(R)) \times F\widetilde{Bp} \\ &= K_0(R, I) \times F\widetilde{Bp} \end{aligned}$$

where $\widetilde{Bp} : \mathrm{BGL}(R \oplus I)^+ \to \mathrm{BGL}(R)^+$. Thus, we have

$$K_0(p) = K_0(R, I) \times \pi_0(F\widetilde{Bp}).$$

We claim that the fiber $F\widetilde{Bp}$ is connected, so that $K_0(p) = K_0(R, I)$. Indeed, observe by homotopy l.e.s. corresponding to $F\mathcal{K}p$ that it suffices to show that $p_* : K_1(R \oplus I) \to K_1(R)$ is a surjection. This follows immediately from functoriality of π_1 and the splitting $R \to R \oplus I \to R$. This shows that

$$K_0(p) = K_0(R, I).$$

We thus reduce to showing that $K_0(p) \cong K_0(\pi)$. **TODO.** □

17.4.4 Finite coefficients

Construction 17.4.4.1 (Moore spectrum)**.** Let G be an abelian group. We have an associated suspension CW-spectrum called Moore spectrum $P^\infty G$ whose n^{th}-term is $P(G, n)$, the unique homotopy type whose n^{th} reduced cohomology is G and rest are 0. Recall that for any compactly generated based spaces X, Y, the based homotopy classes of maps $[\Sigma X, Y]$ is a group and $[\Sigma^2 X, Y]$ is an abelian group. Since $P(G, n) \simeq \Sigma^{n-1} P(G, 1)$, it follows at once that

$$[P(G, n), X] = [\Sigma^{n-2} P(G, 2), X]$$

is a group for $n = 3$ and an abelian group for $n \geq 4$. For $n = 2$ it may not be a group.

Using Moore spectrum, we define mod l homotopy groups for any integer l as follows.

Definition 17.4.4.2 ($\pi_n(X; \mathbb{Z}/l)$)**.** Let X be a based CW-complex and $P^\infty \mathbb{Z}/l$ be the Moore spectrum for group $\mathbb{Z}/l$ for some integer l. Then mod l homotopy groups of X are defined as the following homotopy class of maps

$$\pi_n(X; \mathbb{Z}/l) = [P(\mathbb{Z}/l, n), X].$$

As noted in Construction 17.4.4.1, $\pi_3(X; \mathbb{Z}/l)$ is a group and $\pi_n(X; \mathbb{Z}/l)$ is an abelian group for $n \geq 4$. For a map $f : X \to Y$, we get by composition maps $f_* : \pi_n(X; \mathbb{Z}/l) \to \pi_n(Y; \mathbb{Z}/l)$ which is a group homomorphism for $n \geq 3$. Thus $\pi_n(-; \mathbb{Z}/l)$ is a functor on $\mathbf{CW}_*$.

Remark 17.4.4.3 (Main facts about mod l homotopy)**.** The following three facts is what we shall use in the discussion later, most of which are quite familiar.

1. Every fibration sequence $F \to E \to B$ induces a long exact sequence in mod l homotopy groups.
2. For every complex X, there is a s.e.s.
$$0 \to \pi_n(X) \otimes \mathbb{Z}/l \to \pi_n(X;\mathbb{Z}/l) \to T_l(\pi_{n-1}(X)) \to 0$$
which is split exact if $l \neq 2 \mod 4$, where $T_l(G) = \{g \in G \mid l \cdot g = 0\}$.
3. There is a natural mod l Hurewicz map
$$h_i : \pi_i(X;\mathbb{Z}/l) \longrightarrow H_i(X;\mathbb{Z}/l)$$
which is an isomorphism for all $1 \leq i \leq n$ if X is an $(n-1)$ connected nilpotent space.

All these can results in more generality can be found in [**NeiPHT**].

The following observation tells us why we are looking at mod l-homotopy groups.

Proposition 17.4.4.4. *Let X be an H-space and consider the map*
$$m_l : X \to X$$
which is multiplication by $l \in \mathbb{Z}$[11]*. Denote F to be the homotopy fiber of the map m_l. Then*
$$\pi_n(X;\mathbb{Z}/l) \cong \pi_{n-1}(F)\ \forall n \geq 2.$$

Proof. A model of Moore space $P^{n+1}(\mathbb{Z}/l)$ is obtained by gluing an $n+1$-cell to S^n by a degree l-map. Equivalently, this CW-complex is homeomorphic to the homotopy cofiber
$$S^n \xrightarrow{f_l} S^n \longrightarrow Cf_l \cong P^{n+1}(\mathbb{Z}/l)$$
where f_l is a degree l map. Recall that the inclusion $S^n \to Cf_l$ is a cofibration, thus the above is a short cofiber sequence. We also have a short fiber sequence
$$F \longrightarrow X \xrightarrow{m_l} X$$
where $F \to X$ is a fibration.

It is well-known that $\mathrm{Map}_*(-,X) : \mathbf{Top}_*^{cg} \to \mathbf{Top}_*^{cg}$ takes cofibrations to fibrations and $\mathrm{Map}_*(X,-)$ takes fibrations to fibrations. Furthermore, for any based space X, we have the smash-map duality (akin to $\otimes$-hom duality):
$$\pi_k(\mathrm{Map}_*(S^n,X)) = [S^k, \mathrm{Map}_*(S^n,X)] \cong [S^k \wedge S^n, X] \cong [S^{k+n}, X] = \pi_{k+n}(X).$$
Similarly for $P^{n+1}(\mathbb{Z}/l)$. Now with this information, we apply $\mathrm{Map}_*(-,X)$ onto the first cofiber sequence and $\mathrm{Map}_*(S^n,-)$ onto the second fiber sequence to get two fiber sequences for homotopic maps c(S^n is an H-cogroup and X is an H-group):
$$\mathrm{Map}_*(P^{n+1}(\mathbb{Z}/l),X) \to \mathrm{Map}_*(S^n,X) \xrightarrow{-\circ f_l} \mathrm{Map}_*(S^n,X)$$
$$\mathrm{Map}_*(S^n,F) \to \mathrm{Map}_*(S^n,X) \xrightarrow{m_l\circ -} \mathrm{Map}_*(S^n,X).$$
Hence $\mathrm{Map}_*(P^{n+1}(\mathbb{Z}/l),X) \simeq \mathrm{Map}_*(S^n,F)$, which yields the proof after taking homotopy groups and using the above mentioned fact. $\square$

[11]this makes sense as ΩA has a canonical H-space structure.

Definition 17.4.4.5 ($K_n(R;\mathbb{Z}/l)$)**.** Let $l \in \mathbb{Z}$ be an integer and R be a ring. Then mod l K-groups of R are defined as

$$K_n(R) := \pi_n(\mathrm{BGL}(R)^+;\mathbb{Z}/l).$$

Construction 17.4.4.6 (Bott element in $K_2(R;\mathbb{Z}/l)$)**.** Let R be any ring containing a primitive l^{th}-root of unity ζ. Note that the class of the diagonal matrix consisting of ζ in $\mathrm{GL}(R)/E(R) = K_1(R)$ is l-torsion and is thus in $T_l(K_1(R))$, which we denote by $[\zeta] \in K_1(R)$. Then, by universal coefficients theorem (Remark 17.4.4.3, item 2), we get an l-torsion element $\beta \in K_2(R;\mathbb{Z}/l)$ which maps to $\zeta \in K_1(R)$. This element β is called a Bott element of $K_2(R;\mathbb{Z}/l)$.

By the calculation for finite fields done by Quillen, one can deduce the following by basic algebra.

Theorem 17.4.4.7. *Let p be a prime and consider the algebraic closure $\bar{\mathbb{F}}_p$.*

1. *We have*

$$K_n(\bar{\mathbb{F}}_p) = \begin{cases} (\mathbb{Q}/\mathbb{Z})[1/p] & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

2. *We have for $l \in \mathbb{Z}$ coprime to p*

$$K_n(\bar{\mathbb{F}}_p;\mathbb{Z}/l) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \mathbb{Z}/l & \text{if } n \text{ is even.} \end{cases}$$

17.5 K-theory & étale cohomology

We discuss some fundamental connections between K-theory and algebraic geometry.

17.5.1 The case of fields : K_2 & Galois cohomology

In §4.4, we constructed a representation of $K_2(F)$ in the Brauer group $\mathrm{Br}(F)$ for fields which contains a primitive root of unity. Our goal in this section is to construct Galois symbol for fields not necessarily containing primitive roots of unity. We will thus replace the role of ζ by $\mu_m(F)$, the group of m^{th}-roots of unity in F.

The main idea is to replace Brauer group from the codomain of φ of Theorem 17.3.4.6 to an object which is more cohomological w.r.t. F and thus will have bilinear pairing from $F^\times \otimes F^\times$, which will, hopefully, satisfy Matsumoto relations. Indeed, such an object exists and is the content of Galois cohomology.

Construction 17.5.1.1 (Galois cohomology)**.** Let F be a field. Denote F_{sep} to be the separable closure of F (in the algebraic closure). Thus, F_{sep}/F is a separable normal extension and hence Galois. Let $G = \mathrm{Gal}\left(F_{\text{sep}}/F\right)$ be the Galois group. As F_{sep}/F may not be finite, thus G may not be a finite group. Recall the fundamental Galois theorem for finite case

$$\begin{array}{c} \left\{L \mid F_{\text{sep}}/L/F \text{ is a finite intermediate extension}\right\} \\ F_{\text{sep}}^{(-)}\big\uparrow \quad \big\downarrow \mathrm{Gal}\left(F_{\text{sep}}/-\right) \\ \{H \mid H \le G \text{ is a subgroup}\} \end{array}.$$

Our goal is to study this relationship more "homologically" in the the case of separable closure. In particular, note that G here is acting on field F_{sep} by the obvious action, and the whole classical Galois theory is the study of this action together with its orbits and stabilizers.

Consider the following collection of subgroups of G:

$$G_E = \text{Gal}\left(F_{\text{sep}}/E\right),\ F_{\text{sep}}/E/F \ \&\ E/F \text{ is finite.}$$

Observe that $\{G_E\}_{F_{\text{sep}}/E/F}$ forms a basis of topology on G, as we have

$$\bigcup_{F_{\text{sep}}/E/F} G_E = G \qquad [\text{as } E = F \text{ is possible}]$$
$$G_E \cap G_{E'} \supseteq G_{E\cdot E'} \qquad [\text{by definition}].$$

Thus G is a topological group. We will now study left modules over the group ring $\mathbb{Z}[G]$, also called G-modules. Note that for any G-module M, we will have a left multiplication map $G \times M \to M$. We will call M *discrete* if this map is continuous, where M has discrete topology.

Now fix a G-module M and as usual in Galois theory, consider the G-invariant subgroup $M^G := \{m \in M \mid g \cdot m = m\ \forall g \in G\}$. Observe that invariant subgroup construction is functorial on the category of discrete G-modules:

$$\begin{aligned}(-)^G : \mathbf{Mod}_c(G) &\longrightarrow \mathbf{Ab}\\ M &\longmapsto M^G\end{aligned}$$

and for a map of G-modules $f : M \to N$, we get a map $f^G : M^G \to N^G$ as $f(g \cdot m) = g \cdot f(m)$ for all $m \in M$. An easy observation tells us that $(-)^G$ is also left-exact. Observe that the category of discrete G-modules is abelian with enough injectives (Lemma 6.11.10 of **[WeibHA]**). We may thus right derive the functor $(-)^G$, to obtain the *Galois cohomology groups*

$$H^i_{\text{ét}}(F; M) := (R^i(-)^G)(M).$$

Observe that $H^0_{\text{ét}}(F; M) = M^G$ (Lemma 19.2.3.3).

Remark 17.5.1.2 (G-modules)**.** Consider the notation of Construction 17.5.1.1. Here are some elementary examples of discrete G-modules:

1. $\mathbb{G}_m := F_{\text{sep}}^\times$, the abelian group of units is a discrete G-module as $\sigma \in G$ acts on $x \in \mathbb{G}_m$ by $(\sigma, x) \mapsto \sigma(x)$ and since σ is an automorphism, thus every fiber of the above map is open, thus the action being continuous.
2. μ_n, the subgroup of $\mathbb{G}_m$ of all n^{th}-roots of unity. This is discrete for the same reason as above.
3. $M \otimes_{\mathbb{Z}} N$, the tensor product of two G-modules. Indeed, we define the action of $\sigma \in G$ on a simple tensor $m \otimes n \in M \otimes_{\mathbb{Z}} N$ by $(\sigma, m \otimes n) \mapsto (\sigma \cdot m) \otimes (\sigma \cdot n)$. This is continuous since the individual actions on M and N are continuous. In particular, $\mu_n^{\otimes 2} = \mu_n \otimes \mu_n$ is a discrete G-module.

The cohomology of G-groups $\mathbb{G}_m, \mu_n$ and $\mu_n^{\otimes 2}$ is very interesting. We begin by seeing this for μ_n.

Remark 17.5.1.3. Let F be a field, F_{sep} be its separable closure and $G = \text{Gal}\left(F_{\text{sep}}/F\right)$. Denote $\mu_n \leq \mathbb{G}_m$ to be the group of all n^{th}-roots of unity in F_{sep}. Consider the map $\mathbb{G}_m \to \mathbb{G}_m$ mapping $g \mapsto$

Construction 17.5.1.4 (Kummer sequence)**.** Let F be a field of characteristic $p > 0$ and let n be coprime to p. Denote F_{sep} to be the separable closure of F and $\mathbb{G}_m = F_{\text{sep}}^{\times}$. We obtain the following short exact sequence:

$$1 \longrightarrow \mu_n \longrightarrow \mathbb{G}_m \stackrel{g \mapsto g^n}{\longrightarrow} \mathbb{G}_m \longrightarrow 1.$$

Indeed, the only part that needs to be shown is the surjectivity of the above map. To this end, pick any $a \in \mathbb{G}_m$ and consider the polynomial $x^n - a \in F_{\text{sep}}[x]$. As the derivative of $x^n - a$ is nx^{n-1} which is not zero as $p \not| n$, thus $x^n - a$ is separable. Let $b \in \overline{F}$ be a root of $x^n - a$. By Proposition 16.6.8.7, it follows that $\overline{F}/F_{\text{sep}}$ is purely inseparable. If $b \notin F_{\text{sep}}$, then the minimal polynomial of b in F_{sep} is $m_{b,F_{\text{sep}}}(x) = x^{p^k} - c$ for some $c \in F_{\text{sep}}$ (Theorem 16.6.8.3). As $m_{b,F_{\text{sep}}}(x) | x^n - a$ and $m_{b,F_{\text{sep}}}$ is not separable (in-fact it has only one root which repeats), therefore $x^n - a$ is not separable as well, a contradiction.

The above short-exact sequence is called the Kummer sequence of F.

As the above short exact sequence is that of discrete G-modules, therefore in right derived functors we get a long exact sequence.

Lemma 17.5.1.5 (Kummer cohomology sequence)**.** *Let F be a field of characteristic $p > 0$ and let n be coprime to p. Denote F_{sep} to be the separable closure of F, $\mathbb{G}_m = F_{\text{sep}}^{\times}$ and $G = \mathrm{Gal}\left(F_{\text{sep}}/F\right)$. Then the following is a long exact sequence:*

$$\begin{tikzcd}
1 \arrow[r] & \mu_n(F) \arrow[r] & F^{\times} \arrow[r, "g \mapsto g^n"] & F^{\times} \arrow[dll] \\
& H^1_{\text{ét}}(F; \mu_n) \arrow[r] & H^1_{\text{ét}}(F; \mathbb{G}_m) \arrow[r] & H^1_{\text{ét}}(F; \mathbb{G}_m) \arrow[dll] \\
& H^2_{\text{ét}}(F; \mu_n) \arrow[r] & H^2_{\text{ét}}(F; \mathbb{G}_m) \arrow[r] & H^2_{\text{ét}}(F; \mathbb{G}_m) \arrow[dll, no head] \\
& {} & &
\end{tikzcd}.$$

Proof. Follows from Theorem 19.2.3.5. ∎

Chapter 18

Abstract Analysis

Contents

We will discuss some topics from integration theory and functional analysis (topological vector spaces).

18.1 Introduction

We would like to state and portray the uses of some of the important and highly usable results of integration theory, elucidating in the process the analytical thought which is of paramount importance in any route of exploration in this field[1]. We give bare-bone proofs as all this is standard, but we will highlight the main part of the proof by ♡ or if there are many main parts, then by ♡♡ . . . (!) Let us first begin with some motivation behind modern measure theory.

[1]One may argue, instead, in whole of mathematics.

We know that the class of all Riemann integrable functions on $[a, b]$, denoted $R([a, b])$, is not complete under pointwise limit (a sequential approximation of Dirichlet's function shows that). Further, motivated by Weierstrass approximation, one would like to have commutability results between lim and $\int$, which again $R([a, b])$ lacks. Consequently, one is motivated to find a larger class of "integrable" functions for which these defects would be rectified.

The idea that H. Lebesgue had was quite simple. He continued the idea of Riemann (that is, of partitions) but made sure that the function under investigation is much more intertwined in with it. Indeed, for a bounded function $f : [a, b] \to \mathbb{R}$, we contain the image $\mathrm{Im}\,(f) \subseteq [\alpha, \beta]$ and then consider a partition $\mathcal{P} = \{I_i\}_{i=1}^n$ where I_i is an interval. Now choose $\xi_i \in f^{-1}(I_i) =: J_i$ for each i. Consequently, we may naturally define *Lebesgue sum of f w.r.t. $\mathcal{P}$* as follows

$$L(f, \mathcal{P}) := \sum_{i=1}^{n} f(\xi_i) m(J_i),$$

where $m(J_i)$ is supposed to be some sort of measure of J_i. Note that J_i in general might be very bad (may not even be an interval!). To complete this idea of "integration", we are naturally led to considering more general notions of measures. Indeed, this is what we will pursue in this course.

Remark 18.1.0.1. (***Pseudo-definition of measure***) First, what do we expect from a notion of measure on $\mathbb{R}$? Perhaps the following is the minimum conditions we would require to call a function "measure": A function $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ is said to be a *pseudo-measure* if it satisfies the following

1. (*measure of intervals*) for any interval I, the measure $\mu(I) = l(I)$ where l is the length function,
2. (*measure of disjoint unions*) for any disjoint sequence of subsets $\{A_n\}$, $\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$,
3. (*translation invariance*) for any subset A and $x \in \mathbb{R}$, we have $\mu(A + x) = \mu(A)$.

We will call such a function a *pseudo-measure on* $\mathbb{R}$. Observe that for $A \subseteq B$, we obtain $\mu(A) \leq \mu(B)$ by breaking $B = A \cup B \setminus A$. We call μ a pseudo-measure because it does not exists!

Theorem 18.1.0.2. *(Vitali set) There exists no pseudo-measure on $\mathbb{R}$. In paritcular, there exists a set $V \subseteq \mathbb{R}$ such that for a pseudo-measure μ, $\mu(V) \notin [0, \infty]$.*

Proof. We will construct such a set V. Begin with the closed interval $J = [0, 1]$. Define an equivalence relation $\sim$ on J given as follows:

$$x \sim y \iff x - y \in \mathbb{Q}.$$

This can easily be seen to be an equivalence relation on J. We have first some observations to make about this equivalence relation and the consequent partition of J that it entails.

1. Observe that the class of any rational r in J under $\sim$ is simply $[0]$, as $r - 0 \in \mathbb{Q}$.
2. Every equivalence class is countable in size. Indeed, for any $x \in J$, the class $[x]$ is just translate of x by rationals, which is countable.
3. There are uncountably many equivalence classes. Indeed, if there were atmost countably many equivalence classes, then by statement 2 above, it would follow there are atmost countably many elements in J, which is a contradiction.

Consequently, this equivalence relation partitions J into following classes:

$$J = \bigcup_{\alpha \in \mathcal{I}} [\alpha]$$

where $\mathcal{I}$ is an uncountable set.

We would now construct the set V as follows. First, let us assume axiom of choice, so that for each class $[\alpha]$, we may pick an element $r_\alpha \in [\alpha]$ and would thus obtain a subset of J, denoted $V = \{r_\alpha \mid \alpha \in I\}$. We call this the Vitali set.

Consider the set $Q = [-1,1] \cap \mathbb{Q}$. Since it is countable so consider an enumeration $Q = \{q_n\}$. Now consider the translates $V + q_n$ for all $n \in \mathbb{N}$ and their union $X = \bigcup_n V + q_n$. We now observe the following two facts about X.

1. If $n \neq m$, then $(V+q_n) \cap (V+q_m) = \emptyset$. Indeed, if $x \in (V+q_n) \cap (V+q_m)$, then $x = r_a + q_n = r_b + q_m$. Consequently, $r_a - r_b \in \mathbb{Q}$ and hence $[a] = [b]$. But by single choice of r_c for each $c \in \mathcal{I}$, we get $r_a = r_b$ and thus $q_n = q_m$ from above, which is a contradiction.
2. $J = [0,1] \subseteq X$. Indeed, for any $x \in [0,1]$, consider the class $[a]$ in which x is present. Consequently we have a unique $r_a \in V$ corresponding to x which satisfies $x \in [r_a]$. Thus, $x = r_a + t$ where $t \in \mathbb{Q}$. We may write $t = q_n$ to obtain that $x \in V + q_n$, as desired.
3. $X \subseteq [-1,2]$. Indeed, this follows immediately since $X = \bigcup_n V + q_n$ where q_ns are rationals in $[-1,1]$ and $V \subseteq [0,1]$.

With the above three observations, we obtain the following inclusions:

$$[0,1] \subseteq \bigcup_n V + q_n \subseteq [-1,2].$$

Now, if we apply the pseudo-measure μ on the above inclusions, we will obtain the following:

$$1 \leq \sum_n \mu(V) \leq 3.$$

If $\mu(V) = 0, \infty$, then we have an immediate contradiction. Else if $0 < \mu(V) < \infty$, then $\sum_n \mu(V) = \infty$ and we again have a contradiction. Thus, $\mu(V) \notin [0,\infty]$, a contradiction. □

Remark 18.1.0.3. The main issue in pseudo-measures is that we trying to get a measure on *all* of the subsets of $\mathbb{R}$. By Theorem 18.1.0.2, this is hopeless. What we shall now do instead is to obtain a measure not on all of the subsets of $\mathbb{R}$, but rather on only a subcollection of subsets of $\mathbb{R}$, and we shall choose this subcollection in a manner so that we don't allow sets like Vitali sets. Indeed, this becomes our point of departure for the abstract definition of σ-algebras and measure/measurable spaces, the need for the right domain of a measure function.

18.1.1 Few introductory notions

These are few of the basic definitions that one might remember from real analysis.

- **Limit Points** : $x \in X$ is called a *limit point* of a subset $S \subseteq X$ if $\forall\, r > 0$, $\exists\, a \neq x$ such that $a \in S \cap B_r(x)$. That is, ball of any size r around x contains atleast one point of S.
- **Isolated Points** : $y \in S$ is called an *isolated point* of a subset $S \subseteq X$ if $\exists\, r > 0$ such that $(B_r(y) \setminus \{y\}) \cap S = \Phi$. That is, $B_r(y)$ contains no other point of S apart from y.
  - Also note that every point of closure $\overline{S}$ is either a limit point or an isolated point of S.
  - More specifically, *any subset of $\mathbb{R}^d$ is closed if and only if it contains all of it's limit points*.

- **Perfect Set** : A is called a perfect set if $A = A'$ where A' is the set of all *limit points* of A. More conveniently, if A does not contain any isolated points then it is a perfect set. $\mathbb{R}$ is a perfect set.
- **Symmetric Difference** : A and B are two sets then symmetric difference is $A \Delta B = (A \setminus B) \cup (B \setminus A)$.
- **Power Set** : Collection of all subsets of a set S, written as $P(S)$.
- **Lower Bound** : A *lower bound* of a subset S of a poset $(P, \leq)$ is an element $a \in P$ such that $a \leq x$ for all $x \in S$.
- **Infimum** : A lower bound $p \in P$ is called an *infimum* of S if for all lower bounds y of S in P, $y \leq p$.
- **Limit Infimum** : For a sequence $\{x_n\}$, limit inferior is defined by:

$$\begin{aligned} \liminf_{n\to\infty} x_n &= \varprojlim_{n\to\infty} \left(\inf_{m\geq n} x_m \right) \\ &= \sup_{n\geq 0} \inf_{m\geq n} x_m \\ &= \sup\{\inf\{x_m \mid m \geq n\} \mid n \geq 0\}. \end{aligned} \tag{18.1}$$

- **Upper Bound** : An *upper bound* of a subset S of a poset P is an element $b \in P$ such that $b \geq x$ for all $x \in S$.
- **Supremum** : An upper bound $u \in P$ is called a *supremum* of S if for all upper bounds z of S in P, $z \geq u$.
- **Limit Supremum** : For a sequence $\{x_n\}$, limit supremum is defined by:

$$\begin{aligned} \limsup_{n\to\infty} x_n &= \varprojlim_{n\to\infty} \left(\sup_{m\geq n} x_m \right) \\ &= \inf_{n\geq 0} \sup_{m\geq n} x_m \\ &= \inf\{\sup\{x_m \mid m \geq n\} \mid n \geq 0\} \end{aligned} \tag{18.2}$$

- **Limit** : Consider the sequence $\{x_n\}$ in $[-\infty, +\infty]$, then $\varprojlim_n x_n$ is defined as

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n := \varprojlim_{n\to\infty} x_n.$$

- **Lower Sum** : $l(f, \mathcal{P})$ is the sum of the minimum functional values at the partition. That is,

$$l(f, \mathcal{P}) = \sum_{i=0}^{n-1} m_i (a_{i+1} - a_i)$$

where $m_i = \inf\{f(x) \mid x \in [a_{i-1}, a_i]\}$.
- **Upper Sum** : Similarly,

$$u(f, \mathcal{P}) = \sum_{i=0}^{n-1} M_i (a_{i+1} - a_i)$$

where $M_i = \sup\{f(x) \mid x \in [a_{i-1}, a_i]\}$.
Remember that the function is *Riemann Integrable* if $l(f, \mathcal{P}) = u(f, \mathcal{P})$.

- **Countable Sets** : Note the following,
  1. *Cardinality* : Sets X and Y have the same cardinality if there exists a bijection from X to Y.
  2. *Finite Set* : A set is finite if it is empty or it has the same cardinality as $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$.
  3. *Countably Infinite* : If the set has the same cardinality as $\mathbb{N}$.
  4. *Enumeration* : An enumeration of a countably infinite set X is a bijection of $\mathbb{N}$ onto X. That is, an enumeration is an infinite sequence $\{x_n\}$ such that each of the x_i's are in X and each element of X is x_i for some i.
  5. *Countable* : A set is countable if it is finite or countably infinite. For example, $\mathbb{N}$ is countable, $\mathbb{Q}$ is also countable (!), $\mathbb{R} \setminus \mathbb{Q}$ (irrationals) is not countable, $\mathbb{R}$ is not countable.
- **Totally Bounded** : A subset $B \subseteq X$ is *totally bounded* when it can be covered by a finite number of r-balls for all $r > 0$. That is,

$$\forall\, r > 0,\ \exists N \in \mathbb{N},\ \exists a_1, \ldots, a_N \in X \text{ such that } B \subseteq \bigcup_{n=1}^{N} B_r(a_n)$$

- **Compact Set** : A set K is said to be *compact* when given any cover of balls of possibly unequal radii, there is a finite sub-collection of them that still covers the set K. That is,

$$K \subseteq \bigcup_i B_{r_i}(a_i) \implies \exists\, i_1, \ldots, i_N,\ \ K \subseteq \bigcup_{n=1}^{N} B_{r_{i_n}}(a_{i_n})$$

  Note that *compact metric spaces are totally bounded (!)*. Also, *compact sets are closed*.

The problem begins with Riemann Integrable functions when we see that functions like Dirichlet function (1 on irrational and 0 on rational points) can become *measurable* even when the function is not continuous! This motivates the need of a formal notion of a *measure*.

We begin with some recollections from classical analysis of one real variable.

1. Every open set in $\mathbb{R}$ can be written as disjoint union of open intervals.

   Proof. Let $G \subseteq \mathbb{R}$ be a open subset. Now by definition of an open subset, we have that for any $x \in G$, there exists atleast one open subset U such that $x \in U \subseteq G$. Now consider the following union of all such open subsets of x,

$$U_x = \bigcup_{x \in U \subseteq G} U$$

   It's now easy to see that U_x is the largest such subset of G, as any other $V \subseteq G$ such that $x \in V$ is by definition contained in U_x. Moreover, U_x is an interval as it is an arbitrary union of open intervals. Now, define the following relation on G:

$$y \sim x \iff y \in U_x$$

   Now we clearly have that $x \in U_x$ (reflexive); for $y \sim U_x$ we have $U \subseteq U_x$ such that $x, y \in U$, hence $x \in U_y$ (symmetric); for $x \in U_y$ and $y \in U_z$, we have that $x, y, z \in U_y$, since $z \in U_y \subseteq G$ so $U_y \subseteq U_z$, so $x \in U_z$ (transitive). Hence $\sim$ is an equivalence relation, hence $\sim$ partitions the set G. Denote the set of all equivalence classes as $\mathcal{I}$ so we get

$$G = \bigcup_{I \in \mathcal{I}} I$$

such that $I_1 \cap I_2 = \Phi$ for any $I_1, I_2 \in \mathcal{I}$. Now note that for any $I \in \mathcal{I}$ is open because each I is generated by the relation $\sim$ such that $y \sim x$ iff $y \in U_x$. Hence for any $z \in I$, we have $z \in U_x \subseteq G$ where U_x is open. Therefore, we have $G = \cup_{I \in \mathcal{I}} I$ for disjoint open intervals in $\mathcal{I}$. □

2. Prove that every non-empty perfect subset of $\mathbb{R}$ (or $\mathbb{R}^n$) is uncountable. That is, if $A = A'$ then A is uncountable.

   Proof. Take $A \subseteq \mathbb{R}$ to be a perfect subset. Since A it is perfect, therefore, it must contain all of it's limit points or, equivalently, contains no isolated points. Clearly, then, A cannot be finite, but can only be countably infinite or uncountable. If it is uncountable, then the proof is over. If A is countably infinite, then we can write A as the following :
$$A = \{a_1, a_2, \dots\}.$$
   Construct a ball around a_{i_1} of any radius $r_1 > 0$. Since A is perfect, therefore $\exists\ a_{i_2} \in B_{r_1}(a_{i_1}) \cap A = C_1$. Similarly, for some $r_2 > 0$, we have $a_{i_3} \in B_{r_2}(a_{i_2}) \cap B_{r_1}(a_{i_1}) \cap A = C_2$ such that $a_{i_1} \notin C_2$ and so on. In general, we would have the following,
$$a_{i_{n+1}} \in \left(\bigcap_{j=1}^{n} B_{r_j}(a_{i_j})\right) \cap A = C_n.$$
   Now, consider $C = \cap_n C_n$. Since $C_{n+1} \subseteq C_n$, therefore $C \neq \Phi$. But, $a_i \notin C$ for any $i \in \mathbb{N}$ as $a_i \notin C_{i+1}$. Therefore we have a contradiction. Hence A cannot by countably infinite, it must only be uncountable. □

3. In the definition of Lebesgue Outer measure on $\mathbb{R}$, one can instead take $\mathcal{C}_A$ to be collection of infinite sequences of the any form from $\{[a_n, b_n]\}, \{(a_n, b_n)\}$ or $\{(a_n, b_n]\}$.

   Proof. Refer Proof of Proposition *18.2.5.3*. □

4. Show the following:
$$\bigcup_{n=1}^{N} E_n = \bigcup_{n=1}^{N} \left(E_n \cap \left(\bigcup_{k<n} E_k\right)^c\right)$$
   Proof. Take $x \in \bigcup_{n=1}^{N} E_n$. Then $\exists\ E_k$ for some a such that $x \in E_a$. Now, clearly, $x \in E_a \subseteq (\bigcup_{k<a} E_k)^c$, hence $x \in \left(E_a \cap (\bigcup_{k<a} E_k)^c\right)$. Hence, we have $\bigcup_{n=1}^{N} E_n \subseteq \bigcup_{n=1}^{N} \left(E_n \cap (\bigcup_{k<n} E_k)^c\right)$. The converse is easy to see too. □

18.2 Measures

18.2.1 Algebras & σ-algebras

Definition 18.2.1.1. (**Algebra/Field**) Let X be an arbitrary set. A collection $\mathcal{A} \subseteq P(X)$ of subsets of X is an algebra on X if:

- $X \in \mathcal{A}$.
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$.

- For each finite sequence $A_1, A_2, \ldots, A_n \in \mathcal{A}$ implies that

$$\bigcup_{i=1}^{n} A_i \in \mathcal{A}$$

- For each finite sequence $A_1, A_2, \ldots, A_n \in \mathcal{A}$ implies that

$$\bigcap_{i=1}^{n} A_i \in \mathcal{A}$$

Definition 18.2.1.2. (**σ-Algebra/σ-Field**) Let X be an arbitrary set. A collection $\mathcal{A} \subseteq P(X)$ of subsets of X is a σ-algebra on X if:

- $X \in \mathcal{A}$.
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$.
- For each infinite sequence $\{A_i\}$ such that $A_i \in \mathcal{A}$, it implies that

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$$

- For each infinite sequence $\{A_i\}$ such that $A_i \in \mathcal{A}$, it implies that

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$$

Proposition 18.2.1.3. *Let X be a set. Then the intersection of an arbitrary non-empty collection of σ-algebras on X is a σ-algebra on X.*

Proof. Consider a collection $\mathcal{C}$ of σ-algebras on X. Denote $\mathcal{A} = \bigcap \mathcal{C}$ as intersection of all σ-algebras in $\mathcal{C}$. We can now easily see that any subset in $\mathcal{A}$ would be present in every σ-algebra present in collection $\mathcal{C}$, hence, it would obey all properties of a σ-algebras. Therefore, $\mathcal{A}$ is a σ-algebra. □

Corollary 18.2.1.4. *Let X be a set and let $\mathcal{F} \subseteq P(X)$ be a family of subsets of X. Then there exists a smallest σ-algebra on X that includes $\mathcal{F}$.*

Proof. Consider any given family $\mathcal{F} \subseteq P(X)$ and just take intersection of the family $\mathcal{C}$ of all σ-algebras which contains $\mathcal{F}$ to construct this smallest σ-algebra. □

Definition 18.2.1.5. (**Generated σ-algebra**) The smallest σ-algebra on X containing a given family $\mathcal{F} \subseteq P(X)$ of subsets is called the σ-algebra *generated* by $\mathcal{F}$, denoted as $\sigma(\mathcal{F})$.

Definition 18.2.1.6. (**Borel σ-algebra on $\mathbb{R}^d$**) It is the σ-algebra on $\mathbb{R}^d$ generated by the collection of all open subsets of $\mathbb{R}^d$, denoted as $\mathcal{B}(\mathbb{R}^d)$.

Definition 18.2.1.7. (**Borel Subsets of $\mathbb{R}^d$**) Any $A \subseteq \mathbb{R}^d$ is called a Borel subset of $\mathbb{R}^d$ if $A \in \mathcal{B}(\mathbb{R}^d)$.

Proposition 18.2.1.8. *The Borel σ-algebra on $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$, of Borel subsets of $\mathbb{R}$ is generated by each of the following collection of sets:*

1. *The collection of all closed subsets of $\mathbb{R}$.*
2. *The collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$.*

3. *The collection of all subintervals of* $\mathbb{R}$ *of the form* $(a, b]$.

Proof. To show all of these, consider the three σ-algebras $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ corresponding to conditions 1,2 & 3 respectively and try to prove $\mathcal{A}_3 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_1 \subseteq \mathcal{B}(\mathbb{R})$ together with $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_3$. The first three inclusions are trivial to see. For the case that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_3$, simply note that any open subset can be made by unions of the sets of form $(a, b]$ and by Homework-I,1, each open set is union of open subsets. $\square$

Proposition 18.2.1.9. *The σ-algebra $\mathcal{B}(\mathbb{R}^d)$ of Borel subsets of $\mathbb{R}^d$ is generated by each of the following collections:*
1. *The collection of all closed subsets of* $\mathbb{R}^d$.
2. *The collection of all closed half-spaces in $\mathbb{R}^d$ that have the form $\{(x_1, \dots, x_d) \mid x_i \leq b\}$ for some index i and some $b \in \mathbb{R}$.*
3. *The collection of all rectangles in $\mathbb{R}^d$ that have the form*

$$\{(x_1, \dots, x_d) \mid a_i < x_i \leq b_i \text{ for } i = 1, \dots, d\}$$

Proof. Almost the same as in Proposition 18.2.1.8. $\mathcal{A}_1 \subseteq \mathcal{B}(\mathbb{R}^d)$ trivially by definition. $\mathcal{A}_2 \subseteq \mathcal{A}_1$ as $\{(x_1, \dots, x_d) \mid x_i \leq b\}$ is closed itself. $\mathcal{A}_3 \subseteq \mathcal{A}_2$ by the observation that $\{(x_1, \dots, x_d) \mid a_i < x_i \leq b_i\}$ is made by the difference of two subsets of the form $\{(x_1, \dots, x_d) \mid x_i \leq b_i\}$ and $\{(x_1, \dots, x_d) \mid x_i > a_i\}$, the latter is the complement of a certain subset in $\mathcal{A}_2$, moreover, $\{(x_1, \dots, x_d) \mid a_i < x_i \leq b_i \text{ for } i = 1, \dots, d\}$ is then constructed by intersection of d such subsets. Finally, $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{A}_3$ can be seen via the fact that open subsets in $\mathbb{R}^d$ are made by union of rectangles of type 3 and as such, they are called open subsets. $\square$

Lemma 18.2.1.10. *Let X be a set and $S \subseteq P(X)$ a class of subsets of X. Let $A \subseteq X$ be a subset. Denote by $S \cap A = \{B \cap A \mid B \in S\}$. Then,*

$$\sigma_A(S \cap A) = \sigma(S) \cap A.$$

where $\sigma_A(S \cap A)$ denotes the smallest σ-algebra over A generated by the class $S \cap A \subseteq P(A)$.

Proof. It is easy to see that $\sigma_A(S \cap A) \hookrightarrow \sigma(S) \cap A$ by considering that $S \cap A \subseteq \sigma(S) \cap A$. Conversely, we use the generating set principle. That is, since we wish to show that for any $B \in \sigma(S)$, we have $B \cap A \in \sigma_A(S \cap A)$, therefore we define

$$\mathcal{S} := \{B \in \sigma(S) \mid B \cap A \in \sigma_A(S \cap A)\}$$

and then observe quite easily that $\mathcal{S}$ is a σ-algebra over X inside $\sigma(S)$ containing S. Thus $\mathcal{S} = \sigma(S)$, as needed. $\square$

The following are some conditions for an algebra to become a σ-algebra.

Proposition 18.2.1.11. *Let X be a set and let $\mathcal{A}$ be an algebra on X. Then, $\mathcal{A}$ is a σ-algebra on X if* ***either***
- *$\mathcal{A}$ is closed under the formation of* ***unions*** *of* increasing *sequence of sets, or,*
- *$\mathcal{A}$ is closed under the formation of* ***intersections*** *of* decreasing *sequence of sets.*

Proof. Take any countably infinite collection of subsets $A_1, A_2, \dots \in \mathcal{A}$ where $\mathcal{A}$ is an algebra. Due to the definition of an algebra, we have that $C_n = \bigcup_{i=1}^n A_i \in \mathcal{A}$ for any $n \geq 1 \in \mathbb{Z}_+$. Now note that $C_1 \subseteq C_2 \subseteq \dots$, that is, the sequence $\{C_n\}$ forms an increasing sequence of sets. Hence, by the requirement of the question, we have that $\bigcup_{i=1}^\infty C_i \in \mathcal{A}$. But then we also have that $\bigcup_{i=1}^\infty A_i \subseteq \bigcup_{i=1}^\infty C_i \in \mathcal{A}$. Hence we have the required condition for part 1. For part 2, we can see that $C_1^c \supseteq C_2^c \supseteq \dots$ is a decreasing sequence of sets. Then we must have, by the requirement of the question, that $\bigcap_{i=1}^\infty C_i^c = \left(\bigcup_{i=1}^\infty C_i\right)^c \in \mathcal{A}$. But then by definition of algebra, we must have $\bigcup_{i=1}^\infty C_i \in \mathcal{A}$, which already contains the countably infinite union $\bigcup_{i=1}^\infty A_i$. □

The following are some finiteness conditions we would like to have on measure spaces.

Definition 18.2.1.12. (**Finiteness conditions**) Let $(X, \mathcal{A}, \mu)$ be a measure space. Then,

1. X is said to be *finite* if $\mu(X) < \infty$,
2. X is said to be *σ-finite* if there exists $\{A_n\} \subseteq \mathcal{A}$ such that $\bigcup_n A_n = X$ and $\mu(A_n) < \infty$,
3. X is said to be *semi-finite* if for all $A \in \mathcal{A}$ such that $\mu(A) = \infty$, there exists $B \subseteq A$ such that $B \in \mathcal{A}$ and $\mu(B) < \infty$.

X-indexed $\mathbb{R}$-series

We would now like to make sense of the sum $\sum_{x \in X} f(x)$ where $f : X \to [0, \infty]$ is an arbitrary function.

Definition 18.2.1.13. (**X-indexed $\mathbb{R}$-series**) Let $f : X \to [0, \infty]$ be a function where X is a set. We define the series $\sum_{x \in X} f(x)$ as follows:

$$\sum_{x \in X} f(x) = \sup \left\{ \sum_{x \in F} f(x) \mid F \subseteq X \text{ is finite} \right\}.$$

The following are some basic properties of X-indexed $\mathbb{R}$-series.

Proposition 18.2.1.14. *Let X be a set and $f : X \to [0, \infty]$ be a function. Denote $S = \{x \in X \mid f(x) > 0\}$.*

1. *If S is uncountable then $\sum_{x \in X} f(x) = \infty$.*
2. *If S is countably finite then for any bijection $\varphi : \mathbb{N} \to S$, we have*

$$\sum_{x \in X} f(x) = \sum_{n \in \mathbb{N}} f(\varphi(n)).$$

Proof. 1. Write $S = \bigcup_n S_n$ where $S_n = \{f(x) > 1/n\}$. Note that S_n forms an increasing sequence of sets. As S is uncountable, there exists $N \in \mathbb{N}$ such that S_N is uncountable. Consequently, for any finite set $F \subseteq S_N$, we have $\sum_{x \in F} f(x) \geq \frac{|F|}{N}$. As $\sum_{x \in F} f(x) \leq \sum_{x \in X} f(x)$, therefore

$$\frac{|F|}{N} \leq \sum_{x \in X} f(x). \tag{$\heartsuit$}$$

As $F \subseteq S_N$ is arbitrary finite set and S_N is uncountable, therefore we get the desired result.

2. Pick any bijection $\varphi : \mathbb{N} \to S$ and pick a finite set $F \subseteq X$. We have $\sum_{x \in F} f(x) = \sum_{x \in F \cap S} f(x)$,

so replace $F \subseteq X$ by a finite set $F \subseteq S$. Let $n \in \mathbb{N}$ be large enough so that $\varphi(\{1, \dots, n\}) \supseteq F$. Consequently, we have

$$\sum_{x \in F} f(x) \leq \sum_{k=1}^{n} f(\varphi(k)) \leq \sum_{x \in X} f(x). \tag{$\heartsuit$}$$

Take $n \to \infty$ in the above inequality to obtain

$$\sum_{x \in F} f(x) \leq \sum_{k=1}^{\infty} f(\varphi(k)) \leq \sum_{x \in X} f(x).$$

Take sup over all finite subsets F of X in the above inequality to obtain

$$\sum_{x \in X} f(x) \leq \sum_{n \in \mathbb{N}} f(\varphi(n)) \leq \sum_{x \in X} f(x),$$

which yields the desired result. □

18.2.2 measures

Definition 18.2.2.1. (**Countably Additive Function**) Let X be a set and $\mathcal{A}$ be a σ-algebra on X. Function $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ is said to be *countably additive* **if** it satisfies:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for **each** infinite sequence $\{A_i\}$ of **disjoint** sets in $\mathcal{A}$.

Definition 18.2.2.2. (**measure**) A *measure* on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ that is **countably additive** and satisfies:

$$\mu(\Phi) = 0.$$

Remark 18.2.2.3. This is sometimes also referred as *countably additive measure* on $\mathcal{A}$.

Definition 18.2.2.4. We have following definitions to compactly represent above definitions:

1. (***measure Space***) If X is a set, $\mathcal{A}$ is a σ-algebra on X and if μ is a measure on $\mathcal{A}$, then the triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.
2. (***measurable Space***) If X is a set and $\mathcal{A}$ is a σ-algebra on X, then the pair $(X, \mathcal{A})$ is called a *measurable space*.

Proposition 18.2.2.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $A, B \in \mathcal{A}$ such that $A \subseteq B$. Then,*

- *We have $\mu(A) \leq \mu(B)$.*
- *Additionally, if A satisfies that $\mu(A) < +\infty$, then:*

$$\mu(B - A) = \mu(B) - \mu(A).$$

Proof. Note that A and $B \cap A^c$ are disjoint sets in the sigma algebra $\mathcal{A}$. Hence we can write, by countably additive property of μ, that:

$$\begin{aligned} \mu(A \cup (B \cap A^c)) &= \mu(B) \\ &= \mu(A) + \mu(B \cap A^c) \end{aligned}$$

Since $\mu(B \cap A^c) \geq 0$, hence $\mu(A) \leq \mu(B)$. Moreover, if $\mu(A) < \infty$, then we can additionally write $\mu(B \cap A^c) = \mu(B) - \mu(A)$. □

Definition 18.2.2.6. Let μ be a measure on a measurable space $(X, \mathcal{A})$. Then,

- (***Finite measure***) If $\mu(X) < +\infty$.
- (***σ-Finite measure***) If $X = \bigcup_i A_i$ where $A_i \in \mathcal{A}$ such that $\mu(A_i) < +\infty$ for all $i \in \mathbb{N}$.

Remark 18.2.2.7. In other words, a subset $A \in \mathcal{A}$ is *σ-finite* if it is a union of a countable sequence of sets that are in $\mathcal{A}$ and are of finite measure under μ.

Elementary Properties of measures

Proposition 18.2.2.8. *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $\{A_k\}$ is an **arbitrary** sequence of sets that belong to $\mathcal{A}$, then,*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Proof. Denote $B_1 = A_1$ and $B_i = A_i \cap \left(\bigcup_{k=1}^{i-1} A_k\right)^c$. Note that B_i and B_j are disjoint for distinct i and j. Since $\{A_k\} \in \mathcal{A}$, therefore $\{B_i\} \in \mathcal{A}$. Moreover, $\bigcup_{i=1}^{\infty} B_i = \bigcup_{k=1}^{\infty} A_k$ by construction. We then get,

$$\begin{aligned}\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) &= \sum_{i=1}^{\infty} \mu(B_i) \\ &\leq \sum_{i=1}^{\infty} \mu(A_i) \quad (\because\ B_i \subseteq A_i \text{ by construction.})\end{aligned}$$

Hence proved. $\square$

18.2.3 Basic results on measure spaces

We have the following first result.

Proposition 18.2.3.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

1. *If $A, B \in \mathcal{A}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*
2. *If $A, B \in \mathcal{A}$ and $A \subseteq B$ where $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.*
3. *For any sequence $\{A_n\} \subseteq \mathcal{A}$, we have*

$$\mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n).$$

4. *If $\{A_n\} \subseteq \mathcal{A}$ is an increasing sequence of measurable sets, then*

$$\mu\left(\bigcup_n A_n\right) = \lim_n \mu(A_n).$$

5. *If $\{A_n\} \subseteq \mathcal{A}$ is a decreasing sequence of measurable sets where $\mu(A_1) < \infty$, then*

$$\mu\left(\bigcap_n A_n\right) = \lim_n \mu(A_n).$$

6. *If X is σ-finite, then X is semi-finite.*

Proof. Statements 1. and 2. are immediate from the disjoint decomposition $B = A \amalg (B \setminus A)$. For 3. note that for any $\{A_n\} \subseteq \mathcal{A}$, we can form a disjoint sequence $\{B_n\} \subseteq \mathcal{A}$ such that $\bigcup_n A_n = \coprod_n B_n$ and $\mu(B_n) \leq \mu(A_n)$. Statement 4. also follows from similar reasons, where we can now let $B_n = A_n \setminus A_{n-1}$. Let us do statement 5. in some detail.

Observe that the sequence $C_1 = \emptyset$ and $C_n = A_1 \setminus A_n$ is an increasing sequence of sets. Thus, we have by statement 4. that

$$\mu\left(\bigcup_n C_n\right) = \lim_n \mu\left(C_n\right). \tag{$\heartsuit$}$$

We can write $A_1 = (A_1 \setminus A_n) \amalg A_n$. Using statement 2. we obtain that

$$\begin{aligned} \mu(A_1) &= \mu(C_n) + \mu(A_n) \\ \mu(A_1) - \mu(A_n) &= \mu(C_n). \end{aligned} \tag{$\heartsuit\heartsuit$}$$

We now claim that $\bigcap_n A_n = A_1 \setminus \bigcup_n C_n$. Indeed, for $x \in \bigcap_n A_n$, $x \in A_n \subseteq A_1$ for all n and thus $x \in A_1$. But if $x \in C_n$ for some n, then $x \notin A_n$, consequently a contradiction. Hence $x \in A_1 \setminus \bigcup_n C_n$. Conversely, for $x \in A_1 \setminus \bigcup_n C_n$ and any $n \in \mathbb{N}$, we have that if $x \notin A_n$, then $x \in A_1 \setminus A_n = C_n$, a contradiction. Hence the claim is proved.

As each $C_n \subseteq A_1$, thus $\bigcup_n C_n \subseteq A_1$. Consequently, by statement 2. and above claim we obtain that

$$\begin{aligned} \mu\left(\bigcap A_n\right) &= \mu(A_1) - \mu\left(\bigcup_n C_n\right) \\ &= \mu(A_1) - \lim_n \mu(C_n) \\ &= \mu(A_1) - \lim_n \left(\mu(A_1) - \mu(A_n)\right) \\ &= \lim_n \mu(A_n). \end{aligned}$$

This proves statement 5.

For statement 6. pick any $A \in \mathcal{A}$ with $\mu(A) = \infty$. We wish to construct a subset $B \subseteq A$ with $B \in \mathcal{A}$ and $0 < \mu(B) < \infty$. Let $\{D_n\} \subseteq \mathcal{A}$ be a collection of finite measure sets such that $\bigcup_n D_n = X$. Note that we can assume D_n are disjoint by suitably replacing D_n by $D_n \setminus D_1 \cup \cdots \cup D_{n-1}$. Assume to the contrary, so that for each $B \subseteq A$ with $B \in \mathcal{A}$, either $\mu(B) = 0$ or $\mu(B) = \infty$. Let $D_n \cap A$ be such that $D_n \cap A \neq \emptyset$. Consequently, $\mu(D_n \cap A) = 0$ or ∞. The latter isn't possible, therefore $\mu(D_n \cap A) = 0$ for all $n \in \mathbb{N}$.

Since we have $A = \coprod_n D_n \cap A$, therefore $\mu(A) = \sum_n \mu(D_n \cap A) = 0$, a contradiction to the fact that $\mu(A) = \infty$. $\square$

We now cover an important example of a measure.

Construction 18.2.3.2. (*measures from a positive function*) Let $(X, \mathcal{A})$ be a measurable space and $f : X \to [0, \infty]$ be a function. We construct the following map

$$\{\text{All functions } X \to [0, \infty]\} \longrightarrow \{\text{measures on } (X, \mathcal{A})\}.$$

Indeed, define

$$\begin{aligned} \mu_f : \mathcal{A} &\longrightarrow [0,\infty] \\ A &\longmapsto \sum_{x\in A} f(x). \end{aligned}$$

We claim that μ_f forms a measure.

It is clear that $\mu_f(\emptyset) = 0$. Consequently we need to show that for a disjoint collection $\{A_n\} \subseteq \mathcal{A}$, we have

$$\mu_f\left(\coprod_n A_n\right) = \sum_n \mu_f(A_n).$$

We first have that

$$\mu_f\left(\coprod_n A_n\right) = \sup\left\{\sum_{x\in F} f(x) \mid F \subseteq \coprod_n A_n \text{ is finite}\right\} \tag{1}$$

and

$$\sum_n \mu_f(A_n) = \sum_n \sup\left\{\sum_{x\in G} f(x) \mid G \subseteq A_n \text{ is finite}\right\}. \tag{2}$$

We first show that (1) $\leq$ (2). We need only show that for a finite set $F \subseteq \amalg_n A_n$, we have $\sum_{x\in F} f(x) \leq (2)$. Indeed, as $F_n := F \cap A_n$ is a collection of disjoint finite set where $F_n \subseteq A_n$ and only for finitely many n is F_n non-empty, therefore $\sum_{x\in F} f(x) = \sum_n \sum_{x\in F_n} f(x) \leq (2)$.

Conversely, we now wish to show that (2) $\leq$ (1). We use a standard technique for this. Pick any $\epsilon > 0$. For each $n \in \mathbb{N}$, we obtain a finite set $G_n \subseteq A_n$ such that

$$\mu_f(A_n) - \frac{\epsilon}{2^n} \leq \sum_{x\in G_n} f(x). \tag{$\heartsuit$}$$

Summing this till $N \in \mathbb{N}$, we obtain

$$\sum_{n=1}^{N}\left(\mu_f(A_n) - \frac{\epsilon}{2^n}\right) \leq \sum_{n=1}^{N}\sum_{x\in G_n} f(x) = \sum_{x\in \amalg_{n=1}^{N} G_n} f(x) \leq (1).$$

Now take $N \to \infty$ and $\epsilon \to 0$ to obtain the result[2].

Observe that the map defined above in Construction 18.2.3.2 is neither injective nor surjective, and that's good, otherwise measure theory would have been redundant. We now study completions of a measure space.

Remark 18.2.3.3. The goal of next few sections is to establish a good measure on $\mathbb{R}^n$ through which we can proceed to a theory of integration of measurable functions. Indeed, this goal was achieved by Lebesgue and he constructed what will be called the Lebesgue measure on $\mathbb{R}^n$. Hence, one should view the goal of the next few sections as to construct this measure space $(\mathbb{R}^n, \mathcal{M}, m)$, which is highly usable (as we will see in the integration theory) and is the gold standard of modern analysis.

[2] We call this the *ϵ-wiggle around* inf *and* sup *technique.*

18.2.4 Completion of a measure space

Definition 18.2.4.1. (**Null sets and complete measure spaces**) Let $(X, \mathcal{A}, \mu)$ be a measure space. A null set is an element $A \in \mathcal{A}$ such that $\mu(A) = 0$. The collection of all null sets is written as $\mathrm{Null}(\mathcal{A}) \subseteq \mathcal{A}$. A measure space $(X, \mathcal{A}, \mu)$ is said to be complete if for all $A \in \mathrm{Null}(\mathcal{A})$, $\mathcal{P}(A) \subseteq \mathcal{A}$.

Remark 18.2.4.2. Note that for a measure space $(X, \mathcal{A}, \mu)$, the collection of all null sets $\mathrm{Null}(\mathcal{A})$ contains $\emptyset$ and is closed under countable union. Indeed, for $\{A_n\} \subseteq \mathrm{Null}(\mathcal{A})$, we have $\mu(\cup_n A_n) \leq \sum_n \mu(A_n) = 0$ by Proposition 18.2.3.1, 3.

Definition 18.2.4.3. (**Extension of measure spaces**) Let $(X, \mathcal{A}, \mu)$ and $(X, \mathcal{A}', \mu')$ be two measure spaces. Then we say that $(X, \mathcal{A}', \mu')$ is an extension of $(X, \mathcal{A}, \mu)$ if $\mathcal{A}' \supseteq \mathcal{A}$ and $\mu'|_{\mathcal{A}} = \mu$.

We will now for each measure space $(X, \mathcal{A}, \mu)$ will construct an extension of it which will be complete.

Construction 18.2.4.4. Let $(X, \mathcal{A}, \mu)$ be a measure space. Consider the following collection

$$\hat{\mathcal{A}} := \{A \cup B \mid A \in \mathcal{A}, B \subseteq N, N \in \mathrm{Null}(\mathcal{A})\}.$$

Define $\hat{\mu} : \hat{\mathcal{A}} \to [0, \infty]$ as $A \cup B \mapsto \mu(A)$.

Theorem 18.2.4.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then, $(X, \hat{\mathcal{A}}, \hat{\mu})$ is a complete measure space extending $(X, \mathcal{A}, \mu)$. We call it the completion of $(X, \mathcal{A}, \mu)$.*

Proof. We need to show the following things.
1. $\hat{\mathcal{A}}$ is a σ-algebra,
2. $\hat{\mu}$ is a measure,
3. $\hat{\mu}|_{\mathcal{A}} = \mu$,
4. $(X, \hat{\mathcal{A}}, \hat{\mu})$ is complete.

The first three are straightforward. We show 4. in some detail.

Pick $A \cup B \in \hat{\mathcal{A}}$ such that $\hat{\mu}(A \cup B) = \mu(A) = 0$. Then $A \in \mathrm{Null}(\mathcal{A})$. Further, $B \subseteq N$ where $N \in \mathrm{Null}(\mathcal{A})$. Let $C \subseteq A \cup B$. Then $C = (C \cap A) \cup (C \cap B)$. Since $C \cap A \subseteq A$ and $C \cap B \subseteq N$, therefore $C \subseteq A \cup N$ where $A \cup N \in \mathrm{Null}(\mathcal{A})$. Consequently, we may write $C = \emptyset \cup C$ where C is a subset of a null set. Hence $C \in \hat{\mathcal{A}}$. □

Example 18.2.4.6. Let $X = \{1, 2, 3\}$ and $\mathcal{A} = \{\emptyset, X, \{1\}, \{2, 3\}\}$. Define $\mu : \mathcal{A} \to [0, \infty]$ by $\mu(\emptyset) = 0 = \mu(\{2, 3\})$ and $\mu(\{1\}) = \mu(X)$. Clearly, $(X, \mathcal{A}, \mu)$ is a measure space which is not complete. We calculate its completion $(X, \hat{\mathcal{A}}, \hat{\mu})$. By Construction 18.2.4.4, as the only null set is $\{2, 3\}$, we have

$$\hat{\mathcal{A}} = \{\emptyset, X, \{1\}, \{2, 3\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}\}.$$

Hence $\hat{\mathcal{A}} = \mathcal{P}(X)$. Similarly, $\hat{\mu}$ is easy to find by the definition in Construction 18.2.4.4.

18.2.5 Outer measures

Definition 18.2.5.1. (**Outer measure**) Let X be a set and let $\mathcal{P}(X)$ be the collection of all subsets of X. An *outer measure* on X is a function $\mu^* : \mathcal{P}(X) \longrightarrow [0, +\infty]$ such that:

- For the empty set Φ,

$$\mu^*(\Phi) = 0$$

- If $A \subseteq B \subseteq X$, then
$$\mu^*(A) \leq \mu^*(B).$$
- If $\{A_n\}$ is an infinite sequence of subsets of X, then
$$\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n)$$

Definition 18.2.5.2. (**Lebesgue outer measure on** $\mathbb{R}$) For each subset $A \subseteq \mathbb{R}$, let $\mathcal{C}_A$ be the *set of all infinite sequences* $\{(a_i, b_i)\}$ of *bounded open intervals* such that $A \subseteq \bigcup_i (a_i, b_i)$. That is,

$$\mathcal{C}_A = \{\{(a_i, b_i)\} \mid A \subseteq \cup_i (a_i, b_i) \text{ and } a_i, b_i \in \mathbb{R}\}$$

Then, $\lambda^* : \mathcal{P}(\mathbb{R}) \longrightarrow [0, +\infty]$ is the Lebesgue outer measure, defined by:

$$\lambda^*(A) = \inf\left\{\sum_i (b_i - a_i) \;\middle|\; \{(a_i, b_i)\} \in \mathcal{C}_A\right\} \tag{18.3}$$

To verify that λ^* is indeed an outer measure.

Proposition 18.2.5.3. *Lebesgue outer measure on $\mathbb{R}$ is an outer measure and it assigns to each subinterval of $\mathbb{R}$ it's length.*

Proof. Denote $\mathcal{C}_A = \{\{(a_i, b_i)\} \mid A \subseteq \cup_i (a_i, b_i)\}$. To show that λ^* is an outer measure, we first need to show that $\lambda^*(\Phi) = 0$. For that, consider the set of all infinite sequences $\{(a_i, b_i)\} \in \mathcal{C}_\Phi$, that is (trivially) $\Phi \subseteq \cup_i (a_i, b_i)$, such that $\sum_i (b_i - a_i) < \epsilon$ for all $\epsilon > 0$. Then, if we denote $\mathcal{L}_A = \{\sum_i (b_i - a_i) \mid \{(a_i, b_i)\} \in \mathcal{C}_A\}$, then $\inf \mathcal{L}_\Phi = 0$ as for any lower bound l of $\mathcal{L}_A$, if $l > 0$ then $\exists \{(a_i, b_i)\} \in \mathcal{C}_\Phi$ such that $\sum_i (b_i - a_i) < l$, hence $l \leq 0$, or $\inf \mathcal{L}_\Phi = 0$.
Second, we need to show that if $A \subseteq B \subseteq X$, then $\lambda^*(A) \leq \lambda^*(B)$. For this, consider $A \subseteq B$. Clearly, we have that $\mathcal{C}_B \subseteq \mathcal{C}_A$, therefore $\mathcal{L}_B \subseteq \mathcal{L}_A$ and hence $\inf \mathcal{L}_B \geq \inf \mathcal{L}_A$.
Third, we need to show that for any infinite sequence $\{A_n\}$ of subsets of X,

$$\lambda^*\left(\bigcup_n A_n\right) \leq \sum_n \lambda^*(A_n)$$

For this, consider the Lebesgue outer measure of A_n, that is, $\lambda^*(A_n)$. We must have, that for any infinite sequence $\{(a_{n,i}, b_{n,i})\} \in \mathcal{C}_{A_n}$, that

$$\sum_{i=1}^{\infty} (b_{n,i} - a_{n,i}) \geq \lambda^*(A_n).$$

Hence, consider that the difference is upper bounded according to n, that is the sequence $\{(a_{n,i}, b_{n,i})\} \in \mathcal{C}_{A_n}$ is such that,

$$\sum_{i=1}^{\infty} (b_{n,i} - a_{n,i}) - \lambda^*(A_n) \leq \epsilon/2^n.$$

Now, we can cover the entire $\bigcup_i A_i$ by the union of the above intervals, that is,

$$\bigcup_i A_i \subseteq \bigcup_n \bigcup_i (a_{n,i}, b_{n,i}).$$

Now, we know that

$$\lambda^*\left(\bigcup_i A_i\right) = \inf \mathcal{L}_{\cup_i A_i}.$$

But since

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) \in \mathcal{L}_{\cup_i A_i},$$

and

$$\sum_n \left(\sum_i (b_{n,i} - a_{n,i}) - \lambda^*(A_n)\right) \leq \sum_n \epsilon/2^n$$

which is equal to

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) - \sum_n \lambda^*(A_n) \leq \epsilon \times 1$$

or,

$$\sum_n \sum_i (b_{n,i} - a_{n,i}) \leq \sum_n \lambda^*(A_n) + \epsilon$$

and since $\lambda^*(\bigcup_i A_i) = \inf \mathcal{L}_{\bigcup_i A_i}$, therefore,

$$\lambda^*\left(\bigcup_i A_i\right) \leq \sum_n \sum_i (b_{n,i} - a_{n,i}) \leq \sum_n \lambda^*(A_n)$$

Hence proved.

Now, we need to show that λ^* assigns each subinterval it's length.
For this first show that $\lambda^*([a,b]) \leq b - a$. This is easy to show if we take,

$$[a,b] = \bigcup_i (a_i, b_i)$$

where $(a_1, b_1) = (a, b)$, $(a_i, b_i) = (a - \epsilon/2^i, a)$ for all even i and $(a_j, b_j) = (b, b + \epsilon/2^j)$ for all odd j.
Now,

$$\begin{aligned}\sum_i (b_i - a_i) &= (b - a) + \sum_{i=2,4,\ldots} \epsilon/2^i + \sum_{i=3,5,\ldots} \epsilon/2^i \\ &= b - a + \sum_{i=1,2,\ldots} \epsilon/2^i \\ &= b - a + \epsilon\end{aligned}$$

therefore $\lambda^*([a,b]) = \inf \mathcal{L}_{[a,b]} \leq b - a + \epsilon$ for all $\epsilon > 0$, hence $\lambda^*([a,b]) \leq b - a$.
Now, to show the converse that $b - a \leq \lambda^*([a,b])$, we first note that $[a,b]$ is compact, so for any infinite cover $\{(a_i, b_i)\} \in \mathcal{C}_{[a,b]}$, there exists a finite subcover $\{(a_i, b_i)\}_{i=1}^n$ of $[a,b]$. Now, since λ^* is an outer measure, therefore,

$$b - a \leq \sum_{i=1}^n \lambda^*((a_i, b_i)) \leq \sum_{i=1}^\infty \lambda^*((a_i, b_i)) \in \mathcal{L}_{[a,b]}$$

Therefore, $b-a$ is a lower bound of $\mathcal{L}_{[a,b]}$ and hence $b-a \leq \inf \mathcal{L}_{[a,b]} = \lambda^*([a,b])$.
Hence $\lambda^*([a,b]) = b-a$.
Now since, one can construct subintervals of the form $(a,b]$ or $[a,b)$ from the following manner:

$$(a,b] \subseteq (a,b) \bigcup \left(\bigcup_n [b, b+\epsilon/2^n] \right)$$

from which we get that $\lambda^*((a,b]) \leq b-a$ and also,

$$[a,b] \subseteq (a,b] \bigcup \left(\bigcup_n [a-\epsilon/2^n, a] \right)$$

which yields $b-a \leq \lambda^*((a,b])$. Similarly for $(-\infty, b]$ to show that $\lambda^*((-\infty,b]) = +\infty$. $\square$

Construction 18.2.5.4. (*Lebesgue outer measure on* $\mathbb{R}^n$) Consider $\mathbb{R}^m$ and for an box $I \subseteq \mathbb{R}^n$, by which we mean a product of interval $I = I_1 \times \cdots \times I_m$ for $I_i \subseteq \mathbb{R}$, denote $v(I)$ to be its volume; $v(I) = \prod_{i=1}^m l(I_i)$. For any $A \subseteq \mathbb{R}^n$, we define

$$\mu^*(A) = \inf \left\{ \sum_n v(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n \text{ are boxes} \right\}.$$

We claim that μ^* forms an outer measure on $\mathbb{R}^n$.

Indeed, $\mu^*(\emptyset) = 0$ as $\emptyset \subseteq (-1/k, 1/k)^m$ for all $n \in \mathbb{N}$ so we have $\mu^*(A) \leq 2^m/n^m$. Taking $n \to \infty$ does the job.

Let $A \subseteq B$ in $\mathbb{R}^m$. Observe that to show $\mu^*(A) \leq \mu^*(B)$ we need only show that $\{\sum_n v(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n$ are b $\{\sum_n v(I_n) \mid \bigcup_n I_n \supseteq B,\ I_n$ are boxes$\}$. But this is trivial as and sequence of boxes $\{I_n\}$ covering B also covers A.

Finally we wish to show countable subadditivity. Pick $\{A_n\} \subseteq \mathcal{P}(\mathbb{R}^m)$. We wish to show that

$$\mu^*\left(\bigcup_n A_n\right) \leq \sum_n \mu^*(A_n).$$

We use the ϵ-wiggle around sup and inf technique to show this, as discussed earlier in Construction 18.2.3.2. Pick any $\epsilon > 0$ and observe that we have a sequence of boxes $\{I_{n,k}\}_k$ for each $n \in \mathbb{N}$ such that $\bigcup_k I_{n,k} \supseteq A_n$ and

$$\mu^*(A_n) + \frac{\epsilon}{2^n} \geq \sum_k v(I_{n,k}). \tag{$\heartsuit$}$$

Observe further that $\bigcup_n \bigcup_k I_{n,k} \supseteq \bigcup_n A_n$. Consequently, we have $\sum_n \sum_k v(I_{n,k}) \geq \mu^*(\bigcup_n A_n)$. Hence,

$$\sum_n \left(\mu^*(A_n) + \frac{\epsilon}{2^n}\right) \geq \sum_n \sum_k v(I_{n,k}) \geq \mu^*\left(\bigcup_n A_n\right).$$

Hence μ^* is an outer measure on $\mathbb{R}^n$.

Note that the only place we required knowledge about boxes explicitly was only to show that $\mu^*(\emptyset) = 0$. This motivates the following simple result

Theorem 18.2.5.5. *Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ be a collection of sets containing $\emptyset$ and X. Let $l : \mathcal{S} \to [0, \infty]$ be a function such that $l(\emptyset) = 0$. Then μ^* defined by*

$$\begin{aligned} \mu^* : \mathcal{P}(X) &\longrightarrow [0, \infty] \\ A &\longmapsto \inf \left\{ \sum_n l(I_n) \mid \bigcup_n I_n \supseteq A,\ I_n \in \mathcal{S} \right\} \end{aligned}$$

is an outer measure on X.

Proof. Verbatim to Construction 18.2.5.4, except that $\mu^*(\emptyset) = 0$ follows now by the assumption that $l(\emptyset) = 0$ and $\emptyset \in \mathcal{P}(X)$ so that $\emptyset$ forms its own covering. $\square$

18.2.6 Lebesgue measurability & Carathéodory's theorem

Definition 18.2.6.1. (**μ^*-measurable subset**) Let X be a set and let μ^* be an *outer measure* on X. A subset $B \subseteq X$ is μ^*-measurable if:

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds **for all** subsets $A \subseteq X$.

Definition 18.2.6.2. (**Lebesgue measurable subset of $\mathbb{R}$**) A subset $B \subseteq \mathbb{R}$ is called a Lebesgue measurable subset of $\mathbb{R}$ if B **is λ^*-measurable.** That is, for **any** $A \subseteq \mathbb{R}$, we must have:

$$\lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^c)$$

Remark 18.2.6.3. Important to note are the following:

- Due to sub-additivity of μ^* and $A \subseteq (A \cap B) \cup (A \cap B^c)$, we already have that

$$\mu^*(A) \le \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

  for *any* subsets $A, B \subseteq X$.
- ★ Due to the above fact, *all that remains to be shown to ascertain that $B \subseteq \mathbb{R}$ is μ^*-measurable* is to show the following converse:

$$\mu^*(A) \ge \mu^*(A \cap B) + \mu^*(A \cap B^c).$$

  for all $A \subseteq X$.

Proposition 18.2.6.4. *Let X be a set and let μ^* be an outer measure on X. Then each subset $B \subseteq X$ that satisfies $\mu^*(B) = 0$ **or** that satisfies $\mu^*(B^c) = 0$ is μ^*-measurable.*

Proof. This result actually proves that for subset $B \subseteq X$ which has zero outer measure under μ^*, any other subset $A \subseteq X$ would be such that $\mu^*(A \cap B) = 0(!)$ After proving this, and from the remark above, we would just be left to show that if $\mu^*(B) = 0$, then $\mu^*(A) \ge \mu^*(A \cap B) + \mu^*(A \cap B^c)$. We show the former here, from which the latter follows naturally.

Consider $B \subseteq X$ such that $\mu^*(B) = 0$. It's true that $A \cap B \subseteq B$. Now since μ^* is an outer measure on X, therefore, we must have $\mu^*(A \cap B) \le \mu^*(B) = 0$. This implies that $\mu^*(A \cap B) = 0$.

Now, we would see that the required condition follows naturally from the previous. First, note the following:

$$A \cap B \subseteq A \text{ and } A \cap B^c \subseteq A.$$

Hence, we can write:

$$\mu^*(A \cap B) \leq \mu^*(A) \text{ and } \mu^*(A \cap B^c) \leq \mu^*(A).$$

Now if $\mu^*(B) = 0$, then $\mu^*(A \cap B) = 0$ and then in the second inequality, we would have:

$$\mu^*(A \cap B^c) + \mu^*(A \cap B) \leq \mu^*(A) + 0$$

Or, if $\mu^*(B^c) = 0$, then $\mu^*(A \cap B^c) = 0$ and then in the first inequality, we would have:

$$\mu^*(A \cap B) + \mu^*(A \cap B^c) \leq \mu^*(A) + 0.$$

Hence, B is μ^*-measurable for any $B \subseteq X$ which satisfies that either $\mu^*(B) = 0$ or $\mu^*(B^c) = 0$. □

The following theorem is a fundamental fact about outer measures.

Theorem 18.2.6.5 (Carathéodory). *Let X be a set, let μ^* be an outer measure on X and let $\mathcal{M}_{\mu^*}$ be the collection of all μ^*-measurable subsets of X. Then,*

- *$\mathcal{M}_{\mu^*}$ is a σ-algebra.*
- *The restriction of μ^* to $\mathcal{M}_{\mu^*}$ is a measure on $\mathcal{M}_{\mu^*}$.*

Proof. **Act 1.** *$\mathcal{M}_{\mu^*}$ is an algebra.*
First, it is clear that $X, \Phi \in \mathcal{M}_{\mu^*}$ from Proposition 18.2.6.4, because $\mu^*(\Phi) = \mu^*(X^c) = 0$. Now, if $B \in \mathcal{M}_{\mu^*}$, then $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)\ \forall\ A \subseteq X$. But if we replace B by B^c in the above, we would get the same equation, hence $B^c \in \mathcal{M}_{\mu^*}$. So $\mathcal{M}_{\mu^*}$ is closed under complements. Now, to show closed nature under finite unions, we take any two subsets $B_1, B_2 \in \mathcal{M}_{\mu^*}$ and show that $A \cup B \in \mathcal{M}_{\mu^*}$. First we have

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) \\ &= \mu^*(A \cap B_2) + \mu^*(A \cap B_2^c)\end{aligned}$$

for any $A \subseteq X$. Now, we see that from the fact that $B_1 \in \mathcal{M}_{\mu^*}$,

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_2 \cap B_1^c)\end{aligned}$$

Similarly, we have from the fact $B_2 \in \mathcal{M}_{\mu^*}$,

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)^c) &= \mu^*(A \cap (B_1 \cup B_2)^c \cap B_2) + \mu^*(A \cap (B_1 \cup B_2)^c \cap B_2^c) \\ &= \mu^*(A \cap B_1^c \cap B_2^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c \cap B_2^c) \\ &= \mu^*(\Phi) + \mu^*(A \cap B_1^c \cap B_2^c) \\ &= \mu^*(A \cap (B_1 \cup B_2)^c)\end{aligned}$$

Now, adding the above results yield,

$$\begin{aligned}\mu^*\left(A\cap(B_1\cup B_2)^{\mathrm{c}}\right)+\mu^*\left(A\cap(B_1\cup B_2)\right)&=\mu^*\left(A\cap(B_1\cup B_2)^{\mathrm{c}}\right)+\mu^*\left(A\cap B_1\right)+\mu^*\left(A\cap B_2\cap B_1^{\mathrm{c}}\right)\\&=\mu^*\left(A\cap B_1^{\mathrm{c}}\cap B_2^{\mathrm{c}}\right)+\mu^*\left(A\cap B_1^{\mathrm{c}}\cap B_2\right)+\mu^*\left(A\cap B_1\right)\\&=\mu^*\left(A\cap B_1^{\mathrm{c}}\right)+\mu^*\left(A\cap B_1\right)\\&=\mu^*\left(A\right).\end{aligned}$$

Hence, $B_1\cup B_2$ is μ^*-measurable, so $B_1\cup B_2\in\mathcal{M}_{\mu^*}$. Now, we can, for a finite collection of subsets in $\mathcal{M}_{\mu^*}$, we can proceed like above, to show that $\mathcal{M}_{\mu^*}$ is closed under finite union, hence showing that $\mathcal{M}_{\mu^*}$ is an algebra.

Act 2. *$\mathcal{M}_{\mu^*}$ is a σ-algebra.*
All that is left to show that $\mathcal{M}_{\mu^*}$ is a σ-algebra is to show that it is closed under countable union. We have already proved closed nature under finite union. We extend it via induction principle. Suppose $\{B_i\}$ is a sequence of disjoint subsets in $\mathcal{M}_{\mu^*}$. For this, we first prove[3] using induction that, for all $A\subseteq X$ and $n\in\mathbb{N}$,

$$\text{To Prove}:\mu^*\left(A\right)=\sum_{i=1}^{n}\mu^*\left(A\cap B_i\right)+\mu^*\left(A\cap\left(\bigcap_{i=1}^{n}B_i^{\mathrm{c}}\right)\right)\tag{18.4}$$

For the case when $n=1$, we see that it Eq. 18.4 reduces to $\mu^*\left(A\right)=\mu^*\left(A\cap B_1\right)+\mu^*\left(A\cap B_1^{\mathrm{c}}\right)$. But since $B_i\in\mathcal{M}_{\mu^*}\ \forall\, i\in\mathbb{N}$, therefore this is trivially true. Now, by the induction principle, we assume that Eq. 18.4 is true uptill n and then we try to prove it for $n+1$ step. For this, since $B_{n+1}\in\mathcal{M}_{\mu^*}$ is disjoint to all other B_i's, we have,

$$\begin{aligned}\mu^*\left(A\cap\bigcap_{i=1}^{n}B_i^{\mathrm{c}}\right)&=\mu^*\left(\left(A\cap\bigcap_{i=1}^{n}B_i^{\mathrm{c}}\right)\cap B_{n+1}\right)+\mu^*\left(\left(A\cap\bigcap_{i=1}^{n}B_i^{\mathrm{c}}\right)\cap B_{n+1}^{\mathrm{c}}\right)\\&=\mu^*\left(A\cap B_{n+1}\right)+\mu^*\left(A\cap\bigcap_{i=1}^{n+1}B_i^{\mathrm{c}}\right)\end{aligned}$$

where the last line follows from the fact that each B_i is disjoint to other B_j's, hence each B_j^{c} would contain B_i and therefore $B_{n+1}\subseteq\bigcap_{i=1}^{n}B_i^{\mathrm{c}}$. Now, substituting the above equation in Eq. 18.4 gives,

$$\begin{aligned}\mu^*\left(A\right)&=\sum_{i=1}^{n}\mu^*\left(A\cap B_i\right)+\mu^*\left(A\cap B_{n+1}\right)+\mu^*\left(A\cap\bigcap_{i=1}^{n+1}B_i^{\mathrm{c}}\right)\\&=\sum_{i=1}^{n+1}\mu^*\left(A\cap B_i\right)+\mu^*\left(A\cap\bigcap_{i=1}^{n+1}B_i^{\mathrm{c}}\right)\end{aligned}$$

[3] But why to prove Eq. 18.4? The motivation for Eq. 18.4 comes from Part 1. More specifically, notice in the equation where we added $\mu^*\left(A\cap(B_1\cup B_2)^{\mathrm{c}}\right)$ and $\mu^*\left(A\cap(B_1\cup B_2)\right)$. Note it's 2nd line, this is the case when $n=2$ in Eq. 18.4 combined with the fact that B_i's are disjoint. Now why to take B_i's to be disjoint? The reason for this comes from the fact that for any infinite sequence of subsets $\{A_i\}$, one can construct infinite sequence of disjoint subsets, that is : $A_1, A_2\cap A_1^{\mathrm{c}}, A_3\cap(A_1\cup A_2)^{\mathrm{c}},\ldots$ and it's union is again $\bigcup_n A_n$. Hence if we prove that a disjoint infinite sequence is closed under union, then we could prove that any infinite sequence of subsets is closed under union too!

Hence, by induction principle, Eq. 18.4 is true for all $n \in \mathbb{N}$. Hence, now we can write,

$$\begin{aligned}\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \bigcap_{i=1}^{\infty} B_i^c\right)\\ &= \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c\right)\end{aligned}$$

Now, to prove that $\bigcup_i B_i \in \mathcal{M}_{\mu^*}$, we need to show

$$\text{To Show : } \mu^*(A) \geq \mu^*\left(A \cap \bigcup_i B_i\right) + \mu^*\left(A \cap \left(\bigcup_i B_i\right)^c\right)$$

This comes from previous result as follows:

$$\begin{aligned}\mu^*(A) &\geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c\right)\\ &\geq \mu^*\left(\bigcup_{i=1}^{\infty}(A \cap B_i)\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c\right)\\ &= \mu^*\left(A \cap \bigcup_{i=1}^{\infty} B_i\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c\right)\end{aligned} \tag{18.5}$$

Therefore, $\bigcup_i B_i \in \mathcal{M}_{\mu^*}$. Now, as the previous footnote mentions, for every infinite sequence $\{C_i\}$ in $\mathcal{M}_{\mu^*}$, we have a disjoint sequence of subsets as $C_1, C_2 \cap C_1^c, C_3 \cap C_2^c \cap C_1, \ldots$. Now, this disjoint sequence is closed under union as we just showed and since union of this disjoint sequence is equal to the union of $\{C_i\}$, hence $\bigcup_i C_i \in \mathcal{M}_{\mu^*}$ for any sequence $\{C_i\}$ in $\mathcal{M}_{\mu^*}$. Thus, $\mathcal{M}_{\mu^*}$ is a σ-algebra.

Act 3. *μ^* restricted to $\mathcal{M}_{\mu^*}$ is a measure.*
Consider $\{B_n\}$ be an infinite sequence of subsets in $\mathcal{M}_{\mu^*}$. Now, by finite subadditivity, we trivially have

$$\mu^*\left(\bigcup_i B_i\right) \leq \sum_i \mu^*(B_i)$$

Moreover, from Part 2 and setting $A = \cup_i B_i$, we get:

$$\begin{aligned}\mu^*\left(\bigcup_i B_i\right) &\geq \sum_j \mu^*\left(\bigcup_i B_i \cap B_j\right) + \mu^*\left(\bigcup_i B_i \cap \left(\bigcup_i B_i\right)^c\right)\\ &= \sum_j \mu^*(B_j) + \mu^*(\Phi)\\ &= \sum_j \mu^*(B_j).\end{aligned}$$

We hence have the complete proof. □

Definition 18.2.6.6. (**Lebesgue measure**) The restriction of Lebesgue outer measure on $\mathbb{R}$ to the collection $\mathcal{M}_{\lambda^*}$ of Lebesgue measurable subsets of $\mathbb{R}$ is called *Lebesgue measure*. It would be denoted by λ. Hence, we would work with the measure space $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$[4].

18.2.7 Does $\lambda^*(E) = 0$ implies E is countable?

We would construct today a set which has measure 0, but not countable(!).

1. Take $E_0 = [0, 1]$.
2. Remove $(1/3, 2/3)$ from E_0 to form $E_1 = [0, 1/3] \cup [2/3, 1]$.
3. Proceed in the same way to form $E_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$.
4. At n^{th} step, E_n contains 2^n subintervals and each of which is of length $\frac{1}{3^n}$.
5. We clearly have $E_0 \supset E_1 \supset E_2 \supset \dots$.
6. Here, note that each E_n is a closed and compact subset of $\mathbb{R}$.
7. The set
$$P = \bigcap_{n=0}^{\infty} E_n \text{ is known as } \textbf{Cantor Set}.$$

Properties of Cantor Set

Proposition 18.2.7.1. *Lebesgue measure of Cantor Set is 0.*

Proof. Note that Cantor Set is Lebesgue measurable as it is countable intersection of closed sets, hence it is present in the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ and hence is also in $\mathcal{M}_{\lambda^*}$. Hence, instead of λ^*, we can now write λ as $P \in \mathcal{M}_{\lambda^*}$. Now, measure of Cantor set P can be written as:

$$\begin{aligned}
\lambda(P) &= \lambda\left(\bigcap_n E_n\right) \\
&= \varinjlim_{n\to\infty} \lambda(E_n) \\
&= \varinjlim_{n\to\infty} \frac{2^n}{3^n} \\
&= \varinjlim_{n\to\infty} \frac{1}{1.5^n} \\
&= 0.
\end{aligned}$$

□

Proposition 18.2.7.2. *Cantor set is uncountable(!)*

Proof. We will show that there exists a bijection between Cantor Set and an uncountable set, specifically ternary system. For this, consider the ternary representation of every number in $[0, 1]$. What this means is that every number in $[0, 1]$ can be represented only using the numbers $0, 1$ and 2. Hence, one write $\frac{1}{3}$ as 0.1 and $\frac{2}{3}$ as 0.2. Now, $(1/3, 2/3) = E_1^c \cap [0, 1]$ is the set that has been removed from the process of creating E_1 from E_0. Clearly, every number in this $E_1^c \cap [0, 1]$ is of the

[4] From this point on-wards, whenever this text mentions that a given set is measurable in space $(X, \mathcal{A}, \mu)$, it must be assumed that the given set is in $\mathcal{A}$, given that there is no ambiguity.

form $0.1\ldots$ where $\ldots$ are all combinations of $0, 1$ and 2. Therefore, we are now left with the E_1 that has all the numbers represented as $0.0\ldots$ or $0.2\ldots$.
As we saw in the generation of E_1, the generation of E_2 from E_1 would hence involve removing numbers of the forms $0.01\ldots$ and $0.21\ldots$. And hence E_2 would then be the set of numbers whose first two decimal places are restricted to NOT have the digit 1; that is, E_2 would be of form $0.02\ldots, 0.00\ldots, 0.20\ldots, 0.22\ldots$.
Continuing like this, we see that E_n would have in ternary representation, all those numbers whose first n digits are NOT 1. Hence, for any $p \in P$, p would have the ternary representation constructed only from 0 and 2, but NOT 1.
Now, consider the map $f : P \to [0, 1]$ such that $f(p)$ replaces each occurence of 2 by 1 in the ternary representation of p. We now show that this map is surjective(!) so that P has atleast as many elements as $[0, 1]$. To show this, take any $x \in [0, 1]$ in it's ternary form, and replace all 1 by 2 and denote it as x'. Clearly, x' would be in P as x' has all decimal digits generated by 0 and 2. But $f(x')$ would be opposite action and would be equal to x. Therefore, we showed that for any $x \in [0, 1], \exists\ x' \in P$ such that $f(x') = x$. Hence f is surjective. Therefore P has atleast as many elements as $[0, 1]$. But since $P \subseteq [0, 1]$ therefore P has atmost as many elements as $[0, 1]$. This dichotomy suggests that

Cantor Set has as many elements as in $[0, 1]$ *(!)*

But since $[0, 1]$ is uncountable, therefore, P is uncountable. □

With this, we conclude that **for any set $E \subseteq \mathbb{R}$, if $\lambda^*(E) = 0$, then it's NOT necessarily true that E is countable**.
We now see an extremely interesting example of a Non-measurable set.

18.2.8 A non-measurable set

Theorem 18.2.8.1. *There is a subset of* $\mathbb{R}$ *that is not Lebesgue measurable*[5].

Proof. We construct the proof in the following *Acts*:

Act 1. *Equivalence Relation on* $\mathbb{R}$.
Construct the following relation $\sim$ on $\mathbb{R}$:

$$x \sim y \equiv x - y \in \mathbb{Q}.$$

Clearly, $\sim$ is reflexive as $x - x = 0$ is rational; it is also symmetric as negative of a rational is also a rational number; and it is also transitive as if $x - y$ and $y - z$ is rational, then $x - y + y - z = x - z$ is sum of two rationals, which is also rational. Hence $\sim$ is an equivalence relation. Therefore $\sim$ partitions the whole $\mathbb{R}$ into equivalence classes. Note that each equivalence class of x would consist elements of the form $\mathbb{Q} + x$. But since $\mathbb{Q}$ is dense in $\mathbb{R}$, therefore $\mathbb{Q} + x$, that is each equivalence class, is dense in $\mathbb{R}$.
Now, each equivalence class clearly intersects $(0, 1)$, therefore, inducing the Axiom of Choice on the set of all equivalence classes, we can form a subset $E \subset (0, 1)$ which contains exactly one element from each of the equivalence classes. We will later prove that E is not Lebesgue measurable.

[5] See **[Solovay70]** for more information.

Act 2. *E satisfies certain properties.*
Consider the set $\mathbb{Q} \cap (-1, 1)$. Clearly, this is countable as it's subset of $\mathbb{Q}$. Then, consider $\{r_n\}$ to be the enumeration of $\mathbb{Q} \cap (-1, 1)$. Construct the sequence of subsets $E_n = E + r_n$. We now verify that $\{E_n\}$ satisfies the following properties:

1. The sets E_n are disjoint.
2. $\bigcup_n E_n$ is a subset of the interval $(-1, 2)$.
3. The interval $(0, 1)$ is included in $\bigcup_n E_n$.

Property 1 : Assume that $E_n \cap E_m \neq \Phi$ for some $n, m \in \mathbb{N}$ such that $n \neq m$. Then $\exists e_1, e_2 \in E$ such that $e_1 + r_n = e_2 + r_m$ which means that $e_1 - e_2 = r_m - r_n \in \mathbb{Q}$. But this cannot happen as e_1, e_2 are elements of E and E contains exactly one element from the equivalence class of $\sim$ intersected with $(0, 1)$. Therefore $e_1 - e_2 \notin \mathbb{Q}$. Which is a contradiction. Hence $E_n \cap E_m = \Phi$ for all $n, m \in \mathbb{N}$ such that $n \neq m$.
Property 2 : Take $x \in \bigcup_n E_n$. This implies that $x \in E_m$ for some $m \in \mathbb{N}$. But $E_m = E + r_m = \{e + r_m \mid e \in E\}$. Since $E \subset (0, 1)$ and $r_m \in \mathbb{Q} \cap (-1, 1) \subset (-1, 1)$, therefore $x \in E + r_m \subseteq (-1, 2)$. Hence $\bigcup_n E_n \subseteq (-1, 2)$.
Property 3 : Take any $x \in (0, 1)$. Now take the $e \in E$ such that $x \sim e$, or $x - e \in \mathbb{Q}$. Hence $x \in \mathbb{Q} + e$. That is $x = r + e$. But since $0 < e < 1$ and $0 < x < 1$, therefore $r = x - e \in \mathbb{Q} \cap (-1, 1)$. Hence $x \in E + r$ and if we denote $r = r_n$ for some $n \in \mathbb{N}$, we get $x \in E + r_n = E_n$, therefore $x \in \bigcup_i E_i$. Hence $(0, 1) \subseteq \bigcup_i E_i$.

Act 3. *E is Not Lebesgue measurable.*
Assume that E is in-fact Lebesgue measurable. Now since E_n are disjoint (Property 1), therefore we can write:

$$\lambda\left(\bigcup_n E_n\right) = \sum_n \lambda(E_n).$$

Now, since **Lebesgue measure is translation invariant**[6], therefore $\lambda(E_n) = \lambda(E + r_n) = \lambda(E)$. Two cases now arise for $\lambda(\bigcup_n E_n)$:

1. If $\lambda(E) = 0$: Then $\lambda(\bigcup_n E_n) = 0$. But
$$\lambda((-1, 2)) = 3 \leq \lambda\left(\bigcup_n E_n\right) \quad \text{(Property 3)}.$$
Therefore we have a contradiction.
2. If $\lambda(E) \neq 0$: Then $\lambda(\bigcup_n E_n) = \sum_n \lambda(E) = +\infty$. But
$$\lambda\left(\bigcup_n E_n\right) \leq \lambda((-1, 2)) = 3 \quad \text{(Property 2)}.$$
We again have a contradiction.

Hence, the set E is just not Lebesgue measurable! $\square$

18.2.9 Regularity

First consider the following proposition.

[6] Proof?

Proposition 18.2.9.1. *Consider $E \subseteq \mathbb{R}$. The following statements are equivalent:*

1. *E is Lebesgue measurable.*
2. *$\forall\, \epsilon > 0$, $\exists$ an open set O such that*
$$E \subseteq O \text{ and } \lambda^* (O \setminus E) < \epsilon.$$
3. *$\exists$ a G_δ set G such that*
$$E \subseteq G \text{ and } \lambda^* (G \setminus E) = 0.$$

Proof. The equivalence of each statement is as follows:
1 $\implies$ 2. Consider $E \subseteq \mathbb{R}$ to be Lebesgue measurable. By above, for any $E \subseteq \mathbb{R}$ and any $\epsilon > 0$, there exists open set U such that $E \subseteq U$ which satisfies

$$\lambda^* (U) \leq \lambda^* (E) + \epsilon.$$

Now since $E \subseteq U$, therefore,

$$\begin{aligned} \lambda^* (U \setminus E) &= \lambda^* (U) - \lambda^* (E) \\ &\leq \epsilon \end{aligned}$$

2 $\implies$ 3. Similarly, the above shows that there exists a G_δ set G such that $E \subseteq G$ which satisfies $\lambda^* (E) = \lambda^* (G)$. This directly means that $\lambda^* (G \setminus E) = 0$ because $E \subseteq G$ so $\lambda^* (G \setminus E) = \lambda^* (G) - \lambda^* (E)$.

3 $\implies$ 1. Since G is G_δ set therefore it is intersection of open sets in $\mathbb{R}$. Now since any open set in $\mathbb{R}$ is an union of open intervals (Homework I, 1) which is Lebesgue measurable and therefore G is Lebesgue measurable. Now, we can write E as

$$E = G \setminus (G \setminus E)$$

where $G \setminus E$ is such that (from Statement 3) $\lambda^* (G \setminus E) = 0$, therefore, by Proposition 18.2.6.4, $G \setminus E$ is Lebesgue measurable. Hence E is also Lebesgue measurable.

□

Now, consider the next proposition, which is dual of the above.

Proposition 18.2.9.2. *Consider $E \subseteq \mathbb{R}$. The following statements are equivalent:*

1. *E is Lebesgue measurable.*
2. *$\forall\, \epsilon > 0$, $\exists$ closed set C such that*
$$C \subseteq E \text{ and } \lambda^* (E \setminus C) < \epsilon.$$
3. *$\exists$ a F_σ set F such that*
$$F \subseteq E \text{ and } \lambda^* (E \setminus F) = 0.$$

Proof. Implications are as follows:
1 $\implies$ 2. Suppose $E \subseteq \mathbb{R}$ is Lebesgue measurable. Note that if E is Lebesgue measurable (that is $E \in \mathcal{M}_{\lambda^*}$), then E^c is also Lebesgue measurable as $\mathcal{M}_{\lambda^*}$ is a σ-algebra (Theorem 18.2.6.5). Hence, using Proposition 18.2.9.1 on E^c gives us an open set O for all $\epsilon > 0$ such that $E^c \subseteq O$ and

$\lambda^*(O \setminus E^c) < \epsilon$. Now let's take it's complement. Therefore, $C = O^c \subseteq E$ where C is clearly closed. Now, $E \setminus O^c = O \setminus E^c$[7]. Now,

$$\begin{aligned}\lambda^*(E \setminus O^c) &= \lambda^*(O \setminus E^c)\\ &< \epsilon\end{aligned}$$

which proves the first implication.

2 $\implies$ 3. From Proposition 18.2.9.1, we have that $\exists$ a G_δ set G such that $E^c \subseteq G$ and $\lambda^*(G \setminus E^c) = 0$. Note that the complement of countable intersection of open sets is countable union of closed sets. Therefore, $F = G^c$ is an F_σ set. Now, $G^c \subseteq (E^c)^c = E$. Now, we know that $E \setminus G^c = G \setminus E^c$. Therefore, we have the result as follows:

$$\begin{aligned}\lambda^*(E \setminus G^c) &= \lambda^*(G \setminus E^c)\\ &= 0.\end{aligned}$$

3 $\implies$ 1. Since F is an F_σ set, therefore, $F \in \mathcal{M}_{\lambda^*}$. Moreover, as Statement 2 show, $\lambda^*(E \setminus F) = 0$, thus by Proposition 18.2.6.4, $E \setminus F \in \mathcal{M}_{\lambda^*}$. Since,

$$E = F \cup (E \setminus F)$$

that is E is union of two Lebesgue measurable sets, therefore $E \in \mathcal{M}_{\lambda^*}$, completing the proof. □

Definition 18.2.9.3. (**Complete measure Space**) The measure space $(X, \mathcal{A}, \mu)$ is complete if the for any $A \in \mathcal{A}$ such that $\mu(A) = 0$ implies that for any subset $B \subseteq A$,

$$\mu(B) = 0.$$

Remark 18.2.9.4. Trivial to see are the following:

- Hence, if μ^* is an outer measure defined on X, then the space $(X, \mathcal{M}_{\mu^*}, \mu^*)$ is complete (follows from Proposition 18.2.6.4).
- This means that the Lebesgue outer measure restricted to Lebesgue measurable subsets of $\mathbb{R}$, $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ is complete.

Definition 18.2.9.5. (**Completion of a measure Space**) Let $(X, \mathcal{A})$ be a measurable space and let μ be a measure on $\mathcal{A}$. The completion of $\mathcal{A}$ under μ is the collection $\mathcal{A}_\mu$ of subsets $A \subseteq X$ for which there are sets E and F in $\mathcal{A}$ such that

$$E \subseteq A \subseteq F$$

and

$$\mu(F - E) = 0^{8}.$$

[7] It's not difficult to see as for any $x \in E \setminus O^c$, $x \in E$ but $x \notin O^c$. Therefore, $x \in O$ but $x \notin E^c$, that is $x \in O \setminus E^c$. Similarly for the converse.

[8] Note that, in Exercise III, Q. 2, we proved that for any $A \in \mathcal{A}$, this is trivially true. That is, all $\mathcal{A}$-measurable subsets are $\mathcal{A}_\mu$-measurable. In particular, E was a F_σ set and F was a G_δ set.

18.3 Measurable functions

We now see the definition and basic properties of measurable Functions, which would later be used to define Lebesgue integral.

Definition 18.3.0.1. (**Measurable function**) Let $(X, \mathcal{A})$ be a measurable space and let $A \subseteq X$ which is in $\mathcal{A}$. The function $f : A \to [-\infty, +\infty]$, is called a measurable function[9] if

$$\{x \mid f(x) > \alpha\} \text{ for any } \alpha \in \mathbb{R} \text{ is measurable (belongs in } \mathcal{A}).$$

Remark 18.3.0.2. Please note that the function f defined above has a measurable domain.

Proposition 18.3.0.3. *Let $(X, \mathcal{A})$ be a measurable space and $A \in \mathcal{A}$. Let $f : A \to [-\infty, +\infty]$ be a function. Then, the following statements are equivalent:*

1. *f is a measurable function.*
2. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$.*
3. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) < \alpha\} \in \mathcal{A}$.*
4. *For all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) \leq \alpha\} \in \mathcal{A}$.*

Proof. The equivalence is shown as follows:
1 $\implies$ 2. Since f is a measurable, therefore for all $\alpha \in \mathbb{R}$, the set $\{x \mid f(x) > \alpha\} \in \mathcal{A}$. This means that $C_{\alpha - \frac{1}{n}} = \left\{x \mid f(x) > \alpha - \frac{1}{n}\right\} \in \mathcal{A}$ for all $n \in \mathbb{N}$. Now, the following set

$$C = \bigcap_n C_{\alpha - \frac{1}{n}} = \{x \mid f(x) \geq \alpha\}.$$

is measurable as $C \in \mathcal{A}$ because $C_{\alpha - \frac{1}{n}} \in \mathcal{A}$ for any $n \in \mathbb{N}$, hence the countable intersection would also be in $\mathcal{A}$, hence measurable.
2 $\implies$ 3. Since $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$, therefore it's complement $\{x \mid f(x) < \alpha\} \in \mathcal{A}$ for any $\alpha \in \mathbb{R}$.
3 $\implies$ 4. Since $\{x \mid f(x) < \alpha\} \in \mathcal{A}$ for any $\alpha \in \mathbb{R}$, thus, $C_{\alpha + \frac{1}{n}} = \{x \mid f(x) < \alpha + \frac{1}{n}\} \in \mathcal{A}$ for all $n \in \mathbb{N}$, hence

$$C = \bigcap_n C_{\alpha + \frac{1}{n}} = \{x \mid f(x) \leq \alpha\}$$

and since each $C_{\alpha + \frac{1}{n}} \in \mathcal{A}$, therefore $C \in \mathcal{A}$.
4 $\implies$ 1. Since $\{x \mid f(x) \leq \alpha\} \in \mathcal{A}$ then it's complement $\{x \mid f(x) > \alpha\}$ for all $\alpha \in \mathbb{R}$, making f measurable. □

Proposition 18.3.0.4. *The following are basic examples of measurable functions:*

- *If f is a measurable function, then the set $\{x \mid f(x) = \alpha\}$ is measurable for all $\alpha \in R$.*
- *Constant functions are measurable.*
- *The characteristic function χ_A defined by:*

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

is measurable if and only if A is measurable.

[9] One writes f as $\mathcal{A}$-measurable function to denote the σ-algebra over whose subset the function f is defined.

- *Continuous functions are measurable.*
- *Let $(X, \mathcal{A})$ be a measurable space. If f and g are measurable functions on X, then the sets*

$$\{x \in X \mid f(x) \neq g(x)\}$$
$$\{x \in X \mid f(x) < g(x)\}$$

are measurable (belongs to $\mathcal{A}$).
- [10] *Monotone functions are measurable.*
- [11]*Consider $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function. Then f' is a λ-measurable function.*

Proof. The **first** example is trivial to see in light of Proposition 18.3.0.3 by taking intersection of $\{x \mid f(x) \leq \alpha\}$ and $\{x \mid f(x) \geq \alpha\}$, both of which are measurable.
For **second**, consider the constant function $f(x) = b \ \forall \ x \in \mathbb{R}$. Now, for all $\alpha \in \mathbb{R}$, consider the set $f^{-1}((\alpha, \infty)) = \{x \mid f(x) > \alpha\}$. If $b > \alpha$, then we are done, if $b \leq \alpha$, then by previous result, $\{x \mid f(x) \leq \alpha\}$ is also measurable (equal to $\mathbb{R}$ and $\mathbb{R} \in \mathcal{A}$).
For **third** example, consider the set ${\chi_A}^{-1}(\alpha, \infty) = \{x \mid \chi_A(x) > \alpha\}$ for any $\alpha \in \mathbb{R}$. If $\alpha > 1$, then $f^{-1}(\alpha, \infty) = \Phi \in \mathcal{A}$. If $\alpha = 1$, then $f^{-1}[\alpha, \infty) = A$, since $\chi_A(x)$ is given measurable, hence A is measurable. Now, Assume that A is measurable. Then consider the set ${\chi_A}^{-1}(\alpha, \infty)$ for any $\alpha \in \mathbb{R}$. As we saw previously, the case for $\alpha > 1$ is trivial. For $0 < \alpha \leq 1$, ${\chi_A}^{-1}(\alpha, \infty) = A \in \mathcal{A}$. Finally, for $\alpha \leq 0$, ${\chi_A}^{-1}(-\infty, \alpha] = \Phi \in \mathcal{A}$. Thus, χ_A is measurable.
For **fourth**, since f is continuous (so inverse of open sets is open, by definition), therefore $f^{-1}(\alpha, \infty)$ is open in $\mathbb{R}$, hence it must be Borel, hence measurable for any $\alpha \in \mathbb{R}$.
For **fifth**, since f and g are measurable. Then due to next Proposition 18.3.0.5, we know that $f - g$ is also measurable. This means that for any $\alpha \in \mathbb{R}$,

$$\{x \in X \mid f(x) - g(x) < \alpha\}$$

is measurable. Now set $\alpha = 0$ to get the result. Moreover, from this, we also get that $\{x \in X \mid f(x) - g(x) > 0\}$ is also measurable. Hence,

$$\{x \in X \mid f(x) - g(x) \neq 0\} = \{x \in X \mid f(x) - g(x) < 0\} \bigcup \{x \in X \mid f(x) - g(x) > 0\}$$

is also measurable.
For **sixth**, we proceed as follows:
Consider the function $f : A \to \mathbb{R}$ where $A \in \mathcal{M}_{\lambda^*}$ to be monotone. Now, consider the following two sets for any $\alpha \in \mathbb{R}$:

$$A_1 = \{x \in A \mid f(x) > \alpha\}$$
$$A_2 = \left(f^{-1}(\alpha), \infty\right) \cap A$$

Now, take any $x \in A_1$, then $f(x) > \alpha \implies x > f^{-1}(\alpha)$. Now if $f^{-1}(\alpha) \cap A = \Phi$, then $\{x \in A \mid f(x) > \alpha\} = \Phi$ which is trivially measurable and we would be done. If however $f^{-1}(\alpha) \cap A \neq \Phi$, then $f^{-1}(\alpha) = \{y \in A \mid f(y) = \alpha\}$ so that $f(y) > \alpha$ implies that $y > f^{-1}(\alpha)$ so that

[10]Question 3 of Exercise 3.
[11]Question 4 of Exercise 3.

$f(y) > f(f^{-1}(\alpha)) = \alpha$. Therefore, $x > f^{-1}(\alpha)$, that is $x \in A_2$, proving that $A_1 \subseteq A_2$.
Similarly, take $x \in A_2$, therefore

$$\begin{aligned} x &> f^{-1}(\alpha) \\ f(x) &> f\left(f^{-1}(\alpha)\right) \\ f(x) &> \alpha \\ x &\in \{x \mid f(x) > \alpha\} \\ x &\in A_1. \end{aligned}$$

Therefore $A_2 \subseteq A_1$. Hence, $A_1 = A_2$. But since $\left(f^{-1}(\alpha), \infty\right)$ is an interval, hence measurable and A is given measurable, therefore $A_2 = \left(f^{-1}(\alpha), \infty\right) \cap A$ is measurable, which makes $A_1 = \{x \mid f(x) > \alpha\} = A_2$ measurable for all $\alpha \in \mathbb{R}$.
For **seventh**, the result is simple to see since we are given that f is λ-measurable due to continuity (see Statement 4). Therefore, we can define the sequence of functions $\{f_n\}$ as follows:

$$f_n(x) = \frac{f\left(x + \frac{1}{n}\right) - f(x)}{\frac{1}{n}} \forall\, x \in \mathbb{R}.$$

As we can see, f_n is λ-measurable due to Proposition 18.3.0.5. Hence, we can see that because $f'(x) = \varinjlim_{h \to 0} \frac{f(x+h)-f(x)}{h}$ for any $x \in \mathbb{R}$, and since $\varinjlim_{n\to\infty} f_n(x) = f(x)$, therefore $f_n \to f'$ is λ-measurable (Proposition 18.3.0.9). $\square$

Proposition 18.3.0.5. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$. Consider two measurable functions $f, g : A \longrightarrow [0, +\infty]$ and $c \in \mathbb{R}$. Then,*

1. $f + c$,
2. $f \pm g$
3. cf,
4. fg

are also measurable.

Proof. **1.** Since f is measurable, therefore the set $\{x \mid f(x) > \alpha - c\} = \{x \mid f(x) + c \geq \alpha\}$ is measurable for any $\alpha \in \mathbb{R}$.
2. Both f and g are given measurable. The set $(f+g)^{-1}(\alpha, \infty)$ can be written as:

$$\begin{aligned} (f+g)^{-1}(\alpha, \infty) &= \{x \mid f(x) + g(x) > \alpha\} \\ &= \{x \mid f(x) > \alpha - g(x)\} \\ &= \{x \mid f(x) > b\} \end{aligned}$$

where $b \in [-\infty, \alpha]$. Note that the case where $g(x) = +\infty$ is trivial as $f(x) > \alpha - (+\infty) \equiv f(x) > -\infty$, which is by definition of co-domain of f. Now since $\{x \mid f(x) > b\}$ is measurable for any $b \in \mathbb{R} \supset (-\infty, \alpha]$ for any $\alpha \in \mathbb{R}$, therefore $(f+g)^{-1}(\alpha, \infty)$ is measurable for any $\alpha \in \mathbb{R}$.
3. Note that for $c = 0$, the function becomes constant and hence measurable (Proposition 18.3.0.4). Consider the set $(cf)^{-1}(\alpha, \infty)$. We can write this as follows,

$$\begin{aligned} (cf)^{-1}(\alpha, \infty) &= \{x \mid cf(x) > \alpha\} \\ &= \{x \mid f(x) > \alpha/c\} \end{aligned}$$

where $c \neq 0$. Since f is measurable, therefore $\{x \mid f(x) > \alpha/c\}$ is also measurable for any $\alpha \in \mathbb{R}$. Hence cf is measurable.
4. Consider the set $(f^2)^{-1}(-\infty, \alpha)$ for any $\alpha \in \mathbb{R}$.

$$\begin{aligned}(f^2)^{-1}(-\infty, \alpha) &= \{x \mid f^2(x) < \alpha\} \\ &= \{x \mid -\sqrt{\alpha} < f(x) < \sqrt{\alpha}\} \\ &= \{x \mid f(x) < \sqrt{\alpha}\} \bigcap \{x \mid f(x) > -\sqrt{\alpha}\}\end{aligned}$$

Therefore if f is measurable, then f^2 is measurable. With this, we can simply write fg as:

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}$$

which, by previous results (2 & 3), is measurable. $\square$

Proposition 18.3.0.6. [12] *Let $(X, \mathcal{A})$ be a measurable space. Consider a function $f : A \to \mathbb{R}$ where $A \in \mathcal{A}$. Then the following are equivalent:*
1. *f is a $\mathcal{A}$-measurable function.*
2. *$f^{-1}(U)$ is a measurable set $\forall$ open sets $U \subseteq \mathbb{R}$.*
3. *$f^{-1}(C)$ is a measurable set $\forall$ closed sets $C \subseteq \mathbb{R}$.*
4. *$f^{-1}(B)$ is a measurable set $\forall$ borel sets $B \in \mathcal{B}(R)$.*

Proof. *The proof is exactly the same as of Proposition 18.3.2.2.* $\square$

Definition 18.3.0.7. (**sequence of fuctions**) If $\{f_n\}$ is a sequence of $[-\infty, +\infty]$ valued functions on A, then $\sup_n f_n : A \to [-\infty, +\infty]$ is defined by

$$\left(\sup_n f_n\right)(x) = \sup\{f_n(x) \mid n \in \mathbb{N}\}.$$

Remark 18.3.0.8. One similarly defines the following:
- The infimum:
$$\left(\inf_n f_n\right)(x) = \inf\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit supremum:
$$\left(\limsup_n f_n\right)(x) = \limsup\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit infimum:
$$\left(\liminf_n f_n\right)(x) = \liminf\{f_n(x) \mid n \in \mathbb{N}\}.$$
- The limit:
$$\left(\varinjlim_n f_n\right)(x) = \varinjlim\{f_n(x) \mid n \in \mathbb{N}\}.$$

Proposition 18.3.0.9. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$. Consider $\{f_n\}$ be a sequence of $[-\infty, +\infty]$-valued measurable functions on A. Then,*

[12] Question 1 of Exercise 3.

1. *The functions* $\sup_n f_n$ *and* $\inf_n f_n$ *are measurable.*
2. *The functions* $\limsup_n f_n$ *and* $\liminf_n f_n$ *are measurable.*
3. *The function* $\varinjlim_n f_n$ *(whose domain is* $\{x \in A \mid \limsup_n f_n(x) = \liminf_n f_n(x)\}$*) is measurable.*

Proof. Note that the set $(\sup_n f_n)^{-1}(-\infty, \alpha] = \{x \in A \mid (\sup_n f_n)(x) \leq \alpha\} = \bigcap_n \{x \in A \mid f_n(x) \leq \alpha\}$. Therefore $\sup_n f_n$ is measurable. Similarly, $(\inf_n f_n)^{-1}(-\infty, \alpha) = \{x \in A \mid (\inf_n f_n)(x) < \alpha\} = \bigcup_n \{x \in A \mid f_n(x) < \alpha\}$. Now, denote $g_k = \sup_{n \geq k} f_n$ and $h_k = \inf_{n \geq k} f_n$. But since $\limsup_n f_n = \inf_{n \geq 0} \sup_{k \geq n} f_k = \inf_{n \geq 0} g_n$ and $\{g_n\}$ is measurable by 1^{st} property, therefore $\limsup_n f_n$ is also measurable, similarly for $\liminf_n$. □

18.3.1 Almost everywhere property.

Definition 18.3.1.1. (**μ-almost everywhere**) Let $(X, \mathcal{A}, \mu)$ be a measure space. A property P of points of X is said to hold μ-almost everywhere if the set

$$N = \{x \in X \mid P \text{ does not hold for } x\}$$

has measure zero. That is,

$$\mu(N) = 0.$$

Remark 18.3.1.2. Note that it's not necessary for the set N to belong in $\mathcal{A}$. The only requirement is for the set N to be contained in a set $F \in \mathcal{A}$ and then $\mu(F) = 0$ (which automatically implies that $\mu^*(N) = 0$).

But, if μ is complete then $N \in \mathcal{A}$. See Definition 18.2.9.3.

Definition 18.3.1.3. (**Almost everywhere convergence**) If $\{f_n\}$ is a sequence of functions on X and f is a function on X, then

$$\{f_n\} \longrightarrow f \text{ almost everywhere.}$$

if the set

$$\{x \in X \mid f(x) \neq \varinjlim_n f_n(x)\}$$

is of measure zero.

Proposition 18.3.1.4. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and let* f *and* g *be extended real valued functions on* X *that are equal almost everywhere. If* μ *is **complete** and if* f *is measurable, then* g *is also measurable.*

Proof. Consider the region of non-equality as

$$N = \{x \mid f(x) \neq g(x)\}.$$

Given to us is the fact that $\mu^*(N) = 0$ and since μ is complete, so $N \in \mathcal{A}$. Now, consider the following for any $\alpha \in \mathbb{R}$:

$$\{x \mid g(x) \geq \alpha\} = (\{x \mid g(x) \geq \alpha\} \cap N) \bigcup (\{x \mid g(x) \geq \alpha\} \cap N^c).$$

Denote the set $A = \{x \mid g(x) \geq \alpha\} \cap N$ and $B = \{x \mid g(x) \geq \alpha\} \cap N^c$. Since for any $x \in (\{x \mid g(x) \geq \alpha\} \cap N^c)$, $f(x) = g(x)$, therefore, we can equivalently write $B = (\{x \mid f(x) \geq \alpha\} \cap N^c)$.

Now $N^c \in \mathcal{A}$ and due to Measurability of f, $\{x \mid f(x) \geq \alpha\} \in \mathcal{A}$. Hence $B \in \mathcal{A}$. Finally, due to $\{x \mid g(x) \geq \alpha\} \cap N \subseteq N$ and μ being complete with $\mu(N) = 0$, we get $\{x \mid g(x) \geq \alpha\} \cap N \in \mathcal{A}$, completing the proof. ☐

Proposition 18.3.1.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, let $\{f_n\}$ be sequence of extended real valued functions on X and let f be an extended real valued function on X such that*

$$\{f_n\} \longrightarrow f \text{ almost everywhere.}$$

*If μ is **complete** and if each f_n is measurable, then f is measurable.*

Proof. As Proposition 18.3.0.9 shows, $\liminf_n f_n$ and $\limsup_n f_n$ are measurable. As the given condition shows, $\liminf_n f_n$ is equal to f for almost all X. Hence Proposition 18.3.1.4 implies that f is also measurable. ☐

18.3.2 Cantor set

With the new tool in hand (measurable functions), we now turn back to the ever-interesting Cantor set, this time, to prove the sheer size of the σ-algebra $\mathcal{M}_{\lambda^*}$ in comparison to the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. In particular we show that $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{M}_{\lambda^*}$.
But before that, we look at following results:

Proposition 18.3.2.1. *The function ϕ defined by*

$$\phi : [0,1] \longrightarrow P$$
$$\phi(\alpha) = \sum_{n=1}^{\infty} \frac{2b_n}{3^n} \text{ for } \alpha \in [0,1]$$

where $b_n \in \{0,1\}\ \forall\, n \in \mathbb{N}$ is measurable in $\mathcal{M}_{\lambda^*}$.

Proof. Note that $\phi(\alpha)$ thus maps a decimal number to it's binary representation $\{b_n\}$. First, we define the following function:

$$\phi_n : [0,1] \longrightarrow \{0,1\}$$
$$\phi_n(\alpha) = b_n.$$

That is, ϕ_n maps α to it's n^{th} binary digit. We can see that $\phi_n(\alpha)$ can be written as the following:

$$\phi_n(\alpha) = \chi_{E_n} = \begin{cases} 1 & \text{if } \alpha \in E_n \\ 0 & \text{otherwise.} \end{cases}$$

where E_n is the intersection of countable sequence of sub-intervals of $[0,1]$. Hence E_n is a Lebesgue measurable subset of $\mathbb{R}$, so it is in $\mathcal{M}_{\lambda^*}$. But, as Proposition 18.3.0.4, statement 3 shows, $\chi_{E_n} = \phi_n$

is then a measurable function.
Now, the following arguments:

$$\frac{2}{3^n}\phi_n(\alpha) = \frac{2b_n}{3^n} \text{ is measurable (Proposition 18.3.0.5).}$$
$$\left\{\frac{2\phi_n(\alpha)}{3^n}\right\} \text{ is a sequence of measurale functions.}$$
$$\left\{\sum_{k=1}^{n}\frac{2\phi_n(\alpha)}{3^n}\right\} \text{ is also a sequence of measurable functions (Proposition 18.3.0.5).}$$
$$\underleftarrow{\lim}_{n\to\infty}\sum_{k=1}^{n}\frac{2\phi_n(\alpha)}{3^n} \text{ is a measurable function (Proposition 18.3.0.9).}$$

Hence the function which maps each real from $[0,1]$ to it's binary representation is measurable. $\square$

Proposition 18.3.2.2. *Let $(X, \mathcal{A})$ be a measurable space. If f is a $\mathcal{A}$-measurable function on A and $B \in \mathcal{B}(\mathbb{R})$, then $f^{-1}(B) \in \mathcal{A}$.*

Proof. Denote $\mathcal{D}$ be the following set:

$$\mathcal{D} = \{B \subseteq \mathbb{R} \mid f^{-1}(B) \in \mathcal{A}\}.$$

Now, note that,

1. Since $f^{-1}(\mathbb{R}) = A \in \mathcal{A}$, therefore $\mathbb{R} \in \mathcal{D}$.
2. If $B \in \mathcal{D}$, then

$$B^c = \mathbb{R} \cap B^c$$

and

$$\begin{aligned} f^{-1}(B^c) &= f^{-1}(\mathbb{R} \cap B^c) \\ &= f^{-1}(\mathbb{R}) \cap f^{-1}(B^c) \\ &= A \cap \left(f^{-1}(B)\right)^c \end{aligned}$$

Now since $A \in \mathcal{A}$ and $f^{-1}(B) \in \mathcal{A}$ because $B \in \mathcal{D}$, therefore $f^{-1}(B^c) \in \mathcal{A}$ so that $B^c \in \mathcal{D}$.

3. We know that from the basic results of set functions that

$$f^{-1}\left(\bigcup_n B_n\right) = \bigcup_n f^{-1}(B_n)$$

Hence $\mathcal{D}$ is a σ-algebra on $\mathbb{R}$ (!) Now, due to measurability of f, we know that the set $\{x \mid f(x) > \alpha\}$ is in $\mathcal{A}$, which is equivalent to saying that $f^{-1}(\alpha, \infty) \in \mathcal{A}$. This hence means that $(\alpha, \infty) \in \mathcal{D}$ for any $\alpha \in \mathbb{R}$. Proposition 18.2.1.8 showed that a σ-algebra generated by such subsets of $\mathbb{R}$ is $\mathcal{B}(\mathbb{R})$. Hence, for any $B \in \mathcal{B}(\mathbb{R})$, we have that $B \in \mathcal{D}$. Therefore for any Borel set B, $f^{-1}(B) \in \mathcal{A}$[13]. $\square$

[13] This is a very interesting way to prove such a statement. Notice how we analyzed the set of all possible subsets of $\mathbb{R}$ for which $f^{-1}(B) \in \mathcal{A}$ right from the start!

18.3.3 sequence of functions approximating a measurable function.

We now show that any measurable function can be defined in terms of a simple function and a step function. For this, we first define what we mean by simple functions in Definition 18.3.3.4. Before that, let's see few more interesting-but-basic properties of measurable functions.

Proposition 18.3.3.1. *Let $(X, \mathcal{A})$ be a measurable space and f be an extended real valued function on $A \in \mathcal{A}$. Define the following:*

$$f^+(x) = \max(f(x), 0) \textit{ and } f^-(x) = -\min(f(x), 0).$$

Then, f is measurable if and only if f^+ and f^- both are measurable on A.

Proof. If f is measurable, then $\{x \mid f(x) \geq \alpha\}$ is measurable. Note that $f^+(x) \geq 0$. Hence, for the case when $\alpha > 0$, the set $\{x \mid f^+(x) \geq \alpha\} = \{x \mid f(x) > \alpha\}$ which is measurable due to measurability of f. Similarly, if $\alpha = 0$, then $\{x \mid f^+(x) \geq 0\} = \{x \mid f(x) > 0\} \bigcup \{x \mid f(x) = 0\}$ in which both sets are measurable in view of Proposition 18.3.0.4. Finally, for $\alpha < 0$, we have $\{x \mid f^+(x) > \alpha\} = \{x \mid f^+(x) \geq 0\}$ which again is measurable. Now, $f^-(x) = -\min(f(x), 0) = \max(-f(x), 0)$ and since $-f$ is also measurable (Proposition 18.3.0.5), therefore if f is a measurable function, then f^+ and f^- are both measurable functions too.
To show the converse, note that $f = f^+ - f^-$ and since both are measurable, therefore f is also measurable (Proposition 18.3.0.5). □

Remark 18.3.3.2. Due to the above result, we can hence deduce that if f is a $\mathcal{A}$-measurable function then,

$$|f| = f^+ + f^- \text{ is a measurable function on } A.$$

Proposition 18.3.3.3. *Let $(X, \mathcal{A})$ be a measurable space and $A \in \mathcal{A}$. Let $f : A \to [-\infty, +\infty]$. Then,*

1. *If f is $\mathcal{A}$-measurable and if B is a subset of A, then the restriction f_B of f to B is also $\mathcal{A}$-measurable.*
2. *If $\{B_n\}$ is a sequence of sets that belong to $\mathcal{A}$ such that $A = \bigcup_n B_n$ and f_{B_n} is $\mathcal{A}$-measurable for each n, then f is also $\mathcal{A}$-measurable.*

Proof. The first result follows directly from the following observation:

$$\{x \in B \mid f_b(x) > \alpha\} = B \bigcap \{x \in A \mid f(x) > \alpha\}$$

and the second result follows from the following:

$$\{x \in A \mid f(x) > \alpha\} = \bigcup_n \{x \in B_n \mid f_{B_n}(x) > \alpha\}.$$

both for any $\alpha \in \mathbb{R}$. □

Definition 18.3.3.4. (**Simple Function**) A function is called simple if it has only finitely many values. Equivalently, we say that f is simple if we can write it as the following:

$$f = \sum_{k=1}^{N} \alpha_k \chi_{E_k} \ , \quad \alpha_k \in \mathbb{R}$$

where each E_k is a measurable set of finite measure.

Remark 18.3.3.5. Note that

- If E_k are intervals, then we say f to be a **step function**.

The following Proposition asserts that ***any measurable function can be approximated by an increasing sequence of simple functions.***

Proposition 18.3.3.6. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$ with $f : A \to [0, +\infty]$ be a measurable function on A. Then there exists a sequence $\{f_n\}$ of* ***simple*** *$[0, +\infty)$-valued measurable functions on A that satisfy*

$$f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots$$

and

$$\lim_{\substack{\longleftarrow \\ n \to \infty}} f_n(x) = f(x)$$

for any $x \in A$.

Proof. For the proof, construct the following sequence of sets, by dividing the whole interval $[0, n]$ for any $n \in \mathbb{N}$ into $n2^n$ number of intervals each of length $\frac{1}{2^n}$ and denote the following set:

$$A_{n,k} = \left\{ x \in A \,\middle|\, \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\}$$

for any $n \in \mathbb{N}$ and $k = 1, 2, \dots, n2^n$. With this construction, we can now define the following function for each n:

$$\phi_n : A \to [0, \infty), \text{ defined as}$$

$$\phi_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } x \in A_{n,k} \text{ for any } k = 1, 2, \dots, n2^n \\ n & \text{if } x \in A - \bigcup_k A_{n,k}. \end{cases}$$

Note that we can alternatively write $\phi_n(x)$ as the following (with more clarity):

$$\phi_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } f(x) \leq n \text{ , } \quad \text{where } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \text{ for some k} \in \{1, 2, \dots, n2^n\} \\ n & \text{if } f(x) > n. \end{cases}$$

We now show that $\phi_n(x) \leq \phi_{n+1}(x) \; \forall \; x \in A$. Let's first show this for $f(x) \leq n$.
If $f(x) \leq n$, then,

$$\phi_n(x) = \frac{k_0 - 1}{2^n} \text{ for some } k_0 \in \{1, 2, \dots, n2^n\}$$

such that $\frac{k_0-1}{2^n} \leq f(x) < \frac{k_0}{2^n}$. Now, two cases arises:

- **If** $\frac{k_0-1}{2^n} \leq f(x) < \frac{2k_0-1}{2^{n+1}}$: This is just the case that $f(x)$ lies in the first half of the interval $\left[\frac{k_0-1}{2^n}, \frac{k_0}{2^n}\right]$. Hence, in this case we get that:

$$\frac{k_0 - 1}{2^n} = \frac{2k_0 - 2}{2^{n+1}} \leq f(x) < \frac{2k_0 - 1}{2^{n+1}}$$

  such that $\phi_n(x) = \frac{k_0-1}{2^n} = \phi_{n+1}(x)$.

- **If** $\frac{2k_0-1}{2^{n+1}} \leq f(x) < \frac{k_0}{2^n}$: This is the case when $f(x)$ lies in the second half of the interval. In this case, we see that,
$$\frac{2k_0-1}{2^{n+1}} \leq f(x) < \frac{2k_0}{2^{n+1}} = \frac{k_0}{2^n}$$
so that $\phi_n(x) = \frac{k_0-1}{2^n} = \frac{2k_0-2}{2^{n+1}} < \frac{2k_0-1}{2^{n+1}} = \phi_{n+1}(x)$.

Hence from both the cases, we have $\phi_n(x) \leq \phi_{n+1}(x)$ for all $x \in A$ such that $f(x) \leq n$. One can similarly see the same result for $n < f(x) \leq n+1$ and for $f(x) > n+1$, $\phi_n(x) \leq \phi_{n+1}(x)$ follows trivially. Hence, we have proved that $\forall n \in \mathbb{N}$ and $x \in A$,

$$\phi_n(x) \leq \phi_{n+1}(x). \tag{18.6}$$

Now, one can write the function ϕ_n as the following combination too:

$$\phi_n(x) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\chi_{A_{n,k}} + n\chi_{A-\bigcup_k A_{n,k}} \tag{18.7}$$

Due to the above representation of ϕ_n, the following steps becomes easier (& interesting) to see. Now, first note that $A_{n,k}$ is a measurable set because it's intersection of two measurable sets. Moreover, $A - \bigcup_k A_{n,k}$ is also a measurable set. Hence, in view of Proposition 18.3.0.4, Statement 3, we get that $\phi_n(x)$ is a measurable function for any $n \in \mathbb{N}$. Therefore, $\{\phi_n\}$ is a sequence of measurable functions adhering (18.6). We again find two cases:

- **If f is finite** : Now since f is finite, therefore $\exists n_0 \in \mathbb{N}$ such that $f(x) \leq n_0$. Hence, one can further deduce the following for all $n \geq n_0$ (hence $f(x) \leq n_0 \leq n$),
$$f(x) - \phi_n(x) = f(x) - \frac{k-1}{2^n} \text{ for some } k \in \{1, 2, \dots, n2^n\} \text{ such that } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$$
$$< \frac{1}{2^n}$$
Hence, as $n \to \infty$, $|f(x) - \phi_n(x)| \to 0$.
- **If f is infinite for some $x \in A$** : If f is infinite, then $\forall n \in \mathbb{N}$, $f(x) > n$. Hence,
$$\phi_n(x) = n \text{ for all } n \in N.$$
Therefore $\varinjlim_{n\to\infty} \phi_n(x) = +\infty = f(x)$ for particular $x \in A$ where f is infinity.

Hence, in both cases, $\{\phi_n\}$ converges to f. The proof is therefore complete. □

*The following can be considered as an **important corollary of the above Proposition**.*

Proposition 18.3.3.7. *Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$ with $f : A \to [-\infty, +\infty]$ be a measurable function on A. Then there exists a sequence $\{f_n\}$ of **simple** $(-\infty, +\infty)$-valued measurable functions on A that satisfy*
$$|f_1(x)| \leq |f_2(x)| \leq |f_3(x)| \leq \dots$$
and
$$\varinjlim_{n\to\infty} f_n(x) = f(x)$$
for any $x \in A$.

Proof. Since f is a measurable function, therefore f^+ and f^- are measurable functions too (Proposition 18.3.3.1). Now, since any function f can be written as

$$f = f^+ - f^-$$

therefore, by Proposition 18.3.3.6, we have two sequences $\{f_n^{(1)}\}$ and $\{f_n^{(2)}\}$ such that

$$f_n^{(1)} \longrightarrow f^+ \text{ and } f_n^{(2)} \longrightarrow f^-$$

where $f_1^{(1)}(x) \leq f_2^{(1)}(x) \leq \dots$ and $f_1^{(2)}(x) \leq f_2^{(2)}(x) \leq \dots$. Denote

$$f_n(x) = f_n^{(1)}(x) - f_n^{(2)}(x)$$

Therefore, we see that

$$|f_n(x)| = f_n^{(1)}(x) + f_n^{(2)}(x) \leq f_{n+1}^{(1)}(x) + f_{n+1}^{(2)}(x) = |f_{n+1}(x)|$$

Now, we can deduce that

$$\begin{aligned}
|f(x) - f_n(x)| &= \left| f^+(x) - f^-(x) - f_n^{(1)}(x) + f_n^{(2)}(x) \right| \\
&= \left| f^+(x) - f_n^{(1)}(x) - \left(f^-(x) - f_n^{(2)}(x) \right) \right| \\
&\leq \left| f^+(x) - f_n^{(1)}(x) \right| + \left| f^-(x) - f_n^{(2)}(x) \right| \\
&\to 0 + 0
\end{aligned}$$

Hence proved. $\square$

Replacing *simple* functions by *step* functions[14].

We now prove a similar result akin to Proposition 18.3.3.6, where we show that ***any measurable function can be approximated by a sequence of step functions, almost everywhere***. But before that, we prove **a basic fact about Lebesgue measurable sets with finite measure.**

Proposition 18.3.3.8. *For any λ-measurable set E of finite measure and a given $\epsilon > 0$, there exists a finite sequence of open intervals $\{I_n\}_{n=1}^N$ such that*

$$\lambda\left(E \Delta \left(\bigcup_{n=1}^{N} I_n\right)\right) < \epsilon.$$

Proof. Take any $\epsilon > 0$, then we have for any set $E \subseteq \mathbb{R}$, a sequence of open intervals $\{I_n\}$ such that $E \subseteq \bigcup_n I_n$ and $\lambda(\bigcup_n I_n) \leq \lambda(E) + \epsilon$ or $\lambda(\bigcup_n I_n \setminus E) \leq \epsilon < 2\epsilon$. Now since $\{I_n\}$ is a disjoint sequence, therefore, $\lambda(\bigcup_n I_n) = \sum_n \lambda(I_n)$ and due to the fact that $\lambda(E) < +\infty$, we get that $\sum_n \lambda(I_n) < +\infty$.

[14]Unfortunately, this section onwards, we would tentatively only work with the Lebesgue measure space $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ for pedagogical reasons of the instructor, so the abstract versions of the following results would need to wait.

Now, since $\lambda(E) < +\infty$, therefore the sum $\sum_n \lambda(I_n) < +\infty$, hence, $\exists N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} \lambda(I_n) < \epsilon$. With this N, we now see that:

$$\begin{aligned}
\lambda\left(E\Delta \bigcup_{n=1}^{N} I_n\right) &= \lambda\left(E \setminus \bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n=1}^{N} I_n \setminus E\right) \qquad \text{(both are disjoint.)} \\
&\leq \lambda\left(E \setminus \bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right) \\
&= \lambda\left(E \setminus \bigcup_{n=1}^{N} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right) \\
&= \lambda\left(E \cap \left(\bigcup_{n=1}^{N} I_n\right)^{c}\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right) \\
&\leq \lambda\left(\bigcup_{n=N+1}^{\infty} I_n\right) + \lambda\left(\bigcup_{n} I_n \setminus E\right) \quad \because E \cap \left(\bigcup_{n=1}^{N} I_n\right)^{c} \subseteq \bigcup_{n=N+1}^{\infty} I_n \\
&\leq \epsilon + \epsilon = 2\epsilon
\end{aligned}$$

Hence, we get that for any finite Lebesgue measurable set E, for all $\epsilon > 0$, $\exists$ a sequence of open intervals $\{I_n\}_{n=1}^{N}$ such that their symmetric difference is a set with measure $\leq \epsilon$. $\square$

Proposition 18.3.3.9. *Consider $(\mathbb{R}, \mathcal{M}_{\lambda^*})$ to be the Lebesgue measurable space and $A \in \mathcal{M}_{\lambda^*}$. Let $f : A \to [-\infty, +\infty]$ be a λ-measurable function. Then there exists a sequence of step functions $\{\phi_k\}$ such that*

$$\phi_k \longrightarrow f \text{ almost everywhere.}$$

Proof. We will prove first that for any characteristic function, there exists a sequence of step functions converging to it. Let $g = \chi_A$ be the characteristic function on A. Continuing from Proposition 18.3.3.8, we see that if we write the **step-function** ϕ as

$$\psi = \sum_{k=1}^{N} \chi_{I_k}$$

where $\{I_k\}$ is the set of open intervals such that $\lambda\left(A\Delta\left(\bigcup_{n=1}^{N} I_n\right)\right) < \epsilon$ for a given $\epsilon > 0$, from Proposition 18.3.3.8, then we get that the set $\{x \mid g(x) \neq \psi(x)\}$ has upper bound on it's measure given as follows:

$$\begin{aligned}
&\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) = \psi(x)\right\} \subseteq A \cap \bigcup_{k=1}^{N} I_k \quad \because g(x) = \psi(x) \text{ iff } x \in \cup_{k=1}^{N} I_k \text{ and } 0 \text{ for other } x \in A \\
&\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} \supseteq \left(A \cap \bigcup_{k=1}^{N} I_k\right)^{c} \supseteq A\Delta \bigcup_{k=1}^{N} I_k.
\end{aligned}$$

Similarly, it's easy to see that for any x such that $g(x) \neq \psi(x)$, we have $x \in A\Delta \bigcup_{k=1}^{N} I_k$ so that we get,

$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} \subseteq A\Delta \bigcup_{k=1}^{N} I_k.$$

Hence,

$$\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\} = A \Delta \bigcup_{k=1}^{N} I_k.$$

Therefore,

$$\lambda\left(\left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi(x)\right\}\right) < \epsilon$$

Therefore, for every $n \geq 1$, there exists a step-function ψ_n so that the set $E_n = \left\{x \in A \cup \left(\bigcup_{k=1}^{N} I_k\right) \mid g(x) \neq \psi_n(x)\right\}$ is such that

$$\lambda\left(E_n\right) < \frac{1}{2^n}.$$

Now, **define** the following two sets:

$$\begin{aligned} F_n &= \bigcup_{j=n+1}^{\infty} E_j \quad \text{(a decreasing sequence)} \\ F &= \bigcap_{k=1}^{\infty} F_k. \end{aligned}$$

For the set F_n, observe that

$$\begin{aligned} \lambda\left(F_n\right) = \lambda\left(\bigcup_{j=n+1}^{\infty}\right) &\leq \sum_{j=n+1}^{\infty} \lambda\left(E_j\right) \\ &< \sum_{j=n+1}^{\infty} \frac{1}{2^j} \\ &= \frac{1}{2^n} \end{aligned}$$

and for set F,

$$\begin{aligned} \lambda\left(F\right) = \lambda\left(\bigcap_{k=1}^{\infty} F_k\right) &= \underleftarrow{\lim}_{k\to\infty} \lambda\left(F_k\right) \quad \because \{F_k\} \text{ is measurable \& decreasing.} \\ &= 0. \end{aligned}$$

Note that $\{F_k\}$ is measurable because any E_i is itself measurable because of Proposition 18.3.0.4, Statement 5.

Now,

$$\psi_n(x) \longrightarrow g(x) \; \forall \; x \in F^c$$

because $F^c = F_1^c$ since $\{F_k\}$ is a decreasing sequence, therefore F^c is the set where $g(x)$ satisfies with the limit step function.

Finally, $\psi_n \not\to f \; \forall \; x \in F$, but since $\lambda\left(F\right) = 0$, hence

$$\psi_n \longrightarrow g \text{ almost everywhere}$$

Now, what we have proved so far is that for any characteristic function $g = \chi_A$ on a measurable set, there exists a sequence of step functions converging to it point-wise almost everywhere. Since from Proposition 18.3.3.6, there exists a sequence of simple functions converging to f, and since a simple function $h = \sum_{i=1}^{M} \alpha_i \chi_{E_i}$ is a finite combination of characteristics functions over measurable sets, therefore there exists a sequence of step functions converging to f almost everywhere. In particular if $\psi_n^i \longrightarrow \chi_{E_i}$ almost everywhere, then $\sum_{i=1}^{M} \alpha_i \psi_n^i \longrightarrow \sum_{i=1}^{M} \alpha_i \chi_{E_i} = h$. Now by Proposition 18.3.3.6, there exists the sequence $\{h_n\}$ of simple functions converging to f. Since

$$K_n = \left\{ \sum_{i=1}^{M_n} \alpha_i \psi_n^i \right\} \longrightarrow h_n \text{ almost everywhere,}$$

where note that $K_n = \sum_{i=1}^{M_n} \alpha_i \psi_n^i$ is a step function because ψ_n^i is a step function and there are finitely many (M_n) of them, and

$$\{h_n\} \longrightarrow f$$

therefore,

$$K_n \longrightarrow f \text{ almost everywhere.}$$

Hence proved. □

18.3.4 Egorov's theorem

We now discuss a very important result in the theory of measurable functions named after Dmitri Fyodorovich Egorov, who published this result in 1911, thus establishing **a condition required for uniform convergence of a point-wise convergent sequence of measurable functions.**

Theorem 18.3.4.1. *(**Egorov's theorem**) Let $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ be the Lebesgue measure space on $\mathbb{R}$. Suppose $\{f_k\}$ is a sequence of real-valued, Lebesgue measurable functions on $E \in \mathcal{M}_{\lambda^*}$ where $\lambda(E) < +\infty$. If*

$$f_k \longrightarrow f \text{ pointwise on } E,$$

[15] *Then for each $\epsilon > 0$, there exists a **closed** set $A_\epsilon \subset E$ such that*

1. *$\lambda(E \setminus A_\epsilon) < \epsilon$, and*
2. *$f_k \longrightarrow f$ **uniformly** on A_ϵ.*

Proof. We break down the proof in the following 3 parts.

Act 1. *A Basic Construction.*
For each pair of integers n, k, construct the following set:

$$E_k^n = \left\{ x \in E \ : \ |f_j(x) - f(x)| < \frac{1}{n} \ , \ \ \forall\, j > k \right\} .$$

Now, fix n, so that we have the following observations:

$$E_k^n \subseteq E_{k+1}^n \tag{18.8}$$

[15]From Proposition 18.3.0.9, the limit of a sequence of measurable functions is also measurable, hence there's no point in writing extraneously the requirement for f to be also measurable.

and since $f_k \longrightarrow f$ point-wise, therefore

$$\varinjlim_{k\to\infty} \bigcup_{i=1}^{k} E_i^n = E. \tag{18.9}$$

Hence

$$\lambda\left(E \setminus E_k^n\right) \longrightarrow 0 \text{ as } k \to \infty.$$

Note that the above result utilizes the fact that $\lambda(E) < +\infty$. Now by the above, we can say that $\exists\ k_n$ such that

$$\lambda\left(E \setminus E_{k_n}^n\right) < \frac{1}{2^n}$$

which, by definition of E_k^n implies that

$$|f_j(x) - f(x)| < \frac{1}{n} \text{ whenever } j > k_n \text{ and } \underline{x \in E_{k_n}^n}.$$

Act 2. *Constructing A_ϵ.*
Now choose $N \in \mathbb{N}$ such that

$$\sum_{n=N}^{\infty} \frac{1}{2^n} < \frac{\epsilon}{2}$$

and define

$$\tilde{A}_\epsilon = \bigcap_{n=N}^{\infty} E_{k_n}^n \tag{18.10}$$

We now observe that

$$\begin{aligned}
\lambda\left(E \setminus \tilde{A}_\epsilon\right) &= \lambda\left(E \cap \bigcup_{n=N}^{\infty} \left(E_{k_n}^n\right)^{\mathrm{c}}\right) \\
&= \lambda\left(\bigcup_{n=N}^{\infty} E \cap \left(E_{k_n}^n\right)^{\mathrm{c}}\right) \\
&\leq \sum_{n=N}^{\infty} \lambda\left(E \cap \left(E_{k_n}^n\right)^{\mathrm{c}}\right) \\
&= \sum_{n=N}^{\infty} \lambda\left(E \setminus E_{k_n}^n\right) \\
&< \sum_{n=N}^{\infty} \frac{1}{2^n} \\
&< \frac{\epsilon}{2}
\end{aligned}$$

Act 3. *Finalé.*
We now claim and prove the following:

$$\text{Claim}: f_k \longrightarrow f \textit{ uniformly on } \tilde{A}_\epsilon.$$

For this, let $\delta > 0$ and choose $n' \geq N$ such that $\frac{1}{n'} < \delta$. Then

$$\text{if } x \in \tilde{A}_\epsilon \implies x \in E^{n'}_{k_{n'}} \implies |f_j(x) - f(x)| < \frac{1}{n'} < \delta\ ,\ \ \forall\, j > k_{n'}. \tag{18.11}$$

Note that this is just the definition of uniform convergence.
Finally, note that E^n_k is a Lebesgue measurable set due to Proposition 18.3.0.4, Statement 5. Hence, $\tilde{A}_\epsilon$ is measurable. Now, by Proposition 18.2.9.2, Statement 2, there exists a closed set $A_\epsilon \subset \tilde{A}_\epsilon$ such that

$$\lambda\left(\tilde{A}_\epsilon \setminus A_\epsilon\right) < \frac{\epsilon}{2}.$$

Now,

$$\begin{aligned}
\epsilon &> \lambda\left(E \setminus \tilde{A}_\epsilon\right) + \lambda\left(\tilde{A}_\epsilon \setminus A_\epsilon\right) \\
&\geq \lambda\left(E \setminus \tilde{A}_\epsilon \bigcup \tilde{A}_\epsilon \setminus A_\epsilon\right) \\
&= \lambda\left(E \setminus A_\epsilon\right).
\end{aligned}$$

Now, by (18.11), we see that $f_k \longrightarrow f$ uniformly for all $x \in A_\epsilon \subset \tilde{A}_\epsilon$ such that $\lambda\left(E \setminus A_\epsilon\right) < \epsilon$ and A_ϵ is closed. Proof is now complete. $\square$

18.3.5 Lusin's theorem

The following is the final important result on the basic theory of measurable functions, attributed to Nikolai Nikolaevich Luzin who penned this theorem around 1912.

Theorem 18.3.5.1. *(**Lusin's theorem**) Consider the Lebesgue measure space $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$. Suppose f is a real-valued, Lebesgue measurable function defined over a Lebesgue measurable set E with finite measure. Then for all $\epsilon > 0$, there exists a **closed** set $F_\epsilon \subset E$ with*

1. *$\lambda\left(E \setminus F_\epsilon\right) < \epsilon$, and*
2. *The restriction $f|_{F_\epsilon}$ of f over F_ϵ is **continuous**.*

Proof. From the Proposition 18.3.3.9, we have a sequence $\{f_n\}$ of step functions such that

$$f_n \longrightarrow f \text{ almost everywhere.}$$

Now, consider, for example the characteristic function over an interval $\chi_{[a,b]}$. Then, we can define a function $\phi(x)$ for any $\delta > 0$ as follows:

$$\phi(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{\delta/2} & a \leq x \leq a + \frac{\delta}{2} \\ 1 & a + \frac{\delta}{2} \leq x \leq b - \frac{\delta}{2} \\ \frac{b-x}{\delta/2} & b - \frac{\delta}{2} \leq x \leq b \\ 0 & x > b \end{cases}$$

which then satisfies

$$\{x \in \mathbb{R} \mid \phi(x) \neq \chi_{[a,b]}\} = \left(a, a + \frac{\delta}{2}\right) \bigcup \left(b - \frac{\delta}{2}, b\right)$$

which then implies that,

$$\lambda\left(\{x \in \mathbb{R} \mid \phi(x) \neq \chi_{[a,b]}\}\right) = \lambda\left(\left(a, a+\frac{\delta}{2}\right) \bigcup \left(b-\frac{\delta}{2}, b\right)\right) = \delta.$$

Note that $\phi(x)$ is also continuous over all $\mathbb{R}$. Hence, for any step function (finite sum of $\chi_{[a,b]}$-type functions) and $\delta > 0$, one can construct a continuous function which does not satisfies with the step function on a set with measure $< \delta$.

Hence, for step-functions $\{f_n\}$, corresponding to each f_n, $\exists$ a continuous function ϕ_n and a set E_n such that

$$E_n = \{x \mid \phi_n(x) \neq f_n(x)\} \text{ and } \lambda(E_n) < \frac{1}{2^n}.$$

Now, for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that

$$\sum_{n \geq N} \frac{1}{2^n} < \frac{\epsilon}{3}.$$

With the above fact, construct the set F' as follows:

$$F' = \left(A_{\frac{\epsilon}{3}} \setminus \bigcup_{n \geq N} E_n\right)$$

where $A_{\frac{\epsilon}{3}}$ is the closed subset $A_{\frac{\epsilon}{3}} \subset E$ such that **1.** $\lambda\left(E \setminus A_{\frac{\epsilon}{3}}\right) < \frac{\epsilon}{3}$ and **2.** $f_n \longrightarrow f$ **uniformly** on $A_{\frac{\epsilon}{3}}$. This is guaranteed by Theorem 18.3.4.1 (Egorov's Theorem).

Note that $f_n|_{F'}$ is **continuous** $\forall\ n \geq N$ because for any $x \in F' \implies \phi_n(x) = f_n(x)\ \forall\ n \geq N$ and since ϕ_n are already continuous $\forall n \in \mathbb{N}$.

Furthermore, since $F' \subset E$ and $f_n \longrightarrow f$ uniformly, then the restriction $f_n|_{F'}$ is continuous and converges uniformly to $f|_{F'}$, which **due to uniform convergence, is also continuous!**

Now, note that E_n's are measurable sets due to Proposition 18.3.0.4, Statement 5. Similarly, since $A_{\frac{\epsilon}{3}}$ is closed, therefore it is also measurable. Hence, F' is measurable.

Now by Proposition 18.2.9.2, there exists a closed set $F_\epsilon \subset F'$ such that $\lambda(F' \setminus F_\epsilon) < \frac{\epsilon}{3}$. Note that because $F_\epsilon \subset F'$ and $f|_{F'}$ is continuous, therefore the restriction $f|_{F_\epsilon}$ is also continuous.

Finally, combining

1. $\sum_{n \geq N} \frac{1}{2^n} < \frac{\epsilon}{3}$,
2. $\lambda\left(E \setminus A_{\frac{\epsilon}{3}}\right) < \frac{\epsilon}{3}$,
3. $\lambda(F' \setminus F_\epsilon) < \frac{\epsilon}{3}$

it can be easily seen that

$$\left(E \setminus A_{\frac{\epsilon}{3}}\right) \bigcup (F' \setminus F_\epsilon) = E \setminus F_\epsilon$$

Hence,

$$\begin{aligned}\lambda(E \setminus F_\epsilon) &= \lambda\left(\left(E \setminus A_{\frac{\epsilon}{3}}\right) \bigcup (F' \setminus F_\epsilon)\right) \\ &\leq \lambda\left(\left(E \setminus A_{\frac{\epsilon}{3}}\right)\right) + \lambda\left((F' \setminus F_\epsilon)\right) \\ &< \frac{2\epsilon}{3} \\ &< \epsilon.\end{aligned}$$

which completes the proof. □

18.3.6 Applications-I : measure spaces and measurable functions

We now present applications of the above theory. This is, in particular, to showcase the true strength of abstract analysis. This can also be used to strengthen one's intuition about the topic.

σ-algebras and measure spaces

Lemma 18.3.6.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Prove that μ is σ-finite if and only if there exists a countable disjoint family of measurable sets $\{A_n\}$ such that $X = \coprod_n A_n$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.*

Proof. Note that R $\implies$ L is immediate from definition. Let μ be σ-finite. Then there exists $\{B_n\} \subseteq \mathcal{A}$ such that $\mu(B_n) < \infty$ and $\bigcup_n B_n = X$. Define $A_1 = B_1$ and $A_n = B_n \setminus B_1 \cup \cdots \cup B_{n-1}$. As $\mathcal{A}$ is a σ-algebra, so $\{A_n\} \subseteq \mathcal{A}$. Moreover, $A_n \cap A_m = \emptyset$ for all $n \neq m$ because if $m > n$[16] and $x \in A_m \cap A_n$, then $x \in B_m \setminus B_1 \cup \cdots \cup B_n \cup \ldots B_{m-1}$ and $x \in B_n \setminus B_1 \cup \ldots B_{n-1}$, a contradiction. As $A_n \subseteq B_n$, therefore $\mu(A_n) \leq \mu(B_n) < \infty$. To complete the proof, we need only show that $\bigcup_n A_n = \bigcup_n B_n$.

Pick any $x \in \bigcup_n A_n$. Then $x \in B_n \setminus B_1 \cup \cdots \cup B_{n-1}$ for some $n \in \mathbb{N}$. Thus, $x \in B_n$ and hence $x \in \bigcup_n B_n$. Conversely, pick $x \in \bigcup_n B_n$. Then $x \in B_n$ for some $n \in \mathbb{N}$. Now, either $x \in B_n \setminus B_1 \cup \cdots \cup B_{n-1}$ or $x \in B_n \cap (B_1 \cup \cdots \cup B_{n-1})$. If the former is true, then $x \in A_n$ and we are done. If the latter is true, then we may assume $x \in B_{n-1} \cap B_n$. Now again either $x \in B_{n-1} \setminus B_1 \cup \cdots \cup B_{n-2}$ or $x \in B_n \cap B_{n-1} \cap (B_1 \cup \cdots \cup B_{n-2})$. Repeating this process inductively, we will end up in either of the following cases:

1. $x \in A_k$ for some $1 \leq k \leq n$,
2. $x \in B_1 \cap \cdots \cap B_n$.

As $B_1 = A_1$ by construction, therefore in either case we are done. □

Lemma 18.3.6.2. *Given $\mathcal{S} \subseteq \mathcal{P}(X)$, denote by $\mathcal{A}(\mathcal{S})$ the σ-algebra generated by $\mathcal{S}$. Then,*

$$\mathcal{A}(\mathcal{S}) = \mathcal{A}(\mathcal{A}(\mathcal{S})).$$

Proof. Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ be an arbitrary collection of subsets of X. If X is empty then the statement is vacuously true, so let X be non-empty. Since the σ-algebra generated by $\mathcal{A}(\mathcal{S})$ is the intersection of all σ-algebras containing $\mathcal{A}(\mathcal{S})$, therefore we have that $\mathcal{A}(\mathcal{A}(\mathcal{S})) = \bigcap_{\mathcal{C} \supseteq \mathcal{A}(\mathcal{S})} \mathcal{C}$. Since $\mathcal{A}(S)$ is a σ-algebra containing $\mathcal{A}(\mathcal{S})$, therefore $\mathcal{A}(\mathcal{A}(\mathcal{S})) \subseteq \mathcal{A}(\mathcal{S})$. Since $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{C}$ for all σ-algebras $\mathcal{C}$ containing $\mathcal{A}(\mathcal{S})$, therefore $\mathcal{A}(\mathcal{A}(\mathcal{S})) \supseteq \mathcal{A}(\mathcal{S})$. □

[16] which we may assume wlog.

Lemma 18.3.6.3. *Let $\mathcal{A}(\mathcal{S})$ be the σ-algebra generated by a set $\mathcal{S} \subseteq \mathcal{P}(X)$. Then, $\mathcal{A}(\mathcal{S})$ is the union of the σ-algebras generated by Y as Y ranges over all countable subsets of $\mathcal{S}$.*

Proof. Let X be a non-empty set and $\mathcal{S} \subseteq \mathcal{P}(X)$. We wish to show that

$$\mathcal{A}(\mathcal{S}) = \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}).$$

Let $\mathcal{Y} \subseteq \mathcal{S}$ be a countable subcollection. Then, $\mathcal{A}(\mathcal{Y}) \subseteq \mathcal{A}(\mathcal{S})$. Consequently, $\bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}) \subseteq \mathcal{A}(\mathcal{S})$. Conversely, we need to show that

$$\mathcal{A}(\mathcal{S}) \subseteq \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y}).$$

We claim that $\bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y})$ is a σ-algebra containing $\mathcal{S}$. This would complete the proof as $\mathcal{A}(\mathcal{S})$ is the smallest σ-algebra containing $\mathcal{S}$.

Denote $\mathcal{Z} = \bigcup_{\mathcal{Y} \subseteq \mathcal{S},\ \text{countable}} \mathcal{A}(\mathcal{Y})$. As $\mathcal{A}(\mathcal{Y})$s are σ-algebras, therefore $\mathcal{Z}$ contains X and $\emptyset$. Let $A \in \mathcal{Z}$. Then $A \in \mathcal{A}(\mathcal{Y})$ for some $\mathcal{Y} \subseteq \mathcal{S}$ countable. Consequently, $A^c \in \mathcal{A}(\mathcal{Y})$ and thus $A^c \in \mathcal{Z}$. Let $\{A_n\} \subseteq \mathcal{Z}$ be a countable collection of sets. Then $A_n \in \mathcal{A}(\mathcal{Y}_n)$ for all $n \in \mathbb{N}$. Further, we have that $\mathcal{Y}_k \subseteq \mathcal{A}\left(\bigcup_n \mathcal{Y}_n\right)$ for all $k \in \mathbb{N}$ as $\mathcal{Y}_k \subseteq \bigcup_n \mathcal{Y}_n$. As $\mathcal{Y}_k$ are countable and countable union of countable sets is countable, therefore $\bigcup_n \mathcal{Y}_n$ is countable. Thus, we have

$$A_k \in \mathcal{A}(\mathcal{Y}_k) \subseteq \mathcal{A}\left(\bigcup_n \mathcal{Y}_n\right) \subseteq \mathcal{Z}\ \forall k \in \mathbb{N}.$$

Thus from above, we obtain that

$$\bigcup_k A_k \in \left(\bigcup_n \mathcal{Y}_n\right) \subseteq \mathcal{Z}.$$

Hence, $\mathcal{Z}$ is a σ-algebra. To complete the proof, we need only show that $\mathcal{Z}$ contains $\mathcal{S}$.

Let $A \in \mathcal{S}$. Then since $\{A\}$ is a countable subset of $\mathcal{S}$, therefore $\mathcal{A}(\{A\})$ is contained in $\mathcal{Z}$ and thus $A \in \mathcal{Z}$. $\square$

Lemma 18.3.6.4. *The σ-algebra generated by*
1. $\mathcal{S} = \{(a, b] \mid a < b \in \mathbb{Q}\}$,
2. $\mathcal{S} = \{(a, n] \mid a \in \mathbb{Q}, n \in \mathbb{Z}\}$,

is the Borel σ-algebra on $\mathbb{R}$.

Proof. 1. Let $\mathcal{S} = \{(a, b] \mid a, b \in \mathbb{Q}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the Borel σ-algebra of $\mathbb{R}$. Since $(a, b]$ for $a, b \in \mathbb{Q}$ is contained in $\mathcal{B}$ as $(a, b] = (a, b) \cup \bigcap_{n \in \mathbb{N}}(b - 1/n, b + 1/n)$, therefore $\mathcal{S} \subseteq \mathcal{B}$. Consequently, $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$ as $\mathcal{B}$ is the smallest σ-algebra containing open intervals.

Since we also know that $\mathcal{B}$ is generated by the collection of all closed intervals $[a, b]$ in $\mathbb{R}$, therefore to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$, it would suffice to show $[a, b] \in \mathcal{A}(\mathcal{S})$ where $a < b$ in $\mathbb{R}$. Pick $a < b$ in $\mathbb{R}$. By density of $\mathbb{Q}$, we may pick $\{a_n\}$ to be an increasing sequence such that $a_n \in \mathbb{Q}$, $a_n < a$ and $\lim_{n\to\infty} a_n = a$. Similarly, we may pick a decreasing sequence $\{b_n\}$ such that $b_n \in \mathbb{Q}$, $b_n > b$ and $\lim_{n\to\infty} b_n = b$. Consequently, we claim that

$$[a, b] = \bigcap_n (a_n, b_n]$$

where $(a_n, b_n] \in \mathcal{S}$. Indeed, $(\subseteq)$ is clear. For $(\supseteq)$, take $x \in \bigcap_n (a_n, b_n]$. Hence $a_n < x \leq b_n$. Taking $n \to \infty$, we get $a \leq x \leq b$ as desired. Thus, $[a, b] \in \mathcal{A}(\mathcal{S})$.

2. Let $\mathcal{S} = \{(a, n] \mid a \in \mathbb{Q},\ n \in \mathbb{N}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the Borel σ-algebra of $\mathbb{R}$. Since $(a, n]$ for $a \in \mathbb{Q}$ and $n \in \mathbb{N}$ is contained in $\mathcal{B}$ as $(a, n] = (a, n) \cup \bigcap_{k \in \mathbb{N}} (n - 1/k, n + 1/k)$, therefore $\mathcal{S} \subseteq \mathcal{B}$. Consequently, $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$.

Since we also know that $\mathcal{B}$ is generated by the collection of all open intervals of the form (a, ∞), $a \in \mathbb{R}$, therefore to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$, it would suffice to show $(a, \infty) \in \mathcal{A}(\mathcal{S})$ for all $a \in \mathbb{R}$. Pick (a, ∞) for some $a \in \mathbb{R}$. By density of $\mathbb{Q}$, there exists a decreasing sequence $\{a_n\}$ in $\mathbb{R}$ such that $a_n \in \mathbb{Q}$, $a_n > a$ and $\lim_{n \to \infty} a_n = a$. Consequently, we claim that

$$(a, \infty) = \bigcup_n (a_n, n]$$

where $(a_n, n] \in \mathcal{S}$. Indeed, for $(\subseteq)$, take $x \in (a, \infty)$. We therefore have $a < x < \infty$. As $\lim_{n \to \infty} a_n = a$ and $a_n > a$ for all $n \in \mathbb{N}$, therefore there exists $N \in \mathbb{N}$ such that $a < a_n \leq a_N < x$ for all $n \geq N$. Consequently, for some large $n \in \mathbb{N}$ greater than N such that $x \leq n$, we obtain $a_n < x \leq n$ and hence $x \in (a_n, n]$. For $(\supseteq)$, take $x \in \bigcup_n (a_n, n]$ and thus we get $a < a_n < x \leq n < \infty$. Thus, $(a, \infty) \in \mathcal{A}(\mathcal{S})$. □

Lemma 18.3.6.5. *The Borel σ-algebra on $\mathbb{R}^2$ is generated by*

$$\{(I \times \mathbb{R}) \cup (\mathbb{R} \times J) \mid I, J \subseteq \mathbb{R},\ \textit{open intervals}\}.$$

Proof. Let $\mathcal{S} = \{(I \times \mathbb{R}) \cup (\mathbb{R} \times J) \mid I, J \subseteq \mathbb{R} \text{ is open}\}$. We wish to show that $\mathcal{A}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the σ-algebra of $\mathbb{R}^2$.

As $\mathcal{S}$ is a collection of open sets of $\mathbb{R}^2$ and $\mathcal{B}$ is generated by all open sets of $\mathbb{R}^2$, therefore $\mathcal{S} \subseteq \mathcal{B}$ and thus $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{B}$.

We now wish to show that $\mathcal{B} \subseteq \mathcal{A}(\mathcal{S})$. It would suffice to show that any open set $U \subseteq \mathbb{R}^2$ is in $\mathcal{A}(\mathcal{S})$. Note that $\mathcal{A}(\mathcal{S})$ consists of all open rectanlges $I \times J = (I \times \mathbb{R}) \cap (\mathbb{R} \times J)$. Thus, it would suffice to show that U can be written as countable union of open rectangles. Recall that open rectangles forms a basis for the usual topology on $\mathbb{R}^2$. Consider the collection of all open rectangles K inside U whose vertices have both rational coordinates. We claim that the union of such open rectangles is equal to U. Indeed, their union is inside U and for any $x \in U$, there exists an open ball $x \in B \subseteq U$, so there exists an open rectangle K inside B which contains x and has vertices which have both rational coordinates. Thus U is equal to the union of all such rectangles. Since there are only countably many such open rectangles as they are parameterized by choice of 4 points in $\mathbb{Q}^2 \cap U$ which is atmost countably many, therefore we have obtained a countable cover of U by open rectangles. This completes the proof. □

Lemma 18.3.6.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $A, B \in \mathcal{A}$. Then,*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Proof. Observe that we can write

$$A \cup B = (A \setminus (A \cap B)) \cup B$$

where the right side is a disjoint union. Consequently, we have

$$\mu(A \cup B) = \mu(A \setminus A \cap B) + \mu(B). \tag{6.1}$$

We now have two cases. If $\mu(A \cap B) = \infty$, then since $\mu(A \cap B) \leq \mu(A), \mu(B)$ and $\mu(A) \leq \mu(A \cup B)$, therefore we get $\mu(A \cup B) = \mu(A \cap B) = \mu(A) = \mu(B) = \infty$, so that the statement to be proven is a tautology. Else if $\mu(A \cap B) < \infty$, then we can write

$$\mu(A \setminus A \cap B) = \mu(A) - \mu(A \cap B).$$

Consequently, by Eq. (6.1) and the fact that $\mu(A \cap B) < \infty$, we have

$$\begin{aligned} \mu(A \cup B) &= \mu(A) - \mu(A \cap B) + \mu(B) \\ \mu(A \cup B) + \mu(A \cap B) &= \mu(A) + \mu(B). \end{aligned}$$

This completes the proof. □

Lemma 18.3.6.7. *Let $x \in \mathbb{R}$ and let B be a Borel subset of $\mathbb{R}$. Then, $x + B$ and xB are Borel subsets of $\mathbb{R}$ (that is, Borel subsets of $\mathbb{R}$ are translation and dilation invariant).*

Proof. 1. Let $x \in \mathbb{R}$ and $\mathcal{B}$ be the Borel σ-algebra of $\mathbb{R}$. We wish to show that for all $B \in \mathcal{B}$, the translate $x + B \in \mathcal{B}$. Consequently, we wish to show

$$x + \mathcal{B} \subseteq \mathcal{B}$$

where $x + \mathcal{B} = \{x + B \mid B \in \mathcal{B}\}$. We use the following standard technique to show this.

Consider the following collection

$$\mathcal{C} = \{B \in \mathcal{B} \mid x + B \in \mathcal{B}\}.$$

Our goal is to show that $\mathcal{C} = \mathcal{B}$. Note that $\mathcal{C} \subseteq \mathcal{B}$. Conversely, we wish to show that $\mathcal{B} \subseteq \mathcal{C}$. This would follow immediately if we show that $\mathcal{C}$ is a σ-algebra containing all open intervals, as $\mathcal{B}$ is the σ-algebra generated by all open intervals.

We now establish that $\mathcal{C}$ is a σ-algebra. Since $x + \mathbb{R} = \mathbb{R}$ and $x + \emptyset = \emptyset$, therefore $\mathbb{R}, \emptyset \in \mathcal{C}$. Let $A \in \mathcal{C}$. We wish to show that $A^c \in \mathcal{C}$. Since $x + A \in \mathcal{B}$, therefore $(x + A)^c \in \mathcal{B}$. Thus it suffices to show that $(x + A)^c = x + A^c$. Indeed, we have the following equalities

$$\begin{aligned} (x + A)^c &= \{y \in \mathbb{R} \mid y \notin x + A\} \\ &= \{y \in \mathbb{R} \mid y - x \notin A\} \\ &= \{y \in \mathbb{R} \mid y - x \in A^c\} \\ &= \{y \in \mathbb{R} \mid y \in x + A^c\} \\ &= x + A^c. \end{aligned}$$

Let $\{A_n\} \subseteq \mathcal{C}$. We wish to show that $\bigcup_n A_n \in \mathcal{C}$. We have that for each $n \in \mathbb{N}$, $x + A_n \in \mathcal{B}$. It would thus suffice to show that

$$x + \bigcup_n A_n = \bigcup_n (x + A_n).$$

Indeed, take $x+a \in x+\bigcup_n A_n$. Hence $a \in A_n$ for some $n \in \mathbb{N}$. Consequently, $x+a \in x+A_n$. Thus $x+a \in \bigcup_n (x+A_n)$. Conversely, let $z \in \bigcup_n (x+A_n)$. Then $z = x+y_n$ for $y_n \in A_n$. Consequently, $z \in x+\bigcup_n A_n$. This show that $\mathcal{C}$ is a σ-algebra.

To complete the proof, we now need only show that $\mathcal{C}$ has all open intervals. This is immediate, as we show now. Take any $(a,b) \subseteq \mathbb{R}$. Since $x+(a,b) = (x+a, x+b) \in \mathcal{B}$, therefore $(a,b) \in \mathcal{C}$.

2. Let $x \in \mathbb{R}$ and $\mathcal{B}$ be the Borel σ-algebra of $\mathbb{R}$. We wish to show that for all $B \in \mathcal{B}$, the dilate $x \cdot B \in \mathcal{B}$. Note that $x \cdot B = \{xb \mid b \in B\}$. Consequently, we wish to show

$$x \cdot \mathcal{B} \subseteq \mathcal{B}$$

where $x \cdot \mathcal{B} = \{x \cdot B \mid B \in \mathcal{B}\}$. If $x = 0$, then $x \cdot \mathcal{B} = \{0\}$ and that is trivially inside $\mathcal{B}$ as $\{0\} = \bigcap_n (-1/n, 1/n)$. Thus we now assume that $x \neq 0$. We use the following standard technique to show the above inclusion.

Consider the following collection

$$\mathcal{C} = \{B \in \mathcal{B} \mid x \cdot B \in \mathcal{B}\}.$$

Our goal is to show that $\mathcal{C} = \mathcal{B}$. Note that $\mathcal{C} \subseteq \mathcal{B}$. Conversely, we wish to show that $\mathcal{B} \subseteq \mathcal{C}$. This would follow immediately if we show that $\mathcal{C}$ is a σ-algebra containing all open intervals, as $\mathcal{B}$ is the σ-algebra generated by all open intervals.

We now establish that $\mathcal{C}$ is a σ-algebra. Observe that $x \cdot \mathbb{R} = \mathbb{R}$. Indeed, as $x \cdot \mathbb{R} \subseteq \mathbb{R}$ is clear, we can also write any $a \in \mathbb{R}$ as $x \cdot x^{-1}a$. We also have $x \cdot \emptyset = \emptyset$. Therefore $\mathbb{R}, \emptyset \in \mathcal{C}$. Let $A \in \mathcal{C}$. We wish to show that $A^c \in \mathcal{C}$. Since $x \cdot A \in \mathcal{B}$, therefore $(x \cdot A)^c \in \mathcal{B}$. Thus it suffices to show that $(x \cdot A)^c = x \cdot A^c$. Indeed, we have the following equalities

$$\begin{aligned}
(x \cdot A)^c &= \{y \in \mathbb{R} \mid y \notin x \cdot A\} \\
&= \{y \in \mathbb{R} \mid x^{-1}y \notin A\} \\
&= \{y \in \mathbb{R} \mid x^{-1}y \in A^c\} \\
&= \{y \in \mathbb{R} \mid y \in x \cdot A^c\} \\
&= x \cdot A^c.
\end{aligned}$$

Let $\{A_n\} \subseteq \mathcal{C}$. We wish to show that $\bigcup_n A_n \in \mathcal{C}$. We have that for each $n \in \mathbb{N}$, $x \cdot A_n \in \mathcal{B}$. It would thus suffice to show that

$$x \cdot \bigcup_n A_n = \bigcup_n (x \cdot A_n).$$

Indeed, take $x \cdot a \in x \cdot \bigcup_n A_n$. Hence $a \in A_n$ for some $n \in \mathbb{N}$. Consequently, $x \cdot a \in x \cdot A_n$. Thus $x \cdot a \in \bigcup_n (x \cdot A_n)$. Conversely, let $z \in \bigcup_n (x \cdot A_n)$. Then $z = x \cdot y_n$ for $y_n \in A_n$. Consequently, $z \in x \cdot \bigcup_n A_n$. This show that $\mathcal{C}$ is a σ-algebra.

To complete the proof, we now need only show that $\mathcal{C}$ has all open intervals. This is immediate, as we show now. Take any $(a,b) \subseteq \mathbb{R}$. If $x > 0$, then we have $x \cdot (a,b) = (x \cdot a, x \cdot b) \in \mathcal{B}$, therefore $(a,b) \in \mathcal{C}$. If $x < 0$, then we have $x \cdot (a,b) = (x \cdot b, x \cdot a) \in \mathcal{B}$, therefore $(a,b) \in \mathcal{C}$. $\square$

Lemma 18.3.6.8. *Let $(X, \mathcal{A})$ be a measurable space and let $\{\mu_i\}_{i=1}^n$ be a finite collection of measures on $(X, \mathcal{A})$. If $r_1, \ldots, r_n \in \mathbb{R}_{\geq 0}$, then $\sum_i r_i \mu_i$ is a measure on $(X, \mathcal{A})$ (that is, positive linear combination of measures is a measure).*

Proof. Let $(X, \mathcal{A})$ be a measurable space and $\{\mu_i\}_{i=1}^n$ be a collection of measures on it. Let $\{r_i\}_{i=1}^n \subseteq \mathbb{R}_{\geq 0}$. We wish to show that $\mu = \sum_{i=1}^n r_i\mu_i$ is a measure on $(X, \mathcal{A})$. First we may assume that each $r_i > 0$ as if any $r_j = 0$, then $\mu(A) = \sum_{i=1}^n r_i\mu_i(A) = \sum_{i\neq j} r_i\mu_i(A) + r_j\mu_j(A)$, therefore if $\mu_j(A) < \infty$, then $r_j\mu_j(A) = 0$ and if $\mu_j(A) = \infty$, then since $0 \cdot \infty = 0$, therefore still $r_j\mu_j(A) = 0$. Further, if all $r_i = 0$, then $\mu = 0$, which is the trivial measure. Consequently, we assume that $r_i > 0$ for all $i = 1, \dots, n$.

We now show that μ is a measure on $(X, \mathcal{A})$. We have $\mu(\emptyset) = \sum_{i=1}^n r_i\mu_i(\emptyset) = \sum_{i=1}^n r_i \cdot 0 = 0$. Let $\{A_n\} \subseteq \mathcal{A}$ be a collection of disjoint measurable sets. We wish to show that

$$\mu\left(\coprod_k A_k\right) = \sum_k \mu(A_k).$$

We have

$$\begin{aligned}\mu\left(\coprod_k A_k\right) &= \sum_{i=1}^n r_i\mu_i\left(\coprod_k A_k\right)\\ &= \sum_{i=1}^n r_i \sum_{k=1}^\infty \mu_i(A_k).\end{aligned}$$

We now claim that

$$\sum_{i=1}^n r_i \sum_{k=1}^\infty \mu_i(A_k) = \sum_{k=1}^\infty \sum_{i=1}^n r_i\mu_i(A_k) \tag{8.1}$$

and showing this will complete the proof as

$$\sum_{k=1}^\infty \sum_{i=1}^n r_i\mu_i(A_k) = \sum_{k=1}^\infty \mu(A_k).$$

We have few cases for establishing Eq. (8.1).

1. *If for all $i = 1, \dots, n$, the series $\sum_{k=1}^\infty \mu_i(A_k)$ is finite.* Then, $\sum_{i=1}^n r_i \sum_{k=1}^\infty \mu_i(A_k) = \sum_{i=1}^n \sum_{k=1}^\infty r_i\mu_i(A_k)$. Now, if $\sum_n x_n, \sum_n y_n$ are two convergent positive series, then their linear combination $c\sum_n x_n + d\sum_n y_n$ is equal to $\sum_n cx_n + dy_n$, where $c, d \in \mathbb{R}_{\geq 0}$. Indeed, this follows at once from the equality $c\lim_{n\to\infty} \sum_{k=1}^n x_k + d\lim_{n\to\infty} \sum_{k=1}^n y_k = \lim_{n\to\infty} \sum_{k=1}^n cx_k + dy_k$, which follows from the fact that both the limit exists and $c, d \in \mathbb{R}$. Consequently, we have

   $$\sum_{i=1}^n \sum_{k=1}^\infty r_i\mu_i(A_k) = \sum_{k=1}^\infty \sum_{i=1}^n r_i\mu_i(A_k),$$

   which is what we needed.
2. *If there exists $i_0 = 1, \dots, n$ such that the series $\sum_{k=1}^\infty \mu_{i_0}(A_k) = \infty$.* In this case, in the Eq. (8.1), the left side is ∞. The right side is also infinity as shown below:

   $$\begin{aligned}\sum_{k=1}^\infty \sum_{i=1}^n r_i\mu_i(A_k) &\geq \sum_{k=1}^\infty r_{i_0}\mu_{i_0}(A_k)\\ &= \infty\end{aligned}$$

   where the first inequality follows from $r_i > 0$ for all $i = 1, \dots, n$ and measure being positive by definition. Consequently, Eq. (8.1) follows in this case as well.

This completes the proof. $\square$

Lemma 18.3.6.9. *For any set X and a subset $S \subseteq X$, the collection*

$$\mathcal{A}_S = \{A \subseteq X \mid A \subseteq S \text{ or } A^c \subseteq S\}$$

is a σ-algebra on X.

Proof. Let X be a non-empty set, $S \subseteq X$ and define

$$\mathcal{A}_S := \{A \subseteq X \mid A \subseteq S \text{ or } A^c \subseteq S\}.$$

We claim that this forms a σ-algebra on X. As $X^c = \emptyset \subseteq S$, therefore $X \in \mathcal{A}_S$ and $\emptyset \in \mathcal{A}_S$. Let $A \in \mathcal{A}_S$. If $A \subseteq S$, then A^c is such that $(A^c)^c = A \subseteq S$, so $A^c \in \mathcal{A}_S$. If $A^c \subseteq S$, then A^c is such that $A^c \subseteq S$, so $A^c \in \mathcal{A}_S$. So in both cases $\mathcal{A}_S$ is closed uncer complements.

Let $\{A_n\} \subseteq \mathcal{A}_S$ be a collection of subsets. We wish to show that $\bigcup_n A_n \in \mathcal{A}_S$. We have three cases.

C1. $A_n \subseteq S$ *for all* $n \in \mathbb{N}$. Then $\bigcup_n A_n \subseteq S$ and thus $\bigcup_n A_n \in \mathcal{A}_S$.

C2. $\exists A_m$ *such that* $A_m \nsubseteq S$. Then $A_m^c \subseteq S$. We then observe by De-Morgan's law that

$$\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c \subseteq A_m^c \subseteq S.$$

Consequently, $\bigcup_n A_n \in \mathcal{A}_S$.

C3. $A_n \nsubseteq S$ *for all* $n \in \mathbb{N}$. Then $A_n^c \subseteq S$ for all $n \in \mathbb{N}$. We again observe by De-Morgan's law that

$$\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c \subseteq A_m^c \subseteq S \ \forall m \in \mathbb{N}.$$

Consequently, $\bigcup_n A_n \in \mathcal{A}_S$.

In all three cases, $\bigcup_n A_n \in \mathcal{A}_S$. Hence $\mathcal{A}_S$ is a σ-algebra. $\square$

Lemma 18.3.6.10. *Let $(X, \mathcal{A}, \mu)$ be a semifinite measure space, and let $\mu(A) = \infty$ for some $A \in \mathcal{A}$. If $M > 0$, then there exists $B \subseteq A$ such that $M < \mu(B) < \infty$.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a semi-finite measure space and $A \in \mathcal{A}$ such that $\mu(A) = \infty$. We wish to show that for all $M > 0$, there exists a subset $B \subseteq A$ such that $B \in \mathcal{A}$ and $M < \mu(B) < \infty$.

We wish to show that there exists measurable subsets of A of arbitrarily large size. Therefore, consider the collection

$$S = \{\mu(B) \mid B \subseteq A, B \in \mathcal{A}, \mu(B) < \infty\}.$$

Denote $l = \sup S$. We wish to show that $l = \infty$. Pick a sequence $\{B_n\} \subseteq S$ such that $\lim_{n\to\infty}\mu(B_n) = l$. We first claim that

$$\mu\left(\bigcup_n B_n\right) = l \tag{10.1}$$

Clearly, $\bigcup_n B_n \in \mathcal{A}$. Observe that since

$$\mu(B_k) \le \mu\left(\bigcup_n B_n\right)$$

for all $k \in \mathbb{N}$, therefore taking $k \to \infty$, we easily obtain

$$l \le \mu\left(\bigcup_n B_n\right).$$

Conversely, we wish to show that

$$\mu\left(\bigcup_n B_n\right) \le l.$$

Let $D_1 = B_1$, $D_2 = B_1 \cup B_2$ and in general $D_n = B_1 \cup \dots \cup B_n$. Then we observe that $\{D_n\} \subseteq \mathcal{A}$ forms an increasing sequence of sets with $\bigcup_n D_n = \bigcup_n B_n$. Consequently,

$$\mu\left(\bigcup_n B_n\right) = \mu\left(\bigcup_k D_k\right) = \lim_{k\to\infty}\mu(D_k).$$

Since $D_k \subseteq A$ is such that $\mu(D_k) \le \sum_{i=1}^{k} \mu(B_i) < \infty$ (by subadditivity), therefore $\mu(D_k) \in S$ for all $k \in \mathbb{N}$. Consequently,

$$\lim_{k\to\infty}\mu(D_k) \le l.$$

Therefore we obtain $\mu\left(\bigcup_n B_n\right) \le l$. Hence this completes the proof of Eq. (10.1).

Since we wish to show that $l = \infty$, so assume to the contrary that $l < \infty$. It follows from Eq. (10.1) that $\mu\left(\bigcup_n B_n\right) < \infty$ and therefore $\bigcup_n B_n \in S$. Let $C = \bigcup_n B_n$. Then consider $A_1 = A \setminus C$. Since $\mu(A_1) = \mu(A) - \mu(C)$ as $\mu(C) < \infty$, therefore we have $\mu(A_1) = \infty - \mu(C) = \infty$. It follows from semifiniteness that there exists $C_1 \subseteq A_1$ such that $C_1 \in \mathcal{A}$ and $0 < \mu(C_1) < \infty$. Note that C_1 and C are disjoint. It follows that the disjoint union $C_1 \cup C \subseteq A$ is such that $\mu(C \cup C_1) \in S$. But since $\mu(C_1 \cup C) = \mu(C_1) + \mu(C) > \mu(C) = l$, therefore S contains an element which is strictly larger than its supremum, a contradiction. Hence $l = \infty$ and this completes the proof. □

Lebesgue measure on $\mathbb{R}$

In this section $(\mathbb{R}, \mathcal{M}, m)$ denotes the Lebesgue measure space on $\mathbb{R}$ and m^* denotes the Lebesgue outer measure on $\mathbb{R}$.

Lemma 18.3.6.11. *Every Borel subset of $\mathbb{R}$ is Lebesgue measurable.*

Proof. Let $(\mathbb{R}, \mathcal{M}, m)$ be the Lebesgue measure space over $\mathbb{R}$. We wish to show that the σ-algebra of Borel sets denoted $\mathcal{B}$ is in $\mathcal{M}$. Denote by $\mathcal{A}$ the following:

$$\mathcal{A} = \{\text{disjoint finite union of intervals of form } (-\infty, a], (b, \infty), (a, b] \text{ for } a < b \in \mathbb{R}\}. \tag{1.1}$$

By construction of Lebesgue measure, we know that $\mathcal{A} \subseteq \mathcal{M}$. We thus claim that the σ-algebra generated by $\mathcal{A}$ contains $\mathcal{B}$, that is, $\langle\mathcal{A}\rangle \supseteq \mathcal{B}$. This will conclude the proof.

Indeed, as we know that $\mathcal{B}$ is generated by all closed intervals of the form $(-\infty, a]$ for all $a \in \mathbb{R}$, therefore it suffices to show that $(-\infty, a] \in \langle\mathcal{A}\rangle$, but that is a tautology as $(-\infty, a]$ is in $\mathcal{A}$. Hence $\mathcal{B} \subseteq \langle A\rangle$. □

Lemma 18.3.6.12. *Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. Then,*

1. $m^*(A+c) = m^*(A)$,
2. $A \in \mathcal{M}$ *if and only if* $A + c \in \mathcal{M}$,
3. *if* $A \in \mathcal{M}$, *then* $m(A+c) = m(A)$.

Proof. Consider the Lebesgue measure space $(\mathbb{R}, \mathcal{M}, m)$. Take $A \subseteq \mathbb{R}$ and for $c \in \mathbb{R}$ define $A + c = \{a + c \in \mathbb{R} \mid a \in A\}$. Let us set up some notation. For any $E \subseteq \mathbb{R}$, we denote

$$C(E) = \left\{ \{I_n\} \mid \bigcup_n I_n \supseteq A,\ I_n = (a_n, b_n] \in \mathcal{A} \right\} \tag{$*$}$$

where $\mathcal{A}$ is the algebra defined in Eq. (1.1). Further, let us denote

$$\Sigma C(E) = \left\{ \sum_n l(I_n) \in [0, \infty] \mid \{I_n\} \in C(E) \right\} \tag{$**$}$$

where $l((a,b]) = b - a$ is the length function. By definition, we have $m^*(E) := \inf \Sigma C(E)$.

(i) : We first wish to show that the Lebesgue outer measure m^* is translation invariant. That is, $m^*(A+c) = m^*(A)$. We first show $m^*(A+c) \geq m^*(A)$. Pick any $\{I_n\} \in C(A)$. Then we claim that $\{I_n + c\}$ is an element of $C(A+c)$. Indeed, denoting $I_n = (a_n, b_n]$, we immediately get $I_n + c = (a_n + c, b_n + c]$. Now to see that $\bigcup_n (I_n + c) \supseteq A + c$, pick any $a + c \in A + c$ where $a \in A$. Then, as $\bigcup_n I_n \supseteq A$, therefore $a \in I_n$ for some n and thus $a + c \in I_n + c$. It follows that $\{I_n + c\} \in C(A+c)$. Further note that $l(I_n) = l(I_n + c)$ by definition. Consequently, we have

$$\Sigma C(A) \subseteq \Sigma C(A+c).$$

Taking infima, we yield $m^*(A) = \inf \Sigma C(A) \leq \inf \Sigma C(A+c) = m^*(A+c)$, that is $m^*(A) \leq m^*(A+c)$.

Conversely, we wish to show that $m^*(A) \geq m^*(A+c)$. For this, we use the standard technique of *ϵ-wiggle around* inf. Fix $\epsilon > 0$. By definition of $m^*(A)$, there exists $\{I_n\} \in C(A)$ where $I_n = (a_n, b_n]$ such that

$$m^*(A) + \epsilon > \sum_n b_n - a_n. \tag{2.1}$$

Note that we can write the above as

$$\begin{aligned} m^*(A) + \epsilon &> \sum_n (b_n + c) - (a_n + c) \\ &= \sum_n l((a_n + c, b_n + c]) \\ &= \sum_n l(I_n + c). \end{aligned}$$

We have $\{I_n + c\} \in C(A+c)$ as shown previously, therefore we obtain

$$m^*(A) + \epsilon > \sum_n l(I_n + c) \geq \inf \Sigma C(A+c) = m^*(A+c).$$

Hence we have $m^*(A) + \epsilon > m^*(A+c)$. Taking $\epsilon \to 0$, we obtain $m^*(A) \geq m^*(A+c)$. This completes the proof.

(ii) : We next wish to show that $A + c \in \mathcal{M}$ if and only if $A \in \mathcal{M}$. Observe that it suffices to show that $A \in \mathcal{M} \implies A + c \in \mathcal{M}$. Indeed, for the converse, take $B = A + c \in \mathcal{M}$. To show that $A \in \mathcal{M}$, it would suffice to show that $B - c \in \mathcal{M}$, which would follow at once by previous. Hence, we may only show that $A \in \mathcal{M} \implies A + c \in \mathcal{M}$.

Pick $A \in \mathcal{M}$. Fix $\epsilon > 0$. By regularity theorems, there exists open $U \supseteq A$ such that $m^*(U \setminus A) < \epsilon$. We now claim the following three statements:

1. *$U + c$ is open* : Indeed, pick any $x + c \in U + c$ where $x \in U$. As U is open, there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq U$. Consequently, $(x - \delta + c, x + \delta + c) \subseteq U + c$, hence $U + c$ is open.
2. *$U + c$ contains $A + c$* : Pick any $a + c \in A + c$ where $a \in A$. As $U \supseteq A$, therefore $a + c \in U + c$.
3. *$(U + c) \setminus (A + c)$ equals $(U \setminus A) + c$* : We first show $(U + c) \setminus (A + c) \subseteq (U \setminus A) + c$. Pick any $x + c \in (U + c) \setminus (A + c)$. Then $x + c \in U + c$ and $x + c \notin A + c$. Thus, $x \in U$ and $x \notin A$. Hence $x \in U \setminus A$ and thus $x + c \in U \setminus A + c$.
   Conversely, pick $x + c \in (U \setminus A) + c$. Then $x \in U \setminus A$ and thus $x + c \in U + c$ and $x + c \notin A + c$. Thus $x + c \in (U + c) \setminus (A + c)$. This completes the proof of this claim.

By above three claims, we conclude that $U + c$ is an open set containing $A + c$ such that

$$m^*(U + c \setminus A + c) = m^*((U \setminus A) + c) \overset{(ii)}{=} m^*(U \setminus A) < \epsilon.$$

By regularity theorems, we conclude the proof.

(iii) : We wish to show that if $A \in \mathcal{M}$, then $m(A + c) = m(A)$. This is immediate from *(i)* as $m = m^*|_{\mathcal{M}}$. □

Lemma 18.3.6.13. *Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. Then,*

1. *$m^*(cA) = |c|\, m^*(A)$,*
2. *for $c \neq 0$, $A \in \mathcal{M}$ if and only if $cA \in \mathcal{M}$,*
3. *if $A \in \mathcal{M}$, then $m(cA) = |c|\, m(A)$.*

Proof. Let $(\mathbb{R}, \mathcal{M}, m)$ be the Lebesgue measure space. Take any $A \subseteq \mathbb{R}$ and $c \in \mathbb{R}$.

(i) : We first wish to show that $m^*(cA) = |c|\, m^*(A)$. If $c = 0$, then the equality is immediate as $cA = \{0\}$ and we know that $m^*(\{0\}) = 0$ as $0 \in (-1/n, 1/n]$ for all $n \in \mathbb{N}$ so that $m^*(\{0\}) \leq 2/n$. Taking $n \to \infty$, we get that $m^*(\{0\}) = 0$. So we assume from now on that $c \neq 0$. We first immediately reduce to showing either one of

$$m^*(cA) \geq |c|\, m^*(A) \text{ or } m^*(cA) \leq |c|\, m^*(A)$$

Indeed, the other side follows by replacing A by cA and replacing c by $1/c$ in either of the above. We now have two cases based on c being positive or negative.

If $c > 0$, then we proceed as follows. We follow the convention of Eqns $(*)$ and $(**)$ as set up in Q2. We use the standard technique of ϵ-wiggle around inf. Fix $\epsilon > 0$. By definition of outer measure, there exists $\{I_n\} \in C(cA)$ where $I_n = (a_n, b_n]$ such that

$$m^*(cA) + \epsilon > \sum_n l(I_n). \tag{3.1}$$

As $\bigcup_n I_n \supseteq cA$ and $c > 0$, therefore we claim that $\bigcup_n(\frac{1}{c}I_n) \supseteq A$. Indeed, for any $a \in A$, $cA \in I_n$. Thus $ca \in (a_n, b_n]$. Consequently, $a \in (a_n/c, b_n/c] = (\frac{1}{c}I_n)$. Thus, $\{\frac{1}{c}I_n\} \in C(A)$. Consequently, we have

$$\sum_n l\left(\frac{1}{c}I_n\right) = \sum_n \frac{1}{c}l(I_n) \geq m^*(A).$$

Consequently, $\sum_n l(I_n) \geq cm^*(A)$. Using this in Eq. (3.1), we thus obtain

$$m^*(cA) + \epsilon > \sum_n l(I_n) \geq cm^*(A).$$

Taking $\epsilon \to 0$, we obtain $m^*(cA) \geq cm^*(A)$, as required.

If $c < 0$, then we begin similarly to the previous case. Fix $\epsilon > 0$. There exists $\{I_n\} \in C(A)$ where $I_n = (a_n, b_n]$ such that

$$m^*(A) + \epsilon > \sum_n l(I_n). \tag{3.2}$$

Note that $cI_n = c(a_n, b_n] = [cb_n, ca_n)$ as $c < 0$ and this type of set is not half-open and is thus not in $\mathcal{A}$, the algebra of half-opens of Eq. (1.1). Consequently, we have to use ϵ-wiggle to find a new collection of intervals obtained via cI_n which are half open but their sum of lengths in only in ϵ-neighborhood of those $\{cI_n\}$. Indeed, for each $n \in \mathbb{N}$, we may construct

$$J_n = \left(cb_n - \frac{\epsilon}{2^{n+1}}, ca_n + \frac{\epsilon}{2^{n+1}}\right].$$

Note that $J_n \supseteq cI_n$. As $\bigcup_n cI_n \supseteq cA$, therefore $\bigcup_n J_n \supseteq cA$. Thus $\{J_n\} \in C(cA)$. Consequently,

$$\begin{aligned}
m^*(cA) &\leq \sum_n l(J_n) \\
&= \sum_n c(a_n - b_n) + \frac{2\epsilon}{2^{n+1}} \\
&= \sum_n -c(b_n - a_n) + \sum_n \frac{\epsilon}{2^n} \\
&= -c\sum_n (b_n - a_n) + \epsilon \\
&= -c\sum_n l(I_n) + \epsilon
\end{aligned}$$

where the third line follows from the series being positive and thus we can rearrange such a series. It thus follows by Eq. (3.2) and above that

$$\begin{aligned}
m^*(cA) &< -c(m^*(A) + \epsilon) + \epsilon \\
&= -cm^*(A) + \epsilon(1 - c).
\end{aligned}$$

Taking $\epsilon \to 0$, we obtain ($-c = |c|$ as $c < 0$)

$$m^*(cA) \leq |c|\, m^*(A)$$

as required. This completes the proof.

(ii) : We now wish to show that for $c \neq 0$, $A \in \mathcal{M}$ if and only if $cA \in \mathcal{M}$. Note that this is not true for $c = 0$ as if we take a non-measurable set $V \subseteq \mathbb{R}$, then $cV = \{0\}$ is measurable but V is not.

Pick $c \neq 0$. We first note that showing only $A \in \mathcal{M} \implies cA \in \mathcal{M}$ is sufficient. Indeed, the other side follows by replacing c by $1/c$ in the above. So we reduce to showing $A \in \mathcal{M} \implies cA \in \mathcal{M}$.

Pick $A \in \mathcal{M}$ and $c \neq 0$ in $\mathbb{R}$. Fix $\epsilon > 0$. By regularity theorems, there exists open $U \setminus A$ such that $m^*(U \setminus A) < \frac{\epsilon}{|c|}$. We now claim the following statements:

1. *cU is open* : Pick $cx \in cU$ where $x \in U$. As U is open therefore there exists $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq U$. Consequently, $c(x-\delta, x+\delta) = (c(x+\delta), c(x-\delta)) \subseteq cU$ and contain cx. Hence cU is open.
2. *cU contains cA* : Pick any cx in cA. Then $x \in A$. As $U \subseteq A$, therefore $x \in U$ and hence $cx \in cU$.
3. *$cU \setminus cA$ equals $c(U \setminus A)$* : For $(\subseteq)$, pick any $cx \in cU \setminus cA$. Then $cx \in cU$ and $cx \notin cA$. Thus, $x \in U$ and $x \notin A$, that is $\in U \setminus A$ and thus $cx \in c(U \setminus A)$. Conversely to show $(\supseteq)$, pick any $cx \in c(U \setminus A)$ where $x \in U \setminus A$. Thus, $x \in U$ and $x \notin A$. Thus $cx \in cU$ and $cx \notin cA$. Thus $cx \in cU \setminus cA$.

Following the above three lemmas, we conclude that cU is an open set containing cA such that

$$m^*(cU \setminus cA) = m^*(c(U \setminus A)) \overset{(i)}{=} |c|\, m^*(U \setminus A) < |c| \frac{\epsilon}{|c|} = \epsilon.$$

Thus by regularity theorems, $cA \in \mathcal{M}$ as well.

(iii) : We wish to show that if $A \in \mathcal{M}$, then $m(cA) = |c|\, m(A)$. But this is immediate from *(i)* as $m = m^*|_{\mathcal{M}}$. This completes the whole proof. □

Lemma 18.3.6.14. *For each subseteq $A \subseteq \mathbb{R}$, there exists a Borel subset $B \supseteq A$ such that*

$$m^*(A) = m(B).$$

Proof. We wish to show that for each $A \subseteq \mathbb{R}$, there exists a Borel set $B \supseteq A$ such that $m(B) = m^*(A)$. We divide into two cases based on outer measure of A. We will follow the notations of Eq. $(*)$ and $(**)$.

If $m^*(A) = \infty$. In this case, we claim that $B = \mathbb{R}$ will work. Indeed $\mathbb{R}$ is open and thus Borel. We thus claim that $m(\mathbb{R}) = \infty$. Indeed, for $I_n = (n, n+1]$, $n \in \mathbb{Z}$, we have that $\{I_n\}$ are disjoint and $\coprod_n I_n = \mathbb{R}$. As m is a measure and I_n are measurable, therefore

$$m(\mathbb{R}) = \sum_n m(I_n) = \sum_n 1 = \infty.$$

Hence $B = \mathbb{R}$ will work.

If $m^*(A) < \infty$, then we proceed as follows. For each $N \in \mathbb{N}$, there exists $\{I_n^N\} \in C(A)$ such that

$$m^*(A) + \frac{1}{N} > \sum_n l(I_n^N).$$

Define $U_N = \bigcup_n I_n^N$. As each half open interval $(a,b] = \bigcap_{n\in\mathbb{N}}(a, b+1/n)$ is a Borel set, therefore U_N is a Borel set. Observe that

$$m(U_N) \leq \sum_n m^*(I_n^N) = \sum_n l(I_n^N) < m^*(A) + \frac{1}{N}.$$

Note that in the above we have used the fact that Lebesgue measure restricted to half opens is exactly the length function. We thus have for each $N \in \mathbb{N}$ a Borel set U_N containing A such that

$$m(U_N) < m^*(A) + \frac{1}{N}. \tag{4.1}$$

Denote $B_K = \bigcap_{N=1}^K U_N$. Then each B_K is Borel and $\{B_K\}$ is a decreasing sequence of sets. Furthermore, $\bigcap_{K=1}^\infty B_K = \bigcap_{N=1}^\infty U_N$. Denote $B = \bigcap_{K=1}^\infty B_K$. Observe that $B \supseteq A$ as $B_K \supseteq A$ for each $K \in \mathbb{N}$. Consequently, by continuity of m^* we have

$$m(B) \geq m^*(A).$$

For the converse, first note that by Eq. (4.1), $m(U_1) < \infty$. Thus by monotone convergence property of measures, we obtain that $\lim_{K\to\infty} m(B_K) = m\left(\bigcap_{K=1}^\infty B_K\right)$. It follows from above, $B_K \subseteq U_K$ and Eq. (4.1) that

$$\begin{aligned}
m(B) &= m\left(\bigcap_{K=1}^{\infty} B_K\right) \\
&= \lim_{K\to\infty} m(B_K) \\
&\leq \lim_{K\to\infty} m(U_K) \\
&\overset{(4.1)}{<} \lim_{K\to\infty}\left(m^*(A) + \frac{1}{K}\right) \\
&\leq m^*(A).
\end{aligned}$$

Thus $m(B) \leq m^*(A)$ and we are done. □

Lemma 18.3.6.15. *A bounded set $E \subseteq \mathbb{R}$ is measurable if and only if $m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$ for all bounded subsets $A \subseteq \mathbb{R}$.*

Proof. Let E be a bounded set of $\mathbb{R}$. We wish to show that E is measurable if and only if for all bounded sets $A \subseteq \mathbb{R}$, we get $m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c)$.

The ($\Rightarrow$) is immediate from definitions. For ($\Leftarrow$), we proceed as follows. We wish to show that for any $F \subseteq \mathbb{R}$, we have

$$m^*(F) \geq m^*(F \cap E) + m^*(F \cap E^c).$$

Indeed, if $m^*(F) = \infty$, then there is nothing to show. So we assume $m^*(F) < \infty$. Observe then that $m^*(F \cap E), m^*(F \cap E^c) \leq m^*(F) < \infty$. Fix $\epsilon > 0$. There exists a sequence $\{I_n\}$ of half-opens such that $\bigcup_n I_n \supseteq F$ and

$$m^*(F) + \epsilon > \sum_n m^*(I_n)$$

where we are using the fact that measure of a half-open interval is its length. Observe that for each $n \in \mathbb{N}$, we have $m^*(F) + \epsilon > m^*(I_n)$, thus each I_n is a half-open interval with bounded length, hence I_n is bounded as a set. Consequently, we have

$$\begin{aligned}
m^*(F) + \epsilon &> \sum_n m^*(I_n) \\
\text{(by hypothesis)} \quad &\geq \sum_n m^*(I_n \cap E) + m^*(I_n \cap E^c) \\
\text{(by rearrangement of +ve series)} \quad &= \sum_n m^*(I_n \cap E) + \sum_n m^*(I_n \cap E^c) \\
\text{(by subadditivity)} \quad &\geq m^*\left(\bigcup_n I_n \cap E\right) + m^*\left(\bigcup_n I_n \cap E^c\right) \\
\text{(by } \cup_n I_n \supseteq F\text{)} \quad &\geq m^*(F \cap E) + m^*(F \cap E^c).
\end{aligned}$$

This completes the proof. □

measurable functions

Notation 18.3.6.16. At times, we will write a subset of X as follows:

$$\{x \in X \mid \mathcal{P}_x \text{ is true}\} = \{\mathcal{P}_x \text{ is true}\}.$$

This makes some constructions much more clearer to see and interpret.

Lemma 18.3.6.17. *Let $f : X \to Y$ be a function and $\mathcal{A}$ be an algebra on Y. Then,*

$$\langle f^{-1}(\mathcal{A})\rangle = f^{-1}(\langle\mathcal{A}\rangle).$$

Proof. Let $f : X \to Y$ be a function and $\mathcal{A}$ be an algebra over Y. We wish to show that

$$\langle f^{-1}(\mathcal{A})\rangle = f^{-1}(\langle\mathcal{A}\rangle). \tag{2.1}$$

We first claim that $f^{-1}(\langle\mathcal{A}\rangle)$ is a σ-algebra over X. Indeed, as $Y, \emptyset \in \langle\mathcal{A}\rangle$, we have $f^{-1}(Y) = X$ and $f^{-1}(\emptyset) = \emptyset$. Further, if $B \in f^{-1}(\langle A\rangle)$, then $B = f^{-1}(A)$ for some $A \in \langle\mathcal{A}\rangle$. Hence $B^c = f^{-1}(A)^c = f^{-1}(A^c)$ and $A^c \in \langle A\rangle$ as $\langle A\rangle$ is a σ-algebra. Finally, pick $\{B_n\} \subseteq f^{-1}(\langle A\rangle)$. Then $B_n = f^{-1}(A_n)$ for $A_n \in \langle\mathcal{A}\rangle$. Consequently, $\bigcup_n B_n = \bigcup_n f^{-1}(A_n) = f^{-1}(\bigcup_n A_n)$ and since $\bigcup_n A_n \in \langle A\rangle$, hence this proves that $f^{-1}(\langle\mathcal{A}\rangle)$ is a σ-algebra.

We now show ($\subseteq$) part of Eq. (2.1). Indeed, by above, it would suffice to show that $f^{-1}(\mathcal{A})$ is contained in the σ-algebra $f^{-1}(\langle\mathcal{A}\rangle)$. Pick any $B \in f^{-1}(\mathcal{A})$, so that $B = f^{-1}(A)$ where $A \in \mathcal{A}$. As $\mathcal{A} \subseteq \langle\mathcal{A}\rangle$, therefore $A \in \langle\mathcal{A}\rangle$. It follows that $B = f^{-1}(A) \in f^{-1}(\langle A\rangle)$. This shows that $\langle f^{-1}(\mathcal{A})\rangle \subseteq f^{-1}(\langle\mathcal{A}\rangle)$.

We now show ($\supseteq$) part of Eq. (2.1). We will use the standard technique of *good sets* for this. Consider

$$\mathcal{C} := \{A \in \langle\mathcal{A}\rangle \mid f^{-1}(A) \in \langle f^{-1}(\mathcal{A})\rangle\} \subseteq \langle\mathcal{A}\rangle$$

We now claim the following two statements:

1. *$\mathcal{C}$ is a σ-algebra on Y* : Indeed, $Y = f^{-1}(X)$ and $\emptyset = f^{-1}(\emptyset)$ where $X, \emptyset \in \langle \mathcal{A} \rangle$ and $X, \emptyset \in \langle f^{-1}(\mathcal{A}) \rangle$. Further, for $A \in \mathcal{C}$, we have $f^{-1}(A) \in \langle f^{-1}(\mathcal{A}) \rangle$ and thus $(f^{-1}(A))^c = f^{-1}(A^c) \in \langle f^{-1}(\mathcal{A}) \rangle$. Thus $A^c \in \mathcal{C}$. Finally, pick $\{A_n\} \subseteq \mathcal{C}$. Then $f^{-1}(A_n) \in \langle f^{-1}(\mathcal{A}) \rangle$ for each $n \in \mathbb{N}$. Thus, $\bigcup_n f^{-1}(A_n) = f^{-1}\left(\bigcup_n A_n\right) \in \langle f^{-1}(\mathcal{A}) \rangle$. It then follows that $\bigcup_n A_n \in \mathcal{C}$. This shows that $\mathcal{C}$ is a σ-algebra.
2. $\mathcal{C} \supseteq \mathcal{A}$: Pick any $A \in \mathcal{A}$. As $\langle f^{-1}(\mathcal{A}) \rangle$ contains $f^{-1}(\mathcal{A})$, so $f^{-1}(A) \in \langle f^{-1}(\mathcal{A}) \rangle$.

We now conclude the proof. As $\mathcal{C}$ is a σ-algebra containing $\mathcal{A}$ and inside $\langle \mathcal{A} \rangle$, therefore $\mathcal{C} = \langle \mathcal{A} \rangle$. It follows that for each $A \in \langle \mathcal{A} \rangle$, we have $f^{-1}(A) \in \langle f^{-1}(\mathcal{A}) \rangle$, that is $f^{-1}(\langle \mathcal{A} \rangle) \subseteq \langle f^{-1}(\mathcal{A}) \rangle$, as required. This completes the proof. $\square$

Lemma 18.3.6.18. *Let $(X, \mathcal{M}, m)$ be the Lebesgue measure space. Let $A \in \mathcal{M}$ be a bounded set such that $0 < m(A) < \infty$. For each $0 < M < m(A)$, there exists a $B \subsetneq A$ such that $B \in \mathcal{M}$ and $m(B) = M$.*

Proof. There are two proofs that we wish to present, one uses Lemma 18.3.6.19 and other is independent. The latter uses a nice technique which we would like to write down concretely.

Method 1 : *(Using Lemma 18.3.6.19)* Consider the map

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]). \end{aligned}$$

As A is a bounded set, therefore $m(A) < \infty$ as there exists a bounded interval $I \supseteq A$ where $I = [c, d]$. By Lemma 18.3.6.19, the map f is a continuous map. Let $a \in \mathbb{R}$ be such that $a < c$. Then $f(a) = m(A \cap (-\infty, a]) = m(\emptyset) = 0$. Let $b \in \mathbb{R}$ such that $b > d$. Then, $f(b) = m(A \cap (-\infty, b]) = m(A)$. On the interval $J = [a, b]$ we have $f(a) = 0$ and $f(b) = m(A)$. By intermediate value property of f, there exists $c \in J$ such that $f(c) = M$. Consequently, $A \cap (-\infty, c]$ is a measurable subset of A whose measure is M.

Method 2 : *(Exponential subdivision technique)* We shall explicitly construct $B \subsetneq A$ such that $m(B) = M$. First, we observe that the question is invariant under translation and dilation. Hence we may, after suitable dilation and translation, assume that $A \subseteq [0, 1)$. For each $n \in \mathbb{N}$, consider the following partition of $[0, 1)$

$$P_n : 0 < x_1 = \frac{1}{2^n} < x_2 = 2 \cdot \frac{1}{2^n} < \cdots < x_{2^n - 1} = (2^n - 1) \cdot \frac{1}{2^n} < 1.$$

Denote $I_{n,j} = [\frac{j-1}{2^n}, \frac{j}{2^n})$ for each $j = 1, \dots, 2^n$. Observe that $I_{n,j}$ are disjoint and, denoting $A_{n,j} = A \cap I_{n,j}$, we further have a disjoint collection $\{A_{n,j}\}$ of measurable subsets[17] of A such that

$$\coprod_{j=1}^{2^n} A_{n,j} = A.$$

Further, we have that

$$\begin{aligned} \sum_{j=1}^{2^n} m(A_{n,j}) &= m\left(\coprod_{j=1}^{2^n} A_{n,j}\right) \\ &= m(A) \end{aligned}$$

[17] measurable because A and $I_{n,j}$ are measurable

and that

$$m(A_{n,j}) \leq m(I_{n,j}) = \frac{1}{2^n}.$$

Now, for each $n \in \mathbb{N}$, let N_n be the largest index such that

$$\sum_{j=1}^{N_n} m(A_{n,j}) \leq M.$$

By the choice of index N_n, we observe that

$$\begin{aligned} M &< \sum_{j=1}^{N_n+1} m(A_{n,j}) \\ &= \sum_{j=1}^{N_n} m(A_{n,j}) + m(A_{n,N_n+1}) \\ &\leq \sum_{j=1}^{N_n} m(A_{n,j}) + \frac{1}{2^n}. \end{aligned}$$

Denoting $C_n = \amalg_{j=1}^{N_n} A_{n,j}$, we obtain,

$$M - \frac{1}{2^n} < \sum_{j=1}^{N_n} m(A_{n,j}) = m\left(C_n\right) \leq M. \tag{3.1}$$

We now claim that $\{C_n\}$ is an increasing sequence of measurable subsets of A. First observe that for each $n \in \mathbb{N}$, we have that N_{n+1} is either $2N_n - 1$ or $2N_n$. Indeed, pick any $x \in C_n$. Then $x \in A_{n,j}$ where $j = 1, \ldots, N_n$. Expanding this, we have

$$\begin{aligned} x &\in A_{n,j} \\ &= A \cap \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right) \\ &= A \cap \left(\left[\frac{2(j-1)}{2^{n+1}}, \frac{2j-1}{2^{n+1}}\right) \amalg \left[\frac{2j-1}{2^{n+1}}, \frac{2j}{2^{n+1}}\right)\right) \\ &= A_{n+1,2j-1} \amalg A_{n+1,2j}. \end{aligned} \tag{3.2}$$

As $N_{n+1} = 2N_n - 1$ or $2N_n$, therefore for $j = 1, \ldots, N_n$, $2j = 2, \ldots, 2N_n$, hence in Eq. (3.2), we obtain that $x \in A_{n+1,2j-1}$ or $x \in A_{n+1,2j}$ and as $2j \leq 2N_n$, hence $x \in C_{n+1}$. This shows that $C_n \subseteq C_{n+1}$.

Applying $\lim_{n\to\infty}$ on Eq. (3.1), we thus obtain

$$M \leq \lim_{n\to\infty} m(C_n) \leq M.$$

Thus, by monotone convergence of measures, we conclude

$$\begin{aligned} M &= \lim_{n\to\infty} m(C_n) \\ &= m\left(\bigcup_n C_n\right). \end{aligned}$$

As $C_n \subseteq A$ for each $n \in \mathbb{N}$, therefore $\bigcup_n C_n \subseteq A$. Consequently we have obtained a subset of A whose measure is M. □

Lemma 18.3.6.19. *Let $(X, \mathcal{M}, \mu)$ be the Lebesgue measure space and $A \in \mathcal{M}$ be a bounded set. Then the function*

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]) \end{aligned}$$

is continuous.

Proof. Let $A \in \mathcal{M}$ which has finite measure. We wish to show that

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto m(A \cap (-\infty, x]) \end{aligned}$$

is continuous. Pick any $a \in \mathbb{R}$ and any $\epsilon > 0$. We wish to find a $\delta > 0$ such that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$. We now have two cases with respect to the position of x and a in $\mathbb{R}$.

1. *If $a \leq x$:* then $f(x) - f(a)$ can be rewritten as follows:

$$\begin{aligned} |f(x) - f(a)| &= m(A \cap (-\infty, x]) - m(A \cap (-\infty, a]) \\ &= m(A \cap (-\infty, x] \setminus A \cap (-\infty, a]) \\ &= m(A \cap (a, x]) \\ &\leq m((a, x]) \\ &= x - a. \end{aligned}$$

   Therefore taking $\delta = \epsilon$, we would be done.

2. *If $a > x$:* then $|f(x) - f(a)|$ can be written as

$$\begin{aligned} |f(x) - f(a)| &= |f(a) - f(x)| \\ &= m(A \cap (-\infty, a]) - m(A \cap (-\infty, x]) \\ &= m(A \cap (x, a]) \\ &\leq m((x, a]) \\ &= a - x. \end{aligned}$$

   Thus, again, taking $\delta = \epsilon$ would do the job.

This completes the proof. □

Lemma 18.3.6.20. *Let X be a measurable space and let $f : X \to \mathbb{R}$ be a function. Suppose $\{x \in X \mid a \leq f(x) < b\}$ is measurable for all $a < b$. Then f is a measurable function.*

Proof. As the Borel σ-algebra on $\mathbb{R}$ is generated by sets of the form $[a, \infty)$ for $a \in \mathbb{R}$, therefore for a fixed $a \in \mathbb{R}$ we need only show that $f^{-1}([a, \infty))$ is measurable in X.

We can write

$$\begin{aligned} f^{-1}([a, \infty)) &= \{a \leq f(x)\} \\ &= \bigcup_{n > a \text{ in } \mathbb{N}} \{a \leq f(x) < n\}. \end{aligned}$$

As we are given that $\{a \leq f(x) < b\}$ are measurable for all $a < b \in \mathbb{R}$ and countable union of measurable sets is measurable, therefore $f^{-1}([a,\infty))$ is measurable. $\square$

Lemma 18.3.6.21. *All monotone functions $f : \mathbb{R} \to \mathbb{R}$ are measurable.*

Proof. We wish to show that all monotone functions $f : \mathbb{R} \to \mathbb{R}$ are measurable. Note that we may first reduce to assuming that f is non-decreasing as if f is non-increasing, then $-f$ will be non-decreasing.

Hence let $f : \mathbb{R} \to \mathbb{R}$ is non-decreasing. As Borel σ-algebra of $\mathbb{R}$ is generated by intervals of the form $[a,\infty)$, $a \in \mathbb{R}$, therefore it suffices to check that $f^{-1}([a,\infty))$ is measurable in $\mathbb{R}$. Observe

$$f^{-1}([a,\infty)) = \{a \leq f(x)\}.$$

We now have two cases to handle.

1. *If $a \in f(\mathbb{R})$* : Then there exists $b \in \mathbb{R}$ such that $f(b) = a$. We may write

$$\{a \leq f(x)\} = \{a < f(x)\} \amalg \{a = f(x)\}.$$

Now since f is non-decreasing, therefore $f(x) > f(y)$ implies $x > y$. Further, we have that $f^{-1}(a) = \{a = f(x)\}$ is measurable as singletons are Borel. Consequently, we have

$$\begin{aligned}\{a \leq f(x)\} &= \{f(b) < f(x)\} \amalg \{a = f(x)\}\\ &= (b,\infty) \amalg f^{-1}(a).\end{aligned}$$

Hence $f^{-1}([a,\infty))$ is measurable.

2. *If $a \notin f(\mathbb{R})$* : We further have two cases.
   (a) *If there exists $b \in \mathbb{R}$ such that $b \notin \{a \leq f(x)\}$* : Observe first that in this case $f(b) < a$. We claim that in this case $\{a \leq f(x)\}$ is lower bounded by b. Indeed, suppose not. Then there exists $y < b$ such that $y \in \{a \leq f(x)\}$. Then $a \leq f(y) \leq f(b) < a$, a contradiction. Hence $\{a \leq f(x)\}$ is bounded below.
   Let $c = \inf\{a \leq f(x)\}$, which now exists. Consequently, we have two more cases:
   - *If $f(c) \geq a$* : That is, if $c \in \{a \leq f(x)\}$. Then we claim

$$\{a \leq f(x)\} = [c,\infty).$$

which is clearly a measurable. Indeed, for some $x \in \mathbb{R}$ such that $f(x) \geq a$, then $x \geq c$. Conversely, if $b \geq c$ in $\mathbb{R}$, then $f(b) \geq f(c) \geq a$, so $b \in \{a \leq f(x)\}$. This proves the claim.
   - *If $f(c) < a$* : That is, if $c \notin \{a \leq f(x)\}$. Then we claim

$$\{a \leq f(x)\} = (c,\infty)$$

which is clearly a measurbale set. Indeed, for $x \in \mathbb{R}$ such that $f(x) \geq a$, $x > c$. Further $x \neq c$ as otherwise $f(x) < a$. Conversely, if $b > c$, then there exists $d \in \{a \leq f(x)\}$ such that $c < d < b$ as c is the infimum. Consequently, $a \leq f(d) \leq f(b)$. Hence $b \in \{a \leq f(x)\}$. This proves the claim.
   (b) *If there doesn't exists any $b \in \mathbb{R}$ such that $f(b) < a$* : Then for all $b \in \mathbb{R}$ we have $f(b) \geq a$. Consequently, $f^{-1}([a,\infty)) = \{a \leq f(x)\} = \mathbb{R}$, which is measurable.

Hence in all cases $f^{-1}([a, \infty))$ is a measurable set. This completes the proof. $\square$

Lemma 18.3.6.22. *Let $f : X \to \mathbb{C}$ be a complex measurable function on a measurable space X. Then, there exists a complex measurable function $g : X \to \mathbb{C}$ such that $|g| = 1$ and $f = g\,|f|$.*

Proof. Let $f : X \to \mathbb{C}$ be a measurable function. We wish to find a measurable function $g : X \to \mathbb{C}$ such that $|g| = 1$ and $f = g\,|f|$.

As $|f| = f\chi_{\{f(x)\geq 0\}} - f\chi_{\{f(x)<0\}}$, therefore $|f|$ is a measurable function. Denote $E = \{|f(x)| = 0\}$. Consequently, we define g as follows:

$$g(x) = \begin{cases} \frac{f(x)}{|f(x)|} & \text{if } x \in E^c \\ 1 & \text{if } x \in E. \end{cases}$$

We first wish to show that g is measurable. For this, we shall use the fact that measurability of g can be checked on a cover $\{E_\alpha\}$ of X such that $g|_{E_\alpha}$ is measurable. Thus in our case, we need only show that $g|_E$ and $g|_{E^c}$ are measurable. On E, g is a constant, hence measurable. On E^c, g is $f/|f|$. As $|f|$ is not zero on E^c, therefore by Lemma 18.3.6.24, $f/|f|$ is measurable. Hence, g is measurable.

We now see that $|g|(x) = \left|\frac{f(x)}{|f(x)|}\right| = 1$ on E^c and $|g(x)| = 1$ on E. Thus $|g| = 1$ on X. Further, if $x \in E$, then $f(x) = 0 = g(x)\,|f|\,(x)$. If $x \in E^c$, then $g(x) = \frac{f(x)}{|f(x)|}$ which implies $|f(x)|\,g(x) = f(x)$. This shows that in all cases, $f = g\,|f|$. $\square$

Example 18.3.6.23. It is not true that if $f : [0, 1] \to \mathbb{R}$ is a function whose each fibre is measurable, then f is measurable.

Consider the following function

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} x & \text{if } x \in V^c \\ x + N & \text{if } x \in V \end{cases} \end{aligned}$$

where $V \subseteq [0, 1]$ denotes the Vitali set and $N = 3$. Then, for each $y \in \mathbb{R}$, we have that $f^{-1}(y)$ is atmost a singleton, which is measurable in $[0, 1]$. However, for any $1 < b < N$, we see that $f^{-1}((b, \infty)) = V$, which is not measurable. Hence f is a non-measurable function whose fibres are measurable.

Lemma 18.3.6.24. *Let $f, g : X \to \mathbb{C}$ be a measurable function such that $\{g(x) \neq 0\} = X$. Then f/g is measurable.*

Proof. Let $f, g : X \to \mathbb{C}$ be a measurable function such that $\{g(x) \neq 0\} = X$. Then we wish to show that f/g is measurable.

We first have that $(f, g) : X \to \mathbb{R}^2$ given by $x \mapsto (f(x), g(x))$ is measurable. Consequently, we consider the composite

$$X \xrightarrow{(f,g)} \mathbb{R}^2 \setminus \{y = 0\} \xrightarrow{\Phi} \mathbb{R}$$

where $\Phi(x, y) = \frac{x}{y}$. As Φ is continuous, therefore the composite $\Phi \circ (f, g)$ is measurable. Consequently, we obtain that the map $x \mapsto \frac{f(x)}{g(x)}$ is measurable, but this is exactly f/g over X. This completes the proof. $\square$

Lemma 18.3.6.25. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable functions and pick any $r_0 \in \overline{\mathbb{R}}$. Then the map*

$$\begin{aligned} h : X &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto \begin{cases} r_0 & \textit{if } f(x) = -g(x) = \pm\infty \\ f(x) + g(x) & \textit{else} \end{cases} \end{aligned}$$

is measurable[18].

Proof. Let $f, g : X \to \overline{\mathbb{R}}$ be measurable functions and pick any $r_0 \in \overline{\mathbb{R}}$. Then we wish to show that the map

$$\begin{aligned} h : X &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto \begin{cases} r_0 & \text{if } f(x) = -g(x) = \pm\infty \\ f(x) + g(x) & \text{else} \end{cases} \end{aligned}$$

is measurable.

Define the following sets

$$\begin{aligned} E &= \{f(x) = \infty = -g(x)\} \\ F &= \{f(x) = -\infty = -g(x)\}. \end{aligned}$$

As $E = f^{-1}(\infty) \cap g^{-1}(-\infty)$ and $F = f^{-1}(-\infty) \cap g^{-1}(\infty)$, therefore they are measurable. Observe that E and F are disjoint. We thus need only show that h restricted to E, F and $X \setminus (E \amalg F)$ is measurable.

1. On E : As $h|_E$ is constant r_0, therefore $h|_E$ is measurable.
2. On F : As $h|_F$ is again constant r_0, therefore $h|_F$ is measurable.
3. On $X \setminus (E \amalg F)$: We first deduce that

$$\begin{aligned} X \setminus (E \amalg F) &= X \cap E^c \cap F^c \\ &= E^c \cap F^c \\ &= (\{f(x) \neq \infty\} \cup \{g(x) \neq -\infty\}) \bigcap (\{f(x) \neq -\infty\} \cup \{g(x) \neq \infty\}) \end{aligned}$$

Let $G = \{f(x) \in \mathbb{R}\}$ and $H = \{g(x) \in \mathbb{R}\}$. Then we may write $X \setminus (E \amalg F)$ as

$$\begin{aligned} X \setminus (E \amalg F) &= (G \cup H \cup \{f(x) = -\infty\} \cup \{g(x) = \infty\}) \bigcap (G \cup H \cup \{f(x) = \infty\} \cup \{g(x) = -\infty\}) \\ &= (G \cup H) \cup \Big((\{f(x) = -\infty\} \cup \{g(x) = \infty\}) \bigcap (\{f(x) = \infty\} \cup \{g(x) = -\infty\}) \Big) \\ &= (G \cup H) \cup \underbrace{\{f(x) = -\infty = g(x)\}}_{=:A} \cup \underbrace{\{f(x) = \infty = g(x)\}}_{=:B}. \end{aligned}$$

As $h|_{G \cup H}$ is $(f+g)|_{G \cup H}$ and on $G \cup H$, $f + g : G \cup H \to \mathbb{R}$, therefore h is measurable. We thus reduce to checking that $h|_A$ and $h|_B$ are measurable. On both of them, one immediately observes that h is constant $-\infty$ and ∞ respectively. Hence, $h|_A$ and $h|_B$ are measurable. As h restricted to $G \cup H$, A and B is measurable therefore h restricted to $X \setminus (E \amalg F)$ is measurable.

[18]This question in particular shows that modifying a measurable function at a single point doesn't affect measurability at all.

This completes the proof. □

Example 18.3.6.26. It is not true in general that if for a function $f : X \to \mathbb{R}$, the $|f| : X \to [0, \infty]$ is measurable then f is measurable.

Indeed, consider the following function where $V \subseteq [0, 1]$ denotes the Vitali set:

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} -x & \text{if } x \in V \\ x & \text{if } x \in V^c. \end{cases} \end{aligned}$$

Then, $|f| = \mathrm{id}_{[0,1]}$ which is measurable whereas f is not measurable as $f^{-1}((-\infty, 0)) = V$, which is not a measurable set.

Lemma 18.3.6.27. *Let $(X_1, \mathcal{A}_{X_1}, \mu_1)$ be a measure space, $(X_2, \mathcal{A}_{X_2})$ be a measurable space and $f : X_1 \to X_2$ be a measurable function. Then*

$$\begin{aligned} \mu_2 : \mathcal{A}_{X_2} &\longrightarrow [0, \infty] \\ B &\longmapsto \mu_1(f^{-1}(B)) \end{aligned}$$

is a measure on $(X_2, \mathcal{A}_{X_2})$.

Proof. We first immediately observe that $\mu_2(\emptyset) = \mu_1(f^{-1}(\emptyset)) = \mu_1(\emptyset) = 0$. We thus reduce to showing that for any disjoint collection $\{B_n\} \subseteq \mathcal{A}_{X_2}$, we have $\mu_2(\coprod_n B_n) = \sum_n \mu_2(B_n)$. To this end, observe that

$$\begin{aligned} \mu_2\left(\coprod_n B_n\right) &= \mu_1\left(f^{-1}\left(\coprod_n B_n\right)\right) \\ &= \mu_1\left(\coprod_n f^{-1}(B_n)\right) \\ &= \sum_n \mu_1(f^{-1}(B_n)) \\ &= \sum_n \mu_2(B_n). \end{aligned}$$

This completes the proof. □

Lemma 18.3.6.28. *Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{R}$ be a measurable function such that $\mu(\{|f(x)| \geq \epsilon\}) = 0$ for all $\epsilon > 0$. Then $f = 0$ almost everywhere.*

Proof. We first claim that it suffices to show that $\{|f(x)| > 0\}$ is a null set. Indeed, this is because $\{f(x) \neq 0\} = \{|f(x)| > 0\}$. Hence it suffices to show that $|f| = 0$ a.e.

Define for each $n \in \mathbb{N}$ the following subset of X

$$E_n = \{|f(x)| > 1/n\}.$$

We claim that

$$\{|f(x)| > 0\} = \bigcup_{n \in \mathbb{N}} E_n.$$

Indeed, for ($\subseteq$), pick $x \in X$ such that $|f(x)| > 0$. Then there exists $n \in \mathbb{N}$ such that $|f(x)| > 1/n$. Hence $x \in E_n$. Conversely pick $x \in E_n$, then by way of construction of E_n, we have $|f(x)| > 1/n > 0$.

Observe that $\{E_n\}$ is an increasing sequence of sets as if $x \in E_n$ then $|f(x)| > \frac{1}{n} > \frac{1}{n+1}$, so $x \in E_{n+1}$. It then follows by monotone convergence property of measures that

$$\mu(\{|f(x) > 0|\}) = \mu\left(\bigcup_n E_n\right) = \lim_{n\to\infty}\mu(E_n) = \lim_{n\to\infty}0 = 0.$$

This completes the proof. □

Example 18.3.6.29. The statement of Egoroff's theorem depends crucially on the fact that each function in the sequence $\{f_n\}$ is measurable. Indeed, we show by the way of an example that the conclusion of Egoroff's theorem is not true when f_n's are not measurable.

We wish to show that the statement of Egoroff's theorem fails if we drop the condition that functions be measurable.

Consider the measure space $(\mathbb{Z}, \mathcal{A}, \mu)$ where $\mathcal{A} = \{\emptyset, \mathbb{Z}, 2\mathbb{Z}, \mathbb{Z} \setminus 2\mathbb{Z}\}$ and $\mu(\emptyset) = 0 = \mu(2\mathbb{Z})$, $\mu(\mathbb{Z}) = 1 = \mu(\mathbb{Z} \setminus 2\mathbb{Z})$. Consider the functions $f_n : (\mathbb{Z}, \mathcal{A}, \mu) \to \mathbb{R}$ where $\mathbb{R}$ has the Borel measure, given by

$$f_n(k) = \frac{k}{n}$$

for all $k \in \mathbb{Z}$. Observe that $\{f_n\}$ pointwise converges to the constant 0 function at all points of $\mathbb{Z}$. Further note that f_n is not measurable as $f_n^{-1}(\{k/n\}) = \{k\}$ is not a measurable set in $\mathcal{A}$ but $\{k/n\}$ is Borel measurable.

To show that this is a counterexample, it would suffice to show that there exists an $\epsilon_0 > 0$ such that for all measurable sets $E \in \mathcal{A}$, either $\mu(E^c) \geq \epsilon_0$ or f_n does not converges uniformly to 0 on E. We claim that in our situation, $\epsilon_0 = 1/2$ works. Indeed, for $E = \emptyset, 2\mathbb{Z}$, we have $\mu(E^c) = 1 > 1/2$. Thus we reduce to showing that f_n does not converges uniformly on $\mathbb{Z}$ and $\mathbb{Z}\setminus 2\mathbb{Z}$. Indeed, observe that $\sup_{k\in\mathbb{Z}} |f_n(k)| = \sup_{k\in\mathbb{Z}} k/n = \infty$ for each $n \in \mathbb{N}$. As f_n converges uniformly if and only if $\sup_{k\in\mathbb{Z}} |f_n(k)| \to 0$ as $n \to \infty$, therefore we deduce that f_n does not converge uniformly over $\mathbb{Z}$. Similarly, it doesn't converge uniformly over $\mathbb{Z} \setminus 2\mathbb{Z}$.

Lemma 18.3.6.30. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions. If $f = g$ almost everywhere, then $f = g$.*

Proof. Indeed, consider $h = f - g$. Suppose $h \neq 0$, therefore there exists $x_0 \in \mathbb{R}$ such that $h(x_0) \neq 0$. By continuity of h, there exists $\epsilon > 0$ such that $(x_0 - \epsilon, x_0 + \epsilon) \subseteq \{h(x) \neq 0\}$. Hence, $2\epsilon < m(\{h(x) \neq 0\}) = 0$, which yields $0 < 2\epsilon \leq 0$, a contradiction. □

Lemma 18.3.6.31. *Let $(X, \mathcal{S}, \mu)$ be a measure space and $f_n, f : X \to \overline{\mathbb{R}}$ be measurable functions such that $f_n \to f$ pointwise almost everywhere. Then, there exists measurable functions $g_n : X \to \overline{\mathbb{R}}$ such that $f_n = g_n$ almost everywhere and $g_n \to f$ pointwise.*

Proof. Indeed, as f_n converges pointwise to f almost everywhere, therefore the set $E = \{\lim_{n\to\infty} f_n(x) \neq$

$f(x)\}$ is a zero measure set. Consequently, we may define

$$g_n : X \longrightarrow \overline{\mathbb{R}}$$
$$x \longmapsto \begin{cases} f_n(x) & \text{if } x \notin E \\ f(x) & \text{if } x \in E. \end{cases}$$

We then observe that $\{g_n(x) \neq f_n(x)\} = E$, which is of measure zero, hence $g_n = f_n$ almost everywhere. Furthermore, we see that for any $x \in X$,

$$\lim_{n\to\infty} g_n(x) = \begin{cases} \lim_{n\to\infty} f_n(x) = f(x) & \text{if } x \notin E \\ \lim_{n\to\infty} f(x) = f(x) & \text{if } x \in E. \end{cases}$$

Thus, $\lim_{n\to\infty} g_n = f$ pointwise. This completes the proof. □

Example 18.3.6.32. We wish to show that there exists continuous function $f : \mathbb{R} \to \mathbb{R}$ and a Lebesgue measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $g \circ f : \mathbb{R} \to \mathbb{R}$ is not Lebesgue measurable.

While learning about the existence of a non-Borel measurable set, one learns about the existence of a homeomorphism $\varphi : [0,1] \to [0,2]$ such that $m(\varphi(C)) = 1 > 0$ where $C \subseteq [0,1]$ is the Cantor set. Indeed, if $\mathcal{C} : [0,1] \to [0,1]$ denotes the Cantor function, then φ is constructed by defining $\varphi(x) = \mathcal{C}(x) + x$. As, $\mathcal{C}(0) = 0$ and $\mathcal{C}(1) = 1$, therefore $\varphi(0) = 0$ and $\varphi(1) = 2$. Consequently, we may define a continuous function $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) = \begin{cases} x - 1 & \text{if } x > 2 \\ \varphi^{-1}(x) & \text{if } x \in [0,2] \\ x & \text{if } x < 0. \end{cases}$$

Observe that f is continuous as f is obtained by gluing three continuous functions at points where they agree.

As $m(\varphi(C)) = 1 > 0$ for Cantor set C, therefore there exists a non-measurable set $V \subseteq \varphi(C) \subseteq [0,2]$. But since $f(V) = \varphi^{-1}(V) \subseteq \varphi^{-1}(\varphi(C)) = C$ and C is a null set, therefore by completeness of Lebesgue measure, it follows that $f(V)$ is a Lebesgue measurable set. Consequently, we may define $g = \chi_{f(V)} : \mathbb{R} \to \mathbb{R}$, which is Lebesgue measurable as $f(V)$ is Lebesgue measurable. We thus have

$$\mathbb{R} \xrightarrow[\text{cont.}]{f} \mathbb{R} \xrightarrow[\text{Leb. msble}]{g} \mathbb{R} .$$

We claim that $h := g \circ f$ is not Lebesgue measurable. Indeed, observe that $h^{-1}(\{1\}) = (g \circ f)^{-1}(\{1\}) = f^{-1}(g^{-1}(\{1\})) = f^{-1}(f(V))$. But as f restricted to $[0,2]$ is a homeomorphism from $[0,2]$ to $[0,1]$ because on $[0,2]$, f is equal to φ^{-1}, hence $f^{-1}(f(V)) = V$. Hence $h^{-1}(\{1\}) = V$, where $\{1\}$ is measurable but $V \subseteq [0,2]$ is non-measurable. This shows that h is not measurable. This completes the proof.

18.4 Integration of measurable functions

Let's first remind ourselves of the basic definition of a Riemann Integrable function. If we say that the function $f : \mathbb{R} \to \mathbb{R}$ is Riemann Integrable, then the integral of f on $[a,b]$, written as $\int_a^b f$, is given by the following two constructions on a partition P of $[a,b]$,

- Lower Sum :
$$L(f,P) = \sum_i m_i(a_i - a_{i-1}) \text{ where } m_i = \inf_{x\in[a_{i-1},a_i]} f(x)$$
- Upper Sum:
$$U(f,P) = \sum_i M_i(a_i - a_{i-1}) \text{ where } M_i = \sup_{x\in[a_{i-1},a_i]} f(x)$$

so that
$$\int_a^b f = L(f,P) = U(f,P).$$
The chain of observations begins now. One can easily write the Lower and Upper Sum as the following simple functions (remember that the partition is finitely many)

$$L(f,P) = \sum_i m_i \lambda\left([a_{i-1},a_i]\right)$$
$$U(f,P) = \sum_i M_i \lambda\left([a_{i-1},a_i]\right)$$

Or, equivalently, we can define a lower step function as follows:

$$\phi_P = \sum_i m_i \chi_{[a_{i-1},a_i]}$$

so that the Riemann integral is simply

$$\int_a^b f(x)dx = \sup_P \int_a^b \phi_P(x)dx$$

where supremum is defined over all partitions. But since, by definition, $\phi_P(x) \le f(x)\ \forall\ x \in \mathbb{R}$, we can alternatively define Riemann integral as

$$\int_a^b f(x)dx = \sup_{\phi_P \le f} \int_a^b \phi_P(x)dx \tag{18.12}$$

where the supremum is defined for all step functions on any partition P.

This definition presented in (18.12) provides the motivation for extending the definition of Integration from Riemann to Lebesgue. In particular, note the definition of ϕ_P, usual measure on the intervals is applied in Riemann's definition. But, since we know that Borel σ-algebra is a proper subset of $\mathcal{M}_{\lambda^*}$, then it just makes sense to replace $a_{i-1} - a_i$ by $\lambda\left([a_{i-1},a_i]\right)$ in the motivation that it might generalize the notion of integration.

18.4.1 Integration of non-negative measurable functions

Definition 18.4.1.1. (**Lebesgue integral of a simple function**) Consider $\phi : \mathbb{R} \to [0,+\infty)$ be a simple function as
$$\phi = \sum_{i=1}^{N} \alpha_i \chi_{E_i} \text{ where } \alpha_i \ge 0 \text{ and } \lambda\left(E_i\right) < +\infty$$

Then, the Lebesgue integral of ϕ is defined as

$$\int \phi dx = \sum_{i=1}^{N} \alpha_i \lambda\left(E_i\right)$$

Definition 18.4.1.2. (**Lebesgue integral of a measurable function**) Suppose $f : \mathbb{R} \to [0, +\infty)$ is a λ-measurable function, then the Lebesgue integral of f is defined as

$$\bigstar \quad \boxed{\int f dx = \sup_{\phi \leq f} \int \phi dx \text{ where } \phi \text{ are the } \textbf{simple} \text{ functions} \leq f.} \quad \bigstar$$

Definition 18.4.1.3. (**Lebesgue integral over a measurable set**) Consider $f : \mathbb{R} \to [0, +\infty)$ to be a λ-measurable function and $E \subseteq \mathbb{R}$ is Lebesgue measurable. Then,

$$\int_E f dx = \int f \cdot \chi_E dx$$

Remark 18.4.1.4. Therefore, the integral of a non-negative measurable function over a measurable set is given by the integral[19] of restriction of f to it and zero otherwise.

Proposition 18.4.1.5. *Consider the two λ-measurable functions $f, g : \mathbb{R} \to [0, +\infty)$ and $\phi : \mathbb{R} \to [0, +\infty)$ be a simple-function, then the Lebesgue integral has the following properties:*

1. *Consider two Lebesgue measurable subsets A and B of $\mathbb{R}$ such that $A \cap B = \Phi$. Then,*

$$\int_{A \cup B} \phi dx = \int_A \phi dx + \int_B \phi dx.$$

2. *For any $\alpha \in \mathbb{R}$,*

$$\int \alpha f dx = \alpha \int f dx.$$

3. *Integration for positive valued measurable functions is therefore distributive:*

$$\int (f+g) dx = \int f dx + \int g dx.$$

4. *If $f(x) \leq g(x)$ holds for all $x \in \mathbb{R}$, then*

$$\int f dx \leq \int g dx.$$

5. *Consider A and B be Lebesgue measurable subsets of $\mathbb{R}$ such that $A \subseteq B$. Then,*

$$\int_A f dx \leq \int_B f dx.$$

[19]From now on, any instance of *integral* should be presupposed by Lebesgue integral, of-course, unless otherwise stated, in this text.

Proof. **Part 1** : Since ϕ is simple, therefore we can write

$$\phi = \sum_{i=1}^{N} \alpha_i \chi_{E_i}.$$

Now, by definition

$$\begin{aligned}
\int_{A\cup B} \phi dx &= \sum_{i=1}^{N} \alpha_i \lambda \left(E_i \cap (A \cup B)\right) \\
&= \sum_{i=1}^{N} \alpha_i \lambda \left((E_i \cap A) \cup (E_i \cap B)\right) \\
&= \sum_{i=1}^{N} \alpha_i \lambda \left(E_i \cap A\right) + \alpha_i \lambda \left(E_i \cap B\right) && \because E_i \cap A \text{ and } E_i \cap B \text{ are disjoint.} \\
&= \int_A \phi dx + \int_B \phi dx
\end{aligned}$$

Part 2 & 3 : Can be seen easily from Theorem 18.4.2.1.

Part 4 : Note that we define

$$\int f dx = \sup_{\phi \le f} \int \phi dx \text{ where } \phi \text{ are simple functions.}$$

Therefore, for any $\phi \le f$, due to given condition $f \le g$, we would have $\phi \le g$. Hence,

$$\int \phi dx \le \int g dx$$

Since this is true for all simple $\phi \le f$, therefore $\sup_{\phi \le f} \int \phi dx \le \int g dx$, proving the result.

Part 5 : Consider the following:

$$\begin{aligned}
\int_A f dx &= \int f \chi_A dx \\
&\le \int f \chi_B dx && \because \chi_A \le \chi_B, \text{ then apply 4.} \\
&= \int_B f dx
\end{aligned}$$

Hence proved. ∎

18.4.2 Monotone Convergence Theorem

This is arguably one of the most important theorem in Integration theory,

Theorem 18.4.2.1. *(**Monotone Convergence Theorem**) Consider a sequence $\{f_n\}$ of $\mathbb{R} \to [0, +\infty)$ of λ-measurable functions which satisfies*

$$f_n(x) \le f_{n+1}(x) \; \forall \; x \in \mathbb{R} \text{ and } n$$

and suppose $\varinjlim_{n\to\infty} f_n$ *exists. Then,*

$$\int \varinjlim_{n\to\infty} f_n = \varinjlim_{n\to\infty} \int f_n.$$

Proof. Since $f_n \leq f_{n+1}$, therefore,

$$\int f_n \leq \int f_{n+1} \leq \int \varinjlim_{n\to\infty} f_n.$$

Hence,

$$\varinjlim_{n\to\infty} \int f_n \leq \int \varinjlim_{n\to\infty} f_n.$$

Therefore we have proved one inequality.

Now to prove the other inequality, consider any simple function $\phi \leq \varinjlim_{n\to\infty} f$. If we can show that $\int f_n \geq \int \phi$ for any $n \in \mathbb{N}$, then we are done. To this goal, consider $\alpha \in (0,1)$. Construct the set[20]

$$\boxed{E_n = \{x \mid f_n(x) \geq \alpha\phi(x)\}}.$$

Clearly,

$$E_n \subseteq E_{n+1} \;\forall\; n \in \mathbb{N}.$$

Now,

$$\int f_n \geq \int_{E_n} f_n \geq \alpha \int_{E_n} \phi. \tag{18.13}$$

Moreover, we can see that

$$\text{Claim 1}: \bigcup_n E_n = \mathbb{R}.$$

This is easy to see as follows:

$$\begin{aligned}
\text{Take } x \in \bigcup_n E_n &\implies x \in E_{i_0} \text{ for some } i_0 \in \mathbb{N}. \\
&\implies x \in \mathbb{R} && \because E_n \text{ are subset} \\
\text{Take } x \in \mathbb{R} &\implies \text{Either (1) } x \in \{x \mid f_n(x) - \alpha\phi(x) \geq 0\} \text{ or (2) } x \in \{x \mid f_n(x) - \alpha\phi(x) < 0\} && \text{for any } n \in \mathbb{N}. \\
&\implies \text{If (1), then } x \in E_n\text{, else if (2), then } \because \phi \leq \varinjlim_{n\to\infty} f_n,\ \exists\, n' \ s.t.\ x \in E_{n'} \\
&\implies x \in \bigcup_n E_n. && \text{Hence Claim 1.}
\end{aligned}$$

Next, we can also see that

$$\text{Claim 2}: \int_{E_n} \phi \longrightarrow \int \phi$$

[20] After reading the proof, it should appear striking to the reader on actually how much the proof depends on this construction. Both the claims in the following page utilizes this construction E_n to full extent! Hence, it is advised (by Instructor) to purse such effective constructions in the problem sheets and your own proofs.

This can be seen by expanding the *simplicity* of ϕ as follows:

$$\begin{aligned}
\varliminf_{n\to\infty} \int_{E_n} \phi &= \sum_{i=1}^{N} a_i \varliminf_{n\to\infty} \lambda\left(A_i \cap E_n\right) & \\
&= \sum_{i=1}^{N} a_i \lambda\left(\bigcup_n A_i \cap E_n\right) & \because \ \{A_i \cap E_n\}_n \text{ is increasing.} \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i \cap \bigcup_n E_n\right) & \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i \cap \mathbb{R}\right) & \text{Claim 1.} \\
&= \sum_{i=1}^{N} a_i \lambda\left(A_i\right) = \int \phi & \text{Hence Claim 2.}
\end{aligned}$$

Finally, take limit in (18.13) to get:

$$\begin{aligned}
\varliminf_{n\to\infty} \int f_n \geq \varliminf_{n\to\infty} \int_{E_n} f_n &\geq \alpha \varliminf_{n\to\infty} \int_{E_n} \phi \\
&= \alpha \int \phi & \text{Claim 2.} \\
\varliminf_{n\to\infty} \int f_n \geq \int \phi & & \because \ 0 < \alpha < 1 \text{ is arbitrary.}
\end{aligned}$$

Hence, **for any** simple function $\phi \leq \varliminf_{n\to\infty} f_n$, we have concluded that $\int \phi \leq \varliminf_{n\to\infty} \int f_n$, hence it must be true that

$$\int \varliminf_{n\to\infty} f_n = \sup_{\phi \leq \varliminf_n f_n} \int \phi \leq \varliminf_{n\to\infty} \int f_n.$$

Combining the converse inequality at the beginning, we hence get the desired result. □

Proposition 18.4.2.2. *Consider a Lebesgue measurable function $f : \mathbb{R} \to [0, +\infty)$. Then,*

$$\int f dx = 0 \iff f \equiv 0 \textit{ almost everywhere.}$$

Proof. **L** $\implies$ **R** : Consider f is a non-negative real-valued function whose integral is zero. Construct the set,

$$E_n = \left\{x \mid f(x) \geq \frac{1}{n}\right\}.$$

In order to show that $f \equiv 0$ almost everywhere, it is hence sufficient to show that $\lambda\left(E_n\right) = 0 \ \forall \ n \in \mathbb{N}$ because it equivalently proves that the measure of the set where f is greater than zero is zero. Now, consider the following function

$$g_n = \frac{1}{n}\chi_{E_n}.$$

Clearly, because g_n is a simple function and

$$\frac{1}{n}\chi_{E_n}(x) = \begin{cases} \frac{1}{n} & \text{if } f(x) \geq \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

which clearly means that $\frac{1}{n}\chi_{E_n} \leq f$, therefore,

$$\begin{aligned} \int f = 0 &= \sup_{\phi \leq f} \int \phi \\ &\geq \int \frac{1}{n}\chi_{E_n} \\ &= \frac{1}{n}\lambda\left(E_n\right) \\ \implies \lambda\left(E_n\right) &= 0 \; \forall \, n \in \mathbb{N}. \end{aligned}$$

R $\implies$ L : If a non-negative real-valued measurable function f is 0 almost everywhere, then for any simple function $\phi \leq f$, ϕ must also be 0 almost everywhere, so that

$$\begin{aligned} \int \phi &= \sum_{i=1}^{N} \alpha_i \lambda\left(E_i\right) \\ &= 0 \end{aligned}$$

Since this is true for any simple $\phi \leq f$, therefore the supremum of all such $\int \phi$ must also be zero, to make $\int f = 0$. □

A simple corollary of the MCT tells us an equivalent story for decreasing sequence of maps where first term is L^1, as compared to the statement of MCT.

Corollary 18.4.2.3. *Let (X, M, μ) be a measure space and let $f_n : X \to \mathbb{R}$ be a sequence of positive measurable maps. Suppose*

1. $\lim_n f_n(x)$ *exists and is equal to $f(x)$ for some measurable $f : X \to \mathbb{R}$,*
2. $f_n(x) \geq f_{n+1}(x)$ *for all $x \in X$ and $n \in \mathbb{N}$,*
3. $f_1(x) \in L^1$.

Then,

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \lim_{n\to\infty} f_n d\mu.$$

Proof. Since $f \leq f_n \leq f_1$, therefore $f \in L^1$. Now, consider the (not necessarily positive!) measurable sequence $g_n = f - f_n$. Since f_n decreases, therefore g_n increases. Now, $\lim_n g_n = 0$ as $\lim_n f_n = f$. Since $0 \in L^1$, therefore Hence, by MCT, we get that $\lim_n \int_X g_n dm = \int_X \lim_n g_n dm$. Expanding it and using the fact that f is in L^1 (so you can cancel $\int_X f dm$ both sides!) gives the desired result. □

Another important result which is of tremendous usability is the fact that Riemann and Lebesgue agree on compact domains(!)

Theorem 18.4.2.4. *(Riemann = Lebesgue on $[a, b]$) Let $[a, b] \subseteq \mathbb{R}$ be a closed bounded interval and f : $[a, b] \to \mathbb{R}$ be a Riemann integrable map. Then, the Riemann integral and Lebesgue integral of f agrees on $[a, b]$. That is,*

$$\int_a^b f(x)dx = \int_{[a,b]} f dm$$

where m is the Lebesgue measure of $\mathbb{R}$.

18.4.3 Fatou's Lemma

Theorem 18.4.3.1. *If $\{f_n\}$ is a sequence of Lebesgue measurable functions from $\mathbb{R}$ to $[0, +\infty)$, then,*

$$\int \liminf_n f_n \leq \liminf_n \int f_n.$$

Proof. We will use Monotone Convergence Theorem to prove this result. Define

$$g_k = \inf_{n \geq k} f_n$$

Therefore $g_k \leq g_{k+1}$ with $g_k \leq f_n \forall\, n \geq k$. Then,

$$\int g_k \leq \int f_n \forall\, n \geq k.$$

This implies that

$$\int g_k \leq \inf_{n \geq k} \int f_n.$$

Now, by MCT,

$$\int \varinjlim_k g_k = \varinjlim_k \int g_k$$

Therefore

$$\begin{aligned} \varinjlim_{k \to \infty} \inf_{n \geq k} \int f_n = \liminf_k \int f_k \geq \varinjlim_{k \to \infty} \int g_k \\ = \int \varinjlim_{k \to \infty} g_k \\ = \int \varinjlim_{k \to \infty} \inf_{n \geq k} f_n \\ = \int \liminf_k f_k \end{aligned}$$

Hence Proved. □

Remark 18.4.3.2. In fact,

Fatou's Lemma $\iff$ **Monotone Convergence Theorem**.

18.4.4 Integration of General Real-Valued measurable Functions

With the notion of integration of non-negative measurable function in place, it's not difficult to see how can one extend the same notion to measurable functions which takes value in the whole real line.

Definition 18.4.4.1. (**Lebesgue integral of a Real-Valued measurable Function**) Consider $f : \mathbb{R} \to \mathbb{R}$ to be a measurable function such that

1. $\int f^+ dx < \infty$, and
2. $\int f^- dx < \infty$.

If the above two conditions are satisfied, then f is called Lebesgue Integrable. Then, the Lebesgue integral of f is defined as

$$\bigstar \boxed{\int f dx = \int f^+ dx - \int f^- dx} \bigstar$$

Remark 18.4.4.2. It's important to note that the integral $\int f dx = \int f^+ dx - \int f^- dx$ is easily defined for any measurable function, but f is called Lebesgue integral only when it's value is finite!

Definition 18.4.4.3. (**Lebesgue integral over a measurable set**) Consider $f : \mathbb{R} \to \mathbb{R}$ is measurable, $f \cdot \chi_E$ is an Lebesgue Integrable function and $E \subseteq \mathbb{R}$ is also measurable. Then,

$$\int_E f dx = \int f \cdot \chi_E dx.$$

Basic properties of general Lebesgue integral

The following properties are direct extensions of Proposition 18.4.1.5 to the bigger class of Lebesgue Integrable functions.

Proposition 18.4.4.4. *Consider $f, g : \mathbb{R} \to \mathbb{R}$ to be Lebesgue Integrable functions. Then,*

1. *For any $\alpha \in \mathbb{R}$, we have:*
$$\int \alpha f dx = \alpha \int f dx.$$
2. *$f + g$ is also Lebesgue Integrable, with*
$$\int (f + g) dx = \int f dx + \int g dx.$$
3. *If $f \equiv 0$ almost everywhere on $\mathbb{R}$, then,*
$$\int f dx = 0.$$
4. *If $f \leq g$ almost everywhere on $\mathbb{R}$, then,*
$$\int f dx \leq \int g dx.$$
5. *If A and B are measurable sets such that $A \cap B = \Phi$, then,*
$$\int_{A \cup B} f dx = \int_A f dx + \int_B f dx.$$

Proof. **S1** : Consider the case that $\alpha \geq 0$. Then,

$$\begin{aligned}(\alpha f)^+ &= \max(\alpha f, 0) = \alpha \max(f, 0) = \alpha f^+ \\ (\alpha f)^- &= -\min(\alpha f, 0) = -\alpha \min(f, 0 = \alpha f^-\end{aligned}$$

and since $\int f^+ dx < \infty$ and $\int f^- dx < \infty$, therefore αf is also Lebesgue Integrable, with the integral given as

$$\begin{aligned}\int \alpha f &= \int \alpha f^+ - \int \alpha f^- \\ &= \alpha \left(\int f^+ - \int f^- \right) \\ &= \alpha \int f\end{aligned}$$

Now consider that $\alpha < 0$, then

$$\begin{aligned}(\alpha f)^+ &= \max(\alpha f, 0) = -|\alpha| \min(f, 0) = |\alpha| f^- \\ (\alpha f)^- &= -\min(\alpha f, 0) = |\alpha| \max(f, 0) = |\alpha| f^+\end{aligned}$$

Hence, αf is again Lebesgue Integrable, with the integral calculated as:

$$\int \alpha f = \int (\alpha f)^+ - \int (\alpha f)^- = |\alpha| \left(\int f^- - \int f^+ \right) = -|\alpha| \int f = \alpha \int f.$$

S2 : First,

$$\begin{aligned}(f+g)^+ &\leq f^+ + g^+ \\ (f+g)^- &\leq f^- + g^-\end{aligned}$$

for all $x \in \mathbb{R}$, so that $f+g$ is Lebesgue Integrable. Now,

$$\begin{aligned}f + g &= (f+g)^+ - (f+g)^- \\ &= f^+ - f^- + g^+ - g^-\end{aligned}$$

therefore,

$$\begin{aligned}(f+g)^+ - (f+g)^- &= f^+ - f^- + g^+ - g^- \\ (f+g)^+ + f^- + g^- &= (f+g)^- + f^+ + g^+ \\ \int (f+g)^+ + f^- + g^- &= \int (f+g)^- + f^+ + g^+ \\ \int (f+g)^+ + \int f^- + \int g^- &= \int (f+g)^- + \int f^+ + \int g^+ \qquad (\because \text{ of Proposition 18.4.1.5, S3.}) \\ \int (f+g)^+ - \int (f+g)^- &= \int f^+ - \int f^- + \int g^+ - \int g^- \\ \int (f+g)^+ - \int (f+g)^- &= \int f + \int g\end{aligned}$$

S3 : Given to us is that $f \equiv 0$ almost everywhere. This means that

$$\{x \in \mathbb{R} \mid f(x) \neq 0\} \text{ is of measure } 0.$$

We can write it equivalently as the union of the following two disjoint sets

$$\begin{aligned}\{x \in \mathbb{R} \mid f(x) \neq 0\} &= \{x \mid f(x) > 0\} \cup \{x \in \mathbb{R} \mid f(x) < 0\}.\\ \lambda\left(\{x \in \mathbb{R} \mid f(x) \neq 0\}\right) &= \lambda\left(\{x \mid f^+(x) > 0\}\right) + \lambda\left(\{x \mid f^-(x) < 0\}\right)\\ &= 0\end{aligned}$$

Since measure is positive valued by definition, therefore, these two have to be individually be zero. That is,

$$\lambda\left(\{x \mid f^+(x) > 0\}\right) = \lambda\left(\{x \mid f^-(x) < 0\}\right) = 0$$

Now, by Proposition 18.4.2.2, we get that

$$\int f^+ = \int f^- = 0$$

which implies that

$$\int f = \int f^+ - \int f^- = 0.$$

S4 : □

Proposition 18.4.4.5. *If $f : \mathbb{R} \to \mathbb{R}$ is a Lebesgue Integrable Function, then,*

$$\left|\int f dx\right| \leq \int |f|\, dx$$

Proof. Simply note the following:

$$\begin{aligned}\left|\int f dx\right| &= \left|\int \left(f^+ - f^-\right) dx\right|\\ &= \left|\int f^+ dx - \int f^- dx\right|\\ &\leq \left|\int f^+ dx\right| + \left|\int f^- dx\right|\\ &= \int f^+ dx + \int f^- dx\\ &= \int \left(f^+ + f^-\right) dx\\ &= \int |f|\, dx.\end{aligned}$$

□

18.4.5 Dominated convergence theorem

Theorem 18.4.5.1. *Let $\{f_n\}$ be a sequence of measurable functions such that there exists a Lebesgue Integrable function g which satisfies*

$$|f_n| \leq g \; \forall \; n.$$

Suppose that the limit $\varprojlim_{n\to\infty} f_n$ exists. Then $\varprojlim_{n\to\infty} f_n$ is Lebesgue Integrable and,

$$\varprojlim_{n\to\infty} \int f_n dx = \int \varprojlim_{n\to\infty} f_n dx.$$

Proof. Since $\{f_n\}$ is a sequence of measurable functions, therefore, $\varprojlim_{n\to\infty} f_n = f$ is also measurable and $|f|$ is bounded by g. But since g is Lebesgue Integrable, and f_n and f are bounded by g, then each f_n and also f are also Lebesgue Integrable (Trivial to see).
Now, $\{g + f_n\}$ is a sequence of measurable functions. Moreover, since $f_n \leq g$ for all n, therefore $f_n + g \geq 0$, so that $\{f_n + g\}$ is a sequence of non-negative measurable functions.
Now using Fatou's Lemma (Theorem 18.4.3.1), we get,

$$\begin{aligned}
\int \liminf_n (g + f_n) dx &\leq \liminf_n \int (g + f_n) dx \\
\int \left(g + \liminf_n f_n\right) dx &\leq \int g dx + \liminf_n \int f_n dx \\
\int g dx + \int \liminf_n f_n dx &\leq \int g dx + \liminf_n \int f_n dx \\
\int \liminf_n f_n dx &\leq \liminf_n \int f_n dx && \because g \text{ is L.I., so } \int g dx < \infty \\
\int f dx &\leq \liminf_n \int f_n dx && \because \limsup_n x_n = \liminf_n x_n = \varprojlim_n x_n.
\end{aligned}$$

Similarly, since $\{g - f_n\}$ is also a sequence of non-negative measurable functions, therefore we can use Fatou's Lemma to conclude:

$$\begin{aligned}
\int \liminf_n (g - f_n) dx &\leq \liminf_n \int (g - f_n) dx \\
\int \liminf_n (-f_n) dx &\leq \liminf_n \left(-\int f_n dx\right) \\
-\int \limsup_n f_n dx &\leq -\limsup_n \int f_n dx && \because \liminf_n(-x_n) = -\limsup_n x_n. \\
\int f dx &\geq \limsup_n \int f_n dx
\end{aligned}$$

We hence have that

$$\limsup_n \int f_n dx \leq \int f dx \leq \liminf_n \int f_n dx$$

But it is also true that

$$\liminf_n \int f_n dx \leq \limsup_n \int f_n dx.$$

Hence,

$$\liminf_n \int f_n dx = \limsup_n \int f_n dx = \varprojlim_n \int f_n dx = \int f dx$$

Hence proved. ∎

Proposition 18.4.5.2. *Consider $\{f_n\}$ to be a sequence of Lebesgue Integrable functions such that*

$$\sum_{n=1}^{\infty} \int |f_n|\, dx < \infty.$$

Then,

1. *The series*
$$\sum_{n=1}^{\infty} f_n(x) \text{ converges almost everywhere on } \mathbb{R}.$$
2. *The sum*
$$f = \sum_{n=1}^{\infty} f_n \text{ is Lebesgue Integrable.}$$
3. *The integral is*
$$\int \sum_{n=1}^{\infty} f_n dx = \sum_{n=1}^{\infty} \int f_n dx.$$

Proof. **S1**. Denote the following:

$$\varphi = \sum_{n=1}^{\infty} |f_n|$$

Clearly, φ is a non-negative measurable function. Since we know that Lebesgue integral for non-negative functions is countably additive, therefore,

$$\int \varphi dx = \sum_{n=1}^{\infty} \int |f_n|\, dx < \infty.$$

Now, if $\int \varphi < \infty$, then φ is finite almost everywhere on $\mathbb{R}$[21]. Now, since $\sum_{n=1}^{\infty} f_n$ is absolutely convergent almost everywhere (last line), hence it is convergent almost everywhere too on $\mathbb{R}$.

S2. Since $|\sum_{n=1}^{\infty} f_n| \leq \sum_{n=1}^{\infty} |f_n| = \varphi < \infty$ (almost everywhere) and since we can modify the set where φ is not defined (infinite) arbitrarily to make a new function which would be measurable and equal to $\sum_{n=1}^{\infty} f_n$ almost everywhere, therefore $\sum_{n=1}^{\infty} f_n$ would be measurable.

S3. Define

$$\phi_n = \sum_{i=1}^{n} f_i.$$

Clearly, $\phi_n \leq |\sum_{i=1}^{n} f_i| \leq \sum_{i=1}^{n} |f_i| \leq \sum_{i=1}^{\infty} |f_i| = \varphi$. Therefore, ϕ_n is a sequence of measurable functions and $\phi_n \leq \varphi$ where φ is an Integrable function (given). Therefore, using Dominated

[21] For a non-negative measurable function f with given that $\int f dx < \infty$, the set $E = \{x \in \mathbb{R} \mid f(x) = \infty\}$ together with supposition that $\lambda(E) > 0$ is such that; since $\int f dx = \sup_{\phi \leq f} \int \phi$, therefore, if we take $\phi = n\chi_E$ for any $n > 0$ then $n\chi E < f$. Hence $\int f dx > n\lambda(E)$ for all n, so that $\int f dx = \infty$. But it's a contradiction to $\int f dx < \infty$. Therefore $\lambda(E) = 0$.

Convergence Theorem (18.4.5.1), we get,

$$\begin{aligned}
\varinjlim_n \int \phi_n dx &= \int \varinjlim_n \phi_n dx \\
\varinjlim_n \int \sum_{i=1}^{n} f_i dx &= \int \varinjlim_n \sum_{i=1}^{n} f_i dx \\
\varinjlim_n \sum_{i=1}^{n} \int f_i dx &= \int \sum_{i=1}^{\infty} f_i dx && \because \text{Proposition 18.4.4.4, S2} \\
\sum_{i=1}^{\infty} \int f_i dx &= \int \sum_{i=1}^{\infty} f_i dx.
\end{aligned}$$

□

Hence proved.

18.4.6 Applications-II : Integration

We present important applications of the above results, showcasing the power of their usage. At parts here, we are proving results from Folland's exercises.

Lemma 18.4.6.1. *The Lebesgue integral*

$$\int_0^1 \frac{x^p - 1}{\log x} dx$$

exists for $p > -1$.

Proof. The first idea is to break p into cases. In some cases, it is obvious why the above integral exists, in others, we have to work. Denote $f_p(x) = \frac{x^p - 1}{\log x}$.

Act 1 : $p > 0$

In this regime, we can bound the $\int_0^1 f_p(x)dx$ by a fixed quantity. Indeed, since $f_p(x)$ is positive, it will suffice. Observe that

$$\frac{x^p - 1}{\log x} = \frac{1 - x^p}{-\log x} \leq \frac{1}{-\log x}.$$

Now, $-\log x$ can be lower bounded by $1 - ax$ for some $0 < a < 1$ by an easy graphical observation. Hence, continuing above, we get

$$\frac{x^p - 1}{\log x} \leq \frac{1}{1 - ax}.$$

The integral then translates to

$$\int_0^1 \frac{x^p - 1}{\log x} dx \leq -\int_0^1 \frac{1}{1 - ax} dx = -\frac{\log(1 - a)}{a} < \infty.$$

Act 2 : $-1 < p < 0$

This is the regime in which we got to work a bit. First, from some graphical observations about $x^p - 1$ and $\log x$, we conclude the following:

1. $x^p - 1$ is positive and $\log x$ is negative, so that $\frac{x^p-1}{\log x}$ is negative.
2. Viewing $1/\log x$ as an **attenuating factor**[22], we see that $0 < 1/\log x < -1$ for $0 < x < 1/e$ and $1/\log x \leq -1$ for $1/e \leq x < 1$.
3. On $1/e < x < 1$, $\log x > 1 - x$. Hence $1/\log x < 1/1 - x$.

With this, we write our integral as

$$\begin{aligned}\int_0^1 \frac{x^p - 1}{\log x} dx &= \int_0^{1/e} \frac{x^p - 1}{\log x} dx + \int_{1/e}^1 \frac{x^p - 1}{\log x} dx \\ &< \int_0^{1/e} (1 - x^p) dx + \int_0^1 \frac{x^p - 1}{x - 1} dx\end{aligned}$$

Now the first integral is bounded while the second is bounded as the derivative of x^p exists at $x = 1$.

□

Lemma 18.4.6.2. *Let $f : \mathbb{R} \to \mathbb{R} \cup \{\infty, -\infty\}$ be a measurable map with $(\mathbb{R}, M, m)$ be a measure structure on $\mathbb{R}$. If there exists $M > 0$ such that for all $E \in M$ such that $0 < m(E) < \infty$ we have that*

$$\left| \frac{1}{m(E)} \int_E f dm \right| < M,$$

then

$$|f(x)| \leq M \text{ a.e.}.$$

Proof. Let $A = \{x \in \mathbb{R} \mid |f(x)| > M\}$. We can write it as $A = A_+ \cup A_-$ where $A_+ = \{x \in \mathbb{R} \mid f(x) > M\}$ and $A_- = \{x \in \mathbb{R} \mid f(x) < -M\}$. Clearly these are disjoint and covers A. Hence, we wish to show

$$m(A) = m(A_+) + m(A_-) = 0$$

which is equivalent to showing that $m(A_+) = m(A_-) = 0$ as measures are always positive.

Act 1 : $m(A_+) = 0$.

The way A_+ and A_- are defined, it is natural for the next step to be a consideration of integral of f over these. Indeed, we observe that, due to the fact that $f \in L^1$ and $A_+ \subseteq \mathbb{R}$

$$Mm(A_+) = \int_{A_+} M \leq \int_{A_+} |f| \leq \int_{\mathbb{R}} |f| \, dm < \infty.$$

Thus, $\infty > \int_{A_+} f dm \geq Mm(A_+)$. Note we dropped the absolute sign as f is positive on A_+. Hence $m(A_+) \neq \infty$.

[22] we view $1/\log x$ as an attenuating factor instead of $x^p - 1$ as if we remove $1/\log x$, then we would be left with $x^p - 1$, whose integral is easy to find.

Now suppose $0 < m(A_+) < \infty$. Then by hypothesis, we can write

$$\int_{A_+} f dm < Mm(A_+),$$

which is a contradiction. Hence $m(A_+) = 0$.

Act 2 : $m(A_-) = 0$.

Again using $f \in L^1$ and $A_- \subseteq \mathbb{R}$, we get

$$\int_{A_-} |f| \, dm \leq \int_{\mathbb{R}} |f| \, dm < \infty.$$

Since $\left|\int_{A_-} f dm\right| \leq \int_{A_-} |f| \, dm$ and since $\int_{A_-} f dm < \int_{A_-} -M dm = -Mm(A_-)$ so that $\left|\int_{A_-} f dm\right| > Mm(A_-)$, therefore we get

$$Mm(A_-) < \left|\int_{A_-} f dm\right| \leq \int_{A_-} |f| \, dm < \infty.$$

Hence $m(A_-) \neq \infty$. Now with this, if we assume $\infty > m(A_-) > 0$, then by hypothesis, we obtain

$$\left|\int_{A_-} f dm\right| \leq m(A_-)M,$$

which contradicts the above inequality. Hence $m(A_-) = 0$.

□

Lemma 18.4.6.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable map where the domain $\mathbb{R}$ has a measure structure $(\mathbb{R}, M, m)$. If $f \in L^1$ and $f \geq 0$, then for all $E \in M$*

$$\lim_{n\to\infty} \int_E f^{\frac{1}{n}} dm = m(E).$$

Proof. The fundamental observation that one has to make here is that if $y \in [0, \infty)$, then $y^{1/n}$ increases to 1 on $(0, 1]$ and $y^{1/n}$ decreases to 1 on $(1, \infty)$. Indeed, pick any $E \in M$ and define

$$E_{\leq} := E \cap \{x \in \mathbb{R} \mid f(x) \leq 1\}$$
$$E_{>} := E \cap \{x \in \mathbb{R} \mid f(x) > 1\}.$$

We thus have a disjoint measurable cover of E and hence $m(E) = m(E_{\leq}) + m(E_{>})$. Hence we get that

$$\lim_{n\to\infty} \int_E f^{\frac{1}{n}} dm = \lim_{n\to\infty} \int_{E_{\leq}} f^{\frac{1}{n}} dm + \lim_{n\to\infty} \int_{E_{>}} f^{\frac{1}{n}} dm.$$

Now, we have two integrals to consider.

Act 1 : $\lim_{n\to\infty} \int_{E_{\leq}} f^{\frac{1}{n}} dm = m(E_{\leq})$.

Since $f^{\frac{1}{n}}$ is a sequence of positive measurable maps increasing to 1, therefore by MCT, we get that

$$\begin{aligned}\lim_{n\to\infty}\int_{E_{\leq}} f^{\frac{1}{n}}dm &= \int_{E_{\leq}} \lim_{n\to\infty} f^{\frac{1}{n}}dm\\ &= \int_{E_{\leq}} 1dm\\ &= m(E_{\leq}).\end{aligned}$$

Act 2 : $\lim_{n\to\infty}\int_{E_>} f^{\frac{1}{n}}dm = m(E_>)$.

It is this place where we will have to use the fact that $f \in L^1$. Since $f^{\frac{1}{n}}$ is a sequence of positive measurable maps decreasing to 1 where f is L^1. Hence, by Corollary 18.4.2.3 of MCT, we get that

$$\begin{aligned}\lim_{n\to\infty}\int_{E_>} f^{\frac{1}{n}}dm &= \int_{E_>} \lim_{n\to\infty} f^{\frac{1}{n}}dm\\ &= \int_{E_>} 1dm\\ &= m(E_>).\end{aligned}$$

This completes the proof, as we have showed $\lim_{n\to\infty}\int_E f^{\frac{1}{n}}dm = \lim_{n\to\infty}\int_{E_{\leq}} f^{\frac{1}{n}}dm + \lim_{n\to\infty}\int_{E_>} f^{\frac{1}{n}}dm = m(E_{\leq}) + m(E_>) = m(E)$. ∎

Lemma 18.4.6.4. *Let (X, M, m) be a measure space and $f : X \times [a, b] \to \mathbb{C}$ be a function such that $f(x, t) : X \to \mathbb{C}$ is measurable for all $t \in [a, b]$. Let $F(t) := \int_X f(x, t)dm$. Suppose there exists $g \in L^1$ such that*

$$|f(x,t)| \leq |g(x)|\,\forall x \in X$$

for every $t \in [a, b]$. If $\lim_{t\to t_0} f(x, t) = f(x, t_0)$ for every $x \in X$, then

$$\lim_{t\to t_0} F(t) = F(t_0).$$

Proof. Clearly we should use DCT. However, we first need to get a sequence of functions for it. Indeed, since we know that $\lim_{t_n\to t_0} f(x, t) = f(x, t_0)$, thus for any sequence $t_n \to t_0$, we have $\lim_{n\to\infty} f(x, t_n) = f(x, t_0)$. Hence we may define $f_n(x) = f(x, t_n)$ which are by definition measurable. Moreover, we have $|f_n(x)| \leq |g(x)|$ for all $x \in X$ where $g \in L^1$. Hence, by DCT, we obtain

$$\begin{aligned}\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty}\int_X f_n(x)dm &= \int_X \lim_{n\to\infty} f_n(x)dm\\ &= \int_X f(x, t_0)dm\\ &= F(t_0).\end{aligned}$$

Since $t_n \to t_0$ is arbitrary, therefore $\lim_{t\to t_0} F(t) = F(t_0)$. ∎

18.5 The L^p spaces

We now turn into some more abstract formulation for analysis of measurable functions, by analyzing their class formed under certain definitions.

Definition 18.5.0.1. (**L^p norm of a function**) Consider any function f and $p > 0$. The L^p norm of f, denoted $\|f\|_p$, is defined as:

$$\|f\|_p = \left(\int |f|^p\right)^{\frac{1}{p}}$$

Definition 18.5.0.2. (**The L^p Space**) Consider $(X, \mathcal{A}, \mu)$ to be a measure space. Suppose $p > 0$. Then, the class of measurable functions defined as:

$$L^p\,(X, \mathcal{A}, \mu) = \{f : X \to \mathbb{R} \mid \|f\|_p < \infty\}\,/(f \sim g \iff f = g \text{ a.e.}).$$

Moreover, two measurable functions $f, g \in L^p\,(X, \mathcal{A}, \mu)$ are said to be equal if and only if:

$$f = g \text{ almost everywhere on } \mathbb{R}.$$

Remark 18.5.0.3. Note that $L^p\,(X, \mathcal{A}, \mu)$ is just the class of Integrable functions when $p = 1$.

Remark 18.5.0.4. Note carefully the use of word *class* rather than set. It is because that an element of $L^p\,(X, \mathcal{A}, \mu)$ is not a function, but a class of functions identified by the relation $f \sim g$ if and only if $f = g$ almost everywhere. But for out purposes, one can get away by writing $f \in L^p\,(X, \mathcal{A}, \mu)$ to mean that f is measurable and $\|f\|_p < \infty$ so that $|f|^p$ is Integrable.

18.5.1 Algebraic properties of L^p space

We will now see some of the properties of L^p Spaces which reflects it's algebraic nature. In particular, we would prove that L^p is a vector space for any $p > 0$. But proving that $\|\cdot\|_p$ is actually the norm for functions in L^p ($p \geq 1$) would require a lot of construction.

L^p is a vector space

Proposition 18.5.1.1. *Consider $(X, \mathcal{A}, \mu)$ to be a measure space. Then, the L^p space*

$$L^p\,(X, \mathcal{A}, \mu) \text{ is a } \textbf{\textit{Vector Space}}.$$

Proof. First, let's deal with the scalar multiplication. Note that the ground field here is $\mathbb{R}$. For any $a, b \in \mathbb{R}$ and $f, g \in L^p\,(X, \mathcal{A}, \mu)$, we trivially have:

$$\begin{aligned}
(ab)f &= a(bf) \\
1f &= f \\
a(f+g) &= af + ag \\
(a+b)f &= af + bf
\end{aligned}$$

Now, to show that $L^p(X, \mathcal{A}, \mu)$ is an abelian group under addition, the associativity, commutativity, identity (f such that $f = 0$ a.e.) and inverse (for f, $-f$ is the inverse) follows trivially. What remains to be shown is that for $f, g \in L^p(X, \mathcal{A}, \mu)$, $f + g \in L^p(X, \mathcal{A}, \mu)$ too. To see this, note that we know already, that $f + g$ is measurable, what we need to then show is that

$$\text{To Show} : \|f + g\|_p < \infty$$

for any $p > 0$. All we need to show is therefore,

$$\int |f + g|^p < \infty$$

To see this, note:

$$\begin{aligned} |f + g|^p &\leq (|f| + |g|)^p \\ &\leq 2^p \max(|f|^p, |g|^p) \\ &\leq 2^p (|f|^p + |g|^p) \end{aligned}$$

By Proposition 18.4.4.4 S4,

$$\int |f + g|^p \leq 2^p \left(\int |f|^p + \int |g|^p \right) < \infty$$

Therefore, $L^p(X, \mathcal{A}, \mu)$ is a Vector Space. $\square$

norm on L^p vector space

We first see that the norm defined at the beginning is actually not a norm in the case when $p < 1$. Therefore, L^p Vector Space with norm $\|\cdot\|_p$ would make sense only when $p \geq 1$.

Definition 18.5.1.2. (**Norm on a vector space**) Consider a Vector Space $(V, \mathbb{R})$. A norm $\|\cdot\|$ on V is a function

$$\|\cdot\| : (V, \mathbb{R}) \to [0, \infty)$$

satisfying following three conditions:

1. For any $x \in (V, \mathbb{R})$,
$$\|x\| = 0 \iff x = 0_V$$
2. For any $x \in (V, R)$ and $\alpha \in \mathbb{R}$,
$$\|\alpha x\| = |\alpha| \, \|x\|$$
3. For any $x, y \in (V, \mathbb{R})$
$$\|x + y\| \leq \|x\| + \|y\|$$

Now suppose $0 < p < 1$, then, it is simple to see $\|\cdot\|_p$ does not follow Triangle Inequality on $L^p(X, \mathcal{A}, \mu)$. To see this, note that for any $a, b > 0$ and $p \in (0, 1)$, we have:

$$a^p + b^p > (a + b)^p \tag{18.14}$$

This comes naturally from the relation:

$$t^{p-1} > (a + t)^{1-p}$$

and then it's integration.
Using (18.14), we can see that for any two sets $E, F \in \mathcal{A}$ such that $E \cap F = \Phi$, if we write

$$a = \mu(E)^{1/p} = \left(\int |\chi_E|^p\right)^{\frac{1}{p}} = \|\chi_E\|_p$$
$$b = \mu(F)^{1/p} = \left(\int |\chi_F|^p\right)^{\frac{1}{p}} = \|\chi_F\|_p$$

then,

$$\begin{aligned}
\|\chi_E + \chi_F\|_p &= \left(\int |\chi_E + \chi_F|^p\right)^{\frac{1}{p}} \\
&= \left(\int (\chi_E + \chi_F)^p\right)^{\frac{1}{p}} \\
&= \left(\int (\chi_E^p + \chi_F^p)\right)^{\frac{1}{p}} && \because \chi_E \cdot \chi_F = \chi_{E\cap F} = \chi_\Phi = 0 \\
&= \left(\int \chi_E^p + \int \chi_F^p\right)^{\frac{1}{p}} \\
&= (a^p + b^p)^{\frac{1}{p}} \\
&> a + b && \text{Take power } \frac{1}{p} \text{ both sides of Eq. (18.14)} \\
&= a + b \\
&= \|\chi_E\|_p + \|\chi_F\|_p
\end{aligned}$$

Hence, there exists functions in Vector Space $L^p(X, \mathcal{A}, \mu)$ for $p \in (0, 1)$ such that $\|\cdot\|_p$ does not satisfies the Δ-Inequality, hence $\|\cdot\|_p$ is not a norm on the vector space $L^p(X, \mathcal{A}, \mu)$ for $p \in (0, 1)$.

But what about $p \geq 1$? It turns out we need more revelations, in terms of results, to prove that for $p \geq 1$, $\|\cdot\|_p$ is a norm on the vector space $L^p(X, \mathcal{A}, \mu)$. We now discuss those *revelations*.

Lemma 18.5.1.3. *Consider $a \geq 0$, $b \geq 0$ and $0 < \lambda < 1$, then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b.$$

Proof. Consider the convex function e^x. Since it is convex, therefore,

$$\begin{aligned}
a^\lambda b^{1-\lambda} &= e^{\lambda \ln a + (1-\lambda)\ln b} \\
&\leq \lambda e^{\ln a} + (1-\lambda)e^{\ln b} \\
&= \lambda a + (1-\lambda)b
\end{aligned}$$

□

Hölder's inequality

One of the important & frequently used inequalities which would be a stepping stone to show that $\|\cdot\|_p$ is a norm on $L^p(X, \mathcal{A}, \mu)$ for $p \geq 1$.

Theorem 18.5.1.4. *(**Hölder's inequality**) Consider $1 < p, q < \infty$ such that*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then, for any $f \in L^p(X, \mathcal{A}, \mu)$ and $g \in L^q(X, \mathcal{A}, \mu)$

1.
$$fg \in L^1(X, \mathcal{A}, \mu)$$

2.
$$\int |fg| \le \left(\int |f|^p\right)^{\frac{1}{p}} \cdot \left(\int |g|^q\right)^{\frac{1}{q}}$$

.

OR

$$\|fg\|_1 \le \|f\|_p \cdot \|g\|_q$$

Proof. From the above Lemma, we have that for any $a > 0$ and $b > 0$, the following holds:

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \le \frac{a}{p} + \frac{b}{q}$$

Now, if we set

$$a = \frac{|f|^p}{(\|f\|_p)^p}$$
$$b = \frac{|g|^q}{(\|g\|_q)^q}$$

and then use the inequality in above lemma, we get:

$$\frac{|f|\,|g|}{\|f\|_p \|g\|_q} \le \frac{1}{p} \cdot \frac{|f|^p}{(\|f\|_p)^p} + \frac{1}{q} \cdot \frac{|g|^q}{(\|g\|_q)^q}.$$

Now, because $|f|\,|g| = |fg|$, therefore from above inequality, we see that

$$\int |fg| < \infty$$

hence $fg \in L^1(X, \mathcal{A}, \mu)$. Furthermore, since we know that inequality is preserved in Integration, therefore integrating the above inequality leads to the following:

$$\int \frac{|f|\,|g|}{\|f\|_p \|g\|_q} \le \int \frac{1}{p} \cdot \frac{|f|^p}{(\|f\|_p)^p} + \int \frac{1}{q} \cdot \frac{|g|^q}{(\|g\|_q)^q}$$
$$\frac{1}{\|f\|_p \|g\|_q} \int |fg| \le \frac{1}{p\,(\|f\|_p)^p} \int |f|^p + \frac{1}{q\,(\|g\|_q)^q} \int |g|^q$$
$$\frac{\|fg\|_1^1}{\|f\|_p \|g\|_q} \le \frac{1}{p} + \frac{1}{q} = 1$$
$$\|fg\|_1 \le \|f\|_p \|g\|_q$$

Hence proved. □

Remark 18.5.1.5. With Hölder's Inequality, we are one step closer to proving that $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ for any $f, g \in L^p(X, \mathcal{A}, \mu)$ where $1 \leq p < \infty$, to formally make $\|\cdot\|_p$ a norm on the vector space $L^p(X, \mathcal{A}, \mu)$. **This is finally proved by Minkowski's Inequality** which we prove now:

Minkowski's inequality

Theorem 18.5.1.6. *(**Minkowski's inequality**) : Consider any $f, g \in L^p(X, \mathcal{A}, \mu)$ and $1 \leq p < \infty$. Then*

$$\left(\int |f+g|^p\right)^{\frac{1}{p}} \leq \left(\int |f|^p\right)^{\frac{1}{p}} + \left(\int |g|^p\right)^{\frac{1}{p}}$$

OR,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. Since $|f+g| \leq |f| + |g|$, therefore if $p = 1$, then the result follows immediately. Now consider $p > 1$. Moreover, suppose that $q > 1$ is such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Note that this also leads to following equations

$$\begin{aligned} (p-1)q &= p \\ p\left(1 - \frac{1}{q}\right) &= 1 \end{aligned}$$

Now, with this, we can bound $\|f+g\|_p^p$ as follows:

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p \\
&= \int |f+g| \cdot |f+g|^{p-1} \\
&\leq \int (|f|+|g|) \cdot |f+g|^{p-1} \\
&= \int |f| \cdot |f+g|^{p-1} + \int |g| \cdot |f+g|^{p-1} \\
&= \int \left||f| \cdot |f+g|^{p-1}\right| + \int \left||g| \cdot |f+g|^{p-1}\right| \\
&= \int \left|f \cdot (f+g)^{p-1}\right| + \int \left|g \cdot (f+g)^{p-1}\right| \\
&= \|f \cdot (f+g)^{p-1}\|_1 + \|g \cdot (f+g)^{p-1}\|_1 \\
&\leq \|f\|_p \cdot \|(f+g)^{p-1}\|_q + \|g\|_p \cdot \|(f+g)^{p-1}\|_q \quad \text{Hölder's Inequality (18.5.1.4)} \\
&= (\|f\|_p + \|g\|_p) \cdot \|(f+g)^{p-1}\|_q \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int \left|(f+g)^{p-1}\right|^q\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{(p-1)q}\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{p}\right)^{\frac{1}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \left(\int (f+g)^{p}\right)^{\frac{1}{p}\cdot\frac{p}{q}} \\
&= (\|f\|_p + \|g\|_p) \cdot \|f+g\|_p^{\frac{p}{q}} \\
\frac{\|f+g\|_p^p}{\|f+g\|_p^{\frac{p}{q}}} &\leq (\|f\|_p + \|g\|_p) \\
\|f+g\|_p^{p\left(1-\frac{1}{q}\right)} &\leq (\|f\|_p + \|g\|_p) \\
\|f+g\|_p &\leq \|f\|_p + \|g\|_p
\end{aligned}$$

Hence proved. □

Remark 18.5.1.7. ★ Hence, in continuation of our effort to prove that $\|\cdot\|_p$ **is a norm on the vector space** $L^p(X, \mathcal{A}, \mu)$ **for** $1 \leq p < \infty$, we can now satisfactorily say that it is indeed such, especially by Minkowski's Inequality just proved. One also calls a vector space with norm a norm space.

18.5.2 Properties of L^1 maps

We would in this section quickly portray some of the easy properties of L^1-maps which are good to keep in mind. The first tells us that a high schooler's dream of claiming a map to be zero if integral is zero is *almost* true for L^1 maps.

Lemma 18.5.2.1. *Let $f : X \to \mathbb{C}$ be a measurable map where (X, M, m) is a measure space. Suppose $f \in L^1$. Then, $\int_F f dm = 0$ for all $F \in M$ if and only if $f = 0$ almost everywhere.*

Proof. One side is trivial. For the other, we may reduce to the case when f is real valued. Let $A = \{x \in X \mid f^-(x) > 0\}$. As f^- is measurable, therefore $A \in M$. Since $\int_A f dm = 0$, therefore $\int_A f^+ - f^- dm = \int_A f^+ = \int_A f^- dm$. If $x \in A$, then $f^-(x) > 0$, and hence $f^+(x) = 0$. Hence $\int_A f^+ dm = 0$ and hence $\int_A f^- dm = 0$. Since $f^- \geq 0$, therefore $f^- = 0$ almost everywhere. We thus have $\int_X f dm = \int_X f^+ dm = 0$ as $X \in M$. Since $f^+ \geq 0$, therefore $f^+ = 0$ almost everywhere. □

Lemma 18.5.2.2. *Let (X, M, m) be a measure space and $f : X \to \mathbb{R}$ be a measurable map with $f \geq 0$. Then,*

$$m(\{x \in X \mid f(x) = \infty\}) = 0.$$

Proof. This again uses the standard idea of breaking the set which we wish to measure into sets whose bounds on measure is known. Indeed, observe that

$$E := \{f(x) = \infty\} = \bigcap_{n \in \mathbb{N}} \{f(x) > n\} =: \bigcap_{n \in \mathbb{N}} E_n.$$

Moreover, $\{E_n\}$ is decreasing. Thus,

$$m(E) = \lim_{n \to \infty} m(E_n).$$

Now we obtain bound on $m(E_n)$. Indeed,

$$nm(E_n) = \int_{E_n} ndm \leq \int_{E_n} f(x)dm \leq \int_X f(x)dm =: I < \infty.$$

Thus $m(E_n) \leq I/n$. Hence $\lim_{n \to \infty} m(E_n) = 0$. □

18.5.3 Completeness of norm space $L^p(X, \mathcal{A}, \mu)$

We now see that the norm space $L^p(X, \mathcal{A}, \mu)$ is actually a complete metric space on the metric induced by the norm! But before stating the result, let us revisit the definitions of *series*, *Cauchy sequences* & *completeness* for any arbitrary norm space $(V, \mathbb{R}, \|\cdot\|)$.

General definitions and results in normed spaces

Definition 18.5.3.1. (**Convergent sequence**) Let $(V, \mathbb{R}, \|\cdot\|)$ be a norm space and $\{x_n\}$ be a sequence in it. Then $\{x_n\}$ is said to converge to $x \in (V, \mathbb{R}, \|\cdot\|)$ if

$$\|x_n - x\| \longrightarrow 0 \text{ as } n \to \infty.$$

Definition 18.5.3.2. (**Cauchy sequence**) Let $(V, \mathbb{R}, \|\cdot\|)$ be a norm space and $\{x_n\}$ be a sequence in it. Then $\{x_n\}$ is said to be a Cauchy sequence in $(V, \mathbb{R}, \|\cdot\|)$ if

$$\forall \epsilon > 0, \ \exists \ N \in \mathbb{N} \text{ such that } \|x_n - x_m\| < \epsilon \ \forall \ n, m \geq N.$$

Definition 18.5.3.3. (**Complete norm Space** or **Banach Space**) A norm space $(V, \mathbb{R}, \|\cdot\|)$ is called a Complete Metric Space or a Banach Space if

$$\text{Every Cauchy sequence in } (V, \mathbb{R}, \|\cdot\|) \text{ is convergent in } (V, \mathbb{R}, \|\cdot\|).$$

Definition 18.5.3.4. (**Series in a norm Space**) A series in a norm space $(V, \mathbb{R}, \|\cdot\|)$ is defined as

$$\sum_{n=1}^{\infty} x_n \text{ where } x_n \in (V, \mathbb{R}, \|\cdot\|).$$

Definition 18.5.3.5. (**Convergent series in a norm space**) A series $\sum_{n=1}^{\infty} x_n$ in a norm space $(V, \mathbb{R}, \|\cdot\|)$ is said to be convergent if the sequence

$$\{S_n\} \text{ where } S_n = \sum_{i=1}^{n} x_i \text{ is convergent in } (V, \mathbb{R}, \|\cdot\|).$$

Definition 18.5.3.6. (**Absolutely convergent series**) Consider a series $\sum_{n=1}^{\infty} x_n$ in a norm space $(V, \mathbb{R}, \|\cdot\|)$. Then it is called absolutely convergent if and only if

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

We now see the equivalent condition needed for a norm space to become a complete norm space:

Theorem 18.5.3.7. (**Equivalent condition for a Banach space**) *Suppose that* $(V, \mathbb{R}, \|\cdot\|)$ *is a norm space. Then,*

$(V, \mathbb{R}, \|\cdot\|)$ *is a Complete norm Space (or Banach Space)* $\iff$ *Every Absolutely Convergent Series is also Convergent in*

Proof. **L** $\implies$ **R** : Suppose $(V, \mathbb{R}, \|\cdot\|)$ is a Banach Space. Hence any Cauchy sequence in it converges at a point within it. Now, take any Absolutely Convergent series, say,

$$\sum_{n=1}^{\infty} x_n$$

in $(V, \mathbb{R}, \|\cdot\|)$. This means that

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

Now this also means that if we write $S_n = \sum_{i=1}^{n} x_i$, then

$$\begin{aligned}
\|S_n - S_m\| &= \|\sum_{i=1}^{n} x_i - \sum_{i=1}^{m} x_i\| \\
&= \|\sum_{i=n}^{m} x_i\| \\
&\leq \sum_{i=n}^{m} \|x_i\|
\end{aligned}$$

Now since $\sum_{n=1}^{\infty} x_n$ is absolutely convergent, therefore,

$$\sum_{i=n}^{m} \|x_i\| \leq \sum_{i=1}^{\infty} \|x_i\| < \infty \ \forall \ n \leq m \in \mathbb{N}$$

also with the fact that $\sum_{i=n}^{m} \|x_i\| \longrightarrow 0$ as $n, m \to \infty$. Hence,

$$\|S_n - S_m\| < \infty \text{ and } \|S_n - S_m\| \to 0 \text{ as } n, m \to \infty.$$

Hence, $\{S_n\}$ is a Cauchy sequence in $(V, \mathbb{R}, \|\cdot\|)$ and thus is convergent. Therefore, $\sum_{i=1}^{\infty} x_i$ is also convergent.

R $\implies$ **L** : Suppose that $(V, \mathbb{R}, \|\cdot\|)$ is a norm space with given that every absolutely convergent series converges. Since we need to show that $(V, \mathbb{R}, \|\cdot\|)$ is then a Banach space, hence we now consider any arbitrary Cauchy sequence, say, $\{x_n\}$.
Now, construct a new sequence from the taken Cauchy sequence $\{x_n\}$ as $\{y_n\}$ defined by the following:

$$\begin{aligned}
y_1 &= x_{N_1} \text{ where } N_1 \text{ is such that } \|x_n - x_m\| < \frac{1}{2^1} \ \forall \ n, m \geq N_1 \\
y_2 &= x_{N_2} - x_{N_1} \text{ where } N_2 \text{ is such that } \|x_n - x_m\| < \frac{1}{2^2} \ \forall \ n, m \geq N_2 > N_1 \\
\vdots &= \vdots \\
y_k &= x_{N_k} - x_{N_{k-1}} \text{ where } N_k \text{ is such that } \|x_n - x_m\| < \frac{1}{2^k} \ \forall \ n, m \geq N_k > N_{k-1}.
\end{aligned}$$

Now, with $\{y_n\}$ in hand, we see some peculiar properties of it, such as:

$$\sum_{j=1}^{k} y_j = x_{N_k}.$$

and especially, we see that $\sum y_n$ is absolutely convergent(!) as follows:

$$\begin{aligned}
\sum_{j=1}^{\infty} \|y_j\| &\leq \|y_1\| + \sum_{j=1}^{\infty} \|y_j\| \\
&\leq \|x_{N_1}\| + \sum_{j=1}^{\infty} \frac{1}{2^j} \\
&= \|x_{N_1}\| + 1 < \infty \qquad\qquad \because \{x_n\} \text{ is Cauchy, so } \|x_i\| < \infty \forall i.
\end{aligned}$$

Now, since we are given that every absolutely convergent series in $(V, \mathbb{R}, \|\cdot\|)$ converges, therefore $\sum y_n$ also converges in $(V, \mathbb{R}, \|\cdot\|)$. But convergence of a series means convergence of it's sequence of partial sums $S_n = \sum_{i=1}^{n} y_i$ and $S_n = x_{N_n}$ as shown above. Therefore, we have

$$\{x_{N_n}\} \text{ converges in } (V, \mathbb{R}, \|\cdot\|).$$

Since $\{x_{N_n}\}$ converges, therefore, if we suppose $x_{N_n} \to x$, then:

$$\begin{aligned}\|x_n - x\| &= \|x_n - x_{N_n} + x_{N_n} - x\| \\ &\le \|x_n - x_{N_n}\| + \|x_{N_n} - x\| \qquad\qquad 2^{nd} \text{ term} \to 0. \\ &= \|x_n - x_{N_n}\|\end{aligned}$$

Now, we know that $n < N_n$, therefore, $\exists p \in \mathbb{N}$ such that $N_n > n \ge N_{n-p}$. Hence

$$\|x_n - x_{N_n}\| < \frac{1}{2^{n-p}}$$

and as $n \to \infty$, $\|x_n - x_{N_n}\| \to 0$ too. Therefore,

$$\|x_n - x\| \to 0.$$

Hence, $\{x_n\}$ is a convergent sequence, apart from being Cauchy. Since the choice of $\{x_n\}$ was arbitrary, therefore all Cauchy sequences are convergent. Hence $(V, \mathbb{R}, \|\cdot\|)$ is a Complete norm Space or Banach Space. □

$L^p(X, \mathcal{A}, \mu)$ **is a Banach space!**

We now see that $L^p(X, \mathcal{A}, \mu)$ is a Complete norm Space.

Theorem 18.5.3.8. *The normed vector space $L^p(X, \mathcal{A}, \mu)$ for $1 \le p < \infty$ is a Banach Space.*

Proof. From the Theorem 18.5.3.7, we just need to equivalently show that any absolutely convergent series is convergent.
Now consider $\{f_k\}$ in $L^p(X, \mathcal{A}, \mu)$ to be absolutely convergent, so that

$$\sum_{k=1}^{\infty} \|f_k\|_p = B < \infty.$$

Also consider the following sequence:

$$G_n = \sum_{k=1}^{n} |f_k| \text{ and } G = \sum_{k=1}^{\infty} |f_k|.$$

Clearly, for all $n \in \mathbb{N}$ we have

$$\begin{aligned}\|G_n\|_p &= \|\sum_{k=1}^{n} |f_k|\,\| \\ &\le \sum_{k=1}^{n} \|f_k\|_p \le B < \infty.\end{aligned}$$

Also note that $\{G_n\}$ is an increasing sequence of positive-valued measurable functions.
Since $\varinjlim_n G_n$ exists, therefore, by the Monotone convergence theorem (Theorem 18.4.2.1), we

have:

$$\begin{aligned}\int \varprojlim_n G_n^p &= \int G^p \\ &= \varprojlim_n \int G_n^p \\ &\leq B^p\end{aligned}$$

Therefore, we have that $\int G^p$ is finite almost everywhere on $\mathbb{R}$ (just consider the above result that $\int (G^p - \chi_X) = 0$ where $\mu(X) = B^p$.)
Since G^p is finite almost everywhere, therefore G is finite almost everywhere. Hence, we get

$$\begin{aligned}\sum_{k=1}^{\infty} f_k &\leq \sum_{k=1}^{\infty} |f_k| \\ &= G < \infty \text{ almost everywhere.}\end{aligned}$$

Now write

$$F = \sum_{k=1}^{\infty} f_k.$$

Clearly, we have

$$\begin{aligned}|F| &= \left|\sum_{k=1}^{\infty} f_k\right| \\ &\leq \sum_{k=1}^{\infty} |f_k| \\ &= G < \infty\end{aligned}$$

and since f_k are members of the vector space $L^p(X, \mathcal{A}, \mu)$, we also have that $F \in L^p(X, \mathcal{A}, \mu)$.
Now, we see that

$$\begin{aligned}\left|F - \sum_{k=1}^{n} f_k\right|^p &\leq |F| + \left|\sum_{k=1}^{n} f_k\right| \\ &\leq G + G = 2G \\ &\leq (2G)^p \ \because 1 \leq p < \infty. \\ &< \infty\end{aligned}$$

Now since $|F - \sum_{k=1}^{n} f_k|^p < \infty$, hence it is in $L^1(X, \mathcal{A}, \mu)$.
With the above inequality, we see that $|F - \sum_{k=1}^{n}|$ is finite and is absolutely bounded by another measurable function for each n, hence, we can now use the Dominated Convergence Theorem (Theorem 18.4.5.1) to write

$$\begin{aligned}\left(\varprojlim_n \int \left|F - \sum_{k=1}^{n} f_k\right|^p\right)^{\frac{1}{p}} &= \left(\int \varprojlim_n \left|F - \sum_{k=1}^{n} f_k\right|^p\right)^{\frac{1}{p}} \\ \varprojlim_n \|F - \sum_{k=1}^{n} f_k\|_p = 0 & \qquad \text{Note that } F = \sum_{k=1}^{\infty} f_k \in L^p(X, \mathcal{A}, \mu).\end{aligned}$$

Hence, we have

$$\sum_{k=1}^{\infty} f_k = F \in L^p(X, \mathcal{A}, \mu)$$

that is, the absolutely convergent series $\sum_{k=1}^{f_k}$ is also convergent in the same space! Therefore, $L^p(X, \mathcal{A}, \mu)$ is a Banach Space. $\square$

18.6 Product measure

We now turn to product measure spaces. This concept would help us to formalize the notion of double (or higher) integration over the so defined *product measure spaces*. In fact, this concept actually shows the generality of the concept of measure spaces, which we might discuss afterwards.

To introduce formal notion of product measure space, we need a definition based framework to work in, which we learn now:

Definition 18.6.0.1. (**Premeasure**) Consider an **algebra**[23] $\mathcal{A}$ over a set X. The map

$$\mu_0 : \mathcal{A} \longrightarrow [0, +\infty]$$

is called a premeasure if it satisfies:
1. $\mu_0(\Phi) = 0$, and
2. For $A_1, A_2, \dots$ a sequence of disjoint sets from $\mathcal{A}$,

$$\mu_0\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_0(A_i)$$

Definition 18.6.0.2. (**Outer measure by Premeasure**) Consider an algebra $\mathcal{A}$ defined on set X. Suppose $\mu_0 : \mathcal{A} \to [0, +\infty]$ is a premeasure on it. We then define μ_* as the following:

$$\mu_* : \mathcal{A} \longrightarrow [0, +\infty]$$

defined by, for $A \subseteq X$:

$$\mu_*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu_0(E_n) \mid A \subseteq \bigcup_{n=1}^{\infty} E_n \text{ where } \{E_n\} \text{ is a sequence in } \mathcal{A}\right\}$$

Proposition 18.6.0.3. *For an algebra $\mathcal{A}$ on X, μ_* satisfies the following:*
1. *μ_* is an Outer measure.*
2. *The collection of μ_* measurable sets, $\mathcal{M}_{\mu_*}$, is a σ-algebra.*
3. *The σ-algebra generated by algebra $\mathcal{A}$, $\mathcal{B}$, is a proper subset of $\mathcal{M}_{\mu_*}$. That is,*

$$\mathcal{B} \subsetneq \mathcal{M}_{\mu_*}$$

[23] Note that this just an algebra, not a σ-algebra.

Proof. Clearly, $\mu_*(\Phi) = 0$ as, the ∞ sequence of Φ, $\{A_n\}$ where $A_i = \Phi \ \forall \ i$, is such that

$$\Phi \subseteq \bigcup_i A_i$$

and

$$\sum_i \mu_0(A_i) = 0.$$

The next parts has proof similar to one done for Lebesgue Outer measure. □

18.6.1 Some set theoretic concepts

We would need this concepts for later discussions.

Definition 18.6.1.1. (**Elementary class/family**) A collection of sets denoted by $\mathcal{E}$ is called an elementary class if:

1. $\Phi \in \mathcal{E}$,
2. For any $E, F \in \mathcal{E}$, then
$$E \cap F \in \mathcal{E}$$
3. If $E \in \mathcal{E}$, then
$$\exists \{F_n\}_{n=1}^N \text{ where } F_n\text{'s are disjoint and in } \mathcal{E} \text{ such that } E^c = \bigcup_{n=1}^N F_n$$

Proposition 18.6.1.2. *If $\mathcal{E}$ is an elementary class, then the collection $\mathcal{A}$ defined as:*

$$\text{For any } A \in \mathcal{A}, \exists \{E_n\}_{n=1}^N \text{ where } E_n\text{'s are disjoint and in } \mathcal{E} \text{ such that } A = \bigcup_{n=1}^N E_n$$

is an Algebra.

Definition 18.6.1.3. (**Monotone class**) A collection of subsets of a set X denoted $\mathcal{C} \subseteq \mathcal{P}(X)$ is called a monotone class if:

1. For if $\{E_n\}$ is a sequence of monotonically increasing sets from $\mathcal{C}$, that is,
$$E_1 \subseteq E_2 \subseteq \dots$$
then,
$$\bigcup_{n=1}^{\infty} \in \mathcal{C}.$$
2. For if $\{E_n\}$ is a sequence of monotonically decreasing sets from $\mathcal{C}$, that is,
$$E_1 \supseteq E_2 \supseteq \dots$$
then,
$$\bigcap_{n=1}^{\infty} \in \mathcal{C}.$$

Proposition 18.6.1.4. *Consider a family of monotone class given as* $\{\mathcal{C}_n\}$. *Then,*

$$\bigcap_n \mathcal{C}_n \text{ is a monotone class.}$$

Proof. Take any sequence of sets $\{I_n\}$ from $\bigcap_{n=1}^{\infty} \mathcal{C}_n$ such that they are monotonically increasing,

$$I_1 \subseteq I_2 \subseteq \dots.$$

Now, because each $I_i \in \mathcal{C}_n \forall\ n$ and it is a monotonically increasing sequence, therefore $\bigcup_{i=1}^{\infty} I_i \in \mathcal{C}_n\ \forall\ n$. Hence,

$$\bigcup_{i=1}^{\infty} I_i \in \bigcap_n \mathcal{C}_n.$$

Similarly, suppose $\{J_n\}$ is a monotonically decreasing sequence of sets from $\bigcap_n \mathcal{C}_n$,

$$J_1 \supseteq J_2 \supseteq \dots$$

Hence, each $J_i \in \mathcal{C}_n\ \forall\ n$. Since each $\mathcal{C}_n$ is a monotone class, therefore, $\bigcap_{i=1}^{\infty} J_i \in \mathcal{C}_n\ \forall\ n$. Hence,

$$\bigcap_{i=1}^{\infty} J_i \in \bigcap_n \mathcal{C}_n.$$

Hence proved. □

Definition 18.6.1.5. (**Generated monotone class**) Consider any $\mathcal{S} \subset \mathcal{P}(X)$. Then, $\mathcal{C}(\mathcal{S})$ is called the monotone class generated by $\mathcal{S}$ if $\mathcal{C}(\mathcal{S})$ is the smallest monotone class containing $\mathcal{S}$.

Proposition 18.6.1.6. *Let* $\mathcal{A}$ *be an Algebra. Suppose*

- $\mathcal{C}(\mathcal{A})$ *is the Monotone Class generated by* $\mathcal{A}$, *and*
- $\mathcal{M}$ *is the* σ*-Algebra generated by* $\mathcal{A}$.

Then,

$$\mathcal{M} = \mathcal{C}(\mathcal{A}).$$

18.6.2 Product measure space

Definition 18.6.2.1. (**Measurable rectangle**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure spaces. Suppose $X \times Y$ is the Cartesian Product of the sets X and Y. Then, $A \times B \subseteq X \times Y$ is called a measurable Rectangle if

$$A \in \mathcal{A} \text{ and } B \in \mathcal{B}$$

Definition 18.6.2.2. (**Elementary rectangles**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are measure spaces. Denote by $\mathcal{K}$ the collection of all measurable Rectangles. Then, we define Elementary Rectangles, $\mathcal{E}$, as the collection :

For any $A \in \mathcal{E}$, $\exists\ \{E_n\}_{n=1}^{N}$ where E_n's are disjoint measurable rectangles in $\mathcal{K}$ such that $A = \bigcup_{n=1}^{N} E_n$.

Remark 18.6.2.3. ★ It is important to note that **elementary rectangles** $\mathcal{E}$ **is an algebra**, due to Proposition 18.6.1.2.

Definition 18.6.2.4. (**Product of measurable spaces**) Denote $\mathcal{A} \times \mathcal{B}$ to be the **σ-Algebra generated by** $\mathcal{E}$. Then,

$$(X \times Y, \mathcal{A} \times \mathcal{B})$$

is the product of measurable Spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$.

Definition 18.6.2.5. (**Product measure space**) Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure Spaces. The product of these two measure Spaces is defined as the following triple:

$$(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$$

where

1. $X \times Y$ is the Cartesian Product of X and Y.
2. $\mathcal{A} \times \mathcal{B}$ is the σ-Algebra generated by Elementary Rectangles of the product $X \times Y$ under respective measure spaces.
3. $\mu \times \nu$ is defined as:
$$(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B)$$
where $A \times B$ is a measurable Rectangle.

Remark 18.6.2.6. ★ Note the following:

- $\mu \times \nu$ defines a premeasure on Elementary Rectangles, $\mathcal{E}$, which is an Algebra.
- With the premeasure $\mu \times \nu$ on $\mathcal{E}$, we then construct the outer measure $\mu_* \times \nu_*$ by premeasure as done in Definition 18.6.0.2.
- As Proposition 18.6.0.3 shows, the collection of $\mu_* \times \nu_*$ measurable sets from $\mathcal{E}$ forms a σ-Algebra, that is, the σ-Algebra generated from all Elementary Rectangles. This is exactly what we did now.

Properties of product measure space

Definition 18.6.2.7. (**x & y sections**) Suppose $E \subseteq X \times Y$. Then we define

1. x-section as all y available in E if x is fixed:
$$E_x = \{y \in Y \mid (x, y) \in E\}$$
2. y-section as all x available in E if y is fixed:
$$E^y = \{x \in X \mid (x, y) \in E\}$$

Definition 18.6.2.8. (**x & y sections of a function**) Suppose f is a function on $X \times Y$. Then,

1. x-Section of f given $x \in X$ is just $f_x(y) = f(x, y)$.
2. y-Section of f given $y \in Y$ is just $f^y(x) = f(x, y)$

Proposition 18.6.2.9. *Suppose $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are two measurable spaces and $E \subseteq \mathcal{A} \times \mathcal{B}$. Then,*

1. *$E_x \in \mathcal{B}\ \forall x \in X$, and*
2. *$E^y \in \mathcal{A}\ \forall y \in Y$.*

That is, each section of a subset of product of measurable spaces, $\mathcal{A} \times \mathcal{B}$, is itself measurable.

Proof. Omitted. □

Proposition 18.6.2.10. *Suppose $f : \mathcal{A} \times \mathcal{B} \longrightarrow \mathbb{R}$ is a $\mathcal{A} \times \mathcal{B}$-measurable function. Then,*

1. *$f_x : \mathcal{B} \longrightarrow \mathbb{R}$ is a $\mathcal{B}$-measurable function $\forall x \in X$.*
2. *$f^y : \mathcal{A} \longrightarrow \mathbb{R}$ is a $\mathcal{A}$-measurable function $\forall y \in Y$.*

That is, each section of a measurable Function on product measurable space is itself a measurable function.

Proof. Trivial, same as Proposition 18.6.2.9. □

18.6.3 The Fubini-Tonelli theorem

This is perhaps the most important result of this course, whose proof can be found in any course book, available on webpage.

Theorem 18.6.3.1. ***(Tonelli's theorem)*** *Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two σ-finite*[24] *measure spaces. Consider an $\mathcal{A} \times \mathcal{B}$-measurable function*

$$f : X \times Y \longrightarrow [0, +\infty].$$

Then,

1. *The function:*
   - *$g : X \to [0, +\infty]$ given by:*
     $$g(x) = \int_Y f_x d\nu$$
     is $\mathcal{A}$-measurable.
   - *$h : Y \to [0, +\infty]$ given by:*
     $$h(y) = \int_X f_y d\mu$$
     is $\mathcal{B}$-measurable.
2. *f satisfies:*

$$\begin{aligned}\int_{X \times Y} f d(\mu \times \nu) &= \int_X \left(\int_Y f_x d\nu \right) d\mu \\ &= \int_Y \left(\int_X f_y d\mu \right) d\nu\end{aligned}$$

Theorem 18.6.3.2. ***(Fubini's theorem)*** *Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measurable spaces. Consider an $\mathcal{A} \times \mathcal{B}$-measurable function which is also $\mu \times \nu$-Integrable given as:*

$$f : X \times Y \longrightarrow [-\infty, +\infty]^{25}.$$

Then,

1. *We have that*
   - *f_x is ν-Integrable **almost everywhere** on Y.*
   - *f_y is μ-Integrable **almost everywhere** on X.*

[24] This means that there are finite $\{A_n\}$ sets in $\mathcal{A}$ with finite measure such that $\bigcup_n A_n = X$. Similarly for $(Y, \mathcal{B}, \nu)$.
[25] **Note the target set here!**

2. *The following relation holds:*

$$\begin{aligned}\int_{X\times Y} fd(\mu\times\nu) &= \int_X\left(\int_Y f_x d\nu\right)d\mu\\ &= \int_Y\left(\int_X f_y d\mu\right)d\nu\end{aligned}$$

18.6.4 Applications-III : Product and Fubini-Tonelli

Lemma 18.6.4.1. *Let (X,Σ_1,μ) and (Y,Σ_2,ν) be two σ-finite measure space with $f\in\mathcal{L}^1(\mu)$ and $g\in\mathcal{L}^1(\nu)$. Then the function $h(x,y)=f(x)g(y)$ is in $\mathcal{L}^1(\mu\times\nu)$ and that*

$$\int_{X\times Y} hd\mu\times\nu = \left(\int_X fd\mu\right)\left(\int_Y gd\nu\right). \tag{1.1}$$

Proof. We first show that h is measurable. Indeed, as $X\times Y\to\mathbb{C}$ given by $(x,y)\mapsto f(x)$ and $X\times Y\to\mathbb{C}$ given by $(x,y)\mapsto g(y)$ are measurable as they are composites $X\times Y\xrightarrow{\pi_1}X\xrightarrow{f}\mathbb{C}$ and $X\times Y\xrightarrow{\pi_2}X\xrightarrow{g}\mathbb{C}$ respectively, where we know that the projection π_i are measurable, therefore their pointwise product $h(x,y)=f(x)g(y)$ is measurable as well. This shows that h is measurable.

Now note that we have $\int_X |f|\,d\mu = M<\infty$ and $\int_Y |g|\,d\nu = N<\infty$. Furthermore, we have $|h|_x = (|f|\,|g|)_x = |f(x)|\,|g|$ and similarly $|h|^y = (|f|\,|g|)^y = |g(y)|\,|f|$. Consequently by Fubini-Tonelli for $L^+(\mu\times\nu)$, we obtain

$$\begin{aligned}\int_{X\times Y}|h|\,d\mu\times\nu &= \int_X\int_Y |h|\,d\nu d\mu\\ &= \int_X\int_Y |f|\,|g|\,d\nu d\mu\\ &= \int_X |f|\left(\int_Y |g|\,d\nu\right)d\mu\\ &= \int_X N\,|f|\,d\mu\\ &= NM<\infty.\end{aligned}$$

Hence, $h\in\mathcal{L}^1(\mu\times\nu)$.

We now wish to show Eq. (1.1). Indeed, as $h\in\mathcal{L}^1(\mu\times\nu)$, therefore by Fubini-Tonelli for $\mathcal{L}^1(\mu\times\nu)$, we obtain

$$\begin{aligned}\int_{X\times Y} hd\mu\times\nu &= \int_X\int_Y h_x d\nu d\mu\\ &= \int_X\int_Y f(x)gd\nu d\mu\\ &= \int_X f(x)\left(\int_Y gd\nu\right)d\mu\\ &= \left(\int_X fd\mu\right)\left(\int_Y gd\nu\right)\end{aligned}$$

as needed. ☐

Example 18.6.4.2. For $X = Y = \mathbb{N}$, $\Sigma_1 = \Sigma_2 = \mathcal{P}(\mathbb{N})$ and $\mu = \nu = \#$ the counting measure, we wish to restate the Fubini-Tonelli theorem in this setting.

First of all, we observe that both the spaces (X, Σ_1, μ) and (Y, Σ_2, ν) are σ-finite as $\mathbb{N}$ can be covered by $\{E_n\}$ where $E_n = \{n\}$ is a finite measure subset. Hence the Fubini-Tonelli applies.

For any measurable $h : X \to \mathbb{C}$, we first claim that the integral $\int_X h d\mu = \sum_n h(n)$. Indeed, we first have by definition

$$\int_X h d\mu = \int_X \Re(h)^+ d\mu - \int_X \Re(h)^- d\mu + i\left(\int_X \Im(h)^+ - \int_X \Im(h)^- d\mu\right)$$

where each $\Re(h)^\pm, \Im(h)^\pm$ are measurable functions $X \to [0, \infty)$. Hence we reduce to assuming h is a non-negative measurable function. In this case, we observe the following. Consider $g_n = \sum_{k=1}^n h(k)\chi_{\{k\}}$. Observe that g_n are increasing and converges to f pointwise. Then by MCT, we have

$$\begin{aligned}
\int_X h d\mu &= \lim_{n\to\infty} \int_X g_n d\mu \\
&= \lim_{n\to\infty} \int_X \sum_{k=1}^n h(k)\chi_{\{k\}} d\mu \\
&= \lim_{n\to\infty} \sum_{k=1}^n \int_X h(k)\chi_{\{k\}} d\mu \\
&= \lim_{n\to\infty} \sum_{k=1}^n h(k) \\
&= \sum_{k=1}^\infty h(k)
\end{aligned}$$

as needed.

Now pick any $h \in L^+(\mu \times \nu)$. We first claim that $\int_{X\times Y} h d\mu \times \nu = \sum_{n,m} h(n,m)$. Indeed, we claim that $\int_{X\times Y} h d\mu\times\nu = \sup\{\int_{X\times Y} \varphi d\mu\times\nu \mid 0 \le \varphi \le h,\ \varphi \text{ is simple}\} = \sup\{\sum_{(n,m)\in F} h(n,m) \mid F \subseteq \mathbb{N}\times\mathbb{N} \text{ is finite}\} = \sum_{n,m} h(n,m)$, as needed. Let $A = \{\int_{X\times Y} \varphi d\mu \times \nu \mid 0 \le \varphi \le h,\ \varphi \text{ is simple}\}$ and $B = \{\sum_{(n,m)\in F} h(n,m) \mid F \subseteq \mathbb{N}\times\mathbb{N} \text{ is finite}\}$. To show the above claim, we need only show that

$$\sup A = \sup B.$$

First suppose that B is not bounded. Then there exists a sequence $b_k \in B$ such that $b_k \to \infty$ as $k \to \infty$. Let $b_k = \sum_{(n,m)\in F_k} h(n,m) \to \infty$ as $k \to \infty$, where F_k are finite sets. Hence, construct $\varphi_k = \sum_{(n,m)\in F_k} h(n,m)\chi_{\{(n,m)\}}$. Clearly, $\varphi_k \in A$ is a simple function below h. As $\int_{X\times Y} \varphi_k d\mu \times \nu = \sum_{(n,m)\in F_k} h(n,m) = b_k$, therefore we get that A is unbounded as well.

Now suppose B is bounded. Then, A is bounded as well because for any simple function $0 \le \varphi \le h$, φ cannot be supported on an infinite cardinality set as otherwise B will be unbounded. Hence both $\sup A$ and $\sup B$ exists and we wish to show that they are equal. Note that the above argument shows that for any simple function $0 \le \varphi \le h$ given by $\varphi = \sum_{k=1}^n a_k \chi_{E_k}$, the integral $\int_{X\times Y} \varphi d\mu \times \nu = \sum_{k=1}^n a_k \#(E_k)$ is finite. Hence for any $\varphi \in A$, there exists a finite set F such that $\int_{X\times Y} \varphi d\mu \times \nu \le \sum_{(n,m)\in F} h(n,m)$. Thus, $\sup A \le \sup B$. Conversely, pick any $\sum_{(n,m)\in F} h(n,m) \in B$ for some finite F. Then, the simple function $\varphi = \sum_{(n,m)\in F} h(n,m)\chi_{\{(n,m)\}} \in A$ is such that

$\int_{X\times Y}\varphi d\mu\times\nu=\sum_{(n,m)\in F}h(n,m)$. Hence $\sup B\leq\sup A$. This completes the proof that integral of h over $X\times Y$ is just the double sum.

Now by Fubini-Tonelli for L^+, we obtain that

$$\begin{aligned}\int_{X\times Y}hd(\mu\times\nu)&=\sum_{n,m}h(n,m)\\&=\int_X\int_Y h_n d\nu d\mu\\&=\int_X\sum_m h(n,m)d\mu\\\text{(by MCT)}&=\sum_m\int_X h(n,m)d\mu\\&=\sum_m\sum_n h(n,m).\end{aligned}$$

Similarly, we also yield by an application of MCT that

$$\begin{aligned}\int_{X\times Y}hd(\mu\times\nu)&=\sum_{n,m}h(n,m)\\&=\int_Y\int_X h^m d\mu d\nu\\&=\sum_n\sum_m h(n,m).\end{aligned}$$

Now suppose $h\in\mathcal{L}^1(\mu\times\nu)$. Then by Fubini-Tonelli, we yield that

$$\begin{aligned}\int_{X\times Y}hd\mu\times\nu&=\sum_{n,m}h(n,m)\\&=\int_X\int_Y h_n d\nu d\mu\\&=\int_X\sum_m h(n,m)d\mu\\\text{(by DCT as each }h^m\in\mathcal{L}^1(\mu)\text{ by Fubini)}&=\sum_m\int_X h(n,m)d\mu\\&=\sum_m\sum_n h(n,m).\end{aligned}$$

Similarly, we yield

$$\int_{X\times Y}hd\mu\times\nu=\sum_n\sum_m h(n,m).$$

Hence, we yield the following two statements from this discussion:

1. *Let $\sum_{n,m}a_{n,m}$ be a double series of non-negative real numbers. Then,*

$$\sum_{n,m}a_{n,m}=\sum_n\sum_m a_{n,m}=\sum_m\sum_n a_{n,m}.$$

2. *Let* $\sum_{n,m} a_{n,m}$ *be a double series of complex numbers such that*

$$\sum_{n,m} |a_{n,m}| < \infty.$$

Then,

$$\sum_{n,m} a_{n,m} = \sum_{n}\sum_{m} a_{n,m} = \sum_{m}\sum_{n} a_{n,m}.$$

This completes the analysis.

Example 18.6.4.3. Let $c \in \mathbb{R}$ and define $f : [0, \infty) \to \mathbb{R}$ a map given by

$$f(x) = \frac{\sin x^2}{x} + \frac{cx}{1+x}.$$

Let $a > 0$. Then we wish to show that

$$\lim_{n\to\infty} \int_0^a f(nx)dx = ac.$$

We claim that $\frac{\sin x}{x}$ is a bounded function over $[0, \infty)$. Indeed, fix $\epsilon > 0$. As $\lim_{x\to 0} \frac{\sin x^2}{x} = 0$, therefore there exists $\delta > 0$ such that for $x \in (0, \delta)$, we have $\left|\frac{\sin x^2}{x}\right| < \epsilon$. Furthermore, for $x \geq \delta$ we have $\left|\frac{\sin x^2}{x}\right| \leq \frac{1}{|x|} \leq \frac{1}{\delta}$. Hence taking $M = \max\{\epsilon, 1/\delta\}$, we see that $\left|\frac{\sin x^2}{x}\right| \leq M$ over $[0, \infty)$. Consequently, over $[0, \infty)$, we have

$$\begin{aligned} |f(x)| &= \left|\frac{\sin x^2}{x} + \frac{cx}{1+x}\right| \\ &\leq |M| + \left|\frac{cx}{1+x}\right| \\ &\leq M + |c| . \end{aligned}$$

Thus, the sequence of measurable functions $|f(nx)|$ is upper bounded by $|g(x)| = M + |c|$ over $[0, a]$, which is $\mathcal{L}^1$ over $[0, a]$. Furthermore, we see that $f(nx) \to c$ over $(0, a]$ pointwise as $n \to \infty$. Hence, by DCT, we obtain

$$\begin{aligned} \lim_{n\to\infty} \int_0^a f(nx)dx &= \int_0^a \lim_{n\to\infty} f(nx)dx \\ &= \int_0^a c dx \\ &= ca \end{aligned}$$

as needed.

Example 18.6.4.4. Let $X = Y = [0, 1]$, $\Sigma_1 = \Sigma_2 = \mathcal{B}_{[0,1]}$ the Borel σ-algebra on $[0, 1]$ and $\mu =$ Lebesgue measure over $[0, 1]$ and $\nu =$ counting measure over $[0, 1]$. We wish to show that Fubini-Tonelli doesn't holds here for the function $\chi_D : X \times Y \to \mathbb{R}$ where $D = \{(x, x) \mid x \in X\}$.

Let us first calculate $\int_{X\times Y}\chi_D d\mu\times\nu$. As χ_D is just a characteristic function, therefore we simply have

$$\int_{X\times Y}\chi_D d\mu\times\nu = \mu\times\nu(D).$$

1. We claim that $\mu\times\nu(D)=\infty$. Indeed, by definition, we have

$$\mu\times\nu(D)=\inf\left\{\sum_n \mu(I_n)\nu(J_n)\mid \bigcup_n I_n\times J_n\supseteq D, I_n\times J_n\in\mathcal{R}\right\}$$

where $\mathcal{R}$ is the elementary family of all rectangles. We claim that for any such cover $D\subseteq\bigcup_n I_n\times J_n$, we have $\sum_n\mu(I_n)\nu(J_n)=\infty$. Indeed, it suffices to show that there is an $n\in\mathbb{N}$ such that $\mu(I_n)\neq 0$ and J_n is infinite. Suppose there is no such n. It then follows that if $\mu(I_n)\neq 0$, then J_n is finite. Further, if $\mu(I_n)=0$, then J_n can be finite or infinite. Let

$$K:=\{n\in\mathbb{N}\mid \mu(I_n)\neq 0\}$$

and

$$L:=\{n\in\mathbb{N}\mid \mu(I_n)=0\}.$$

Consequently, $K\cup L=\mathbb{N}$.

Pick $n\in K$. Then, $\mu(I_n)\neq 0$ and J_n is finite. It follows that $(I_n\times J_n)\cap D$ is atmost a finite set. Thus, $\bigcup_{n\in K} I_n\times J_n$ covers atmost a countable subset of D. Hence, it follows that $\bigcup_{n\in L} I_n\times J_n$ covers an uncountable subset of D. Furthermore,

$$\begin{aligned} V &:= D\setminus\left(\bigcup_{n\in L}(I_n\times J_n)\cap D\right)\\ &= \bigcup_{n\in K}(I_n\times J_n)\cap D \text{ is countable.} \end{aligned} \tag{4.1}$$

For any $n\in\mathbb{N}$, observe that

$$(I_n\times J_n)\cap D=\{(x,x)\in D\mid x\in I_n\cap J_n\}. \tag{4.2}$$

From the preceding remark, it is thus clear that the set $\bigcup_{n\in L}(I_n\times J_n)\cap D=\{(x,x)\in D\mid x\in I_n\cap J_n \text{ for some } n\in L\}$ is uncountable, which further makes $A:=\bigcup_{n\in L} I_n\cap J_n\subseteq[0,1]$ uncountable. We claim that $[0,1]\setminus A$ is countable. Indeed, by (4.1), we first see that

$$\begin{aligned} V &= \{(x,x)\mid x\in I_n\cap J_n \text{ for some } n\in K\}\\ &\cong \bigcup_{n\in K} I_n\cap J_n. \end{aligned}$$

Thus, $\bigcup_{n\in K} I_n\cap J_n$ is countable.

Observe that

$$[0,1]=\left(\bigcup_{n\in K} I_n\cap J_n\right)\cup\left(\bigcup_{n\in L} I_n\cap J_n\right)$$

because $\{I_n \times J_n\}_{n\in\mathbb{N}}$ covers D. Consequently, as A is uncountable, therefore

$$[0,1] \setminus A \subseteq \bigcup_{n\in K} I_n \cap J_n$$

is countable by Eq. (4.3), as required.

As $A \subseteq [0,1]$ is such that $[0,1] \setminus A$ is countable therefore $\mu(A) = 1$. But, $A \subseteq \bigcup_{n\in L} I_n$, therefore $1 = \mu(A) = \sum_{n\in L} m(I_n) = \sum_n 0 = 0$ as I_n for $n \in L$ is of measure 0. Hence we have $1 = \mu(A) \leq 0$, a contradiction. This shows that $\sum_n \mu(I_n)\nu(J_n) = \infty$ for each $\{I_n \times J_n\} \subseteq \mathcal{R}$ such that $\bigcup_n I_n \times J_n \supseteq D$. Thus,

$$\mu \times \nu(D) = \infty.$$

2. We claim that $\int_Y \int_X \chi_D d\mu d\nu = 0$. Indeed, we have

$$\begin{aligned}
\int_Y \int_X (\chi_D)^y \, d\mu d\nu &= \int_Y \int_X \chi_{D^y} d\mu d\nu \\
&= \int_Y \mu(\{(y,y)\}) d\nu \\
&= \int_Y 0 d\nu \\
&= 0,
\end{aligned}$$

as required.

3. We claim that $\int_X \int_Y \chi_D d\nu d\mu = 1$. Indeed, we have

$$\begin{aligned}
\int_X \int_Y (\chi_D)_x d\nu d\mu &= \int_X \int_Y \chi_{D_x} d\nu d\mu \\
&= \int_X \nu(\{(x,x)\}) d\mu \\
&= \int_X 1 d\mu \\
&= \mu(X) \\
&= 1,
\end{aligned}$$

as needed.

Hence, we have shown that for Fubini-Tonelli to work, we require both spaces to be σ-finite (which is not the case here as Y is not σ-finite).

Example 18.6.4.5. We wish to construct an example of a monotone class of subsets of a non-empty set X such that it is not a σ-algebra. Indeed, consider $X = \{1.,2,3\}$. Define $\mathcal{C} := \{\emptyset, \{1\}, X\}$. Then $\mathcal{C}$ is a monotone class as the only non-trivial increasing sequence of sets is $\emptyset \subseteq \{1\}$ and their union is clearly $\{1\}$ which is in $\mathcal{C}$. Furthermore the only non-trivial decreasing sequence is $X \supseteq \{1\}$, whose intersection is $\{1\}$, which is in $\mathcal{C}$. However, $\mathcal{C}$ is not a σ-algebra as $\{1\}^c = \{2,3\} \notin \mathcal{C}$.

Lemma 18.6.4.6. *Let (X, Σ, μ) be a measure space and $f : X \to \mathbb{C}$ be an $\mathcal{L}^1(\mu)$ map. For each $E \in \Sigma$, define*

$$\nu(E) = \int_E f d\mu.$$

1. *If $\mu(E) = 0$, then $\nu(E) = 0$.*
2. *If $\{E_n\} \subseteq \Sigma$ is a disjoint collection, then*

$$\nu\left(\coprod_n E_n\right) = \sum_n \nu(E_n).$$

3. *For all $\epsilon > 0$, there exists $\delta > 0$ such that*

$$\mu(E) < \delta \implies |\nu(E)| < \epsilon.$$

Proof. 1. Note that $\nu(E) = 0$ iff $|\nu(E)| = 0$. Consequently, we see that

$$\begin{aligned} |\nu(E)| = \left|\int_E f d\mu\right| &\leq \int_E |f| \, d\mu \\ &\leq \infty \cdot \int_E d\mu \\ &= \infty \mu(E) \\ &= \infty \cdot 0 = 0, \end{aligned}$$

as needed.

2. Pick $\{E_n\} \subseteq \Sigma$ to be a disjoint collection. Consider the sequence of measurable functions $g_n = f\chi_{\coprod_{k=1}^n E_k}$. Observe that $g_n \to f\chi_{\coprod_{k=1}^\infty E_k}$ pointwise as $n \to \infty$. Furthermore, observe that $|g_n| \leq |f|$ and as $f \in \mathcal{L}^1(\mu)$, therefore we may apply DCT on $\{g_n\}$.

Applying DCT, we yield

$$\begin{aligned} \int_{\coprod_{k=1}^\infty E_k} f d\mu = \int_X f\chi_{\coprod_{k=1}^\infty E_k} d\mu &= \lim_{n\to\infty} \int_X f\chi_{\coprod_{k=1}^n E_k} d\mu \\ &= \lim_{n\to\infty} \int_{\coprod_{k=1}^n E_k} f d\mu \\ &= \lim_{n\to\infty} \sum_{k=1}^n \int_{E_k} f d\mu \\ &= \sum_{k=1}^\infty \int_{E_k} f d\mu \\ &= \sum_{k=1}^\infty \nu(E_k), \end{aligned}$$

as needed.

3. As $f \in \mathcal{L}^1(\mu)$, therefore there exists a sequence of bounded functions $g_n \in \mathcal{L}^1(\mu)$ such that $g_n \to f$ pointwise as $n \to \infty$ and $|g_n| \leq |f|$ over X. Fix $E \in \Sigma$ of finite measure. It follows from DCT applied on g_n over X that

$$\lim_{n\to\infty} \int_E |f - g_n| \, d\mu \leq \lim_{n\to\infty} \int_X |f - g_n| \, d\mu = 0.$$

Fix $\epsilon > 0$. The convergence of above limit yields that there exists $N \in \mathbb{N}$ such that

$$\int_E |f - g_n| \, d\mu < \epsilon/2$$

for all $n \geq N$. Thus, in particular,

$$\int_E |f| - |g_N| \, d\mu \leq \int_E |f - g_N| \, d\mu < \epsilon/2.$$

Now, from above, we yield that

$$\begin{aligned}|\nu(E)| = \left| \int_E f d\mu \right| &\leq \int_E |f| \, d\mu \\ &\leq \epsilon/2 + \int_E |g_N| \, d\mu.\end{aligned}$$

As $|g_n|$ is bounded, therefore let $|g_N| \leq M_n$ for some $M_N \in [0, \infty)$. Consequently,

$$\begin{aligned}|\nu(E)| &\leq \epsilon/2 + \int_E |g_N| \, d\mu \\ &\leq \epsilon/2 + \int_E M_N d\mu \\ &\leq \epsilon/2 + M_N \mu(E).\end{aligned}$$

Hence, letting $\delta = \epsilon/2M_N$, we yield that for any $E \in \Sigma$ such that $\mu(E) < \delta$ we have

$$\begin{aligned}|\nu(E)| &< \epsilon/2 + \epsilon/2 \\ &= \epsilon.\end{aligned}$$

This completes the proof. □

Example 18.6.4.7. Let $X = Y = \mathbb{N}$, $\Sigma_1 = \Sigma_2 = \mathcal{P}(\mathbb{N})$ and $\mu = \nu =$ counting measure. Further, define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ given by

$$f(m,n) = \begin{cases} 1 & \text{if } m = n, \\ -1 & \text{if } m = n+1, \\ 0 & \text{otherwise.} \end{cases}$$

We wish to show that
1. $\int_{X \times Y} |f| \, d(\mu \times \nu) = \infty$,
2. $\int_X \int_Y f d\nu d\mu = 1$,
3. $\int_Y \int_X f d\mu d\nu = 0$.

Before proving, we would first like to show that f is indeed measurable. Indeed, we may write $f = \chi_D - \chi_S$ where $D = \{(m,m) \mid m \in \mathbb{N}\}$ is the diagonal and $S = \{(n+1,n) \mid n \in \mathbb{N}\}$. Both are measurable subsets of $\Sigma_1 \otimes \Sigma_2$ as $D = \bigcup_m \{(m,m)\}$ and $S = \bigcup_n \{(n+1,n)\}$. Note that singletons of $X \times Y$ are measurable as singletons in X and Y are measurable.

1. Observe that $|f| = \chi_D + \chi_S = \chi_{D \amalg S}$ where D and S are two disjoint subsets as defined above. Consequently

$$\begin{aligned}\int_{X\times Y} |f|\, d\mu \times \nu &= \mu \times \nu(D \amalg S) \\ &= \mu \times \nu(D) + \mu \times \nu(S).\end{aligned}$$

We claim that both $\mu \times \nu(D)$ and $\mu \times \nu(S)$ are ∞.

Indeed, for any $\{I_n \times J_n\}$ for $I_n \times J_n \in \mathcal{R}$ rectangles such that $\bigcup_n I_n \times J_n \supseteq D$, we see that if $(I_n \times J_n) \cap D \neq \emptyset$, then $\mu(I_n)\nu(J_n) \geq 1$ as in this case $I_n \cap J_n \neq \emptyset$. As D is an infinite set and $\bigcup_n (I_n \times J_n) \cap D = D$, therefore $\sum_n \mu(I_n)\nu(J_n) \geq \mu \times \nu\left(\bigcup_n I_n \times J_n\right) \geq \mu \times \nu(D) = \infty$. This shows $\mu \times \nu(D) = \infty$.

Similarly, if $\{I_n \times J_n\}$ for $I_n \times J_n \in \mathcal{R}$ is a collection of rectangles such that $\bigcup_n I_n \times J_n \supseteq S$, then for each n for which $(I_n \times J_n) \cap D \neq \emptyset$ we deduce that $\mu(I_n)\nu(J_n) \geq 1$. Hence, as above, we again get that $\sum_n \mu(I_n)\nu(J_n) = \infty$. This proves that $\int_{X\times Y} |f|\, d\mu \times \nu = \infty$.

2. We simply observe that by definition we have $D_m = \{m\}$ and $S_m = \{m-1\}$. Consequently,

$$\begin{aligned}\int_X \int_Y f_m d\nu d\mu &= \int_X \int_Y \chi_{D_m} - \chi_{S_m} d\nu d\mu \\ &= \int_X \nu(D_m) - \nu(S_m) d\mu \\ &= \int_{X\setminus\{1\}} (1-1) d\mu + \int_{\{1\}} (1-0) d\mu \\ &= 1.\end{aligned}$$

3. We simply observe that by definition $D^n = \{n\}$ and $S^n = \{n+1\}$. Consequently,

$$\begin{aligned}\int_Y \int_X f^n d\mu d\nu &= \int_Y \int_X (\chi_{D^n} - \chi_{S^n}) d\mu d\nu \\ &= \int_Y \mu(D^n) - \mu(S^n) d\nu \\ &= \int_Y 1 - 1 d\nu \\ &= 0.\end{aligned}$$

This completes the proof.

18.7 Differentiation

We now study some of the interconnections between integration and differentiation and related notions.

18.7.1 Differentiability

Definition 18.7.1.1. (**Upper/Lower left & Upper/Lower right Derivatives**) Suppose $f : \mathbb{R} \longrightarrow [-\infty, +\infty]$ is a function such that for all $x \in \mathbb{R}$, f is defined on some open interval around x, then we define the following quantities:

- ***Upper Right Derivative*** :

$$D^+ f(x) = \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h}$$

- ***Lower Right Derivative*** :

$$D_+ f(x) = \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h}$$

- ***Upper Left Derivative*** :

$$D^- f(x) = \limsup_{h \to 0^-} \frac{f(x+h) - f(x)}{h}$$

- ***Lower Left Derivative*** :

$$D_- f(x) = \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h}$$

Definition 18.7.1.2. (**Differentiable function**) The function f is said to be differentiable at x if and only if:

$$D^+ f(x) = D_+ f(x) = D^- f(x) = D_- f(x).$$

Hence, a function is said to be differentiable if it is differentiable at all points of it's domain.

18.7.2 Functions of bounded variation

We now study those functions which do not change too erratically over an interval. We already have the notion of differentiability for the same, so we would see connections between such type of functions and there differential character.

Definition 18.7.2.1. (**Variations of a Function**) Suppose we are given a function on an interval

$$f : [a, b] \longrightarrow \mathbb{R}$$

and any partition $\mathcal{P}_{[a,b]} = \{a = x_0, x_1, x_2, \ldots, x_{k-1}, x_k = b\}$ where $x_i < x_{i+1}$. Now, define the following the following three quantities:

$$p_{\mathcal{P}} = \sum_{i=1}^{k} (f(x_i) - f(x_{i-1}))^+$$
$$n_{\mathcal{P}} = \sum_{i=1}^{k} (f(x_i) - f(x_{i-1}))^-$$
$$t_{\mathcal{P}} = p_{\mathcal{P}} + n_{\mathcal{P}} = \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})|$$

where $_{\mathcal{P}}$ denotes the partition over which the sum is defined and it's simple to observe that $p_{\mathcal{P}} - n_{\mathcal{P}} = f(b) - f(a)$. Also, $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.
Then, we finally define the following three quantities:

- ***Positive Variation of f*** :

$$P_f = \sup_{\mathcal{P}} p_{\mathcal{P}}.$$

- ***Negative Variation of f*** :
$$N_f = \sup_{\mathcal{P}} n_{\mathcal{P}}.$$
- ***Total Variation of f*** :
$$T_f = \sup_{\mathcal{P}} t_{\mathcal{P}}.$$

Definition 18.7.2.2. (**Function of bounded variation**) Suppose $f : \mathbb{R} \to \mathbb{R}$ is a given function. Then f is said to be of bounded variation over interval $[a, b]$ if

$$T_f[a, b] = T_f < \infty.$$

The class of functions on a given interval $[a, b]$ which are of bounded variation is denoted by:

$$\mathcal{BV}\left([a, b]\right).$$

So that for any $f \in \mathcal{BV}\left([a, b]\right)$, $T_f < \infty$.

Remark 18.7.2.3. A function f is said to belong to $\mathcal{BV}\left((-\infty, \infty)\right)$ if f belongs to each $\mathcal{BV}\left([a, b]\right)$ for each interval $[a, b]$. Clearly, in this case $T_f(-\infty, \infty) = \sup_{[a,b]} T_f[a, b]$.

Proposition 18.7.2.4. *Suppose $f \in \mathcal{BV}\left([a, b]\right)$. Then,*
1. $f(b) - f(a) = P_f - N_f$.
2. $T_f = P_f + N_f$.

Proof. Take any $f \in \mathcal{BV}\left([a, b]\right)$. Then we have $T_f[a, b] < \infty$. Now, we know that for any partition $\mathcal{P}$ of $[a, b]$, $f(b) - f(a) = p_{\mathcal{P}} - n_{\mathcal{P}}$. Now, take supremum over all partitions of $[a, b]$, both sides of the above, to write:

$$\begin{aligned}
\sup_{\mathcal{P}} \left(f(b) - f(a)\right) &= \sup_{\mathcal{P}} \left(p_{\mathcal{P}} - n_{\mathcal{P}}\right) \\
f(b) - f(a) &= \sup_{\mathcal{P}} p_{\mathcal{P}} - \sup_{\mathcal{P}} n_{\mathcal{P}} \qquad \text{Known result} : \sup_n \left(x_n - y_n\right) = \sup_n x_n - \sup_n y_n. \\
&= P_f - N_f
\end{aligned}$$

For the 2^{nd} part, we have

$$\begin{aligned}
T_f &= \sup_{\mathcal{P}} t_{\mathcal{P}} \\
&= \sup_{\mathcal{P}} (p_{\mathcal{P}} + n_{\mathcal{P}}) \\
&= \sup_{\mathcal{P}} p_{\mathcal{P}} + n_{\mathcal{P}} \qquad \text{Known result} : \sup_n \left(x_n + y_n\right) = \sup_n x_n + \sup_n y_n. \\
&= P_f + N_f
\end{aligned}$$

Hence proved. □

The following theorem is important as it characterizes the functions in $\mathcal{BV}\left([a, b]\right)$.

Proposition 18.7.2.5. *The following result holds:*

$f \in \mathcal{BV}\left([a, b]\right) \iff \exists\, g,\ h$ *which are* ***monotonically increasing*** *and* ***finite*** *on* $[a, b]$, *such that* $f = g - h$.

Proof. **L** $\implies$ **R** : Consider the functions

$$\begin{aligned} g(x) &= P_f[a,x] + f(a) \\ h(x) &= N_f[a,x] \end{aligned}$$

For any $a \leq x_0 \leq x_1 \leq b$, we observe that:

$$\begin{aligned} g(x_0) &= P_f[a,x_0] + f(a) \leq P_f[a,x_1] + f(a) = g(x_1) \\ h(x_0) &= N_f[a,x_0] \leq N_f[a,x_1] = h(x_1) \end{aligned}$$

because we are adding another partitioning point. Hence, g, h are monotonically increasing functions on $[a,b]$. Hence, $g(b) = P_f[a,b] + f(a) < \infty$ as f is of bounded variation, so g is finite. Similarly h is finite on $[a,b]$. Finally, we note that:

$$\begin{aligned} g(x) - h(x) &= P_f[a,x] + f(a) - N_f[a,x] \\ &= P_f[a,x] - N_f[a,x] + f(a) \\ &= f(x) - f(a) + f(a) \\ &= f(x) \end{aligned}$$

Hence proved that if $f \in \mathcal{BV}([a,b])$, then there exists two monotonically increasing, finite functions on $[a,b]$ such that f is their difference.

R $\implies$ **L** : Take any partition $\mathcal{P}[a,b] = a = x_0 < x_1 < x_2 < \cdots < x_k = b$. Now we see that

$$\begin{aligned} t_{\mathcal{P}}^f &= \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \\ &= \sum_{i=1}^{k} |g(x_i) - h(x_i) - g(x_{i-1}) + h(x_{i-1})| \\ &\leq \sum_{i=1}^{k} |g(x_i) - g(x_{i-1})| + \sum_{i=1}^{k} |h(x_i) - h(x_{i-1})| \\ &= t_{\mathcal{P}}^g + t_{\mathcal{P}}^h \\ &< \infty \end{aligned}$$

Hence $T_f = \sup_{\mathcal{P}} t_{\mathcal{P}}^f < \infty$. So $f \in \mathcal{BV}([a,b])$. □

18.7.3 Differentiability of monotone functions & Lebesgue's theorem

Definition 18.7.3.1. (**Vitali covering**) A collection $\mathcal{C}$ of closed, bounded, nondegenerate[26] intervals is said to cover a given set E in the sense of Vitali if:

For any $x \in E$ and any $\epsilon > 0$, $\exists\, I \in \mathcal{C}$ such that $x \in I$ & $\lambda(I) < \epsilon$.

[26] An interval $[a,b]$ is said to be nondegenerate if $a < b$.

Theorem 18.7.3.2. (***The Vitali covering lemma***) *Suppose $E \subset \mathbb{R}$ is of finite outer measure, that is $\lambda^*(E) < \infty$. Also consider a collection $\mathcal{C}$ of closed, bounded intervals which covers E in the sense of Vitali. Then,*

$$\forall\, \epsilon > 0,\ \exists \text{ disjoint \& finite subcollection } \{I_k\}_{k=1}^n \text{ of } \mathcal{C} \text{ such that}$$

$$\lambda^*\left(E \setminus \bigcup_{k=1}^n I_k\right) < \epsilon.$$

The following is a generalization of mean value theorem from Calculus.

Proposition 18.7.3.3. *Let f be an increasing function on a closed, bounded interval $[a,b]$. Then, $\forall\, \alpha > 0$, we have*

$$\lambda^*\left(\{x \in (a,b) \mid D^+f(x) = D^-f(x) \geq \alpha\}\right) \leq \frac{1}{\alpha} \cdot (f(b) - f(a))$$

Proof. We first begin be denoting $E_\alpha = \{x \in (a,b) \mid D^+f(x) = D^-f(x) \geq \alpha\}$. Now, construct the following collection $\mathcal{F}$ of closed and bounded intervals $[c,d]$ for which,

$$f(d) - f(c) \geq \alpha'(d-c)$$

where $0 < \alpha' \leq \alpha$. Now take any $x \in E_\alpha$. We hence see that $D^+f(x) \geq \alpha$. Now for any $\epsilon > 0$, we can construct a closed bounded interval $I = \left[x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}\right]$ for which $\lambda^*(I) = \epsilon$ with $x \in I$. But moreover, we have that

$$\begin{aligned} Df(x) = \frac{f(x+\epsilon/2) - f(x-\epsilon/2)}{\epsilon} &\geq \alpha \\ f(x+\epsilon/2) - f(x-\epsilon/2) &\geq \epsilon\alpha \geq \epsilon\alpha' \end{aligned}$$

Hence the interval $\left[x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}\right] \in \mathcal{F}$. Therefore, ***$\mathcal{F}$ covers E_α in the sense of Vitali***(!)

Now, by Vitali Covering Lemma (Theorem 18.7.3.2), we get that

$$\forall\, \epsilon > 0,\ \exists \text{ finite disjoint } \{I_k\}_{k=1}^n \text{ from } \mathcal{F} \text{ such that } \lambda^*\left(E_\alpha \setminus \bigcup_{k=1}^n I_k\right) < \epsilon$$

Now, observe that

$$E_\alpha \subseteq \bigcup_{k=1}^n I_k \cup \left(E_\alpha \setminus \bigcup_{k=1}^n I_k\right)$$

Hence, by finite sub-additivity of outer measures, we get the following:

$$\begin{aligned} \lambda^*(E_\alpha) &\leq \lambda^*\left(E_\alpha \setminus \bigcup_{k=1}^n\right) + \lambda^*\left(\bigcup_{k=1}^n I_k\right) \\ &< \epsilon + \sum_{k=1}^n \lambda^*(I_k) \\ &\leq \epsilon + \sum_{k=1}^n \frac{f(d_k) - f(c_k)}{\alpha'} && \text{Suppose } I_k = [c_k, d_k]. \end{aligned}$$

Now since f is an increasing function and $I_k \subset [a, b] \ \forall \ k$, therefore we have:

$$\sum_{k=1}^{n} f(d_k) - f(c_k) \leq f(b) - f(a)$$

That is,

$$\lambda^* (E_\alpha) < \epsilon + \frac{1}{\alpha'} \cdot (f(b) - f(a))$$

But since $\epsilon > 0$ and $\alpha' \in (0, \alpha]$ are arbitrary, therefore,

$$\lambda^* (E_\alpha) \leq \frac{1}{\alpha} \cdot (f(b) - f(a))$$

Hence proved. □

Lebesgue's differentiation theorem

This is also one of the most important theorems of this course. This theorem portrays that monotonicity of a function is much better attribute of *niceness* of a function than the usual belief of continuity, because we know example of continuous functions which is not differentiable, that is the **Weierstrass function**. But with this theorem, if we are given a monotone function on an open interval, then it ought to be differentiable almost everywhere on that interval. The same is obviously not true for just continuous functions.

Theorem 18.7.3.4. *(**Lebesgue's Differentiation Theorem**) Suppose f is a **monotone** function on open interval (a, b) to $\mathbb{R}$. Then,*

$$f \text{ is differentiable on } (a, b) \text{ almost everywhere (!)}$$

Corollary 18.7.3.5. *A function f of bounded variation over an interval $[a, b]$ is differentiable almost everywhere in (a, b).*

Proof. Lebesgue's Differentiation Theorem (18.7.3.4) and the fact that any function of bounded variation is a difference of two increasing functions (Proposition 18.7.2.5). □

18.7.4 Integration & differentiation in context

We now learn some relationships between differentiation and integration. But let us begin with the following basic proposition.

Proposition 18.7.4.1. *Let $f : X \to [0, +\infty)$ be a measurable function which is Lebesgue Integrable on a set $E \subseteq X$. Then,*

$$\forall \ \epsilon > 0, \ \exists \ \delta > 0 \ \text{ such that } \forall \ A \subset E \ \text{ with } \ \lambda(A) < \delta \ , \ \int_A f < \epsilon.$$

Proof. Consider the following sequence $\{f_n\}$ of functions:

$$f_n(x) = \begin{cases} f(x) & \text{if } f(x) \leq n \\ n & \text{if } f(x) \geq n \end{cases}$$

Now, we see that $f_n(x) \leq f_{n+1}(x)$ because if $f_n(x) = f(x)$ then $f_n(x) = f(x) \leq n < n+1$ so that $f_{n+1}(x) = f(x)$. Hence $\{f_n\}$ is an increasing sequence. Therefore

$$f_n \longrightarrow f \text{ almost everywhere}$$

Hence, by Monotone Convergence Theorem (18.4.2.1), we get that

$$\int \varinjlim_n f_n = \varinjlim_n \int f_n.$$

Now, observe the following:

$$\int_E f - \varinjlim_n \int f_n = 0$$
$$\varinjlim_n \int_E f - \varinjlim_n \int f_n = 0$$
$$\varinjlim_n \int_E (f - f_n) = 0$$

where we see that

$$(f - f_n)(x) = \begin{cases} 0 & \text{if } f(x) - n \leq 0 \\ f(x) - n & \text{if } f(x) - n \geq 0. \end{cases}$$

From this, we can construct the following sequence of sets:

$$E_n = \{x \in E \mid f(x) - n \geq 0\}.$$

Again, we see that for any $x \in E_n$, we would have $f(x) \geq n > n-1$, so that $x \in E_{n-1}$. Hence $\{E_n\}$ is a decreasing sequence of subsets of E.
We now observe that

$$\int_{E_n} f \geq \int_{E_n} n = n\lambda(E_n).$$

Hence, we get that, for any $n \in \mathbb{N}$,

$$\lambda(E_n) \leq \frac{1}{n} \int_{E_n} f$$

So that, we can choose n corresponding to any $\delta = \epsilon/n$ such that

$$\lambda(E_n) \leq \frac{1}{n} \int_{E_n} f < \delta = \epsilon/n$$

and

$$\int_{E_n} f < n\delta = \epsilon.$$

Hence proved. $\square$

Indefinite integral

The Indefinite integral of a Lebesgue Integrable function forms a sort of *bridge* between Integration and Differentiation.

Definition 18.7.4.2. (**Indefinite integral**) Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function. We then define the indefinite integral of f as the following function:

$$F(x) = \int_a^x f(t)dt$$

Proposition 18.7.4.3. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function. Then,*

1. *$F(x)$ is a **continuous** function on $[a, b]$.*
2. *$F(x)$ is of **bounded variation** on $[a, b]$.*

Proof. **1.** Take any $\epsilon > 0$. By Proposition 18.7.4.1, we get that

$$F(x) = \int_a^x f(t)dt < \epsilon \implies \exists\, \delta > 0 \;\&\; \exists A \subset [a, x] \text{ such that } \lambda^* (A) < \delta.$$

In more precise words, $\forall \epsilon > 0$, $\exists \delta > 0$ such that whenever

$$|a - x_0| = \lambda^* ([a, x_0]) < \delta$$

then we would have

$$\begin{aligned} |F(x_0) - F(a)| &< \epsilon \\ \left| \int_a^{x_0} f(t)dt - \int_a^a f(t)dt \right| &< \epsilon \\ \left| \int_a^{x_0} f(t)dt \right| &< \epsilon \end{aligned}$$

which is just the definition of continuity.

2. Take any partition of $[a, b]$, say, $\mathcal{P}([a, b]) = a = x_0 < x_1 < x_2 < \ldots x_{k-1} < x_k = b$. Now, we see that,

$$\begin{aligned} t_{\mathcal{P}} &= \sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| \\ &= \sum_{i=1}^{k} \left| \int_a^{x_i} f(t)dt - \int_a^{x_{i-1}} f(t)dt \right| \\ &= \sum_{i=1}^{k} \left| \int_{x_{i-1}}^{x_i} f(t)dt \right| \\ &< \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} |f(t)|\, dt \\ &< \infty \end{aligned}$$

where last line follows because f is Lebesgue Integrable. Since our choice of partition $\mathcal{P}$ was arbitrary, therefore $\sup_{\mathcal{P}} t_{\mathcal{P}} < \infty$ hence $F(x)$ is of bounded variation. □

Corollary 18.7.4.4. *The Indefinite integral of a Lebesgue Integrable function is Differentiable almost everywhere.*

Proof. By-product of Lebesgue's Differentiation Theorem (18.7.3.4), or more succinctly, Corollary 18.7.3.5. □

Proposition 18.7.4.5. *Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function and*

$$F(x) = \int_a^{x_0} f(t)dt = 0 \; \forall \; x \in (a, b).$$

Then, $f = 0$ almost everywhere on (a, b).

Proof. Suppose to the contrary that $\exists \; E \subset [a, b]$ such that $f(x) \neq 0 \; \forall \; x \in E$ with $\lambda^* (E) > 0$. By Proposition 18.2.9.2, we get that $\exists \; G \subset E$ which is closed such that $\lambda^* (G) > 0$ and $\lambda^* (E \setminus G) = 0$. Hence $(a, b) \setminus G$ is open. Now consider the integral:

$$\int_G f = \int_{(a,b)} f - \int_{(a,b)\setminus G} f$$

We know that $\int_{(a,b)} f = 0$ as $F(x) = 0 \; \forall \; x \in (a, b)$. In a similar tone, we have $f|_{(a,b)\setminus G} = 0$ almost everywhere because $\lambda^* (E \setminus G) = 0$ and $f = 0$ on $(a, b) \setminus E$ anyways. Therefore, we have $\int_G f = 0$. But $f|_G \neq 0$ by definition of $G \subset E$. Hence we have a contradiction. Therefore such a set E cannot exist. Thus $f = 0$ almost everywhere on (a, b) if $F = 0 \; \forall \; x \in (a, b)$. □

Theorem 18.7.4.6. *Let $[a, b]$ be a finite interval and let $f : [a, b] \longrightarrow \mathbb{R}$ be a Lebesgue Integrable function over it. Then,*

$$F' = f \text{ almost everywhere in } [a, b].$$

Proof. Omitted □

Absolutely continuous functions

This is a more general form of continuity, and since it has connections with indefinite integral, we then learn them here.

Definition 18.7.4.7. (**Absolutely Continuous Function**) A function $f : [a, b] \longrightarrow \mathbb{R}$ is said to be Absolutely Continuous if

$\forall \; \epsilon > 0 \; , \;\; \exists \; \delta > 0$ such that $\forall$ **finite** & **disjoint** collection of open intervals $\{(a_k, b_k)\}_{k=1}^n$ each subset of (a, b) whic

$$\sum_{k=1}^{n} (b_k - a_k) < \delta,$$

also satisfies

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \epsilon.$$

Remark 18.7.4.8. Some straightforward results are:

- Any Absolutely Continuous function is Continuous in usual sense. This follows trivially from their definitions.
- Any Absolutely Continuous function is Uniformly Continuous. This also follows from the definition.

Proposition 18.7.4.9. *Suppose $f : [a,b] \longrightarrow \mathbb{R}$ is an Absolutely Continuous function. Then, f is of Bounded Variation over $[a,b]$.*

Proof. Since f is an absolutely continuous function, therefore, for any fixed $\epsilon > 0$, we can construct the partition of $[a,b]$, say $\mathcal{P}' = a = x_0 < x_1 < \cdots < x_k = b$ such that each (x_{i-1}, x_i) is of length $< \delta$. Clearly, we would then have that

$$\sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| < k\epsilon.$$

Now, consider any arbitrary partition say $\mathcal{P} \equiv a = y_0 < y_1 < \cdots < y_N = b$ of $[a,b]$. Collect the partition points of $\mathcal{P}$ as the open disjoint intervals $\{(y_{i-1}, y_i)\}_{i=1}^{N}$. Then, for each i^{th} interval in this partition, we can further partition it into k_i open disjoint intervals such that each has length $< \delta$. In particular, we would have the following partition of $[y_{i-1}, y_i]$:

$$\left\{(z^i_{j-1}, z^i_j)\right\}_{j=1}^{k_i} \text{ where } z^i_0 = y_{i-1}\ ,\ \ z^i_{k_i} = y_i.$$

Now, note that the variation of f over the $[y_{i-1}, y_i]$ would then be:

$$\begin{aligned}
|f(y_i) - f(y_{i-1})| &= \left|\sum_{j=1}^{k_i} f(z^i_j) - f(z^i_{j-1})\right| \\
&\leq \sum_{j=1}^{k_i} \left|f(z^i_j) - f(z^i_{j-1})\right| \\
&< \sum_{j=1}^{k_i} \epsilon = k_i\epsilon && \because z^i_j - z^i_{j-1} < \delta \text{ by construction}
\end{aligned}$$

Now, the variation over whole of $\mathcal{P}$ would then be:

$$\begin{aligned}
t_{\mathcal{P}} &= \sum_{i=1}^{N} |f(y_i) - f(y_{i-1})| \\
&< \sum_{i=1}^{N} k_i\epsilon \\
&< \infty
\end{aligned}$$

as k_i is finite for all i. Hence proved. □

This theorem relates Indefinite integral of a Lebesgue integral and Absolute Continuity.

Theorem 18.7.4.10. *Suppose $f : [a,b] \longrightarrow \mathbb{R}$ is a Lebesgue Integrable function and it's Indefinite integral is denoted by the function $F(x)$. Then,*

$$F \text{ is an Indefinite integral} \iff F \text{ is Absolutley Continuous.}$$

Proof. Omitted. □

18.8 Signed measures and derivatives

The concept of signed measures is the next generalization that we seek to understand. So far we have only encountered measures on a space which maps subsets to $[0, \infty]$. But what happens when we *increase* the co-domain to the whole $[-\infty, \infty]$? First of all, we can see clearly that a measure shall not map some subsets to $+\infty$ and some other subset to $-\infty$ as we would then have the problem of $\infty - \infty$, and since we are not doing set theory here, hence we would refrain ourselves only to such *signed* measures which either maps to $(-\infty, \infty]$ or $[-\infty, \infty)$ but not both.

Later we would see that having such a notion of *signed* measure actually leads to some very striking results!

Definition 18.8.0.1. (**Signed measure**) Suppose $(X, \mathcal{M})$ is a measurable Space. A function

$$\nu : \mathcal{M} \longrightarrow [-\infty, +\infty) \ \ \textbf{OR} \ \nu : \mathcal{M} \longrightarrow (-\infty, +\infty]$$

is called a Signed measure if it satisfies:

1. ν atmost maps sets either to $+\infty$ or $-\infty$, but not both[27].
2. ν maps null-set to 0:
$$\nu(\Phi) = 0$$
3. ν follows countable additivity:
$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i)$$
where $\{A_i\}$ is any sequence of disjoint sets from $\mathcal{M}$.

Definition 18.8.0.2. (**Positive set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $A \in \mathcal{M}$ is said to be a positive set w.r.t. ν if:

$$\forall\ S \subseteq A \text{ such that } S \in \mathcal{M}\ ,\ \ \nu(S) \geq 0.$$

Definition 18.8.0.3. (**Negative set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $B \in \mathcal{M}$ is said to be a negative set w.r.t. ν if:

$$\forall\ S \subseteq B \text{ such that } S \in \mathcal{M}\ ,\ \ \nu(S) \leq 0.$$

Remark 18.8.0.4. One could alternatively say that a set is a negative set if it is positive w.r.t. $-\nu$.

Definition 18.8.0.5. (**Null set**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then a set $N \in \mathcal{M}$ is said to be a null set w.r.t. ν if

$$N \text{ is both a Positive and Negative set w.r.t. } \nu$$

Proposition 18.8.0.6. *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Let $\{A_i\}$ be a sequence of positive sets w.r.t. ν. Then,*

$$A = \bigcup_i A_i \textit{ is a Positive Set w.r.t. } \nu.$$

[27] Hence the two possible choices for the ν above.

Proof. We know that we can write the sequence $\{A_i\}$ in the following form:

$$\{B_i\} \text{ where } B_i\text{'s are disjoint } \& \ B_i \subseteq A_i.$$

This can be easily be seen by $B_1 = A_1$ and $B_i = A_i \setminus B_{i-1}$. Hence $\{B_i\}$ is a sequence of disjoint positive sets. Moreover, we can see that

$$A = \bigcup_i A_i = \bigcup_i B_i.$$

Now, take any subset $E \subseteq A$, which we can simply write as:

$$\begin{aligned}\nu(E) &= \nu(E \cap A) \\ &= \nu\left(\bigcup_i E \cap B_i\right) \\ &= \sum_i \nu(E \cap B_i) \\ &> 0\end{aligned}$$

because $E \cap A_i \subseteq B_i$ and B_i is a positive set. Hence proved. □

Remark 18.8.0.7. This is clearly also true for negative sets and null sets. That is, countable union of negative (null) sets is also a negative (null) set.

Proposition 18.8.0.8. *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. If $E \in \mathcal{M}$ is such that $\nu(E) \geq 0$, then*

$$\exists\ A \subseteq E \text{ such that } A \text{ is a positive Set w.r.t. } \nu,\ A \in \mathcal{M}\ \&\ \nu(A) > 0.$$

Proof. Written in Diary at 26^{th} September, 2018. Typeset it here when time allows. □

18.8.1 The Hahn decomposition theorem

Theorem 18.8.1.1. ***(Hahn decomposition theorem)*** *Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then,*

$$\exists \text{ positive Set } A \in \mathcal{M} \text{ and negative Set } B \in \mathcal{M} \text{ such that } \boxed{A \cup B = X\ \&\ A \cap B = \Phi}$$

Moreover, any two such pairs (A, B) and (A', B') are unique upto the fact that

$$A \Delta A'\ \&\ B \Delta B' \text{ are } \nu\text{-Null Sets}$$

18.8.2 The Jordan decomposition of a signed measure

We now, in a sense, generalize the Hahn Decomposition Theorem (18.8.1.1), but to the signed measure ν itself. As usual, let's first familiarize ourselves with some definitions.

Definition 18.8.2.1. (**Mutual singularity of signed measures**) Let ν_1 and ν_2 be two measures (NOT signed!) over measurable space $(X, \mathcal{M})$. Then ν_1 and ν_2 are called mutually singular if

$$\exists\ A \in \mathcal{M} \text{ such that } \nu_1(A) = \nu_2(A^c) = 0$$

and is then denoted by:

$$\nu_1 \perp \nu_2.$$

Theorem 18.8.2.2. *(**Jordan decomposition theorem**) Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure on it. Then,*

$$\exists \text{ measures } \nu^+ \ \&\ \nu^- \text{ on } (X, \mathcal{M}) \text{ such that } \boxed{\nu = \nu^+ - \nu^- \ \&\ \nu^+ \perp \nu^-}$$

and such a decomposition of ν is unique.

18.8.3 The Radon-Nikodym theorem

This is one of the final and most important theorems of this course. As we will see, this theorem gives us a notion of the derivative of a signed measure. However, we would not go more deeper into that fact.
As usual, we first introduce some definitions.

Definition 18.8.3.1. (**Total variation of a signed measure**) The total variation of a signed measure ν over some measurable space is defined by

$$|\nu| = \nu^+ + \nu^-$$

where $\nu = \nu^+ - \nu^-$ is the Jordan Decomposition (Theorem 18.8.2.2) of ν.

Remark 18.8.3.2. Since ν^+ and ν^+ are the usual measures on the measurable space, therefore $|\nu|$ is also a usual measure on the same measurable space.

Definition 18.8.3.3. (**σ-finite signed measure**) Suppose ν is a signed measure on measurable space $(X, \mathcal{M})$. Then ν is called σ-Finite if

$$\exists\ \{X_n\}_{n=1}^{\infty} \text{ where } X_i \in \mathcal{M} \text{ and } |\nu(X_i)| < \infty \text{ such that } \bigcup_{n=1}^{\infty} X_n = X$$

Remark 18.8.3.4. ν is σ-Finite $\iff$ $|\nu|$ is σ-Finite.

Definition 18.8.3.5. (**Absolute continuity of usual measures**) Suppose λ and γ are usual measures over a measurable space $(X, \mathcal{M})$. If,

$$\lambda(E) = 0 \text{ for some } E \in \mathcal{M} \implies \gamma(E) = 0$$

always, then γ is said to be absolutely continuous w.r.t. λ. This is denoted by $\gamma \ll \lambda$.

Definition 18.8.3.6. (**Absolute continuity of signed measures**) Suppose μ and ν are signed measures over a measurable space $(X, \mathcal{M})$. If,

$$|\mu|(E) = 0 \text{ for some } E \in \mathcal{M} \implies \nu(E) = 0$$

always, then ν is called absolutely continuous w.r.t. μ. This is denoted by $\nu \ll \mu$.

Theorem 18.8.3.7. *(**Radon-Nikodym theorem**) Suppose $(X, \mathcal{M})$ is a measurable space and λ & γ are two σ-finite measures on it such that $\gamma \ll \lambda$. Then,*

$$\exists \text{ measurable Function w.r.t. } \lambda\ f : X \longrightarrow [0, \infty) \text{ such that}$$

$$\boxed{\gamma(E) = \int_E f d\lambda \ \forall\ E \in \mathcal{M}}$$

Moreover, f is unique up-to almost everywhere equality, that is, if $\gamma(E) = \int_E g d\lambda \ \forall\ E \in \mathcal{M}$, then,

$$f = g \text{ almost everywhere on } X \text{ w.r.t. } \lambda.$$

18.8.4 Applications-IV : Signed spaces

Lemma 18.8.4.1. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two signed measures on it. Then $\nu \ll \mu$ and $\mu \perp \nu$ if and only if $\nu = 0$.*

Proof. ($\Rightarrow$) As $\mu \perp \nu$, therefore there exists a μ-null set A and a ν-null set B such that $A \amalg B = X$. For any measurable set $E \subseteq X$, we have $E = (E \cap A) \amalg (E \cap B)$. As $E \cap A \subseteq A$, therefore $\mu(E \cap A) = 0$. As $\nu \ll \mu$, therefore $\nu(E \cap A) = 0$. Furthermore, since $E \cap B \subseteq B$, therefore $\nu(E \cap B) = 0$. Hence,

$$\begin{aligned}\nu(E) &= \nu(E \cap A) + \nu(E \cap B) \\ &= 0,\end{aligned}$$

as needed.

($\Leftarrow$) As for any measurable set $E \subseteq X$, we have $\nu(E) = 0$, hence $\nu \ll \mu$. Further, as X is now ν-null and $\emptyset$ is μ-null, therefore $X = X \amalg \emptyset$ gives us the required decomposition to claim that $\mu \perp \nu$. □

Lemma 18.8.4.2. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two positive measures on it. The following are equivalent.*

1. *$\nu \perp \mu$,*
2. *there exists a sequence $\{E_n\} \subseteq \mathcal{A}$ such that $\mu(E_n) \to 0$ and $\nu(X \setminus E_n) \to 0$ as $n \to \infty$.*

Proof. (1. $\Rightarrow$ 2.) As $\nu \perp \mu$, therefore there exists a ν-null set A and a μ-null set B such that $X = A \amalg B$. Hence, we may take $E_n = A$ and $X \setminus E_n = B$ for each $n \in \mathbb{N}$. This provides the required sequence.

(2. $\Rightarrow$ 1.) We wish to construct $A, B \subseteq X$ such that $A \amalg B = X$ and A is ν-null and B is μ-null. To construct A and B, we proceed as follows.

We first observe that since $\mu(E_n) \to 0$, therefore there exists a subsequence of $\mu(E_n)$ say $\mu(E_{n_k})$ such that $\sum_k \mu(E_{n_k}) < \infty$. Indeed, this is a consequence of a general result : for any positive sequence $\{a_n\}$ such that $\lim_n a_n = 0$, we have that there exists a subsequence $\{a_{n_k}\}$ such that $\sum_k a_{n_k} < \infty$. Indeed, for each $k \in \mathbb{N}$ there exists an $n_k \in \mathbb{N}$ such that $a_n \leq 1/2^k$ for all $n \geq n_k$. Consequently, we see that $\sum_{k=1}^{\infty} a_{n_k} \leq \sum_{k=1}^{\infty} 1/2^k < \infty$, as required.

We apply the above result to $\{\mu(E_n)\}$ to obtain a subsequence $\{E_{n_k}\}$. We now replace $\{E_n\}$ by

$\{E_{n_k}\}$ so that we may assume $\sum_n \mu(E_n) < \infty$.

Consider the sequence

$$F_n = X \setminus \bigcup_{k=n}^{\infty} E_k$$
$$= \bigcap_{k=n}^{\infty} X \setminus E_k.$$

Observe that F_n is an increasing sequence and that each $F_n \subseteq X \setminus E_n$. Hence,

$$\nu(F_n) \leq \nu(X \setminus E_n). \tag{2.1}$$

Moreover, observe that since $\lim_{n\to\infty}\nu(X \setminus E_n) = 0$, therefore $\lim_{n\to\infty}\nu(F_n) = 0$. Hence, we deduce by monotone property of measures that

$$\lim_{n\to\infty}\nu(F_n) = \nu\left(\bigcup_n F_n\right).$$

Hence, by previous discussion, we further deduce that

$$\lim_{n\to\infty}\nu(F_n) = 0 = \nu\left(\bigcup_n F_n\right).$$

Thus $A := \bigcup_n F_n$ is a ν-null set. It now suffices to show that $X \setminus A$ is a μ-null set.

Observe that $X \setminus A$ can be written as

$$X \setminus A = \bigcap_{n=1}^{\infty} X \setminus F_n$$
$$= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k.$$

We claim that $X \setminus A$ is a μ-null set. Indeed, denote

$$S_n = \bigcup_{k=n}^{\infty} E_k.$$

We wish to show that

$$\mu\left(\bigcap_{n=1}^{\infty} S_n\right) = \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k\right) = 0.$$

Observe that S_n is a decreasing sequence. Furthermore, as $\mu(S_n) \leq \sum_{k=n}^{\infty} \mu(E_k) < \infty$, therefore we may apply the monotone property of measures. Consequently, we yield the following

$$\lim_{n\to\infty}\mu(S_n) = \mu\left(\bigcap_{n=1}^{\infty} S_n\right). \tag{2.2}$$

We now show that $\lim_{n\to\infty}\mu(S_n) = 0$. Indeed, denoting $l = \sum_k \mu(E_k)$ and $l_n = \sum_{k=1}^{n-1}\mu(E_k)$, we first see that $\lim_{n\to\infty} l_n = l$. Now observe that

$$\mu(S_n) \leq \sum_{k=n}^{\infty} \mu(E_k) = l - l_{n-1}$$

where the last equality follows from rearrangement of a positive convergent series. Hence, taking $\lim_{n\to\infty}$, we obtain that

$$\lim_{n\to\infty}\mu(S_n) \leq l - l = 0,$$

that is, $\lim_{n\to\infty}\mu(S_n) = 0$. By Eq. (2.2), we deduce that $X \setminus A = \bigcap_n S_n$ is a μ-null set, as needed. This completes the proof. $\square$

Example 18.8.4.3. We wish to find Lebesgue decomposition of $\nu = m + \delta_0$ where m is the Lebesgue measure and δ_0 is the Dirac delta measure at $0 \in \mathbb{R}$.

Indeed, as $(\mathbb{R}, \mathcal{M})$ is a σ-algebra, therefore the Lebesgue decomposition theorem holds. We see an immediate candidate for Lebesgue decomposition of ν with respect to m as follows:

$$\nu = \nu_a + \nu_s$$

where we set $\nu_a = m$ and $\nu_s = \delta_0$. Indeed, this works as $m \ll m$ holds trivially and $\delta_0 \perp m$ because of the decomposition $\mathbb{R} = \{0\} \amalg (\mathbb{R}^2 \setminus \{0\})$ where we see immediately that $\{0\}$ is m-null and $\mathbb{R}^2 \setminus \{0\}$ is δ_0-null.

Example 18.8.4.4. Let $p(x) = x^2 - 6x + 1$ be a function $\mathbb{R} \to \mathbb{R}$. Consider the signed measure

$$\nu(E) = \int_E p dm$$

on $(\mathbb{R}, \mathcal{M})$.

1. We first wish to show that $(\mathbb{R}, \mathcal{M}, \nu)$ is σ-finite. Indeed, let $X_n = [n, n+1]$. We claim that $-\infty < \nu(X_n) < \infty$ for each $n \in \mathbb{N}$. Now, observe that over X_n, the polynomial is a continuous function supported on a compact interval, hence it achieves a maxima and a minima, say M_n and m_n respectively. Consequently, we have $m_n \leq p \leq M_n$ over X_n.

$$\int_{X_n} m_n dm \leq \int_{X_n} p dm \leq \int_{X_n} M_n dm$$

and thus $-\infty < m_n \leq \nu(X_n) \leq M_n < \infty$ for each n. Hence ν is σ-finite.

2. We wish to find the Hahn-decomposition of $\mathbb{R}$ w.r.t. ν. That is, we wish to find a decomposition $\mathbb{R} = P \amalg N$ such that P is a ν-positive set and N is a ν-negative set.

Observe that $p(x)$ has two real roots $c_1, c_2 \in \mathbb{R}$. Consequently, we see that over $N = [c_1, c_2]$ the polynomial $p(x)$ is negative and hence $\nu(E) = \int_E p dm \leq 0$ for any measurable $E \subseteq N$. Thus N is a negative set. Similarly, define $P = (-\infty, c_1) \cup (c_2, \infty)$. Then observe that $p(x)$ is positive over p, thus $\nu(E) \geq 0$ for any measurable $E \subseteq P$.

3. We now wish to find the Jordan decomposition of ν. Indeed, define $\nu^+(E) := \nu(E \cap P)$ and

$\nu^-(E) := -\nu(E \cap N)$ where $X = P \amalg N$ is the Hahn decomposition. These are positive measures such that $\nu = \nu^+ - \nu^-$. Furthermore, $\nu^+ \perp \nu^-$ as P is ν^--null and N is ν^+-null by construction.

4. We wish to find the Lebesgue decomposition of ν with respect to the Lebesgue measure m. Indeed, we claim that $\nu \ll m$, which will immediately show that the Lebesgue decomposition of ν with respect to m is simply $\nu = \nu + 0$ where $\nu \ll m$ and $0 \perp m$. Indeed, take any measurable set $E \subseteq X$ such that $m(E) = 0$. As p is measurable therefore

$$\nu(E) = \int_E p dm = 0.$$

Hence $\nu \ll m$, completing the proof.

Lemma 18.8.4.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, $\{E_n\}_{n=1}^N \subseteq \mathcal{A}$ and $\{c_n\}_{n=1}^N \subseteq \mathbb{R}_{\geq 0}$. Consider the positive measure*

$$\nu(E) = \sum_{n=1}^N c_n \mu(E \cap E_n)$$

for some fixed $E_n \in \mathcal{A}$. Then,

1. *$\nu \ll \mu$,*
2. *$d\nu/d\mu = \sum_{n=1}^N c_n \chi_{E_n}$.*

Proof. 1. We wish to show that $\nu \ll \mu$. Indeed, pick any $E \in \mathcal{A}$ such that $\mu(E) = 0$. As μ is positive, consequently $\mu(E \cap E_n) = 0$ for each $n = 1, \ldots, N$ as $E \cap E_n \subseteq E$. Hence, we deduce that $\nu(E) = 0$. Thus $\nu \ll \mu$.

2. We now wish to find the Radon-Nikodym derivative $d\nu/d\mu$, which exists as $\nu \ll \mu$. Indeed, this means we need to find a measurable function $f : X \to [0, \infty]$ such that

$$\nu(E) = \int_E f d\mu$$

for each $E \in \mathcal{A}$. We claim that the following simple function

$$f = \sum_{n=1}^N c_n \chi_{E_n}$$

is the required derivative. Indeed, observe that

$$\begin{aligned}
\int_E f d\mu &= \int_E \sum_{n=1}^N c_n \chi_{E_n} d\mu \\
&= \sum_{n=1}^N c_n \int_E \chi_{E_n} d\mu \\
&= \sum_{n=1}^N c_n \int_X \chi_{E_n \cap E} d\mu \\
&= \sum_{n=1}^N c_n \mu(E \cap E_n) \\
&= \nu(E),
\end{aligned}$$

as required. □

Lemma 18.8.4.6. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two positive measures. Suppose $\nu \ll \mu$. Then,*

1. *if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$,*
2. *Assuming both μ and ν are σ-finite, if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$ and*

$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1} \mu\text{-almost everywhere.}$$

Proof. 1. We first wish to show that if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$.

Denote $f = d\nu/d\mu$. Suppose $E \in \mathcal{A}$ is such that $\nu(E) = 0$. Thus $\nu(E) = \int_E f d\mu = 0$. As $f > 0$ μ-almost everywhere, therefore consider the sequence $E_n = \{f(x) \geq 1/n\}$. Clearly, $\bigcup_n E_n = X \setminus N$ as $f > 0$ over X, where $N = \{f(x) = 0\}$ is a μ-null set. Hence, $\bigcup_n E \cap E_n = E \setminus N$. Thus, $\mu(E \setminus N) \leq \sum_n \mu(E \cap E_n)$. Now,

$$\frac{1}{n}\mu(E \cap E_n) \leq \int_{E \cap E_n} f d\mu \leq \int_E f d\mu = 0.$$

Thus, $\mu(E \cap E_n) = 0$ for each $n \in \mathbb{N}$. Hence,

$$\mu(E \setminus N) \leq \sum_n \nu(E \cap E_n) = 0$$

and thus $\mu(E) = \mu(E \cap N) + \mu(E \setminus N) = 0 + 0 = 0$.

2. Assuming both μ and ν are σ-finite, we now wish to show that if the derivative $d\nu/d\mu > 0$ μ-almost everywhere, then $\mu \ll \nu$ and

$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1} \mu\text{-almost everywhere.}$$

We have shown that $\mu \ll \nu$ in the item 1 above. By Radon-Nikodym theorem, we have the derivative $g = d\mu/d\nu$ which is a measurable function $g : X \to [0, \infty]$ such that

$$\mu(E) = \int_E g d\nu.$$

Denote $f = d\nu/d\mu : X \to [0, \infty]$ which is such that

$$\nu(E) = \int_E f d\mu.$$

We are given that $f > 0$ μ-almost everywhere. We wish to show that $g = 1/f$ μ-almost everywhere.

As we have seen that for an L^+ function h, we obtain a positive measure given by $\mu_h = \int_E h d\mu$, therefore we deduce that $\mu = \nu_g$ and $\nu = \mu_f$. Consequently, denoting $N = \{f(x) = 0\}$ to be the μ-null set, we obtain

$$\begin{aligned}
\int_E \frac{1}{f} d\nu &= \int_{E \setminus N} \frac{1}{f} f d\mu + \int_{E \cap N} \frac{1}{f} d\nu \\
&= \int_{E \setminus N} d\mu + 0 \\
&= \mu(E \setminus N).
\end{aligned}$$

As $\mu(E \cap N) = 0$, therefore adding this to above we add

$$\int_E \frac{1}{f} d\nu = \mu(E \setminus N) + \mu(E \cap N) = \mu(E).$$

Thus by almost everywhere uniqueness of Radon-Nikodym derivative of μ w.r.t. ν, we see that $1/f = g$ μ-almost everywhere. $\square$

Lemma 18.8.4.7. *Let $(X, \mathcal{A})$ be a measurable space with μ and ν be two finite positive measures. Suppose*

$$f = \frac{d\nu}{d(\mu + \nu)}.$$

Assume that $1 - f > 0$. Then,

$$\nu(E) = \int_E \frac{f}{1 - f} d\mu,$$

equivalently, that

$$\frac{d\nu}{d\mu} = \frac{f}{1 - f}.$$

Proof. We first show that $g := 1 - f$ is equal to the derivative $d\mu/d(\mu + \nu)$. Observe that $g > 0$. Indeed, for this, we need to show that for any $E \in \mathcal{A}$, we have

$$\mu(E) = \int_E g d(\mu + \nu).$$

To this end, we see that by definition of f and finiteness of μ, ν and thus $\mu + \nu$ as measures, we may deduce

$$\begin{aligned}
\int_E g d(\mu + \nu) &= \int_E (1 - f) d(\mu + \nu) \\
&= \int_E d(\mu + \nu) - \int_E f d(\mu + \nu) \\
&= \mu(E) + \nu(E) - \nu(E) \\
&= \mu(E),
\end{aligned}$$

as required. We may therefore write $\mu = (\mu + \nu)_g$ as the notation introduced in the class for positive measures defined by positive measurable functions.

Next, we claim that the function f/g is the derivative $d\nu/d\mu$. For this, we wish to show that for any measurable $E \in \mathcal{A}$, we have that

$$\nu(E) = \int_E \frac{f}{g} d\mu.$$

As $\mu = (\mu + \nu)_g$, hence we see that

$$\begin{aligned}
\int_E \frac{f}{g} d\mu &= \int_E \frac{f}{g} g d(\mu + \nu) \\
&= \int_E f d(\mu + \nu) \\
&= \nu(E),
\end{aligned}$$

as required. $\square$

Example 18.8.4.8. Let $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ be a measurable space and ν be a σ-finite signed measure. Further, let μ be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

1. We wish to show that $\nu \ll \mu$. This is immediate, as $\mu(E) = 0$ if and only if $E = \emptyset$, and hence $\nu(E) = 0$ by definition.

2. We now wish to compute the derivative $d\nu/d\mu$. This is straightforward, for we first observe that the following function

$$\begin{aligned} f : \mathbb{N} &\longrightarrow [0, \infty] \\ n &\longmapsto \nu(\{n\}) \end{aligned}$$

is measurable. Indeed, this is because the σ-algebra on $\mathbb{N}$ is the power set $\mathcal{P}(\mathbb{N})$. We thus claim that

$$f = \frac{d\nu}{d\mu}.$$

Indeed, pick any measurable set $E \subseteq \mathbb{N}$. Note that it is countable in size. Observe that

$$\begin{aligned} \int_E f d\mu &= \sum_{n \in E} f(n) \\ &= \sum_{n \in E} \nu(\{n\}) \\ &= \nu\left(\coprod_{n \in E} \{n\}\right) \\ &= \nu(E) \end{aligned}$$

where the second-to-last equality is obtained from the fact ν is a measure. This completes the proof.

Lemma 18.8.4.9. *Let $(X, \mathcal{A})$ be a measurable space and μ, ν be two σ-finite positive measures on $(X, \mathcal{A})$. Let $\lambda = \mu + \nu$. Then the following are equivalent*

1. *$\mu \perp \nu$,*
2. *if $f = d\mu/d\lambda$ and $g = d\nu/d\lambda$, then*

$$fg = 0 \ \lambda\text{-almost everywhere.}$$

Proof. (1. $\Rightarrow$ 2.) As $\mu \perp \nu$, therefore there exists a ν-null set A and a μ-null set B such that

$$A \amalg B = X. \tag{9.1}$$

For any measurable $E \subseteq X$, we have

$$\begin{aligned} \mu(E) &= \int_E f d\lambda \\ \nu(E) &= \int_E g d\lambda. \end{aligned}$$

We first observe that if any of the μ or ν is the zero measure, then we are done. Indeed, for if $\mu = 0$, then we deduce that $\mu(X) = 0$ and hence $f = 0$ λ-a.e. Consequently, $fg = 0$ λ-a.e. Hence, we may now assume that none of the μ and ν are 0 measures.

Observe that since $\mu(B) = 0$, therefore $\int_B f d\lambda = 0$. As $\lambda(B) = \mu(B) + \nu(B) = \nu(B)$, therefore we deduce from the fact that $\nu \neq 0$ and $\nu(A) = \nu(X \setminus B) = 0$ that $\nu(B) \neq 0$. Hence,

$$\lambda(B) \neq 0. \tag{9.2}$$

For exactly the same reasoning applied on $\nu(A) = 0$, we deduce that

$$\lambda(A) \neq 0. \tag{9.3}$$

Hence, we have that $\int_B f d\lambda = 0 = \int_A g d\lambda$. By Eqns (9.2) and (9.3), we conclude that $f = 0$ λ-a.e. over B and $g = 0$ λ-a.e. over A.

Consider the set $N = \{f(x) \neq 0\} \cap \{g(x) \neq 0\}$. Writing

$$N = (N \cap A) \amalg (N \cap B),$$

we observe that

1. $N \cap A$ is ν-null as A is ν-null,
2. $N \cap A$ is μ-null as $\{g(x) \neq 0\} \cap A$ is λ-null and over A, we have $\lambda = \mu$,
3. $N \cap B$ is μ-null as B is μ-null,
4. $N \cap B$ is ν-null as $\{f(x) \neq 0\} \cap B$ is λ-null and over B, $\lambda = \nu$.

Hence, we see that $N \cap A$ and $N \cap B$ both are λ-null. Consequently, N is λ-null.

(2. $\Rightarrow$ 1.) For any measurable $E \subseteq X$, we have

$$\begin{aligned} \mu(E) &= \int_E f d\lambda \\ \nu(E) &= \int_E g d\lambda. \end{aligned}$$

Consider the following measurable sets

$$\begin{aligned} A &= \{g(x) = 0\} \\ B &= \{g(x) \neq 0\} \cap \{f(x) = 0\} \\ N &= \{g(x) \neq 0\} \cap \{f(x) \neq 0\}. \end{aligned}$$

Clearly, $X = A \amalg B \amalg N$. Furthermore, as $fg = 0$ λ-a.e, therefore N is λ-null. Over A we see that ν is 0 and over B we see that μ is 0. As N is λ-null, therefore it is both μ and ν-null as well. Consequently, we have

$$X = A \amalg (B \amalg N)$$

where A is ν-null and $B \amalg N$ is μ-null, as required. □

Lemma 18.8.4.10. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. $\frac{d\nu^+}{d|\nu|} = \chi_P$,

2. $\frac{d\nu^-}{d|\nu|} = \chi_N$.

Proof. First, observe that these derivatives exists because $\nu^+ \ll |\nu|$ and $\nu^- \ll |\nu|$. By Jordan decomposition of ν, we have

$$\nu = \nu^+ - \nu^-$$

where $\nu^+(E) = \nu(P \cap E)$ and $\nu^-(E) = -\nu(N \cap E)$, where $X = P \amalg N$ is the Hahn-decomposition of X into a positive set P and a negative set N obtained by ν and $E \in \mathcal{A}$.

1. We claim that $\frac{d\nu^+}{d|\nu|}$ is given by χ_P. To this end, we need only show that

$$\nu^+(E) = \int_E \chi_P d\,|\nu|$$

as by Radon-Nikodym theorem, we know that the derivatives are unique $|\nu|$-almost everywhere, and therefore ν-almost everywhere.

Now, we see that

$$\begin{aligned}
\int_E \chi_P d\,|\nu| &= |\nu|\,(E \cap P) \\
&= \nu^+(E \cap P) + \nu^-(E \cap P) \\
&= \nu(E \cap P \cap P) - \nu(E \cap P \cap N) \\
&= \nu(E \cap P) - \nu(\emptyset) \\
&= \nu^+(E),
\end{aligned}$$

as needed.

2. We proceed similarly as above and claim that χ_N is the derivative $\frac{d\nu^-}{d|\nu|}$. Indeed, we see that

$$\begin{aligned}
\int_E \chi_N d\,|\nu| &= |\nu|\,(E \cap N) \\
&= \nu^+(E \cap N) + \nu^-(E \cap N) \\
&= \nu(E \cap N \cap P) - \nu(E \cap N \cap N) \\
&= \nu(\emptyset) - \nu(E \cap N) \\
&= \nu^-(E),
\end{aligned}$$

as required. □

Lemma 18.8.4.11. *Let $(X, \mathcal{A}, \nu)$ be a signed space and let $f : X \to \mathbb{C}$ be a measurable function. Define*

$$\int_X f d\nu = \int_X f d\nu^+ - \int_X f d\nu^-$$

where $\nu = \nu^+ - \nu^-$ is the Jordan decomposition of ν. Then,

1. *we have*

$$\left| \int_X f d\nu \right| \leq \int_X |f|\, d\,|\nu|\,,$$

2. *for any $E \in \mathcal{A}$, we have*

$$|\nu|(E) = \sup\left\{\left|\int_E f d\nu\right| \mid |f| \leq 1\right\}.$$

Proof. 1. We may write

$$\begin{aligned}\left|\int_X f d\nu\right| &= \left|\int_X f d\nu^+ - \int_X f d\nu^-\right| \\ &\leq \left|\int_X f d\nu^+\right| + \left|\int_X f d\nu^-\right| \\ &\leq \int_X |f| d\nu^+ + \int_X |f| d\nu^-. \end{aligned} \tag{11.1}$$

We now claim that $\int_X |f| d\nu^+ + \int_X |f| d\nu^- = \int_X |f| d|\nu|$. Indeed, we first observe that for any $E \in \mathcal{A}$, we have $\nu^+(E) = \int_E \chi_P d|\nu|$ and $\nu^-(E) = \int_E \chi_N d|\nu|$. Consequently, we get

$$\begin{aligned}\int_X |f| d\nu^+ + \int_X |f| d\nu^- &= \int_X |f| \chi_P d|\nu| + \int_X |f| \chi_N d|\nu| \\ &= \int_X |f| (\chi_P + \chi_N) d|\nu| \\ &= \int_X |f| \cdot 1 d|\nu| \\ &= \int_X |f| d|\nu|, \end{aligned} \tag{11.2}$$

as required. Hence we conclude by Eqns (11.1) and (11.2).

2. Let $\mathcal{Z} := \{|\int_E f d\nu| \mid |f| \leq 1\}$. We first see that for any measurable $f : X \to \mathbb{C}$ with $|f| \leq 1$, we have the following by item 1 above

$$\begin{aligned}\left|\int_E f d\nu\right| &\leq \int_E |f| d|\nu| \\ &\leq \int_E d|\nu| \\ &\leq |\nu|(E). \end{aligned}$$

Hence, $\sup \mathcal{Z} \leq |\nu|(E)$.

For the converse, we wish to show that $|\nu|(E) \leq \sup \mathcal{Z}$. If $\sup \mathcal{Z} = \infty$, then there is nothing to be shown. So we may assume $\sup \mathcal{Z} < \infty$. As the constant function 1 is in the collection, therefore

$$|\nu(E)| \leq \sup \mathcal{Z} < \infty. \tag{11.3}$$

In order to show $|\nu|(E) \leq \sup \mathcal{Z}$, it suffices to find a measurable function $f : X \to \mathbb{C}$ such that $|f| \leq 1$ and $|\nu|(E) \leq |\int_E f d\nu|$. Indeed, denoting by $X = P \amalg N$ to be the Hahn-decomposition of X obtained by ν, we consider $f = \chi_P - \chi_N$. Clearly, image of f is $\{-1, 0, 1\}$ as $A \cap B = \emptyset$, hence $|f| \leq 1$. Moreover, we observe that

$$\begin{aligned}\left|\int_E f d\nu\right| &= \left|\int_E f d\nu^+ - \int_E f d\nu^-\right| \\ &= \left|\int_E (\chi_P - \chi_N) d\nu^+ - \int_E (\chi_P - \chi_N) d\nu^-\right|. \end{aligned} \tag{11.4}$$

By Eq. (11.3), we deduce that $\nu^+(E)$ and $\nu^-(E)$ are finite. Furthermore, over E we have that χ_P and χ_N are both in $\mathcal{L}^1(\nu^+)$ and $\mathcal{L}^1(\nu^-)$. With this, we may continue Eq. (11.4) as follows

$$\begin{aligned}\left|\int_E f d\nu\right| &= \left|\int_E \chi_P d\nu^+ - \int_E \chi_N d\nu^+ - \int_E \chi_P d\nu^- + \int_E \chi_N d\nu^-\right| \\ &= \left|\int_E \chi_P d\nu^+ - 0 - 0 + \int_E \chi_N d\nu^-\right| \\ &= \nu^+(E\cap P) + \nu^-(E\cap N) \\ &= \nu^+(E) + \nu^-(E) \\ &= |\nu|(E).\end{aligned}$$

where in the second equality we have used the fact the fact that $\nu^+(E) := \nu(E\cap P)$ and $\nu^-(E) := \nu(E\cap N)$. This shows that for some $f : X \to \mathbb{C}$ measurable with $|f| \leq 1$ we have $|\int_E f d\nu| = |\nu|(E)$, which consequently shows that $|\nu|(E) \leq \sup \mathcal{Z}$. This completes the proof. $\square$

Example 18.8.4.12. We wish to find those signed spaces $(X, \mathcal{A}, \nu)$ which satisfies property 1 below. Further, we also wish to find those which satisfies 2 as below:

1. For c the counting measure on $(X, \mathcal{A})$, we have $c \ll \nu$.
2. For $x_0 \in X$ and the Dirac measure δ_{x_0}, we have $\delta_{x_0} \ll \nu$.

1. Let $E \in \mathcal{A}$. We know that $c(E) = 0$ iff $E = \emptyset$. Consequently, if $\nu(E) = 0$, then $c(E) = 0$ iff $E = \emptyset$. That is, $\nu(E) = 0$ iff $E = \emptyset$. Hence all those signed spaces $(X, \mathcal{A}, \nu)$ whose only null set is $\emptyset$ can only be such that $c \ll \nu$.

2. Let $E \in \mathcal{A}$. We know that $\delta_{x_0}(E) = 0$ iff $x_0 \notin E$. Thus if $\nu(E) = 0$ and $\delta_{x_0} \ll \nu$, then $x_0 \notin E$. Hence, $(X, \mathcal{A}, \nu)$ is a signed space such that all its null sets does not contain x_0. This completes the characterizations.

Lemma 18.8.4.13. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. *If $\{E_n\} \subseteq \mathcal{A}$ be an increasing collection of measurable sets, then*

$$\nu\left(\bigcup_n E_n\right) = \lim_{n\to\infty}\nu(E_n).$$

2. *If $\{E_n\} \subseteq \mathcal{A}$ be a decreasing collection of measurable sets such that $\nu(A_1)$ is finite, then*

$$\nu\left(\bigcap_n E_n\right) = \lim_{n\to\infty}\nu(E_n).$$

Proof. 1. Denote $F_1 = E_1$ and $F_n = E_n \setminus E_{n-1}$ for $n \geq 2$. Observe that $\{F_n\}$ are disjoint, but

$$\bigcup_n E_n = \coprod_n F_n. \tag{9.1}$$

Now observe that $E_n = F_n \amalg E_{n-1}$. This is recursive relation, which when unravelled, yields

$$E_n = F_n \amalg F_{n-1} \amalg \cdots \amalg F_1.$$

Applying ν yields

$$\nu(E_n) = \sum_{k=1}^{n} \nu(F_k). \tag{9.2}$$

It follows from Eqns (9.1) and (9.2) that

$$\begin{aligned}\nu\left(\bigcup_n E_n\right) &= \nu\left(\coprod_n F_n\right)\\ &= \sum_{k=1}^{\infty} \nu(F_k)\\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \nu(F_k)\\ &= \lim_{n\to\infty} \nu(E_n),\end{aligned}$$

as needed.

2. Consider the sequence $F_1 = E_1$ and $F_n = E_1 \setminus E_n$ for $n \geq 2$. Note that $\{F_n\}$ is increasing. Hence by item 1, we have

$$\nu\left(\bigcup_n F_n\right) = \lim_{n\to\infty} \nu(F_n). \tag{9.3}$$

Now observe that

$$E_1 = F_n \amalg E_n$$

for each $n \in \mathbb{N}$. Hence, applying ν we yield

$$\nu(E_1) = \nu(F_n) + \nu(E_n).$$

As $\nu(E_1)$ is finite, therefore the RHS in above equation is finite. Consequently, each term in the above equation is finite. Hence we may write it as

$$\nu(E_1) - \nu(F_n) = \nu(E_n).$$

Taking $n \to \infty$ yields

$$\nu(E_1) - \lim_{n\to\infty} \nu(F_n) = \lim_{n\to\infty} \nu(E_n)$$

which by Eq. (9.3), yields

$$\nu(E_1) - \nu\left(\bigcup_n F_n\right) = \lim_{n\to\infty} \nu(E_n). \tag{9.4}$$

We now claim that if $A \in \mathcal{A}$ and $B \subseteq A$ in $\mathcal{A}$ is such that $\nu(B)$ is finite, then $\nu(A \setminus B) = \nu(A) - \nu(B)$. Indeed, we may write $A = (A \setminus B) \amalg B$ where $A \setminus B$ is measurable as well. Applying ν, we yield

$\nu(A) = \nu(A \setminus B) + \nu(B)$. As $\nu(B)$ is finite, therefore we may subtract both sides by $\nu(B)$ to yield $\nu(A \setminus B) = \nu(A) - \nu(B)$, as desired.

Using the above proved statement on Eq. (9.4), we obtain

$$\begin{aligned}\lim_{n\to\infty} &= \nu\left(E_1 \setminus \bigcup_n F_n\right)\\ &= \nu\left(E_1 \cap \bigcap_n F_n^c\right)\\ &= \nu\left(\bigcap_n E_1 \cap F_n^c\right)\\ &= \nu\left(\bigcap_n E_n\right),\end{aligned}$$

as desired. □

Lemma 18.8.4.14. *Let (X, Σ, μ) be a measure space and $f, g : X \to [0, \infty)$ be two non-negative measurable functions such that $f(x)g(x) = 0$ for almost all $x \in X$. Suppose for each $E \in \Sigma$ we have*

$$\mu(E) = \int_E f d\mu.$$

Define for each $E \in \Sigma$

$$\nu(E) = \int_E g d\mu.$$

Then $\mu \perp \nu$.

Proof. We know that ν as defined is a positive measure. Let $N = \{f(x)g(x) \neq 0\}$. This is a null-set. Consequently, we wish to find A and B measurable subsets such that $X = A \amalg B$ with A being μ-null and B being ν-null.

Define $A = \{f(x) = 0\}$ and $B = \{g(x) = 0 \ \& \ f(x) \neq 0\}$. Observe that $X = A \amalg B \amalg N$. Let $X_1 = A \amalg N$ and $X_2 = B$. Consequently $X = X_1 \amalg X_2$. Now, for any measurable $A' \subseteq X_1$, we may write $A' = (A' \cap A) \amalg (A' \cap N)$

$$\mu(A') = \int_{A'\cap A} f d\mu + \int_{A'\cap N} f d\mu = \int_{A'\cap A} 0 d\mu + \int_{A'\cap N} f d\mu = 0 + 0 = 0$$

where the latter term is zero because it is an integral over a measure 0 subset. Similarly, for any measurable $B' \subseteq X_2$, we see that

$$\nu(B') = \int_{B'} g d\mu = \int_{B'} 0 d\mu = 0.$$

Hence we have shown that X_1 is μ-null and X_2 is ν-null, as required. □

Lemma 18.8.4.15. *Let $(X, \mathcal{A}, \nu)$ be a signed space. Then,*

1. *If $A \in \mathcal{A}$ is a positive set, then $B \subseteq A$ such that $B \in \mathcal{A}$ is also a positive set.*
2. *If $\{A_n\} \subseteq \mathcal{A}$ is a sequence of positive sets, then $\bigcup_n A_n$ is a positive set.*

Proof. 1. Pick any $C \subseteq B$. As $B \subseteq A$, therefore $C \subseteq A$. As A is positive, thus $\nu(C) \geq 0$, as needed.

2. Let $B_1 = A_1$ and $B_n = A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$. Then observe that $\{B_n\}$ is a disjoint sequence of sets in $\mathcal{A}$, each positive as well by item 1. Furthermore, observe that

$$\coprod_n B_n = \bigcup_n A_n.$$

Now pick any $E \subseteq \bigcup_n A_n$ and denote $E_n = E \cap B_n$. Then, since B_n are disjoint, thus so is $\{E \cap B_n\}$. Furthermore $E = \coprod_n E_n$. Hence, we obtain

$$\mu(E) = \mu\left(\coprod_n E_n\right) = \sum_n \mu(E_n).$$

As each B_n is a positive set, so $E_n = E \cap B_n$ is a positive set as well by item 1. Consequently, $\mu(E_n) \geq 0$ for all $n \in \mathbb{N}$. Hence, from above, we deduce that

$$\mu(E) = \sum_n \mu(E_n) \geq 0,$$

as needed. □

18.9 The dual of $L^p(\mathbb{R}^n)$: Riesz Representation theorem

Definition 18.9.0.1. (**Linear Functional**) Suppose $(V, \mathbb{R}, \|\cdot\|)$ is a Banach Space. A **linear**[28] map $f : V \longrightarrow \mathbb{R}$ is called a linear functional.

Definition 18.9.0.2. (**Bounded linear functional**) A linear functional $\phi : V \longrightarrow \mathbb{R}$ where $(V, \mathbb{R}, \|\cdot\|)$ is a Banach space is called bounded if

$$\exists\, c \geq 0 \text{ such that } |\phi(x)| \leq c\|x\| \ \forall\, x \in V.$$

The space of all such bounded linear functionals is denoted by

$$\mathcal{B}(V).$$

That is, any $\phi \in \mathcal{B}(V)$ is a bounded linear functional.

Proposition 18.9.0.3. *Suppose $(V, \mathbb{R}, \|\cdot\|_V)$ is a Banach Space and $\mathcal{B}(V)$ is the space of bounded linear functionals over V. Then,*

$$\mathcal{B}(V) \text{ forms a Vector Space}$$

and the map

$$\|\cdot\| : \mathcal{B}(V) \longrightarrow [0, \infty)$$

defined by

$$\begin{aligned} \|\phi\| &= \sup\left\{\frac{|\phi(x)|}{\|x\|_V} \ : \ x \in V\right\} \\ &= \inf\left\{c \ : \ |\phi(x)| \leq c\|x\|_V \ , \ \ x \in V\right\} \end{aligned}$$

for any $\phi \in \mathcal{B}(V)$ forms a norm on the Vector Space $\mathcal{B}(V)$.

[28] $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)\ \forall\, v_1, v_2 \in V$ and $\alpha, \beta \in \mathbb{R}$. Or more simply, a morphism in the Category of Vector Spaces **Vect** :):

Proof. Take any two $\phi_1, \phi_2 \in \mathcal{B}(V)$ and $\alpha, \beta \in \mathbb{R}$. By the very nature of their existence, ϕ_1 & ϕ_2 have to be bounded linear functionals. Suppose

$$\begin{aligned}|\phi_1(x)| &\leq c_1\|x\|_V \\ |\phi_2(x)| &\leq c_2\|x\|_V\end{aligned}$$

$\forall\, x \in V$. Then $\alpha\phi_1$ is a function such that:

$$|\alpha\phi_1(x)| \leq \alpha c_1\|x\|_V$$

Hence $\alpha\phi_1 \in \mathcal{B}(V)$. Similarly, $\beta\phi_2 \in \mathcal{B}(V)$. Now, since we have that

$$|\phi_1 + \phi_2| \leq |\phi_1| + |\phi_2|$$

Therefore, $\phi_1 + \phi_2 \in \mathcal{B}(V)$. Hence, $\mathcal{B}(V)$ is a Vector Space.

To see that $\|\cdot\|$ is a norm over $\mathcal{B}(V)$, we see that for any $\alpha \in \mathbb{R}$ and $\phi \in \mathcal{B}(V)$, we trivially get that

$$\|\alpha\phi\| = |\alpha|\,\|\phi\|$$

and, for $f_1, f_2 \in \mathcal{B}(V)$, we also note that

$$\|f_1 + f_2\| \leq \|f_1\| + \|f_2\|$$

Hence, $\|\cdot\|$ is a norm on Vector Space $\mathcal{B}(V)$. □

18.9.1 $\mathcal{B}(V)$ is a Banach Space

Proposition 18.9.1.1. *Suppose V is a Banach Space. Then $\mathcal{B}(V)$ is a Banach Space.*

Proof. Take any Cauchy sequence $\{\phi_n\}$ in $\mathcal{B}(V)$. Now, since ϕ_n's are bounded linear functionals, therefore,

$$\exists\, c_n \geq 0 \text{ such that } |\phi_n(x)| \leq c_n\|x\|_V \;\forall\, x \in V$$

Now take any $x \in V$. Since $\phi_n(x) \in V$, we therefore have a sequence $\{\phi_n(x)\}$ in $\mathbb{R}$. We now note that

$$\begin{aligned}|\phi_n(x) - \phi_m(x)| &\leq \|x\|_V \times \sup\left\{\frac{|\phi_n - \phi_m|}{\|x\|_V} \;:\; x \in V\right\} \\ &= \|\phi_n - \phi_m\|\|x\|_V.\end{aligned}$$

Hence, $\{\phi_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$. Now write

$$\phi(x) = \varinjlim_n \phi_n(x).$$

Since $\mathbb{R}$ is complete, therefore $\phi(x) \in \mathbb{R}$. But our choice of x was arbitrary, hence $\phi(x) = \varinjlim_n \phi_n(x) < \infty \;\forall\, x \in V$. Hence $\phi \in \mathcal{B}(V)$.

Moreover,

$$\begin{aligned}
\varprojlim_n \|\phi_n\| &= \varprojlim_n \sup\left\{\frac{|\phi_n(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\varprojlim_n \frac{|\phi_n(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\frac{\left|\varprojlim_n \phi_n(x)\right|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \sup\left\{\frac{|\phi(x)|}{\|x\|_V} \;:\; x \in V\right\} \\
&= \|\phi\|
\end{aligned}$$

Hence proved. □

18.10 Remarks on Banach spaces

Following are some exercises, examples and remarks on Banach spaces.

18.10.1 normed linear spaces

Remark 18.10.1.1. a) We claim that any linear space could be normed. Let X be a linear space and $\{b_j\}$ be a Hamel basis. Then for each $x \in X$ there are unique finitely many non-zero elements $c_{x_1}, \ldots, c_{x_k} \in \mathbb{K}$ such that $x = c_{x_1} b_{j_1} + \ldots c_{x_k} b_{j_k}$. Define the following map

$$\begin{aligned}
\| - \| : X &\longrightarrow \mathbb{R}_{\geq 0} \\
x &\longmapsto \max\{|c_{x_1}|, \ldots, |c_{x_k}|\}.
\end{aligned}$$

We claim that $\| - \|$ is a norm. Indeed, if $\|x\| = 0$, then $c_{x_i} = 0$ for all $i = 1, \ldots, k$. Consequently, $x = 0$. If $x = 0$, then it is clear by uniqueness of c_{x_i} that all $c_{x_i} = 0$.

Consider $c \in \mathbb{K}$ and $x \in X$. Then $\|cx\| = \max |cc_{x_1}|, \ldots, |cc_{x_k}| = |c| \max\{|c_{x_1}|, \ldots, |c_{x_k}|\} = |c| \, \|x\|$.

We finally wish to show triangle inequality. Pick $x, y \in X$. Then, (we allow c_{x_i} and c_{y_i} to be zero)

$$\begin{aligned}
\|x + y\| &= \max\{|c_{x_1} + c_{y_1}|, \ldots, |c_{x_k} + c_{y_k}|\} \\
&\leq \max\{|c_{x_1}| + |c_{y_1}|, \ldots, |c_{x_k}| + |c_{y_k}|\} \\
&\leq \max\{|c_{x_1}|, \ldots, |c_{x_k}|\} + \max\{|c_{y_1}|, \ldots, |c_{y_k}|\} \\
&= \|x\| + \|y\|.
\end{aligned}$$

Hence every linear space is normable.

b) We claim that not all metric on a linear space X comes from a norm on X. Indeed, consider the following metric:

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R}_{\geq 0} \\ (x, y) &\longmapsto \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases} \end{aligned}$$

Indeed it is clear that $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$. For triangle inequality, we need only consider the case when $x \neq y$ and to show that for any $z \in X$ we have

$$1 = d(x, y) \leq d(x, z) + d(y, z).$$

It is clear that we need only show that $d(x, z)$ and $d(y, z)$ are both not simultaneously 0. Indeed, if both are simultaneously 0, then $x = z = y$, a contradiction. Hence d is indeed a metric.

We claim that d is not induced by any norm. Indeed, assume to the contrary it is induced by a norm $\| - \|$. It follows that

$$\|x\| = d(x, 0) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Since $\| - \|$ is a norm, it follows that for any $c \neq 1$ in $\mathbb{K}$ and $x \neq 0$ in X, we must have $\|cx\| = 1$ as $cx \neq 0$. We now have the following contradiction

$$1 = \|cx\| = |c| \, \|x\| = |c| \neq 1.$$

This completes the proof.

Remark 18.10.1.2. We wish to show that the following are equivalent for a linear space X with a function $\| : \|X \to \mathbb{R}_{\geq 0}$ satisfying $\|x\| = 0$ iff $x = 0$ and $\|cx\| = |c| \, \|x\|$ for all $c \in \mathbb{K}$ and $x \in X$:

1. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.
2. The closed unit ball $B_1[0] = \{x \in X \mid \|x\| \leq 1\}$ is convex.

(1. $\Rightarrow$ 2.) Pick $x, y \in B_1[0]$ and $c \in [0, 1]$. We wish to show that $cx + (1 - c)y \in B_1[0]$. Indeed, since $\|x\|, \|y\| \leq 1$, therefore we have

$$\begin{aligned} |cx + (1 - c)y| &\leq |c| \, \|x\| + |1 - c| \, \|y\| \\ &\leq c + (1 - c) \\ &= 1. \end{aligned}$$

(2. $\Rightarrow$ 1.) Pick $x, y \in X$. If any of the x or y is 0, then triangle inequality is immediate. Hence we may assume x and y are both not 0. Then $\frac{x}{\|x\|}, \frac{y}{\|y\|} \in B_1[0]$. Let $c = \frac{\|x\|}{\|x\| + \|y\|}$ so that $1 - c = \frac{\|y\|}{\|x\| + \|y\|}$. It is clear that $c \in [0, 1]$. By convexity of $B_1[0]$, it follows that

$$c\frac{x}{\|x\|} + (1 - c)\frac{y}{\|y\|} \in B_1[0].$$

But we have

$$c\frac{x}{\|x\|} + (1 - c)\frac{y}{\|y\|} = \frac{x}{\|x\| + \|y\|} + \frac{y}{\|x\| + \|y\|}$$

hence the RHS above is in $B_1[0]$. Taking norm, we see

$$\|\frac{x}{\|x\|+\|y\|}+\frac{y}{\|x\|+\|y\|}\|=\frac{\|x+y\|}{\|x\|+\|y\|}\leq 1$$

from which we get

$$\|x+y\|\leq\|x\|+\|y\|,$$

as required.

Example 18.10.1.3. Consider $C[a,b]$ be the $\mathbb{R}$-vector space of all continuous functions on $[a,b]$. Define for any $1\leq p<\infty$

$$\|f\|_p:=\left(\int_a^b|f(t)|^p\,dt\right)^{\frac{1}{p}}.$$

a) We wish to show that $\|-\|_p$ is a norm on $C[a,b]$. Indeed, if $f\in C[a,b]$ such that $\|f\|_p=0$, then we have

$$\int_a^b|f(t)^p|\,dt=0.$$

We wish to show that $f=0$. Suppose not, so that $f(t_0)\neq 0$ at a point $t_0\in[a,b]$. If $t_0=a$ or b, then by continuity there is a point in (a,b) where f is non-zero. Replace t by that point in (a,b). It follows by continuity that there exists $\delta>0$ such that f is non-zero on $I=[t_0-\delta,t_0+\delta]\subseteq(a,b)$. Let $m=\min_{t\in I}|f(t)|^p>0$ which exists as f is continuous on compact I and $f\neq 0$ on I. Then

$$0=\int_a^b|f(t)|^p\,dt\geq\int_{t_0-\delta}^{t_0+\delta}m dt=m\cdot(2\delta)>0,$$

a contradiction. It follows that $f=0$ on $[a,b]$.

We now wish to show triangle inequality. For this, we invoke the fact that $C[a,b]$ is contained inside the linear space $L^p[a,b]$ of $\mathbb{R}$-valued Lebesgue measurable functions on $[a,b]$. Moreover, the function

$$\|f\|_p:=\left(\int_{[a,b]}|f|^p\,dm\right)^{1/p}$$

for $f\in L^p[a,b]$ defines a norm. Moreover if f is continuous, then the above Lebesgue integral on $[a,b]$ agrees with the usual Riemann integral. So we may conclude that there is an inclusion of linear spaces

$$(C[a,b],\|-\|_p)\subseteq(L^p[a,b],\|-\|_p).$$

We know that $(L^p[a,b],\|-\|_p)$ forms a normed linear space, where triangle inequality is established by Minkowski's inequality. Using the same theorem on the subspace $(C[a,b],\|-\|_p)$, we get the desired result.

b) We claim that $(C[0,2], \| - \|_1)$ is not complete. It suffices to show a Cauchy sequence which is not convergent. Indeed consider $f_n(x)$ as follows:

$$f_n(x) = \begin{cases} x^n & \text{if } x \in [0,1] \\ 1 & \text{if } x \in (1,2]. \end{cases}$$

We first claim that (f_n) is Cauchy in $C[0,2]$. Indeed, for $n \geq m$, we have

$$\begin{aligned} \|f_n - f_m\|_1 &= \int_0^2 |f_n(x) - f_m(x)| \, dx \\ &= \int_0^1 |x^n - x^m| \, dx \\ &= \int_0^1 x^m - x^n dx \\ &= \int_0^1 x^m dx - \int_0^1 x^n dx \\ &= \frac{1}{m+1} - \frac{1}{n+1} \\ &\leq \frac{1}{m+1}. \end{aligned}$$

So for a fixed $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $1/N < \epsilon$. Then for all $n, m \geq N$, we have

$$\|f_n - f_m\|_1 \leq \frac{1}{m+1} \leq \frac{1}{N+1} < \epsilon,$$

as needed. Next, we claim that (f_n) doesn't converge in $C[0,2]$. Indeed, it would suffice to show that it converges in $L^1[0,2]$ to a non-continuous function. Indeed, consider the following simple function

$$f = \chi_{[1,2]}.$$

This is not continuous in $[0,2]$. We claim that $f_n \to f$ in $L^1[0,2]$. Indeed, we have

$$\begin{aligned} \|f_n - f\|_1 = \int_{[0,2]} |f_n - f| \, dm &= \int_{[0,1]} |f_n - f| \, dm + \int_{[1,2]} |f_n - f| \, dm \\ = \int_{[0,1]} |f_n - f| \, dm &= \int_0^1 |x^n| \, dx, \end{aligned}$$

where the last equality comes from Riemann and Lebesgue integrals being equal on compact intervals for Riemann integrable functions. Consequently, we have

$$\|f_n - f\|_1 = \frac{1}{n+1}$$

which converges to 0 as $n \to \infty$. Hence in $L^1[0,2]$, $f_n \to f$. As $C[0,2] \subseteq L^1[0,2]$ with the given norm, it follows that $(f_n) \subseteq C[0,2]$ does not converge in $C[0,2]$.

Example 18.10.1.4. Let $X = (C[0,1], \| \cdot \|_\infty)$. We wish to calculate the following:

1. $d(f_1, C)$ where $f_1(t) = t$ and C is the linear subspace of all constant functions,
2. $d(f_2, P)$ where $f_2(t) = t^2$ and P is the linear subspace of polynomials of degree at most 1.

1. We claim that $d(f_1, C) = 1/2$. Indeed, we have

$$\begin{aligned} d(f_1, C) &= \inf_{c \in C} \|f_1 - c\| = \inf_{c \in C} \sup_{t \in [0,1]} |t - c| \\ &= \inf_{c \in C} \begin{cases} c & \text{if } \frac{1}{2} \leq c < \infty \\ 1 - c & \text{if } -\infty < c < \frac{1}{2}. \end{cases} \\ &= \frac{1}{2}, \end{aligned}$$

as needed.

2. We claim that $d(f_2, P) = 1/8$. Pick any $at + b \in P$ for $a, b \in \mathbb{R}$. We first show that

$$\sup_{t \in [0,1]} \left| t^2 - at - b \right| = \max \left\{ -b, 1 - a - b, \frac{a^2}{4} + b \right\}. \tag{$*$}$$

Indeed, consider the discriminant $a^2 + 4b$ of $f(t) = t^2 - at - b$. There are two cases to be had here:

1. If $a^2 + 4b \leq 0$: Then the maximum of $|f(t)|$ is equal to that of $f(t)$ and is achieved only on the boundary at $t = 0$ or 1 because $f(t) \geq 0$ for all $t \in [0, 1]$. Consequently, $\sup_{t \in [0,1]} |f(t)| = -b$ or $1 - a - b$.
2. If $a^2 + 4b > 0$: Then the maximum of $|f(t)|$ is either on boundary at $t = 0, 1$ or at the point of minima of $f(t)$ at $t = a/2$, which thus becomes a point of maxima for $|f(t)|$. It follows that $\sup_{t \in [0,1]} |f(t)| = -b, 1 - a - b$ or $\frac{a^2}{4} + b$.

These two cases shows the claim in Eqn ($*$).

Consider now $f(a, b) = \max \left\{ -b, 1 - a - b, \frac{a^2}{4} + b \right\}$ as a function $f : \mathbb{R}^2 \to \mathbb{R}$. We wish to find $\inf_{(a,b) \in \mathbb{R}^2} f(a, b)$. First we observe the following three regions:

1. The region R_1 : This is

$$R_1 = \{(a, b) \in \mathbb{R}^2 \mid f(a, b) = -b\}.$$

2. The region R_2 : This is

$$R_2 = \{(a, b) \in \mathbb{R}^2 \mid f(a, b) = 1 - a - b\}.$$

3. The region R_3 : This is

$$R_3 = \left\{ (a, b) \in \mathbb{R}^2 \mid f(a, b) = \frac{a^2}{4} + b \right\}.$$

We now analyze bounds on a point $(a, b) \in R_i$ as follows.

1. If $(a, b) \in R_1$: Then we have

$$\begin{aligned} -b &> 1 - a - b \\ -b &> a^2/4 + b \end{aligned}$$

solving which, we get bounds

$$\begin{aligned} a &< 1 \\ b &< -\frac{a^2}{8}. \end{aligned}$$

Hence, to minimize b, we need to maximize a, thus to get that $b < -1/8$. So we have $(a,b) = (1, -1/8)$ as a point of minima for $-b$.

2. If $(a,b) \in R_2$: Then we have

$$\begin{aligned} 1 - a - b &> -b \\ 1 - a - b &> \frac{a^2}{4} + b \end{aligned}$$

solving which we get bounds

$$\begin{aligned} a &< 1 \\ b &< \frac{1}{2} - \frac{a}{2} - \frac{a^2}{8}. \end{aligned}$$

Hence to minimize $1 - a - b$, we have to maximize a and b. Doing so yields $a = 1$ and $b = -1/8$. Hence $(a,b) = (1, -1/8)$ is a point of minima for $1 - a - b$.

3. If $(a,b) \in R_3$: Then we have

$$\begin{aligned} \frac{a^2}{4} + b &< -b \\ \frac{a^2}{4} + b &< 1 - a - b \end{aligned}$$

solving which, we get bounds

$$\begin{aligned} b &> -\frac{a^2}{8} \\ b &> -\frac{a^2}{8} - \frac{a}{2} + \frac{1}{2}. \end{aligned}$$

Hence to minimize $\frac{a^2}{4} + b$, we have to minimize b and a. Doing so, we obtain $b = -a^2/8$ which thus yields

$$a > 1.$$

Hence to minimize a, we have to take $a = 1$. Consequently, $(a,b) = (1, -1/8)$ is a point of minima for $a^2/4 + b$.

From all the three cases above, we see that f minimizes at the point $(a,b) = (1, -1/8)$. Indeed, we see that $(1, -1/8) \in R_1 \cap R_2 \cap R_3$ as all three functions $-b, 1 - a - b$ and $a^2/4 + b$ are equal at it. Consequently, the $\inf_{(a,b)\in\mathbb{R}^2} f(a,b) = 1/8$, as required.

18.10.2 Properties

Proposition 18.10.2.1. *Let X be a normed linear space. The following are equivalent:*

1. *X is a Banach space.*
2. *$S^1(X) = \{x \in X \mid \|x\| = 1\}$ is a complete subset of X.*

Proof. content... $\square$

Proposition 18.10.2.2. *Let X be a normed linear space. Then the following are equivalent:*

1. *X is a Banach space.*
2. *Absolutely convergent series in X are convergent in X.*

Proof. 1. $\Rightarrow$ 2. Pick an absolutely convergent series $\sum_n x_n$ in X so that

$$\sum_n \|x_n\| < \infty.$$

It follows that $T_n = \sum_{k=1}^n \|x_k\|$ is a Cauchy sequence in $\mathbb{R}$. We wish to show that $\sum_n x_n$ converges in X. It suffices to show that the sequence $S_n = \sum_{k=1}^n x_k$ converges in X. We reduce to showing that (S_n) is Cauchy. Fix $\epsilon > 0$. For any $n \geq m$, we have

$$\begin{aligned}\|S_n - S_m\| &= \|x_{m+1} + \cdots + x_n\| \\ &\leq \|x_{m+1}\| + \cdots + \|x_n\| \\ &= \left|\left(\sum_{k=1}^n \|x_k\|\right) - \left(\sum_{k=1}^m \|x_k\|\right)\right| \\ &= |T_n - T_m| < \epsilon\end{aligned}$$

some $N \in \mathbb{N}$ and $n, m \geq N$ since (T_n) is Cauchy in $\mathbb{R}$. This shows that (S_n) is Cauchy, as required.

2. $\Rightarrow$ 1. Pick a Cauchy sequence $(x_n) \subseteq X$. We wish to show that there is a convergent subsequence of (x_n). We first find a subsequence of (x_n) which is better behaved. Indeed, by Cauchy condition, we find for each $k \geq 0$ a positive integer $N_k \in \mathbb{N}$ such that

$$\|x_n - x_m\| < \frac{1}{2^k}$$

for all $n, m \geq N_k$. We may assume N_k to be the least such possible by well-ordering on $\mathbb{N}$. Then we see that $N_{k+1} \geq N_k$ by minimality hypothesis. Thus consider the subsequence (x_{N_k}) of (x_n). Observe that

$$\|x_{N_{k+1}} - x_{N_k}\| < \frac{1}{2^k}$$

as $N_{k+1}, N_k \geq N_k$. We replace (x_n) by the subsequence (x_{N_k}) so that we may assume

$$\|x_{n+1} - x_n\| < \frac{1}{2^n} \ \forall n \in \mathbb{N}. \tag{3}$$

We now find the limit to which (x_n) converges. Indeed, define the following sequence in X:

$$\begin{aligned}y_{n-1} &= \sum_{k=1}^{n-1} x_{k+1} - x_k \\ &= x_n - x_1.\end{aligned}$$

We claim that $\sum_n x_{n+1} - x_n$ is an absolutely convergent series. Indeed, denote

$$\begin{aligned} S_{n-1} &:= \sum_{k=1}^{n-1} \|x_{k+1} - x_k\| \\ &\leq \sum_{k=1}^{n-1} \frac{1}{2^k} \end{aligned}$$

where the latter bound follows from Eqn. (3). Then, we see that for any $n \in \mathbb{N}$

$$S_n \leq \sum_{k=1}^{n} \frac{1}{2^k} < \sum_{k=1}^{\infty} \frac{1}{2^k} = M < \infty. \tag{4}$$

That is, (S_n) is a monotonically increasing positive bounded sequence in $\mathbb{R}$, therefore (S_n) is convergent. This shows that the series $\sum_n x_{n+1} - x_n$ is absolutely convergent. By our hypothesis, it follows that $\sum_n x_{n+1} - x_n$ is convergent in X. That is, the sequence

$$y_{n-1} = \sum_{k=1}^{n-1} x_k$$

of partial sums is convergent in X. But since $y_{n-1} = x_n - x_1$, it follows that (x_n) is a convergent sequence in X, as required. □

18.10.3 Bases & quotients

Lemma 18.10.3.1. *If X is a normed linear space with a Schauder basis, then X is separable.*

Proof. Let $(b_n) \subseteq X$ be a Schauder basis. Consider the following subset

$$D = \left\{ \sum_{k=1}^{n} q_k b_k \mid q_k \in E,\ n \in \mathbb{N} \right\}$$

where $E \subseteq \mathbb{K}$ is a countable dense subset. It is clear that D is countable. We claim that D is dense in X.

Pick any point $x \in X$. Since (b_n) is a Schauder basis, there exists $(c_k) \subseteq \mathbb{K}$ such that

$$x = \sum_{k=1}^{\infty} c_k b_k$$

where the series converges in X. Pick a ball $B_\epsilon(x)$ around x. We wish to show that $B_\epsilon(x) \cap D \neq \emptyset$. Indeed, consider $N \in \mathbb{N}$ large enough such that

$$\|x - \sum_{k=1}^{N} c_k b_k\| < \frac{\epsilon}{2}. \tag{9}$$

Moreover, for each $k = 1, \dots, N$, consider $q_k \in E$ such that

$$|c_k - q_k| < \frac{\epsilon}{2 \cdot 2^k \|b_k\|} \tag{10}$$

which exists by density of E in $\mathbb{K}$. Hence, we have by Eqns (9) and (10) the following inequalities:

$$\begin{aligned}\|x - \sum_{k=1}^{N} q_k b_k\| &\leq \|x - \sum_{k=1}^{N} c_k b_k\| + \|\sum_{k=1}^{N}(c_k - q_k)b_k\| \\ &< \frac{\epsilon}{2} + \sum_{k=1}^{N} |c_k - q_k|\,\|b_k\| \\ &< \frac{\epsilon}{2} + \frac{1}{2}\sum_{k=1}^{N}\frac{\epsilon}{2^k} \\ &= \frac{\epsilon}{2} + \frac{\epsilon}{2}\left(1 - \frac{1}{2^N}\right) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,\end{aligned}$$

as needed. This shows that $\sum_{k=1}^{N} q_k b_k \in B_\epsilon(x)$, that is D is dense in X. □

Proposition 18.10.3.2 (2 out of 3 property)**.** *Let X be a normed linear space and $Y \subseteq X$ be a closed linear subspace. Then,*

1. *X, Y Banach implies X/Y Banach.*
2. *$X, X/Y$ Banach implies Y Banach.*
3. *$Y, X/Y$ Banach implies X Banach.*

Proof. 1. We have done in class that if X is Banach then for any closed linear subspace Y, X/Y is Banach.

2. We need the following lemma here:

Lemma 18.10.3.3. *Let X be a Banach space and $Y \subseteq X$ be a linear subspace. Then the following are equivalent:*

1. *Y is complete.*
2. *Y is closed.*

Proof of Lemma 18.10.3.3. 1. ⇒ 2. Take $(y_n) \subseteq Y$ be a convergent sequence in X such that it converges to $x \in X$. We wish to show that $x \in Y$. Indeed, as $(y_n) \subseteq Y$ is convergent, so it is Cauchy. Since Y is complete, it follows that (y_n) converges to a point in Y. By uniqueness of point of convergence in a Hausdorff space, $x \in Y$.

2. ⇒ 1. Pick a Cauchy sequence $(y_n) \subseteq Y$. We wish to show that it converges in Y. Indeed, (y_n) as a sequence in X is Cauchy and thus by completeness of X, we deduce that $y_n \to x$ in X. But since Y is closed, therefore by uniqueness of point of convergence, we must have $x \in Y$, as required. □

Since X is Banach and Y is closed, it follows from Lemma 18.10.3.3 that Y is complete.

3. Pick a Cauchy sequence $(x_n) \subseteq X$. We wish to show that that it converges. We have a sequence $(x_n + Y) \subseteq X/Y$. We first claim that $(x_n + Y)$ is Cauchy. Indeed, we have

$$\begin{aligned}\|x_n - x_m + Y\| &= \inf_{y\in Y} \|x_n - x_m + y\| \\ &\leq \|x_n - x_m\| < \epsilon\end{aligned}$$

for all $n, m \geq N$ for some $N \in \mathbb{N}$ as $(x_n) \subseteq X$ is Cauchy. As X/Y is Banach, it follows that $(x_n + Y) \to (x + Y)$ in X/Y. Consequently, for a fixed $\epsilon > 0$, we get

$$\|x_n - x + Y\| = \inf_{y\in Y} \|x_n - x + y\| < \epsilon/2 < \epsilon$$

for all $n \geq N$ for some $N \in \mathbb{N}$. It follows from above that there is a sequence $(y_n) \subseteq Y$ such that

$$\|x_n - x + y_n\| \leq \epsilon/2 < \epsilon. \tag{1}$$

We claim that $(y_n) \subseteq Y$ is Cauchy. Indeed, we first see from Eqn. (1) that

$$\|x_n + y_n - x\| < \epsilon$$

for all $n \geq N$. Consequently, the sequence $(x_n + y_n) \subseteq X$ converges to $x \in X$. Hence, $(x_n + y_n) \subseteq X$ is Cauchy, from which we get $N \in \mathbb{N}$ such that

$$\|x_n + y_n - x_m - y_m\| = \|x_n - x_m - (y_m - y_n)\| < \epsilon$$

for each $n, m \geq N$. We may write by triangle inequality the following:

$$|\|x_n - x_m\| - \|y_n - y_m\|| \leq \|x_n - x_m - (y_m - y_n)\| < \epsilon$$

so that

$$\|y_n - y_m\| < \|x_n - x_m\| + \epsilon \tag{2}$$

for all $n, m \geq N$. As $(x_n) \subseteq X$ is Cauchy, so for some $N' \in \mathbb{N}$ we have $\|x_n - x_m\| < \epsilon$ for all $n, m \geq N'$. Replacing N by maximum of N' and N, we obtain from Eqn. (2) the following:

$$\|y_n - y_m\| < 2\epsilon \; \forall n, m \geq N.$$

This shows that $(y_n) \subseteq Y$ is Cauchy. As Y is complete, therefore $y_n \to y \in Y$. As $x_n + y_n \to x$ in X, therefore $x_n \to x - y$ in X, thus showing that X is complete. $\square$

Proposition 18.10.3.4. *The Banach space ℓ^p is separable for all $1 \leq p < \infty$.*

Proof. Recall that

$$\ell^p = \left\{(x_n) \mid x_n \in \mathbb{K} \;\&\; \sum_n |x_n|^p < \infty\right\}$$

with the norm being $\|(x_n)\|_p = (\sum_n |x_n|^p)^{1/p}$. Let $D \subseteq \mathbb{K}$ be a countable dense subset of $\mathbb{K}$ (which exists as $\mathbb{R}$ and $\mathbb{C}$ are separable in their usual topology). Using D we will construct a countable dense subset $F \subseteq \ell^p$. Indeed, consider the following subset of ℓ^p:

$$F = \bigcup_{N \geq 0} F_N$$

where

$$F_N = \{(x_n) \in \ell^p \mid x_n \in D,\ x_n = 0\ \forall n \geq N\}.$$

We see that F_N is countable as finite product of countable sets is countable and thus F is a countable union of countable sets, showing that F is countable. We next claim that F is dense in ℓ^p.

Pick any open set $B_r(y) \subseteq \ell^p$. Note that

$$B_r(y) = \left\{(x_n) \in \ell^p \mid \sum_n |x_n - y_n|^p < r^p\right\}.$$

As $y = (y_n) \in \ell^p$, therefore $\sum_n |y_n|^p = M < \infty$. Now observe that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sum_{k=n}^{\infty} |y_n|^p < \epsilon \tag{5}$$

for all $n \geq N$. As $D \subseteq \mathbb{K}$ is dense and each $y_n \in \mathbb{K}$, therefore choose

$$x_n \in B_{r_n}(y_n) \cap D \subseteq \mathbb{K}$$

where $r_n = \frac{r}{2^{\frac{n+1}{p}}}$ for all $n \in \mathbb{N}$. Hence,

$$|x_n - y_n|^p < \frac{r^p}{2 \cdot 2^n}$$

for all $n \in \mathbb{N}$. As $\sum_{n=1}^{\infty} r^p/2^{n+1} = r^p/2$, therefore

$$\sum_{n \geq 1} |x_n - y_n|^p < \frac{r^p}{2}. \tag{6}$$

This shows that the element $(x_n) \in B_r(y) \subseteq \ell^p$.

Now, fix $\epsilon > 0$ so that there exists $K \in \mathbb{N}$ large enough using Eqn. (5) such that

$$\sum_{n=K}^{\infty} |y_n|^p < \epsilon. \tag{7}$$

Using Eqn. (6) and (7), we can write

$$\begin{aligned}
\sum_{n=1}^{K-1} |x_n - y_n|^p + \sum_{n=K}^{\infty} |y_n|^p &< \sum_{n=1}^{K-1} \frac{r^p}{2 \cdot 2^n} + \sum_{n=K}^{\infty} |y_n|^p \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^K}\right) + \sum_{n=K}^{\infty} |y_n|^p \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^K}\right) + \epsilon \\
&< \frac{r^p}{2}\left(1 - \frac{1}{2^N}\right) + \epsilon
\end{aligned}$$

for all $N \geq K$. So let $N \to \infty$ so that we obtain

$$\sum_{n=1}^{K-1} |x_n - y_n|^p + \sum_{n=K}^{\infty} |y_n|^p \leq \frac{r^p}{2} + \epsilon.$$

Thus taking $\epsilon = \frac{r^p}{4}$, we get $\tilde{K} \in \mathbb{N}$ such that

$$\sum_{n=1}^{\tilde{K}-1} |x_n - y_n|^p + \sum_{n=\tilde{K}}^{\infty} |y_n|^p \leq \frac{3r^p}{4} < r^p. \tag{8}$$

Define $\tilde{x} \in \ell^p$ as follows:

$$\tilde{x}_n = \begin{cases} x_n & \text{if } n \leq \tilde{K} - 1 \\ 0 & \text{if } n \geq \tilde{K}. \end{cases}$$

Then $\tilde{x} \in F_{\tilde{K}}$ and by Eqn. (8) it follows that

$$\sum_{n=1}^{\infty} |\tilde{x}_n - y_n|^p < r^p.$$

Consequently, $\tilde{x} \in F \cap B_r(y)$, as needed. □

Example 18.10.3.5 (ℓ^∞ is not separable)**.** We wish to show that ℓ^∞ does not have a Schauder basis. By Lemma 18.10.3.1, it suffices to show that ℓ^∞ is not separable. Suppose to the contrary that $D \subseteq \ell^\infty$ is a countable dense set. We will derive a contradiction to countability of D. Indeed, consider $\kappa = \{0, 1\}$ and the subset $\kappa^\infty \subseteq \ell^\infty$ of all sequences formed by 1 and 0. Observe that κ^∞ is uncountable.

Pick any $x \in \kappa^\infty$. We first claim that $B_{1/2}(x) \cap \kappa^\infty = \{x\}$. Indeed, if $y \in B_{1/2}(x)$, then $\sup_n |x_n - y_n| < 1/2$. It follows that there exists $0 < \epsilon < 1/2$ such that

$$|x_n - y_n| < \epsilon \; \forall n \in \mathbb{N}.$$

As $x_n = 0$ or 1, therefore

$$\begin{cases} -\epsilon < y_n < \epsilon & \text{if } x_n = 0 \\ 1 - \epsilon < y_n < 1 + \epsilon & \text{if } x_n = 1. \end{cases} \tag{9}$$

Hence, if $y \in \kappa^\infty$, then by Eqn. (9) it follows that $y_n = x_n$ for all $n \in \mathbb{N}$ and thus $x = y$.

We next show that for $x \neq x' \in \kappa^\infty$, the open balls $B_{1/2}(x) \cap B_{1/2}(x') = \emptyset$. Since $x \neq x'$, we may assume WLOG that there exists $m \in \mathbb{N}$ such that $x_m = 0$ and $x'_m = 1$. Thus, if $y \in B_{1/2}(x) \cap B_{1/2}(x')$, then by Eqn. (9), it follows that

$$-\epsilon < y_m < \epsilon$$
$$1 - \epsilon < y_m < 1 + \epsilon.$$

Since $\epsilon = 1/2$, therefore the above inequalities give a contradiction. Hence $B_{1/2}(x) \cap B_{1/2}(x') = \emptyset$.

We now complete the proof. As $D \subseteq \ell^\infty$ is dense, therefore $D \cap B_{1/2}(x) \neq \emptyset$ for all $x \in \kappa^\infty$. Pick one $d_x \in D \cap B_{1/2}(x)$ for each $x \in \kappa^\infty$. By above two claims, it follows that we have an injective map

$$f : \kappa^\infty \hookrightarrow D,$$

but κ^∞ is uncountable and D is countable, a contradiction. This completes the proof.

18.10.4 Continuous linear transformations

Example 18.10.4.1. We wish to show that the inverse of a bounded linear operator may not be bounded. Indeed consider $X = (P[0,1]_1, \|\cdot\|_{\sup})$ to be the normed linear space of all polynomials whose least degree term is of degree 1. Similarly, consider $Y = (P[0,1]_2, \|\cdot\|_{\sup})$ to be the normed linear space of all polynomials whose least degree term is of degree 2. We consider the following linear map

$$\begin{aligned} T : X &\longrightarrow Y \\ p &\longmapsto \int p dx \end{aligned}$$

so that if $p(x) = a_n x^n + \dots a_1 x$, then $T(p) = \frac{a_n}{n+1}x^{n+1} + \dots + \frac{a_1}{2}x^2$. We claim that T is bounded. Indeed,

$$\begin{aligned} \|T(p)\| &= \|\frac{a_n}{n+1}x^{n+1} + \dots + \frac{a_1}{2}x^2\| \\ &= \sup_{x\in[0,1]} \left| \frac{a_n}{n+1}x^{n+1} + \dots + \frac{a_1}{2}x^2 \right| \\ &= \sup_{x\in[0,1]} \left| x \cdot \left(\frac{a_n}{n+1}x^n + \dots + \frac{a_1}{2}x \right) \right| \\ &\leq \sup_{x\in[0,1]} |x| \sup_{x\in[0,1]} \left| \left(\frac{a_n}{n+1}x^n + \dots + \frac{a_1}{2}x \right) \right| \\ &\leq 1 \cdot \sup_{x\in[0,1]} |a_n x^n + \dots a_1 x| \\ &= \sup_{x\in[0,1]} |p(x)| \\ &= \|p\|. \end{aligned}$$

Thus indeed, T is a bounded linear transformation. We next claim that the following linear transform is an inverse of T:

$$\begin{aligned} U : Y &\longrightarrow X \\ q &\longmapsto q' \end{aligned}$$

so that if $q(x) = a_n x^n + \dots a_2 x^2$, then $U(q) = n a_n x^{n-1} + \dots + 2a_2 x$. Indeed, we see that

$$\begin{aligned} T \circ U(q) &= T\left(n a_n x^{n-1} + \dots + 2a_2 x\right) \\ &= n a_n \frac{x^n}{n} + \dots 2a_2 \frac{x^2}{2} \\ &= q. \end{aligned}$$

Similarly, for $p(x) = a_n x^n + \dots a_1 x$, we see that

$$\begin{aligned} U \circ T(p) &= U\left(\frac{a_n}{n+1}x^{n+1} + \dots + \frac{a_1}{2}x^2\right) \\ &= \frac{a_n}{n+1}(n+1)x^n + \dots + \frac{a_1}{2}(2)x \\ &= p. \end{aligned}$$

This shows that U is inverse of T. We now show that U is unbounded. Indeed,

$$\begin{aligned}\|U(x^n)\| &= \|nx^{n-1}\| \\ &= \sup_{x\in[0,1]} \left|nx^{n-1}\right| \\ &= n \cdot 1 \\ &= n \cdot \|x^n\|.\end{aligned}$$

This shows that for all $n \geq 2$, there exists $q_n(x) \in Y$ given by $q_n(x) = x^{n+1}$ such that

$$\|U(q_n)\| = n + 1 > n = n\|q_n\|,$$

making U unbounded. This completes the proof.

18.10.5 Miscellaneous applications

Example 18.10.5.1. We wish to construct an additive function $f : \mathbb{R} \to \mathbb{R}$ which is not continuous. Indeed, consider the Hamel basis of $\mathbb{R}$ over $\mathbb{Q}$ and denote it by $\mathcal{B}$. We know that $\mathcal{B}$ is not finite. Observe that any additive map $f : \mathbb{R} \to \mathbb{R}$ is $\mathbb{Q}$-linear as

$$f\left(\frac{p}{q}x\right) = pf\left(\frac{1}{q}x\right)$$

and since $qf\left(\frac{1}{q}x\right) = f(x)$, thus,

$$f\left(\frac{p}{q}x\right) = \frac{p}{q}f(x).$$

Since any function $\mathcal{B} \to \mathbb{R}$ can be extended $\mathbb{Q}$-linearly to $\mathbb{R} \to \mathbb{R}$, therefore we now construct a function $f : \mathcal{B} \to \mathbb{R}$ and show that its $\mathbb{Q}$-linear extension $\tilde{f} : \mathbb{R} \to \mathbb{R}$ cannot be continuous at 0.

Pick any sequence $(b_n) \subseteq \mathcal{B}$ and consider the following sequence in $\mathbb{R}$

$$x_n = \frac{b_n}{n^{\lceil |b_n| \rceil + n}}$$

where $\lceil z \rceil$ is the ceiling function (smallest integer larger than z). Note that the denominator of x_n is a positive integer. Observe that $x_n \to 0$ as $n \to \infty$.

Define the following function $f : \mathcal{B} \to \mathbb{R}$:

$$f(b) = \begin{cases} n^{\lceil |b_n| \rceil + n} & \text{if } b = b_n \\ 1 & \text{else.} \end{cases}$$

Extend this function to a $\mathbb{Q}$-linear map $\tilde{f} : \mathbb{R} \to \mathbb{R}$, so that it is additive. We claim that $\tilde{f}$ is not continuous at 0. Indeed, we have $x_n \to 0$ as $n \to \infty$, but

$$\tilde{f}(x_n) = \tilde{f}\left(\frac{b_n}{n^{\lceil |b_n| \rceil + n}}\right) = \frac{1}{n^{\lceil |b_n| \rceil + n}}\tilde{f}(b_n) = \frac{1}{n^{\lceil |b_n| \rceil + n}}n^{\lceil |b_n| \rceil + n} = 1$$

and thus $\tilde{f}(x_n) = 1 \not\to \tilde{f}(0) = 0$ as $n \to \infty$, making $\tilde{f}$ discontinuous at 0, as needed.

Proposition 18.10.5.2. *Let X be a normed linear space over field $\mathbb{K}$ and $T : X \to \mathbb{K}$ be a linear functional. If T is unbounded, then* $\mathrm{Ker}\,(T) \subseteq X$ *is dense.*

Proof. Since T is unbounded, therefore we first claim that T is unbounded on each $B_{1/n}[0]$. Indeed, if there exists $n_0 \in \mathbb{N}$ such that T is bounded on $B_{1/n_0}[0]$, then for any $x \in X$, we have $\frac{x}{n_0\|x\|} \in B_{1/n_0}[0]$. Thus, by boundedness of T on $B_{1/n_0}[0]$, there exists $K \in \mathbb{R}_{>0}$ such that

$$\left|T\left(\frac{x}{n_0\|x\|}\right)\right| \leq K.$$

By linearity it follows from above that

$$|Tx| \leq Kn_0\|x\|$$

for all $x \in X$. This makes T bounded, a contradiction. Hence T is unbounded on each $B_{1/n}[0]$.

Consequently, for each $n \in \mathbb{N}$, there exists $y_n \in B_{1/n}[0]$ such that $\|Ty_n\| \geq n$. It follows that $y_n \to 0$ as $n \to \infty$ since $y_n \in B_{1/n}[0]$. Further, observe that

$$z_n = \frac{y_n}{Ty_n} - \frac{x}{Tx} \in \mathrm{Ker}\,(T).$$

Now we claim that $z_n \to \frac{x}{Tx}$ as $n \to \infty$. Indeed, since

$$\|\frac{y_n}{Ty_n}\| = \frac{1}{|Ty_n|}\|y_n\| \leq \frac{1}{n}\|y_n\| < \|y_n\|$$

and since $\|y_n\| \to 0$ as $n \to \infty$, therefore this shows that $\frac{y_n}{Ty_n} \to 0$ as $n \to \infty$. It follows that $z_n \to \frac{x}{Tx}$ as $n \to \infty$, as required.

As $z_n \in \mathrm{Ker}\,(T)$, therefore $T(x)z_n \in \mathrm{Ker}\,(T)$ by linearity. Thus $T(x)z_n \to x$ as $n \to \infty$. This shows the density of $\mathrm{Ker}\,(T)$, thus completing the proof. □

The following is a generalization of Riesz lemma to $r = 1$.

Proposition 18.10.5.3. *Let X be a normed linear space and $Y \subseteq X$ be a finite dimensional proper linear subspace. Then there exists $x_1 \in S^1(X) = \{x \in X \mid \|x\| = 1\}$ such that*

$$d(x_1, Y) = 1.$$

Proof. Pick $x \in X \setminus Y$. As Y is finite-dimensional in X, therefore it is closed in X. Hence, $d(x, Y) > 0$. We first claim that there exists $\tilde{y} \in Y$ such that

$$d(x, Y) = d(x, \tilde{y}). \tag{10}$$

Indeed, since $d(x, Y) = \inf_{y \in Y} d(x, y) = M$, therefore there exists a sequence $(y_n) \subseteq Y$ such that $d(x, y_n) \to M$ as $n \to \infty$. Fix $\epsilon > 0$. Thus, there exists $N \in \mathbb{N}$ such that $|d(x, y_n) - M| < \epsilon$ for all $n \geq N$. That is, $0 < d(x, y_n) < M + \epsilon$ for all $n \geq N$. Since $g(y) := d(x, y)$ is a continuous map on Y, therefore we have that

$$(y_n)_{n \geq N} \subseteq K = g^{-1}([0, M + \epsilon])$$

where $K \subseteq Y$ is a closed subset of Y. We now claim that K is bounded. Pick $y \in K$. Then

$$\begin{aligned}\|y\| = d(y,0) &\leq d(y,x) + d(x,0) \\ &< M + \epsilon + d(x,0).\end{aligned}$$

This shows that K is bounded. As Y is finite-dimensional normed linear space, therefore generalized Heine-Borel holds and we deduce that K is a compact subset of Y. Since in a metric space compactness is equivalent to sequentially compact, therefore K is sequentially compact. It follows that $(y_n)_{n\geq N} \subseteq K$ has a subsequence which converges, say to $\tilde{y} \in K \subseteq Y$. Replace (y_n) by that subsequence so that we may write $y_n \to \tilde{y}$ and $d(x, y_n) \to M$. By continuity of g, it follows that $g(y_n) = d(x, y_n) \to g(\tilde{y}) = d(x, \tilde{y})$, but $d(x, y_n) \to M$, thus by uniqueness of limits in a Hausdorff space, it follows that $d(x, \tilde{y}) = M$, as needed. This completes the proof of claim in Eqn. (10).

We now complete the proof. Consider the vector

$$x_1 = \frac{x - \tilde{y}}{\|x - \tilde{y}\|} \in X.$$

We claim that $d(x_1, Y) = 1$. Indeed,

$$\begin{aligned}d(x_1, Y) &= \inf_{y\in Y} \|\frac{x - \tilde{y}}{\|x - \tilde{y}\|} - y\| \\ &= \frac{1}{\|x - \tilde{y}\|} \inf_{y\in Y} \|x - (\tilde{y} + \|x - \tilde{y}\| y)\| \\ &= \frac{1}{\|x - \tilde{y}\|} \inf_{y\in Y} \|x - y\|\end{aligned}$$

where the last equality follows from the bijection provided by affine transformations $Y \to Y$ mapping as $y \mapsto ay + x$ for $a \in \mathbb{K}$ and $x \in Y$, using the linearity of Y. From above equalities, it follows from Eqn. (10) that

$$\begin{aligned}d(x_1, Y) &= \frac{1}{\|x - \tilde{y}\|} \inf_{y\in Y} \|x - y\| \\ &= \frac{1}{\|x - \tilde{y}\|} d(x, Y) \\ &= \frac{1}{d(x, \tilde{y})} d(x, Y) = 1,\end{aligned}$$

as required to complete the proof. □

18.11 Main theorems of functional analysis

There are four major theorems in basic functional analysis, which we discuss now.

Theorem 18.11.0.1 (Uniform boundedness principle)**.** *Let X be a Banach space and Y be a normed linear space. Consider a collection of bounded linear transformations $(T_i)_{i\in I} \subseteq B(X, Y)$ such that for each $x \in X$, the subset $(T_i x)_{i\in I} \subseteq Y$ is bounded. Then, $(\|T_i\|)_{i\in I}$ is bounded in $\mathbb{R}$, that is, $(T_i)_{i\in I} \subseteq B(X, Y)$ is a bounded set.*

Theorem 18.11.0.2 (Open mapping & bounded inverse theorem)**.** *Let X and Y be Banach spaces and $T : X \to Y$ be a surjective bounded linear map. Then,*

1. *T is an open map.*
2. *If T is a bijection, then T is a homeomorphism.*

Theorem 18.11.0.3 (Closed graph theorem)**.** *Let X and Y be Banach spaces and $T : X \to Y$ be a linear transformation. Then the following are equivalent:*

1. *T is continuous/bounded.*
2. *The graph $\Gamma(T) = \{(x, Tx) \in X \times Y \mid x \in X\}$ is closed in $X \times Y$.*

We now show how all three are equivalent.

Theorem 18.11.0.4. *Let X and Y be Banach spaces. Then the following implications are true:*

1. *CGT $\implies$ UBP.*
2. *BIT $\implies$ OMP.*
3. *CGT $\implies$ OMP.*

Proof. 1. Closed graph theorem (CGT) states that a linear map $T : X \to Y$ is bounded if and only if $\Gamma(T) = \{(x, Tx) \in X \times Y \mid x \in X\}$ is a closed set in $X \times Y$. We wish to show that uniform boundedness principle (UBP) holds, that is, if $(T_i)_{i\in I}$ is a non-empty collection of bounded linear maps from X to Y such that for each $x \in X$, the set $(T_i(x))_{i\in I} \subseteq Y$ is bounded, then the set $(\|T_i\|)_{i\in I} \subseteq \mathbb{R}$ is a bounded set.

Pick any collection $(T_i)_{i\in I} \subseteq B(X, Y)$ such that for all $x \in X$, there exists $M_x \in \mathbb{R}_+$ such that $\sup_{i\in I} \|T_i x\| \le M_x$. We wish to show that $(\|T_i\|)_{i\in I}$ is bounded. Indeed, to this end, we contstruct a new norm on X, using which, we will show the above.

Define the following for each $x \in X$:

$$\|x\|_1 := \|x\| + \sup_{i\in I} \|T_i x\|.$$

This is well-defined, as $(T_i x)$ is a bounded set in Y. We now make the following claims:

C1. $(X, \|\cdot\|_1)$ is a normed linear space.

C2. $(X, \|\cdot\|_1)$ is a Banach space.

Assuming the above two claims to be true, let us first show how this will complete the proof. We consider the map

$$\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|_1).$$

We claim that this is a continuous linear transformation. Indeed, by CGT, we need only show that $\Gamma(\mathrm{id})$ is closed. That is, (denote $X_1 = (X, \|\cdot\|_1)$)

$$\Gamma(\mathrm{id}) = \{(x, x) \in X \times X_1 \mid x \in X\} \subseteq X \times X_1$$

is closed. Indeed, consider any sequence $(x_n, x_n) \subseteq \Gamma(\mathrm{id})$ which is convergent in $X \times X_1$. Then suppose $x_n \to x$ in X and $x_n \to x'$ in X_1. We claim that $x = x'$, so that $(x_n, x_n) \to (x, x)$ and since $(x, x) \in \Gamma(\mathrm{id})$, so this will show that $\Gamma(\mathrm{id})$ is closed.

Indeed, we have $x_n \to x$ in X, so $\|x_n - x\| \to 0$ as $n \to \infty$. Similarly, $\|x_n - x'\|_1 \to 0$ as $n \to \infty$. Since

$$\|x_n - x'\|_1 = \|x_n - x'\| + \sup_{i\in I} \|T_i x_n - T_i x'\| \to 0$$

as $n \to \infty$, therefore $\sup_{i\in I} \|T_i x_n - T_i x'\| \to 0$ and $\|x_n - x'\| \to 0$ as well. The latter says that $x_n \to x$ in X. By uniqueness of limits, we conclude that $x = x'$, as required. This shows that $\mathrm{id} : X \to X_1$ is continuous linear transform by CGT, hence bounded.

We wish to bound $\sup_{\|x\|=1} \|T_i x\|$. Pick any $x \in X$ with $\|x\| = 1$. Then we have for each $i \in I$ that

$$\begin{aligned} \|x\|_1 &= \|x\| + \sup_{i\in I} \|T_i x\| \\ &\geq 1 + \|T_i x\|. \end{aligned}$$

Thus, for each $i \in I$, we have

$$\|T_i x\| \leq \|x\|_1 - 1 \leq \|x\|_1.$$

It follows that

$$\sup_{\|x\|=1} \|T_i x\| \leq \sup_{\|x\|=1} \|x\|_1 = \|\mathrm{id}\| < \infty,$$

as required. Hence we now need only prove the claims C1 and C2.

To see claim C1, proceed as follows. Observe that if $\|x\|_1 = 0$, then $\|x\| = 0$, so $x = 0$. Further we have for any $c \in \mathbb{K}$ that $\|cx\|_1 = \|cx\| + \sup_{i\in I} \|T_i(cx)\| = |c| \, \|x\| + |c| \sup_{i\in I} \|T_i x\| = |c| \, \|x\|_1$. Finally, to see triangle inequality, we see that

$$\begin{aligned} \|x + y\|_1 &= \|x + y\| + \sup_{i\in I} \|T_i x + T_i y\| \\ &\leq \|x\| + \|y\| + \sup_{i\in I} (\|T_i x\| + \|T_i y\|) \\ &\leq \|x\| + \|y\| + \sup_{i\in I} \|T_i x\| + \sup_{i\in I} \|T_i y\| \\ &= \|x\|_1 + \|y\|_1, \end{aligned}$$

as required. This shows claim C1.

To see claim C2, proceed as follows. Take any Cauchy sequence $(x_n) \subseteq X_1$. We wish to show that it converges. We claim that (x_n) is Cauchy in X. Indeed, for any $\epsilon > 0$, we have $N \in \mathbb{N}$ such that for any $n, m \geq N$ we have

$$\|x_n - x_m\| \leq \|x_n - x_m\|_1 < \epsilon$$

and for each $j \in I$, we have

$$\|T_j x_n - T_j x_m\| \leq \sup_{i\in I} \|T_i x_n - T_i x_m\| \leq \|x_n - x_m\|_1 < \epsilon/2.$$

Thus, we get by former that (x_n) is Cauchy, so convergent to say $x \in X$. We claim that (x_n) converges to x in X_1. In the latter, by letting $m \to \infty$, we obtain that for each $j \in I$ and each $n \geq N$, we have

$$\|T_j x_n - T_j x\| \leq \epsilon/2 < \epsilon.$$

Thus, taking $\sup_{i \in I}$, we further obtain that for each $n \geq N$ we have

$$\sup_{i \in I} \|T_i x_n - T_i x\| \leq \epsilon/2 < \epsilon.$$

Now, we may write

$$\begin{aligned} \|x_n - x\|_1 &= \|x_n - x\| + \sup_{i \in I} \|T_i x_n - T_i x\| \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

for $n \geq N$, as requird. This completes the proof.

2. Consider any bounded linear map $T : X \to Y$ which is surjective. We wish to show that T is an open mapping using bounded inverse theorem. Indeed, as T is bounded, therefore $\mathrm{Ker}\,(T)$ is a closed linear subspace. Going modulo $\mathrm{Ker}\,(T)$, we get a linear transformation $\tilde{T} : X/\mathrm{Ker}\,(T) \to Y$ such that the following commutes:

$$\begin{tikzcd} X \arrow[r, two heads, "T"] \arrow[d, two heads, "\pi"'] & Y \\ X/\mathrm{Ker}\,(T) \arrow[ur, two heads, "\tilde{T}"'] & \end{tikzcd}$$

We first claim that $\tilde{T}$ is bounded. Indeed, as for any $x + \mathrm{Ker}\,(T) \in X/\mathrm{Ker}\,(T)$ we have $\tilde{T}(x + \mathrm{Ker}\,(T)) = Tx$, therefore

$$\|\tilde{T}(x + \mathrm{Ker}\,(T))\| = \inf_{z \in \mathrm{Ker}(T)} \|T(x+z)\| = \inf_{z \in \mathrm{Ker}(T)} \|Tx\| = \|Tx\|.$$

This shows that $\tilde{T}$ is a bounded linear map which is injective and surjective. Thus, $\tilde{T}$ is a bijection and thus by BIT, we get that $\tilde{T}$ is a homeomorphism. In particular, we see that $\tilde{T}$ is an open map. Now consider the map $\pi : X \to X/\mathrm{Ker}\,(T)$. We wish to show that π is an open map. Let $U \subseteq X$ be an open set and pick any point $x + \mathrm{Ker}\,(T) \in \pi(U) \subseteq X/\mathrm{Ker}\,(T)$ where $x \in U$. As there exists $B_\epsilon(x) \subseteq U$, thus we claim that $B_\epsilon(x + \mathrm{Ker}\,(T)) \subseteq \pi(U)$. Indeed, if $y + \mathrm{Ker}\,(T) \in B_\epsilon(x + \mathrm{Ker}\,(T))$, then $\|x - y + \mathrm{Ker}\,(T)\| < \epsilon$. As

$$\|x - y + \mathrm{Ker}\,(T)\| = \inf_{z \in \mathrm{Ker}(T)} \|x - y + z\| < \epsilon,$$

thus there exists $z \in Z$ such that $\|x - y + z\| < \epsilon$. Thus, $y - z \in B_\epsilon(x) \subseteq U$. Hence, $y - z + \mathrm{Ker}\,(T) = y + \mathrm{Ker}\,(T) \subseteq \pi(U)$, as needed.

3. We first show that closed graph theorem (CGT) implies bounded inverse theorem (BIT). Indeed, this combined with item 2 above will show that CGT $\implies$ OMP. Let $T : X \twoheadrightarrow Y$ be a surjective bounded linear transformation which is a bijection. We then wish to show that the inverse linear transformation of T, $T^{-1} : Y \to X$, is also bounded. By CGT, it is equivalent to showing that the graph $\Gamma(T^{-1}) \subseteq Y \times X$ is a closed set. Since T is a bijection, we get

$$\begin{aligned} \Gamma(T^{-1}) &= \{(y, T^{-1}y) \in Y \times X \mid y \in Y\} \\ &= \{(Tx, x) \in Y \times X \mid x \in X\} \\ &\cong \{(x, Tx) \in X \times Y \mid x \in X\} \end{aligned}$$

where the last homeomorphism is induced by restricting the natural homeomorphism $Y \times X \to X \times Y$. It follows that $\Gamma(T^{-1})$ is closed in $Y \times X$ since $\Gamma(T)$ is closed in $X \times Y$ by CGT (as it is continuous), as required. $\square$

We next see that it is important in closed graph theorem for X and Y to be Banach.

Example 18.11.0.5. We wish to show that there exists a linear map $T : X \to Y$ where X and Y are normed linear spaces such that T is unbounded and the graph $\Gamma(T) \subseteq X \times Y$ is closed.

Indeed, consider $X = C^1[0,1]^*$ to be the subspace of $C^1[0,1]$ of those functions f such that $f(a) = 0$ and $Y = C[0,1]^*$ both with sup norm. Define

$$\begin{aligned} T : X &\longrightarrow Y \\ f(x) &\longmapsto f'(x) \end{aligned}$$

to be the derivative map. We know that T is unbounded as $f_n(x) = x^n \in C^1[0,1]$ has norm 1 but its derivative has unbounded norm. We wish to show that $\Gamma(T)$ is closed in $X \times Y$. Indeed, consider any sequence $(f_n) \subseteq X$ such that $(f_n, Tf_n) \subseteq \Gamma(T)$ is convergent in $X \times Y$. As projection map are continuous, it follows that $(f_n) \subseteq X$ and $(Tf_n) = (f_n') \subseteq Y$ are convergent. Let $f_n \to f$ in X and $f_n' \to g$ in Y. As X and Y are in sup norm, it follows that $f_n \to f$ and $f_n' \to g$ uniformly. As $f_n(0) = 0$, it follows by the theorem on uniform convergence and derivatives that f_n converges uniformly to a differentiable function which we know is f and $f' = g$. That is $Tf = g$. This shows that $(f_n, Tf_n) \to (f, Tf)$ in $X \times Y$, that is, (f_n, Tf_n) converges in $\Gamma(T)$. This shows that $\Gamma(T)$ is closed. Yet, T is unbounded, as required.

Similarly, the hypothesis of completeness is essential in uniform boudnedness principle.

Example 18.11.0.6. We wish to show that the hypothesis of completeness of the domain in uniform boundedness principle is essential.

Indeed, let $X = \mathbb{R}^\infty \subseteq (\ell^2, \|\cdot\|_2)$ of all eventually zero sequences in ℓ^2 with the induced norm. Then X is not Banach as $(x_k^{(n)}) = (1, 1/2, \dots, 1/n, 0, \dots)$ is a sequence in X which is Cauchy but it is not convergent. We now construct a sequence of functionals $f_n : X \to \mathbb{K}$ such that for all $(x_k) \in X$, the sequence $(f_n((x_k)))_n$ is bounded in $\mathbb{K}$ but still $(\|f_n\|)_n \subseteq \mathbb{R}$ is unbounded.

Consider

$$\begin{aligned} f_n : X &\longrightarrow \mathbb{K} \\ (x_k) &\longmapsto \sum_{k=1}^{n} x_k. \end{aligned}$$

Pick any $(x_k) \in X$. Then,

$$|f_n((x_k))| = \left|\sum_{k=1}^{n} x_k\right| \leq \left|\sum_{k=1}^{\infty} x_k\right| < \infty$$

as there are only finitely many non-zero elements, thus for each $(x_k) \in X$, $(f_n((x_k)))_n$ is bounded. Moreover,

$$\|f_n\| = \sup_{(x_k) \in X} \frac{|f_n((x_k))|}{\|(x_k)\|} \geq \frac{|\sum_{k=1}^{n} x_k|}{\left(\sum_{k=1}^{\infty} |x_k|^2\right)^{1/2}}$$

for any (x_k) in X. We claim that $\|f_n\| \to \infty$ as $n \to \infty$. Indeed, consider $(x_k^{(n)}) = (1, 1/2, \dots, 1/n, 0, \dots)$. Then, $\|(x_k^{(n)})\| = 1 + 1/2^2 + \dots 1/n^2 < M$ for a fixed $M > 0$ and for all n. Further, by above we have

$$\begin{aligned} \|f_n\| &\geq \frac{|\sum_{k=1}^{n} 1/k|}{\|(x_k^{(n)})\|} \\ &> \frac{1}{M} \sum_{k=1}^{n} \frac{1}{k} \to \infty \end{aligned}$$

as $n \to \infty$, as required.

We wish to next prove the main theorems using an important technical lemma.

Theorem 18.11.0.7 (Zabreiko's lemma)**.** *Let X be a Banach space and $p : X \to \mathbb{R}_{\geq 0}$ be a seminorm. If p is countably subadditive, then p is continuous.*

Proof. Let us first define a seminorm on a Banach space.

Definition 18.11.0.8 (**Seminorm and countably subadditive functions**)**.** Let X be a normed linear space. A function $p : X \to \mathbb{R}_{\geq 0}$ is said to be a seminorm if $p(\alpha x) = |\alpha|\, p(x)$ for all $\alpha \in \mathbb{K}$ and $x \in X$ and $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.

The function p is said to be countably subadditive if for every convergent series $\sum_n x_n$ in X, we have

$$p\left(\sum_{n=1}^{\infty} x_n\right) \leq \sum_{n=1}^{\infty} p(x_n)$$

In proving Zabreiko's lemma, we would need a notion of absorbing sets.

Definition 18.11.0.9 (**Absorbing set**)**.** Let X be a normed linear space. A subset $A \subseteq X$ is said to be absorbing if for all $x \in X$, there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$.

Note that if A is absorbing, then $-A$ is also absorbing. We now state the following proposition, which will be used in proving Zabreiko's lemma.

Proposition 18.11.0.10. *Let X be a normed linear space, $p : X \to \mathbb{R}_{\geq 0}$ be a function and $A \subseteq X$.*

1. *If A is absorbing, then $0 \in A$.*
2. *If X is Banach and A is closed convex and absorbing, then A contains a neighborhood of 0.*
3. *If p is a seminorm, then if p is continuous at 0, then p is continuous on X.*

Proof. 1. As A is absorbing, therefore for all $x \in X$, there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$. Let $x = 0$. Then, there exists $s_0 \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_0$. Pick any $t \geq s_0$, we get $0 = ta$ for some $a \in A$. As $t \neq 0$, it follows that $a = 0$, as required.

2. Let $A \subseteq X$ be closed convex and absorbing. Then first observe that

$$D = A \cap (-A) \subseteq A$$

is non-empty as $0 \in A$ (and thus so is in $-A$). We claim that for any $S \subseteq D$, we have

$$\frac{1}{2}S + \frac{1}{2}(-S) \subseteq D.$$

Indeed, pick any $\frac{s_1 - s_2}{2} \in \frac{1}{2}S + \frac{1}{2}(-S)$. We wish to show that $\frac{s_1-s_2}{2} \in A$ and $\frac{s_1-s_2}{2} \in -A$. Thus, we reduce to showing that $\frac{s_1-s_2}{2}, \frac{s_2-s_1}{2} \in A$. It is easy to see that $A \cap -A$ is convex as A and $-A$ are convex. As $s_1, s_2 \in S \subseteq A \cap -A$ thus $-s_1, -s_2 \in S \subseteq A \cap -A$ as well. Now, by convexity of $A \cap -A$, we get

$$\frac{s_1 - s_2}{2}, \frac{s_2 - s_1}{2} \in A \cap -A,$$

as required.

We claim that D° is non-empty. This will complete the proof as by above we will have that $\frac{1}{2}D^\circ + \frac{1}{2}(-D^\circ) \subseteq D$ is open in D and since it contains 0, we would have shown that A contains an open set containing 0.

Suppose to the contrary that $D^\circ = \emptyset$. We wish to derive a contradiction to the fact that A is an absorbing set. Indeed, first observe that for all $n \in \mathbb{N}$, we have $(nD)^\circ = \emptyset$ and nD is closed. This gives us that for each $n \in \mathbb{N}$, the set $Y_n = X - (nD)$ is an open dense subset of X. Pick any $x \in X - D$. As $X - D$ is open, there exists $B_1 = \overline{B_{r_1}(x)} \subseteq X - D$ where $r_1 < 1$. As $X - 2D$ is dense, therefore $(X - 2D) \cap (B_1)^\circ$ is non-empty and thus we get a closed ball B_2 of radius less than $1/2$ in B_1. Continuing this, we have a sequence of closed balls $B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n \supseteq \dots$ with radius of B_n less than $1/n$ and $B_n \cap nD = \emptyset$. Let x_n be the center of B_n. We claim that (x_n) is a Cauchy sequence. Indeed, for any $1/k$ we have

$$\|x_n - x_m\| < 2/k$$

for all $n, m \geq k$. As X is complete therefore there exists $x \in X$ such that $x_n \to x$. Thus $x \in B_n$ for all $n \in \mathbb{N}$, that is, $x \notin nD$ for all $n \in \mathbb{N}$. As A is absorbing, therefore there exists $s_x \in \mathbb{R}_{>0}$ such that $x \in tA$ for all $t \geq s_x$. As $-A$ is also absorbing, thus we get $s'_x \in \mathbb{R}_{>0}$ such that $x \in -tA$ for all $t \geq s'_x$. Let $n \in \mathbb{N}$ be larger than both s_x, s'_x. Then we have that $x \in nA$ and $x \in -nA$. It follows that $x \in nA \cap (-nA) = nD$, a contradiction to the fact that $x \notin D$. This completes the proof of item 2.

3. Let $x_n \to x$ in X where $x \neq 0$. We wish to show that $p(x_n) \to p(x)$. Indeed, since $x_n - x \to 0$ and p is continuous at 0, we get $p(x_n - x) \to p(0) = 0$. Thus for any $\epsilon > 0$, we have $p(x_n - x) = |p(x_n - x)| < \epsilon$ for all $n \geq N$. As $p(x_n) - p(x) \leq p(x_n - x)$ by seminorm crieterion, we get $|p(x_n) - p(x)| < \epsilon$ for all $n \geq N$. It follows that $p(x_n) \to p(x)$, as required. □

Using the above proposition, we now prove Zabreiko's lemma.

Proof of Theorem 18.11.0.7. By Proposition 18.11.0.10, 3, we reduce to proving that p is continuous at 0. We claim that it is sufficient to show that there is an open ball $B_r(0)$ of radius $r > 0$ at 0 such that $p(B_r(0))$ is a bounded set in $\mathbb{R}_{\geq 0}$. Indeed, for any sequence (x_n) in X converging to 0, which we may assume to be contained in $B_r(0)$, we get that $p(x_n) \in p(B_r(0))$ for all $n \in \mathbb{N}$. We wish to show that $p(x_n) \to p(0) = 0$. Indeed, if $p(B_r(0))$ is upper bounded by $M > 0$, we thus get for any $x \in B_r(0)$ the following bound:

$$p(x) = \|x\| p\left(\frac{x}{\|x\|}\right) \leq M\|x\|.$$

Consequently, we have

$$p(x_n) \leq M\|x_n\|.$$

As $\|x_n\| \to 0$ as $n \to \infty$, it follows by above that $p(x_n) \to 0$ as $n \to \infty$, as required.

So we reduce to showing that there exists an open $B_r(0)$ of 0 in X such that $p(B_r(0))$ is a bounded set. Consider $A_! = \{x \in X \mid p(x) < 1\}$. We claim that A is an absorbing set. Indeed, for any $x \in X$, we have $\|x\|$ such that for all $t \geq \|x\|$ we have $x \in tA_!$ since $p(x/t) = p(x)/t < 1/t$, so $p(t \cdot x/t) < 1$, as required. This shows that $A_!$ is absorbing. We claim that $A = \overline{A_!}$ is absorbing as well. Indeed, observe that since A contains an absorbing set, namely $A_!$, then A is absorbing as well.

We next show that A is convex. Note that since closure of convex set is convex and $A_!$ is convex since if $x, y \in A_!$, then $p((1-t)x + ty) \leq (1-t)p(x) + tp(y) < (1-t) + t = 1$, therefore A is convex. Thus, A is closed convex absorbing set in a Banach space. By Proposition 18.11.0.10, 2, it follows that A has a neighborhood of 0.

We now find the required ball $B_r(0)$ so that $p(B_r(0))$ is bounded. Indeed consider $r > 0$ such that $\overline{B_r(0)} \subseteq \bar{A}$ and fix a point $x \in B_r(0)$. Pick a point $x_1 \in A$ such that $\|x - x_1\| < r/2$, that is, $x_1 \in B_{r/2}(x) \cap B_r(0) \subseteq \frac{1}{2}A$. Thus $x - x_1 \in \frac{1}{2}B_r(0) \subseteq \frac{1}{2}A \subseteq \frac{1}{2}\bar{A}$. Now there exists $x_2 \in \frac{1}{2}A$ such that $\|x - x_1 - x_2\| \leq r/2^2$, that is, $x_2 \in B_{r/2^2}(x - x_1) \cap B_{r/2}(0) \subseteq \frac{1}{2^2}A$. Continuing this, we get a sequence (x_n) in A such that $x_n \in \frac{1}{2^{n-1}}A$ and $\|x - \sum_{k=1}^n x_k\| < \frac{r}{2^n}$. It follows that $\sum_{k=1}^n x_k \to x$ as $n \to \infty$.

By countable sub-additivity of p, it follows that

$$p(x) = p\left(\sum_{k=1}^{\infty} x_k\right) \leq \sum_{k=1}^{\infty} p\left(x_k\right).$$

As $x_k \in \frac{1}{2^k}A$, therefore $p(x_k) < \frac{1}{2^k}$ by definition of A. Thus, $\sum_{k=1}^{\infty} p(x_k) \leq 1$, and thus $p(x) \leq 1$. As $x \in B_r(0)$ was arbitrary, we have thus shown that $p(B_r(0)) \leq 2$, as required. □

Theorem 18.11.0.11. *One can derive OMT, UBP, CGT from Zabreiko's lemma (Theorem 18.11.0.7).*

Proof. (Zabreiko ⇒ OMT) Let $T : X \twoheadrightarrow Y$ be a surjective linear transformation between Banach spaces. By translation and scaling homeomorphism, we reduce to showing that $T(B_1(0))$ is open. Define

$$\begin{aligned} p : Y &\longrightarrow \mathbb{R}_{\geq 0} \\ y &\longmapsto \inf\{\|x\| \mid Tx = y\}. \end{aligned}$$

We claim that p is a countably subadditive semi-norm, so that by Theorem 18.11.0.7, we will get p is continuous. This is sufficient as

$$T(B_1(0)) = p^{-1}([0,1))$$

which is easy to see. So we reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n y_n$ be a covergent series in Y. We wish to show that $p\left(\sum_n y_n\right) \leq \sum_n p(y_n)$. Indeed, fix $\epsilon > 0$. We get the following

$$p(y_n) + \frac{\epsilon}{2^n} \geq \|x_n\|$$

for each $n \in \mathbb{N}$ where $x_n \in X$ is such that $Tx_n = y_n$. Summing till N we get

$$\sum_{n=1}^{N} p(y_n) + \sum_{n=1}^{N} \frac{\epsilon}{2^n} \geq \sum_{n=1}^{N} \|x_n\| \geq \|\sum_{n=1}^{N} x_n\|$$

and since $T(x_1 + \cdots + x_n) = \sum_{n=1}^{N} y_n$, we get that $\|x_1 + \cdots + x_n\| \geq p(\sum_{n=1}^{N} y_n)$. This yields that

$$\sum_{n=1}^{N} p(y_n) + \sum_{n=1}^{N} \frac{\epsilon}{2^n} \geq p\left(\sum_{n=1}^{N} y_n\right).$$

Taking $N \to \infty$ and then $\epsilon \to 0$, the result follows.
2. p is a seminorm : Fact that $p(cy) = |c|\, y$ is immediate from definition. Subadditivity follows from item 1.

(Zabreiko $\Rightarrow$ UBP) Let X, Y be Banach and $(T_i)_{i\in I} \subseteq B(X,Y)$ be a family of bounded linear transformations such that for all $x \in X$, the set $(T_i(x))_{i\in I} \subseteq Y$ is bounded. We wish to show that $(\|T_i\|)_{i\in I}$ is bounded in $\mathbb{R}$.

Consider

$$\begin{aligned} p : X &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \sup_{i\in I} \|T_i(x)\|. \end{aligned}$$

We claim that p is a countably subadditive seminorm. Indeed, then it would follow by Theorem 18.11.0.7 that p is continuous. Then there exists $\delta > 0$ such that $\|x\| < \delta$ implies $|p(x)| \leq 1$. As p is a seminorm, therefore we would obtain

$$\|x\| < 1 \implies p(x) < 1/\delta.$$

As $\|T_i\| = \sup_{\|x\|<1} \|T_i x\|$ and $p(x) < 1/\delta$ for $\|x\| < 1$ where

$$p(x) = \sup_{i\in I} \|T_i x\| < 1/\delta$$

therefore $\|T_i x\| < 1/\delta$ for all $\|x\| < 1$, which would thus tield $\|T_i\| \leq 1/\delta$, as required. So we reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n x_n$ be a convergent series in X. We wish to show that $p\left(\sum_n x_n\right) \leq \sum_n p(x_n)$. Indeed, we have

$$p\left(\sum_n x_n\right) = \sup_{i\in I} \|T_i\left(\sum_n x_n\right)\| = \sup_{i\in I} \|\sum_n T_i x_n\| \leq \sup_{i\in I} \sum_n \|T_i x_n\| \leq \sum_n \sup_{i\in I} \|T_i x_n\| = \sum_n p(x_n)$$

where $\sup_{i\in I} \|T_i x_n\|$ exists and is bounded as by hypothesis, the set $(T_i x)_{i\in I}$ is bounded for any $x \in X$. This shows that p is countably subadditive.

2. p is a seminorm : Fact that $p(cy) = |c|\,y$ is immediate from definition. Subadditivity follows from item 1.

(Zabreiko $\Rightarrow$ CGT) Let $T : X \to Y$ be a linear transformation between Banach spaces. We wish to show that T is bounded if and only if $\Gamma(T) \subseteq X \times Y$ is closed.

($\Rightarrow$) is immediate by considering the inverse image at 0 of $X \times Y \to Y$ of $(x, y) \mapsto Tx - y$.

($\Leftarrow$) Consider the following function

$$\begin{aligned} p : X &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \|Tx\|. \end{aligned}$$

We claim that p is a countably subadditive seminorm. Indeed, this would imply that p is continuous by Theorem 18.11.0.7. Note that it is sufficient to show that $\{\|Tx\| \mid \|x\| < 1\}$ is bounded. But this set is same as $p(B_1(0))$. Thus, we reduce to showing that $p(B_1(0))$ is bounded. Indeed, this follows as there exists $\delta > 0$ such that

$$\|x\| < \delta \implies p(x) < 1$$

which by seminorm property is equivalent to

$$\|x\| < 1 \implies p(x) < 1/\delta.$$

This shows that $p(B_1(0)) < 1/\delta$, as needed. We thus reduce to showing that p is a countably subadditive seminorm.

1. p is countably subadditive : Let $\sum_n x_n$ be a convergent series in X. We wish to show that $p\left(\sum_n x_n\right) \leq \sum_n p(x_n)$. Indeed, we have

$$p\left(\sum_n x_n\right) = \|T\left(\sum_n x_n\right)\|$$

   where since $(\sum_{k=1}^n x_k, \sum_{k=1}^n Tx_k)$ is in the graph and is convergent where graph is closed, therefore $T(\sum_n x_n) = \sum_n Tx_n$. Thus,

$$\|T\left(\sum_n x_n\right)\| = \|\sum_n Tx_n\| \leq \sum_n \|Tx_n\| = \sum_n p(x_n).$$

   This shows that p is countably subadditive.
2. p is a seminorm : Fact that $p(cy) = |c|\,y$ is immediate from definition. Subadditivity follows from item 1.

This completes the proof of Theorem 18.11.0.7. □

This completes the proof. □

18.12 Strong & weak convergence

These are important definitions as these protray that how fundamental importance this topic gives to functionals, anyways, its *functional* analysis so we must be very comfortable with constructing and manipulating functionals on a normed linear space.

Definition 18.12.0.1 (Strong & weak convergence). Let X be a normed linear space and $(x_n) \subseteq X$ be a sequence in X. Then, (x_n) is said to be strongly convergent if there exists $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. Further (x_n) is said to be weakly convergent if there exists $x \in X$ such that for all functionals $f \in X^*$, the sequence $(f(x_n)) \to f(x)$ in $\mathbb{K}$. In the former case x is said to be the strong limit and in the latter case x is said to be the weak limit.

The following showcases a nice property of weak convergence.

Proposition 18.12.0.2. *Let X be a normed linear space and $x_n \to x$ weakly in X. Then*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Proof. As $f(x_n) \to f(x)$ for all $f \in X^*$, therefore we will construct a functional using Hahn-Banach through which the desrired inequality is straightforward. Indeed, by separation theorem applied on point x, we get that there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$. Consequntly, we get by weak convergence that

$$f(x_n) \to f(x) = \|x\|.$$

Now, for each $n \in \mathbb{N}$ we have

$$|f(x_n)| \leq \|f\|\|x_n\| = \|x_n\|.$$

Taking liminf both sides, we obtain

$$\liminf_{n\to\infty} |f(x_n)| \leq \liminf_{n\to\infty} \|x_n\|.$$

As $f(x_n) \to f(x)$, therefore $\liminf_{n\to\infty} |f(x_n)| = |f(x)| = \|x\|$. Thus we get

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|,$$

as required. □

Definition 18.12.0.3 (Weakly Cauchy and complete). A normed linear space X is *weakly complete* if every weakly Cauchy sequence is weakly convergent, where a sequence (x_n) in X is weakly Cauchy if for all $f \in X^*$, the sequence $(f(x_n))$ is Cauchy. Thus, unravelling this, we have that X is weakly complete if for any sequence (x_n) in X such that $(f(x_n))$ is Cauchy in $\mathbb{K}$ for each $f \in X^*$, there exists $x \in X$ such that $f(x_n) \to f(x)$ for each $f \in X^*$.

Proposition 18.12.0.4. *Any reflexive normed linear space X is weakly complete.*

Proof. Recall X is reflexive if the James map $\mathrm{ev} : X \to X^{**}$ is surjective. Since we have seen that ev is an isometric embedding, therefore reflexivity tells us ev is an isometric isomorphism.

To show that X is weakly complete, pick any weakly Cauchy sequence (x_n) in X. Then, for each $f \in X^*$, the sequence $f(x_n)$ is Cauchy in $\mathbb{K}$. As $\mathbb{K}$ is complete, it follows that $f(x_n)$ converges and let $f(x_n) \to c_f$ where $c_f \in \mathbb{K}$. We claim that the mapping

$$\begin{aligned} \varphi : X^* &\longrightarrow \mathbb{K} \\ f &\longmapsto c_f \end{aligned}$$

is a bounded linear map. This will complete the proof as by reflexivity we will have a unique $x \in X$ such that $\mathrm{ev}_x = \varphi$ and thus $\mathrm{ev}_x(f) = f(x) = c_f = \varphi(f)$, that is,

$$f(x) = \varprojlim_{n\to\infty} f(x_n)$$

for all $f \in X^*$, which shows that (x_n) weakly convergent, as required. We thus reduce to proving that φ is a bounded linear map.

To see linearity, pick any $f, g \in X^*$ and $\alpha \in \mathbb{K}$ to observe that

$$\varphi(f + \alpha g) = c_{f+\alpha g} = \varprojlim_n (f + \alpha g)(x_n) = \varprojlim_n f(x_n) + \alpha \varprojlim_n g(x_n) = c_f + \alpha c_g$$

since each $f(x_n)$ and $g(x_n)$ converges because they are Cauchy. To see boundedness, we first show that the set $\{x_n\} \subseteq X$ is a bounded set. Indeed, by a corollary of uniform boundedness principle we have that a set $Y \subseteq X$ is bounded if and only if $f(Y) \subseteq \mathbb{K}$ is bounded for each $f \in X^*$. For $Y = \{x_n\}$ and any $f \in X^*$, we see that $f(Y) = (f(x_n))$ is bounded as $f(x_n) \to c_f$. It follows that $\{x_n\}$ is a bounded set, as required. Consequently, let $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. We thus have

$$|\varphi(f)| = |c_f| = \varprojlim_n |f(x_n)| \leq \limsup_n \|f\|\|x_n\| \leq \|f\| \cdot M.$$

Hence, φ is a bounded linear map, as required. $\square$

Chapter 19

Homological Methods

Contents

Methods employed in homological algebra comes in handy to attack certain type of local-global problems in geometry. We would like to discuss some foundational homological algebra in this chapter in the setting of additive and abelian categories. The main goal is not to illuminate foundations but to quickly get to the working theory which can allow us to develop deeper results elsewhere in this notebook. Using the Freyd-Mitchell embedding theorem, we can always assume that any (small) abelian category $\mathbf{A}$ is a full subcategory of $\mathbf{Mod}(R)$ over some ring R. Thus we will freely do the technique of diagram chasing in the following, implicitly assuming $\mathbf{A}$ to be embedded in a module category. Consequently, the main example to keep in mind throughout this chapter is of-course the category of R-modules, $\mathbf{Mod}(R)$.

19.1 The setup : abelian categories

Let us begin with the basic definitions. Let $\mathbf{A}$ be a category. Then $\mathbf{A}$ is said to be *preadditive* if for any $x, y \in \mathbf{A}$, the homset $\mathrm{Hom}\,(x, y)$ is an abelian group and the composition $\mathrm{Hom}\,(x, y) \times \mathrm{Hom}\,(y, z) \to \mathrm{Hom}\,(x, z)$ is a bilinear map. For two preadditive categories $\mathbf{A}, \mathbf{B}$ a functor $F : \mathbf{A} \to \mathbf{B}$ is called *additive* if for all $x, y \in \mathbf{A}$, the function $\mathrm{Hom}_{\mathbf{A}}\,(x, y) \to \mathrm{Hom}_{\mathbf{B}}\,(Fx, Fy)$ is a group homomorphism.

Let $\mathbf{A}$ be a preadditive category and $f : x \to y$ be an arrow. This mean for any two object $w, z \in \mathbf{A}$, there is a zero arrow $0 \in \operatorname{Hom}(w, z)$. Then, we can define the usual notions of algebra as follows.

1. $i : \operatorname{Ker}(f) \to x$ is defined by the following universal property w.r.t. $fi = 0$:

$$\begin{tikzcd} x' \arrow[d, dashed, "\exists!"'] \arrow[dr, "\forall g \text{ s.t. } fg=0"] & & \\ \operatorname{Ker}(f) \arrow[r, "i"] & x \arrow[r, "f"] & y \end{tikzcd}$$

2. $j : y \to \operatorname{CoKer}(f)$ defined by the following universal property w.r.t. $jf = 0$:

$$\begin{tikzcd} & & y' \\ x \arrow[r, "f"] & y \arrow[ur, "\forall g \text{ s.t. } gf=0"] \arrow[r, "j"] & \operatorname{CoKer}(f) \arrow[u, dashed, "\exists!"'] \end{tikzcd} .$$

3. $k : x \to \operatorname{CoIm}(f)$ is defined to be the cokernel of the kernel map $i : \operatorname{Ker}(f) \to x$.
4. $l : \operatorname{Im}(f) \to y$ is defined to be the kernel of the cokernel map $j : y \to \operatorname{CoKer}(f)$.

Hence, for each $f : x \to y$ in a preadditive category $\mathbf{A}$, we can contemplate the following four type of maps:

$$\begin{tikzcd} \operatorname{Ker}(f) \arrow[d] & \operatorname{CoKer}(f) \\ x \arrow[r, "f"] \arrow[d] & y \arrow[u] \\ \operatorname{CoIm}(f) & \operatorname{Im}(f) \arrow[u] \end{tikzcd}$$

Lemma 19.1.0.1. *In a preadditive category, if a coproduct $x \oplus y$ exists, then so does the product $x \times y$ and vice versa. In such a case, $x \oplus y \cong x \times y$.* □

A preadditive category $\mathbf{A}$ is said to be *additive* if it contains all finite products, including the empty ones. By the above lemma, we require zero objects and sums of objects to exist.

An additive category $\mathbf{A}$ is said to be *abelian* if all kernels and cokernels exist and the natural map for each $f : x \to y$ in $\mathbf{A}$

$$\operatorname{CoIm}(f) \to \operatorname{Im}(f)$$

is an isomorphism. This intuitively means that the first isomorphism theorem holds in abelian categories by definition.

19.2 Homology, resolutions and derived functors

In this section, we shall discuss basic topics of homological algebra in abelian categories, which we shall need to setup the sheaf cohomology in geometry and Lie group cohomology in algebra and etcetera, etcetera.

19.2.1 Homology

We first define cochain complexes and maps, cohomology and homotopy of such. Since this section is mostly filled with *trivial matters*, therefore we shall allow ourselves to be a bit sketchy with proofs.

Definition 19.2.1.1. (**Cochain complexes, maps and cohomology**) A cochain complex $A^\bullet$ is a sequence of object $\{A^i\}_{i\in\mathbb{Z}}$ with a map $d^i : A^i \to A^{i+1}$ called the coboundary maps which satisfies $d^i \circ d^{i-1} = 0$ for all $i \in \mathbb{Z}$. A map $f : A^\bullet \to B^\bullet$ of cochain complexes is defined as a collection of maps $f^i : A^i \to B^i$ such that the following commutes

$$\begin{tikzcd} A^i \arrow[r, "d^i"] \arrow[d, "f^i"'] & A^{i+1} \arrow[d, "f^{i+1}"] \\ B^i \arrow[r, "d^i"'] & B^{i+1} \end{tikzcd}.$$

That is, $d^i f^i = f^{i+1} d^i$ for each $i \in \mathbb{Z}$. For a cochain complex $A^\bullet$, we define the i^{th} cohomology object as the quotient

$$h^i(A^\bullet) := \mathrm{Ker}\left(d^i\right)/\mathrm{Im}\left(d^{i-1}\right).$$

With the obvious notion of composition, we thus obtain a category of cochain complexes $\mathrm{coCh}\,(\mathbf{A})$ over the abelian category $\mathbf{A}$.

We now show that h^i forms a functor over $\mathrm{coCh}\,(\mathbf{A}\ )$.

Lemma 19.2.1.2. *Let $\mathbf{A}$ be an abelian category. The i^{th}-cohomology assignment is a functor*

$$\begin{aligned} h^i : \mathrm{coCh}\,(\mathbf{A}\ ) &\longrightarrow \mathbf{A} \\ A^\bullet &\longmapsto h^i(A^\bullet). \end{aligned}$$

Proof. For a map of complexes $f : A^\bullet \to B^\bullet$, we first define the map $h^i(f)$

$$\begin{aligned} h^i(f) : h^i(A^\bullet) &\longrightarrow h^i(B^\bullet) \\ a + \mathrm{Im}\left(d^{i-1}\right) &\longmapsto f^i(a) + \mathrm{Im}\left(d^{i-1}\right). \end{aligned}$$

This is well defined group homomorphism. Further, it is clear that this is functorial. □

With this, we obtain the cohomology long-exact sequence.

Lemma 19.2.1.3. *Let* $\mathbf{A}$ *be an abelian category and* $0 \to A^\bullet \to B^\bullet \to C^\bullet \to 0$ *a short-exact sequence in* $\mathrm{coCh}\,(\mathbf{A})$. *Then there is a map* $\delta^i : h^i(C^\bullet) \to h^{i+1}(A^\bullet)$ *for each* $i \in \mathbb{Z}$ *such that the following is a long exact sequence*

$$\begin{tikzcd}
 & & & {} \arrow[llld, dotted] \\
h^i(A^\bullet) \arrow[r] & h^i(B^\bullet) \arrow[r] & h^i(C^\bullet) \arrow[lld, "\delta^i" description] & \\
h^{i+1}(A^\bullet) \arrow[r] & h^{i+1}(B^\bullet) \arrow[r] & h^{i+1}(C^\bullet) \arrow[lld, dotted] & \\
{} & & &
\end{tikzcd} .$$

Proof. (*Sketch*) The proof relies on chasing an element $c \in \mathrm{Ker}\,(d)$ of C^i till we obtain an element $a \in A^{i+1}$ in $\mathrm{Ker}\,(d)$, in the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & A^{i-1} \arrow[r] \arrow[d, "d"'] & B^{i-1} \arrow[r] \arrow[d, "d"'] & C^{i-1} \arrow[r] \arrow[d, "d"'] & 0 \\
0 \arrow[r] & A^{i} \arrow[r] \arrow[d, "d"'] & B^{i} \arrow[r] \arrow[d, "d"'] & C^{i} \arrow[r] \arrow[d, "d"'] & 0 \\
0 \arrow[r] & A^{i+1} \arrow[r] & B^{i+1} \arrow[r] & C^{i+1} \arrow[r] & 0
\end{tikzcd}$$

The chase is straightforward and is thus omitted. The resultant map is indeed a well-defined group homomorphism. □

We now define homotopy of maps of complexes

Definition 19.2.1.4. (**Homotopy between maps**) Let $\mathbf{A}$ be an abelian category and $f, g : A^\bullet \to B^\bullet$ be two maps of cochain complexes. Then a homotopy between f and g is defined to be a collection of maps $k := \{k^i : A^i \to B^{i-1}\}_{i \in \mathbb{Z}}$ such that $f^i - g^i = dk^i + k^{i+1}d$ for each $i \in \mathbb{Z}$:

$$\begin{tikzcd}
 & A^i \arrow[r, "d"'] \arrow[dl, "k^i"'] \arrow[d, shift right, "g^i"'] \arrow[d, shift left, "f^i"] & A^{i+1} \arrow[dl, "k^{i+1}"] \\
B^{i-1} \arrow[r, "d"'] & B^i &
\end{tikzcd} .$$

As one might expect, homotopic maps induces *same* (not isomorphic, but actually same) maps on cohomology.

Lemma 19.2.1.5. *Let* $\mathbf{A}$ *be an abelian category and* $f, g : A^\bullet \to B^\bullet$ *be two maps of cochain complexes. If* $k : f \sim g$ *is a homotopy between* f *and* g, *then* $h^i(f) = h^i(g)$ *as maps* $h^i(A^\bullet) \to h^i(B^\bullet)$ *for all* $i \in \mathbb{Z}$.

Proof. (*Sketch*) Pick any $a \in \mathrm{Ker}\,(d)$ in A^i. We wish to show that $f^i(a) - g^i(a) \in \mathrm{Im}\,(d)$. This follows from unravelling the definition of homotopy $k : f \sim g$. □

We now define the notion of exact functors between two abelian categories.

Definition 19.2.1.6. (**Exactness of functors**) Let $\mathbf{A}$ and $\mathbf{B}$ be abelian categories. A functor $F : \mathbf{A} \to \mathbf{B}$ is said to be

1. *additive* if the map $\mathrm{Hom}_{\mathbf{A}}(A, B) \to \mathrm{Hom}_{\mathbf{B}}(FA, FB)$ is a group homomorphism,
2. *left exact* if it is additive and for every short exact sequence $0 \to A' \to A \to A'' \to 0$ the sequence $0 \to FA' \to FA \to FA''$ is exact,
3. *right exact* if it is additive and for every short exact sequence $0 \to A' \to A \to A'' \to 0$ the sequence $FA' \to FA \to FA'' \to 0$ is exact,
4. *exact* if it is additive and for every short exact sequence $0 \to A' \to A \to A'' \to 0$ the sequence $0 \to FA' \to FA \to FA'' \to 0$ is exact,
5. *exact at middle* if it is additive and for every short exact sequence $0 \to A' \to A \to A'' \to 0$ the sequence $FA' \to FA \to FA''$ is exact.

Remark 19.2.1.7. It is important to keep in mind that all the above definitions are made for *short* exact sequences; a left exact A functor may not map a long exact sequence $0 \to A_1 \to \dots$ to a long exact sequence $0 \to FA_1 \to \dots$.

There are two prototypical examples of such functors in the category of R-modules.

Example 19.2.1.8. ($- \otimes_R M$ and $\mathrm{Hom}_R(M, -)$) Let R be a commutative ring and M be an R-module. It is a trivial matter to see that the functor $- \otimes_R M : \mathbf{Mod}(R) \to \mathbf{Mod}(R)$ is right exact but not left exact as applying $- \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z}$ on the following shows where $\gcd(n, m) = 1$:

$$0 \to n\mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \to 0.$$

Indeed, $n\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ is not injective as the former is an infinite ring whereas the latter is finite.

Consider the covariant hom-functor $\mathrm{Hom}_R(M, -) : \mathbf{Mod}(R) \to \mathbf{Mod}(R)$. This can easily be seen to be left exact. This is not right exact as applying $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, -)$ to the above exact sequence would yield (note that $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, \mathbb{Z}) = 0$).

We next dualize the above theory study the dual notion of homology, without much change. **TODO.**

19.2.2 Resolutions

We begin with injective objects, resolutions and having enough injectives.

Definition 19.2.2.1. (**Injective objects and resolutions**) Let $\mathbf{A}$ be an abelian category. An object $I \in \mathbf{A}$ is said to be injective if the functor thus represented, $\mathrm{Hom}_{\mathbf{A}}(-, I) : \mathbf{A}^{\mathrm{op}} \to \mathbf{AbGrp}$ is exact. An injective resolution of an object $A \in \mathbf{A}$ is an exact cochain complex

$$A \xrightarrow{\epsilon} I^0 \to I^1 \to \dots$$

where each I^i is an injective object. We denote an injective resolution of A by $\epsilon : A \to I^\bullet$.

The following are equivalent characterizations of injective objects.

Proposition 19.2.2.2. *Let $\mathbf{A}$ be an abelian category and $I \in \mathbf{A}$. Then the following are equivalent*

1. *The functor $\mathrm{Hom}_{\mathbf{A}}(-, I)$ is exact.*

2. *For any monomorphism $i : A \to B$ and any map $f : A \to I$, there is an extension $\tilde{f} : B \to I$ to make following commute*

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, "i"] \arrow[d, "f"'] & B \arrow[dl, dashed, "\tilde{f}"] \\ & I & \end{tikzcd}.$$

3. *Any exact sequence*

$$0 \to I \to A \to B \to 0$$

splits.

Proof. 1. $\Rightarrow$ 2. is immediate from definition. 2. $\Rightarrow$ 3. follows from using the universal property of item 2 on $\mathrm{id} : I \to I$ and monomorphism $0 \to I \to A$. For 3. $\Rightarrow$ 1., we need only check right exactness of $\mathrm{Hom}_{\mathbf{A}}(-, I)$, which follows immediately from item 3. □

The following are some properties of injective objects.

Proposition 19.2.2.3. *Let* $\mathbf{A}$ *be an abelian category. If* $\{I_i\}_i$ *is a collection of injective objects of* $\mathbf{A}$ *and* $\prod_i I_i$ *exists, then it is injective.*

Proof. As $\mathrm{Hom}_{\mathbf{A}}(-, \prod_i I_i) \cong \prod_i \mathrm{Hom}_{\mathbf{A}}(-, I_i)$ and arbitrary product of surjective maps is surjective, therefore the claim follows. □

We see that any two injective resolutions of an object are homotopy equivalent.

Lemma 19.2.2.4. *Let* $\mathbf{A}$ *be an abelian category and* $A \in \mathbf{A}$ *be an object with two injective resolutions* $\epsilon : A \to I^\bullet$ *and* $\eta : A \to J^\bullet$*. Then there exists a homotopy* $k : \epsilon \sim \eta$*.*

Proof. Comparison Theorem 2.3.7, pp 40, [cite Weibel Homological Algebra]. □

We then define when an abelian category has enough injectives.

Definition 19.2.2.5. (**Enough injectives**) An abelian category $\mathbf{A}$ is said to have enough injectives if for each object $A \in \mathbf{A}$, there is an injective object $I \in \mathbf{A}$ such that A is a subobject of I, $A \leq I$.

In such abelian categories, all objects have injective resolutions.

Lemma 19.2.2.6. *Let* $\mathbf{A}$ *be an abelian category with enough injectives. Then all objects* $A \in \mathbf{A}$ *admit injective resolutions* $\epsilon : A \to I^\bullet$*.*

Proof. Pick any object $A \in \mathbf{A}$. As $\mathbf{A}$ has enough injectives, therefore we have $0 \to A \xrightarrow{\epsilon} I^0$. Consider $\mathrm{CoKer}(\epsilon)$ and let it be embedded in some injective object I^1, which yields the following diagram

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, hook, "\epsilon"] & I^0 \arrow[r, "d"] \arrow[d, two heads] & I^1 \\ & & \mathrm{CoKer}(\epsilon) \arrow[ur, hook] & \end{tikzcd}.$$

Continue this diagram by considering $\mathrm{CoKer}\,(d)$ which embeds in some other injective I^2 to further yield the following diagram

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, hook, "\epsilon"] & I^0 \arrow[rr, "d"] \arrow[d, two heads] & & I^1 \arrow[r, "d"] \arrow[d, two heads] & I^2 \\ & & \mathrm{CoKer}\,(\epsilon) \arrow[urr, hook] & & \mathrm{CoKer}\,(d) \arrow[ur, hook] & \end{tikzcd}.$$

This builds the required injective resolution. □

We now give examples of abelian categories with enough injectives. Recall that a *divisible group* G is an abelian group such that for any $g \in G$ and ay $n \in \mathbb{Z}$ there exists $h \in G$ such that $g = nh$ (see Definition 16.13.2.1).

Theorem 19.2.2.7. *Let R be a commutative ring with 1. Then,*

1. *Any divisible group in* **AbGrp** *is an injective object.*
2. *If G is an injective abelian group, then $\mathrm{Hom}_{\mathbb{Z}}\,(R, G)$ is an injective R-module.*
3. **AbGrp** *is an abelian category which has enough injectives.*
4. $\mathbf{Mod}(R)$ *is an abelian category which has enough injectives.*

Proof. The main idea of the proofs of the later parts is to use injective objects constructed in a bigger category and an adjunction to a lower category to construct injectives in the smaller subcategory. Further, embedding each object in a large enough product of injectives (which would remain injective by Proposition 19.2.2.3) would show enough injectivity.

1. By Corollary 16.13.2.3, the statement follows.

2. Recall that $F(-) : \mathbf{Mod}(R) \rightleftarrows \mathbf{AbGrp} : \mathrm{Hom}_{\mathbb{Z}}\,(R, -)$ is an adjunction, where F is the forgetful functor. Consequently $\mathrm{Hom}_{\mathbb{Z}}\,(F(M), G) \cong \mathrm{Hom}_R\,(M, \mathrm{Hom}_{\mathbb{Z}}\,(R, G))$. It then follows that $\mathrm{Hom}_{\mathbb{Z}}\,(R, G)$ is injective.

3. Observe that $\mathbb{Q}/\mathbb{Z}$ is a divisible, thus injective abelian group by item 1. Let G be an abelian group. Consider the abelian group

$$I = \prod_{\mathrm{Hom}_{\mathbb{Z}}(G,\mathbb{Q}/\mathbb{Z})} \mathbb{Q}/\mathbb{Z}.$$

By Proposition 19.2.2.3, I is an injective abelian group. We now construct an injection $\varphi : G \to I$, which would complete the proof. We have the canonical map

$$\begin{aligned} \theta : G &\longrightarrow I \\ g &\longmapsto (\varphi(g))_{\varphi \in \mathrm{Hom}_{\mathbb{Z}}(G,\mathbb{Q}/\mathbb{Z})}. \end{aligned}$$

For this to be well-defined, we need to show that $\mathrm{Hom}_{\mathbb{Z}}\,(G, \mathbb{Q}/\mathbb{Z})$ is non-zero. Indeed, we claim that for any element $g \in G$, there is a $\mathbb{Z}$-linear map $\varphi_g : G \to \mathbb{Q}/\mathbb{Z}$ such that $\varphi_g(g) \neq 0$. This would suffice as if $\theta(g) = 0$ for some $g \in G$, then $\varphi(g) = 0$ for all $\varphi \in \mathrm{Hom}_{\mathbb{Z}}\,(G, \mathbb{Q}/\mathbb{Z})$. Consequently, $\varphi_g(g) = 0$, which cannot happen, hence θ is injective. So we need only show the existence of

φ_g. Indeed, if $|g| = \infty$, then we have an injection $\mathbb{Z} \hookrightarrow G$ taking $1 \mapsto g$. Pick any non-zero map $f : \mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$. By injectivity of $\mathbb{Q}/\mathbb{Z}$, f extends to $\varphi_g : G \to \mathbb{Q}/\mathbb{Z}$ which is non-zero at g. On the other hand, if $|g| = k < \infty$, then consider the inclusion $\mathbb{Z}/k\mathbb{Z} \hookrightarrow G$ taking $\bar{1} \mapsto g$. Then, consider the map $f : \mathbb{Z}/k\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$ taking $\bar{1} \mapsto \frac{1}{k}$. Then, by injectivity of $\mathbb{Q}/\mathbb{Z}$, it extends to $\varphi_g : G \to \mathbb{Q}/\mathbb{Z}$ which is non-zero at g.

4. Pick any R-module M. We wish to find an injective R-module I such that $M \leq I$. By items 1 and 2, we know that $\mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z})$ is an injective R-module. By the proof of item 2, we also know that

$$\mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z}) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z})).$$

Conseuqently, by Proposition 19.2.2.3, we have an injective module

$$I = \prod_{\mathrm{Hom}_{\mathbb{Z}}(M, \mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z}))} \mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z}),$$

We claim that the following map

$$\begin{aligned} \theta : M &\longrightarrow I \\ m &\longmapsto (\varphi(m))_{\varphi \in \mathrm{Hom}_{\mathbb{Z}}(M, \mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z}))} \end{aligned}$$

is injective. Indeed, we claim that for each $m \in M$, there exists $\varphi_m \in \mathrm{Hom}_{\mathbb{Z}}(M, \mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q}/\mathbb{Z}))$ such that $\varphi_m(m) \neq 0$. By the above isomorphism, we equivalently wish to show the existence of $g_m \in \mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z})$ such that $g_m(m) \neq 0$. This is immediate from the proof of item 3. $\square$

We next dualize the above theory and study projective objects, projective resolutions and having enough projectives to define homology. **TODO**.

19.2.3 Derived functors and general properties

First, for each covariant left exact functor $F : \mathbf{A} \to \mathbf{B}$ between abelian categories, we will produce a sequence of functors R^iF for each $i \geq 0$. We will then dualize it.

Definition 19.2.3.1. (**Right derived functors of a left-exact functor**) Let $F : \mathbf{A} \to \mathbf{B}$ be a left exact functor of abelian categories where $\mathbf{A}$ has enough injectives. Then, define for each $i \geq 0$ the following

$$\begin{aligned} R^iF : \mathbf{A} &\longrightarrow \mathbf{B} \\ A &\longmapsto h^i(F(I^\bullet)) \end{aligned}$$

where $\epsilon : A \to I^\bullet$ is any injective resolution of A. We call R^iF the i^{th} right derived functor of the left exact functor F.

Remark 19.2.3.2. Indeed the above definition is well-defined, by Lemmas 19.2.1.5 and 19.2.2.4. Further, keep in mind the Remark 19.2.1.7.

Some of the basic properties of right derived functors are as follows. First, the 0^{th}-right derived functor of F is canonically isomorphic to F.

Lemma 19.2.3.3. *Let* $\mathbf{A}, \mathbf{B}$ *be two abelian categories where* $\mathbf{A}$ *has enough injectives. Let* $F : \mathbf{A} \to \mathbf{B}$ *be a left-exact functor. Then, there is a natural isomorphism*

$$R^0F \cong F.$$

Proof. Pick any object $A \in \mathbf{A}$ with an injective resolution $0 \to A \stackrel{\epsilon}{\to} I^\bullet$. Consequently, $R^0F(A)$ is the cohomology of

$$0 \to F(I^0) \stackrel{Fd^0}{\to} F(I^1),$$

that is, $R^0F(A) = \mathrm{Ker}\left(Fd^0\right)$. But since F is left-exact and we have the following exact sequence

$$0 \to A \stackrel{\epsilon}{\to} I^0 \stackrel{d^0}{\to} \mathrm{Im}\left(d^0\right),$$

therefore we get that $\mathrm{Ker}\left(Fd^0\right) = \mathrm{Im}\left(F\epsilon\right)$. This also needs the observation that if F is left-exact, then for any map $f \in \mathbf{A}$, we have $F(\mathrm{Im}\,(f)) \cong \mathrm{Im}\,(Ff)$. Since ϵ is injective, then so is $F\epsilon$ and thus $\mathrm{Im}\,(F\epsilon) \cong FA$. $\square$

Remark 19.2.3.4. Let $I \in \mathbf{A}$ be an injective object. Then we claim that $R^iF(I) = 0$ for all $i \geq 1$. Indeed, this follows immediately because we have $0 \to I \stackrel{\mathrm{id}}{\to} I \to 0$ as a trivial injective resolution of I.

The following is an important property of right derived functors which makes them ideal for defining the general notion of cohomology, because they always have long exact sequene in cohomology.

Theorem 19.2.3.5. *Let* $\mathbf{A}, \mathbf{B}$ *be two abelian categories where* $\mathbf{A}$ *has enough injectives. Let* $F : \mathbf{A} \to \mathbf{B}$ *be a left-exact functor. If*

$$0 \to A \to B \to C \to 0$$

is a short exact sequence in $\mathbf{A}$*, then we have a long exact sequence in right derived functors of* F *as in*

$$\begin{tikzcd} 0 \arrow[r] & R^0FA \arrow[r] & R^0FB \arrow[r] & R^0FC \arrow[dll, "\delta_0"'] \\ & R^1FA \arrow[r] & R^1FB \arrow[r] & R^1FC \arrow[dll, dotted, "\delta_1"'] \\ & {} & & \end{tikzcd}.$$

It follows from above theorem that if F is exact, then R^iF are trivial for $i \geq 1$.

Corollary 19.2.3.6. *Let* $\mathbf{A}, \mathbf{B}$ *be two abelian categories where* $\mathbf{A}$ *has enough injectives. Let* $F : \mathbf{A} \to \mathbf{B}$ *be an exact functor. Then,*

$$R^iF = 0$$

for all $i \geq 1$.

Proof. Pick any object $A \in \mathbf{A}$ and let $I \in A$ be an an injective object such that $0 \to A \to I$ is an injective map. Then we have a short-exact sequence

$$0 \to A \to I \to B \to 0$$

where $B = I/A$. By Theorem 19.2.3.5, Lemma 19.2.3.3 and Remark 19.2.3.4, it follows that we have a long exact sequence in right derived functors of F as in

$$\begin{tikzcd} R^iFA \arrow[r] & 0 \arrow[r] & R^iFB \arrow[dll, "\delta_i"'] \\ R^{i+1}FA \arrow[r] & 0 \arrow[r] & R^{i+1}FB \end{tikzcd}.$$

It follows from exactness of the above sequence that $R^iFB \cong R^{i+1}FA$ for all $i \geq 1$. Repeating the same process for B (embedding B into an injective object and observing the resultant long exact sequence), we obtain that

$$R^{i+1}FA \cong R^1FC$$

for some object $C \in \mathbf{A}$. Replacing A by C, it thus suffices to show that $R^1FA = 0$.

In the beginning of the above long exact sequence we have

$$\begin{tikzcd} 0 \arrow[r] & FA \arrow[r] & FI \arrow[r, two heads] & FB \arrow[dll, "\delta_0"'] \\ & R^1FA \arrow[r] & 0 \arrow[r] & R^1FB \end{tikzcd}$$

from which it follows via exactness that δ_0 is surjective and $\operatorname{Ker}(\delta_0) = FB$. We then deduce that $R^1FA = 0$, as required. $\square$

Injective resolutions might be hard to find in general, but given a left exact functor F, it would be somewhat easier to find objects J such that $R^iF(J) = 0$ for all $i \geq 1$. The remarkable property of such objects is that it can help to calculate the value of right derived functors of F for objects admitting resolutions by them.

Definition 19.2.3.7 (Acyclic resolution). Let $\mathbf{A}, \mathbf{B}$ be two abelian categories where $\mathbf{A}$ have enough injectives. Let $F : \mathbf{A} \to \mathbf{B}$ be a left-exact functor. An object $J \in \mathbf{A}$ is said to be acyclic if $R^iF(J) = 0$ for all $i \geq 1$. An acyclic resolution of $A \in \mathbf{A}$ is an exact sequence of the form

$$0 \to A \xrightarrow{\epsilon} J^0 \to J^1 \ldots$$

where each J^i is acyclic.

The name "acyclic" is justified since they have zero cohomology, so all cocycles are coboundaries, so there are no cycles for that object.

Remark 19.2.3.8. Note that for an acyclic resolution $0 \to A \xrightarrow{\epsilon} J^\bullet$, we have $h^0(F(J^\bullet)) \cong FA$ by following the steps as in the proof of Lemma 19.2.3.3.

We then have the following useful theorem.

Proposition 19.2.3.9. *Let* $\mathbf{A}, \mathbf{B}$ *be two abelian categories where* $\mathbf{A}$ *have enough injectives. Let* $F : \mathbf{A} \to \mathbf{B}$ *be a left-exact functor. For* $A \in \mathbf{A}$*, let* $0 \to A \xrightarrow{\epsilon} J^\bullet$ *be an acyclic resolution. Then for all* $i \geq 0$*, there is a natural isomorphism*

$$R^i F(A) \cong h^i(F(J^\bullet)).$$

Derived functors are equivalent to datum of what is defined to be a universal δ-functor. In the rest of this section we setup the definitions and only state the result.

TODO : Universal δ-functors.

19.3 Results for Mod(R)

When the abelian category is that of modules over a commutative ring R, then we have some special results which is very useful in homotopy theory.

19.3.1 Universal coefficients

19.3.2 Künneth theorem

19.3.3 $\otimes$-Hom adjunction

Chapter 20

Foundational Sheaf Theory

Contents

The notion of sheaves plays perhaps the most important role in modern viewpoint of geometry. It is thus important to understand the various constructions that one can make on them. We assume the reader knows the definition of a sheaf on a space X and morphism of sheaves. We begin with some recollections.

20.1 Recollections

Remark 20.1.0.1 (*Map on stalks*)**.** Recall that a map of sheaves $\varphi : \mathcal{F} \to \mathcal{G}$ on X defines for each point $x \in X$ a map of stalks $\varphi_x : \mathcal{F}_x \to \mathcal{G}_x$ given by $s_x \mapsto \varphi_U(s)_x$ where s is a section of $\mathcal{F}$ over $U \subseteq X$. One can check quite easily that this is well-defined and that this map φ_x is in-fact the unique map given by the universal property of the colimit in the diagram below:

$$\begin{tikzcd} \mathcal{G}(U) \arrow[r] & \varinjlim_{V \ni x} \mathcal{G}(V) \\ \mathcal{F}(U) \arrow[r] \arrow[u, "\varphi_U"] & \varinjlim_{V \ni x} \mathcal{F}(V) \arrow[u, dashed, "\varphi_x"'] \end{tikzcd}.$$

Hence φ_x is the unique map which makes the above diagram commute.

Remark 20.1.0.2 (*Subsheaves*)**.** Recall that $\mathcal{F} \hookrightarrow \mathcal{G}$ is a subsheaf if $\mathcal{F}(U) \subseteq \mathcal{G}(U)$ such that for $U \hookrightarrow V$, the restriction map $\rho_{V,U} : \mathcal{G}(V) \to \mathcal{G}(U)$ restricts to $\rho_{V,U} : \mathcal{F}(V) \to \mathcal{F}(U)$.

Remark 20.1.0.3 (*Constant sheaves*)**.** For an abelian group A and a space X, one defines the constant sheaf $\mathcal{A}$ as the sheaf which for each open set $U \subseteq X$ assigns $\mathcal{A}(U) = \{s : U \to A \mid s \text{ is continuous}\}$, where A is given the discrete topology. One sees instantly that this is a sheaf. Further one observes that if $U = U_1 \amalg \cdots \amalg U_k$ where U_i are components of open set U and U_i are open, then $\mathcal{A}(U) \cong A \oplus \cdots \oplus A$ k-times. In particular, for any open connected subset U, we get $\mathcal{A}(U) = A$.

We now begin by showing how to construct a sheaf out of a presheaf over X.

20.2 The sheafification functor

Let X be a topological space, denote the category of presheaves on X by $\mathbf{PSh}(X)$ and denote the category of sheaves over X by $\mathbf{Sh}(X)$. We have a canonical inclusion functor $i : \mathbf{Sh}(X) \hookrightarrow \mathbf{PSh}(X)$. We construct it's left adjoint commonly known as the process of sheafifying a presheaf.

Theorem 20.2.0.1. *(Sheafification) Let X be a topological space and let F be a presheaf over X. Then there exists a pair $(\mathcal{F}, i_F)$ of a sheaf $\mathcal{F}$ and a map $i_F : F \to \mathcal{F}$ such that for any sheaf $\mathcal{G}$ and a morphism of presheaves $\varphi : F \to \mathcal{G}$, there exists a unique morphism of sheaves $\tilde{\varphi} : \mathcal{F} \to \mathcal{G}$ such that the following commutes*

$$\begin{tikzcd} \mathcal{F} \arrow[r, dashed, "\tilde{\varphi}"] & \mathcal{G} \\ F \arrow[u, "i_F"] \arrow[ur, "\varphi"'] & \end{tikzcd},$$

that is, we have a natural bijection

$$\operatorname{Hom}_{\mathbf{PSh}(X)}(F, \mathcal{G}) \cong \operatorname{Hom}_{\mathbf{Sh}(X)}(\mathcal{F}, \mathcal{G}).$$

Moreover:

1. *$(\mathcal{F}, i_F)$ is unique upto unique isomorphism.*
2. *For every $x \in X$, the map on stalks $i_{F,x} : F_x \to \mathcal{F}_x$ is bijective.*
3. *For any map of presheaves $\varphi : F \to G$, we get a map of sheaves $\tilde{\varphi} : \mathcal{F} \to \mathcal{G}$ which is unique w.r.t. the commuting of the following natural square:*

$$\begin{tikzcd} \mathcal{F} \arrow[r, dashed, "\tilde{\varphi}"] & \mathcal{G} \\ F \arrow[u, "i_F"] \arrow[r, "\varphi"'] & G \arrow[u, "i_G"'] \end{tikzcd}.$$

Hence we have a functor

$$\begin{aligned}(-)^{++} : \mathbf{PSh}(X) &\longrightarrow \mathbf{Sh}(X)\\ F &\longmapsto F^{++} := \mathcal{F}.\end{aligned}$$

Proof. We explicitly construct the sheaf $\mathcal{F}$ out of F. We define the local sections of $\mathcal{F}$ by using germs and turning the gluing condition of sheaf definition onto itself. In particular, define

$$\mathcal{F}(U) := \left\{ ((s_{i_x})_x) \in \prod_{x\in U} F_x \mid \forall x \in U,\ \exists \text{ open } W \ni x \ \&\ t \in F(W) \text{ s.t. } \forall p \in W,\ t_p = (s_{i_p})_p \right\}.$$

The restriction map for $U \hookrightarrow V$ of $\mathcal{F}$ is given by $\rho_{V,U} : \mathcal{F}(V) \to \mathcal{F}(U)$, $((s_{i_x})_x) \mapsto ((s_{i_x})_x)_{x\in U}$, that is, $\rho_{V,U}$ is just the projection map. Next, we show that $\mathcal{F}$ satisfies the gluing criterion and that is where we will see how the above definition of sections of $\mathcal{F}$ came about. Take an open set $U \subseteq X$ and an open cover $U = \bigcup_{i\in I} U_i$. Let $s_i \in \mathcal{F}(U_i)$ be a corresponding collection of sections such that for all $i, j \in I$, we have $\rho_{U_i,U_i\cap U_j}(s_i) = \rho_{U_j,U_i\cap U_j}(s_j)$. We wish to thus construct a section $t \in \mathcal{F}(U)$ such that $\rho_{U,U_i}(t) = s_i$ for all $i \in I$. Indeed let $((t_{i_x})_x) \in \prod_{x\in U} F_x$ where $t := (t_{i_x})_x = (s_i)_x$ if $x \in U_i$. Then since for any $x \in U$, there exists $U \supseteq U_i \ni x$ and $s_i \in \mathcal{F}(U_i)$ such that $\rho_{U,U_i}(t) = ((t_{i_x})_x)_{x\in U_i} = ((s_i)_x)_{x\in U_i}$, we thus conclude that $t \in \mathcal{F}(U)$. So $\mathcal{F}$ satisfies the gluing condition. The locality is quite simple. Next the map i_F is given as follows on sections:

$$\begin{aligned}i_{F,U} : F(U) &\longrightarrow \mathcal{F}(U)\\ s &\longmapsto (s_x).\end{aligned}$$

Now, it can be seen by definition of colimits that $\mathcal{F}_x = F_x$. Finally, let $\mathcal{G}$ be a sheaf and let $\varphi : F \to \mathcal{G}$ be a morphism of presheaves, then we can define $\tilde{\varphi}$ by gluing the germs as follows:

$$\begin{aligned}\tilde{\varphi}_U : \mathcal{F}(U) &\longrightarrow \mathcal{G}(U)\\ ((s_{i_x})_x) &\longmapsto [\varphi_{W_x}(s_{i_x})]\end{aligned}$$

where $[\varphi_{W_x}(s_{i_x})]$ denotes the unique section in $\mathcal{G}(U)$ that one gets by considering the open cover $\bigcup_{x\in U} W_x$ where $s_{i_x} \in \mathcal{F}(W_x)$ and considering the gluing of corresponding sections $\varphi_{W_x}(s_{i_x}) \in \mathcal{G}(W_x)$. These sections agree on intersections because φ is a natural transformation and (s_{i_x}) agree on intersections as sections of $\mathcal{F}(U)$. Hence we have the unique map $\tilde{\varphi}$. Moreover, it is clear that $\tilde{\varphi} \circ i_F = \varphi$. □

Corollary 20.2.0.2. *Let F be a presheaf over a topological space X, then for all $x \in X$, $F_x = (F^{++})_x$.*

Proof. By construction of F^{++}. □

Corollary 20.2.0.3. *If $\mathcal{F}$ is a sheaf over a topological space X, then $\mathcal{F}^{++} = \mathcal{F}$.*

Proof. Follows immediately from the universal property of the sheafification, Theorem 20.2.0.1. □

Remark 20.2.0.4. The sections of sheaf $\mathcal{F}$ in an open set U containing x is defined in such a manner so that $f \in \mathcal{F}(U)$ can be constructed locally out of sections of F. In particular, we can write $\mathcal{F}(U)$ more clearly as follows

$$\mathcal{F}(U) = \left\{ s : U \to \coprod_{x \in U} F_x \mid \forall x \in U,\ s(x) \in F_x \ \&\ \exists \text{ open } x \in V \subseteq U \ \&\ \exists t \in F(V) \text{ s.t. } s(y) = t_y \ \forall y \in V \right\}.$$

Note that this is exactly the realization that $\mathcal{F}(U)$ is the set of section of the étale space of the sheaf $\mathcal{F}$ (see Section 20.4). Most of the time in practice, we would work with the universal property of $\mathcal{F}$ in Theorem 20.2.0.1 as it is much more amenable, but the above must be kept in mind as it is used, for example, to make sure that certain algebraic constructions of $\mathcal{O}_X$-modules remains $\mathcal{O}_X$-modules (no matter how trivial they may sound).

We note that sheafification and restrictions to open sets commute.

Lemma 20.2.0.5. *Let X be a space, $U \subseteq X$ be an open subset and F be a presheaf over X. Then,*

$$(F|_U)^{++} \cong (F^{++})|_U .$$

Proof. Immediate from universal property of sheafification (Theorem 20.2.0.1). □

20.3 Morphisms of sheaves

All sheaves are abelian sheaves in this section. One of the most important aspects of using sheaves is that the injectivity and bijectivity of φ_x can be checked on sections. We first show that taking stalks is functorial

Lemma 20.3.0.1. *Let X be a topological space, $\mathcal{F}, \mathcal{G}$ be two sheaves over X and $x \in X$ be a point. Then the following mapping is functorial:*

$$\begin{aligned} \mathbf{Sh}(X) &\longrightarrow \mathbf{AbGrp} \\ \mathcal{F} &\longmapsto \mathcal{F}_x \\ \mathcal{F} \xrightarrow{f} \mathcal{G} &\longmapsto \mathcal{F}_x \xrightarrow{f_x} \mathcal{G}_x. \end{aligned}$$

Proof. Immediate, just remember how composition of two natural transforms is defined. □

Another simple lemma about sheaves and stalks is that equality of two sections can be checked at the stalk level.

Lemma 20.3.0.2. *Let X be a topological space and $\mathcal{F}$ be a sheaf over X. If $s, t \in \mathcal{F}(U)$ for some open $U \subseteq X$ such that $(U, s)_x = (U, t)_x \ \forall x \in U$, then $s = t$ in $\mathcal{F}(U)$.*

Proof. By equality on stalks, it follows that we have an open set $W_x \ni x$ in U for all $x \in U$ such that $\rho_{U,W_x}(s) = \rho_{U,W_x}(t)$. The result follows from the unique gluing property of sheaf $\mathcal{F}$. □

The above result therefore show why almost all the time it is enough to work with stalks in geometry. Let us now define an injective and surjective map of sheaves.

Definition 20.3.0.3. (**Injective & surjective maps**) Let X be a topological space and $\mathcal{F}, \mathcal{G}$ be two sheaves on X. A map of sheaves $f : \mathcal{F} \to \mathcal{G}$ is said to be

1. *injective* if for all opens $U \subseteq X$, the local homomorphism $f_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective,
2. *surjective* if for all opens $U \subseteq X$ and all $s \in \mathcal{G}(U)$, there exists an open covering $\{U_i\}_{i \in I}$ such that $\rho_{U,U_i}(s) \in \operatorname{Im}(f_{U_i})$,
3. *bijective* if f is injective and surjective.

Heuristically, one may understand the notion of f being surjection by saying that every local section of $\mathcal{G}$ is locally constructible by the image of $\mathcal{F}$ under the map f.

For each map of sheaves, we can also define two corresponding sheaves which are global algebraic analogues of the local algebraic constructions.

Definition 20.3.0.4. (**Quotient sheaf**) Let X be a topological space and $\mathcal{F}$ be a sheaf on X. For a subsheaf $\mathcal{S} \subseteq \mathcal{F}$, one defines the quotient sheaf $\mathcal{F}/\mathcal{S}$ as the sheafification of the presheaf F/S defined on open sets $U \subseteq X$ by

$$F/S(U) := \mathcal{F}(U)/\mathcal{S}(U).$$

Definition 20.3.0.5. (**Image & kernel sheaves**) Let X be a topological space and $\mathcal{F}, \mathcal{G}$ be two sheaves over X and $f : \mathcal{F} \to \mathcal{G}$ be a morphism. Then,

1. *image sheaf* is the sheafification of the presheaf $\operatorname{Im}(f)$ defined on open sets $U \subseteq X$ by

$$(\operatorname{Im}(f))(U) := \operatorname{Im}(f_U),$$

   and we denote it by the same symbol, $\operatorname{Im}(f)$,
2. *kernel sheaf* is the sheafification of the presheaf $\operatorname{Ker}(f)$ defined on open sets $U \subseteq X$ by

$$(\operatorname{Ker}(f))(U) := \operatorname{Ker}(f_U)$$

   and we denote it by the same symbol, $\operatorname{Ker}(f)$.

In both the above definitions, the important aspect is the sheafification of the canonical presheaves.

The main point is that one can check all the three notions introduced in Definition 20.3.0.3 for $f : \mathcal{F} \to \mathcal{G}$ by checking on stalks $f_x : \mathcal{F}_x \to \mathcal{G}_x$ for all $x \in X$.

Theorem 20.3.0.6. [1] *Let X be a topological space and $\mathcal{F}, \mathcal{G}$ be two sheaves over X. Then, a map $f : \mathcal{F} \to \mathcal{G}$ is*

1. *injective if and only if $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is injective for all $x \in X$,*
2. *surjective if and only if $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is surjective for all $x \in X$,*
3. *bijective if and only if $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is bijective for all $x \in X$,*
4. *an isomorphism if and only if $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is bijective for all $x \in X$*[2].
5. *an isomorphism if and only if $f : \mathcal{F} \to \mathcal{G}$ is bijective.*

[1] Exercise II.1.2, II.1.3 and II.1.5 of Hartshorne.

[2] In general, we should write "... *if and only if $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is an isomorphism*", but since we are in the setting of abelian sheaves and bijective homomorphism of abelian groups is an isomorphism, so we can get away with this.

Proof. The proof is more of an exercise to get a familiarity with the flexibility of sheaf language. The main idea almost everywhere is to do some local calculations and use sheaf axioms to construct a unique section out of local sections.

1. (L $\Rightarrow$ R) We wish to show that f_x is injective. Suppose for two $(U,s)_x, (V,t)_y \in \mathcal{F}_x$ we have $f_x((U,s)_x) = f_x((V,t)_x) \in \mathcal{G}_x$, which translates to $(U, f_U(s))_x = (V, f_U(t))_x$. We wish to show that $(U,s)_x = (V,t)_y$. By definition of equality on stalks, we obtain open $W \subseteq U \cap V$ containing x such that

$$\rho_{U,W}(f_U(s)) = \rho_{V,W}(f_V(t)).$$

By the fact that f is a natural transformation, we further translate the above equality to

$$f_W(\rho_{U,W}(s)) = f_W(\rho_{V,W}(t)).$$

By injectivity of homomorphism f_W, we obtain

$$\rho_{U,W}(s) = \rho_{V,W}(t)$$

in $\mathcal{F}(W)$. Hence by the definition of equality on stalks, we obtain $(U,s)_x = (V,t)_x$.

(R $\Rightarrow$ L) Pick any open $U \subseteq X$. We wish to show that $f_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective. Let $s \in \mathcal{F}(U)$ be such that $f_U(s) = 0$. Thus for all $x \in U$, we have $(U, f_U(s))_x = 0$. Further, by definition of the map f_x, we obtain $f_x((U,s))_x = (U, f_U(s))_x = 0$. By injectivity of f_x, we obtain $(U,s)_x = 0$ for all $x \in U$[3]. By definition of equality on stalks, we obtain an open cover $\{W_x\}_{x\in U}$ such that $x \in W_x$ and $s|_{W_x} := \rho_{U,W_x}(s) = 0$. Since f is a natural transformation, we therefore obtain that $\{s|_{W_x}\}_{x\in U}$ is a matching family, i.e. on intersections of W_x, W_y, the corresponding sections agree. Hence, there is a unique glue of $\{s|_{W_x}\}_{x\in U}$ denote $t \in \mathcal{F}(U)$. Since each $s|_{W_x} = 0$, therefore we have two glues of the family over U, one is 0 and the other is s. By uniqueness of the glue, it follows that $s = 0$.

2. (L $\Rightarrow$ R) Pick any $x \in X$. We wish to show that $f_x : \mathcal{F}_x \to \mathcal{G}_x$ is surjective. Pick any $(V,t)_x \in \mathcal{G}_x$. We wish to show that for some open $U \ni x$, we have $(U,s)_x \in \mathcal{F}_x$ such that

$$(V,t)_x = (U, f_U(s))_x.$$

Since $t \in \mathcal{G}(V)$, therefore by surjectivity of f that there exists an open cover $\{V_i\}_{i\in I}$ of V such that

$$\rho_{V,V_i}(t) \in \mathrm{Im}\,(f_{V_i}).$$

Therefore we may pick $s_i \in \mathcal{F}(V_i)$ such that

$$\begin{aligned}\rho_{V,V_i}(t) &= f_{V_i}(s_i)\\ &= f_{V_i}(\rho_{V_i,V_i}(s_i))\\ &= \rho_{V_i,V_i}(f_{V_i}(s_i)).\end{aligned}$$

Thus, $(V,t)_x$ and $(V_i, f_{V_i}(s_i))_x$ are same.

[3]We could be done right here by Lemma 20.3.0.2.

(R $\Rightarrow$ L) We wish to show that $f : \mathcal{F} \to \mathcal{G}$ is surjective. Pick any open set $V \subseteq X$ and $t \in \mathcal{G}(V)$. We wish to find an open cover $\{W_i\}$ of V such that $s_i \in \mathcal{F}(V_i)$ and $f_{V_i}(s_i) = \rho_{V,V_i}(t)$. Since we have $(V,t)_x \in \mathcal{G}_x$ for all $x \in V$, therefore by surjectivity of each $f_x : \mathcal{F}_x \to \mathcal{G}_x$, we obtain germs $(W_x, s_x)_x \in \mathcal{F}_x$ such that $(W_x, f_{W_x}(s_x))_x = (V,t)_x$ for all $x \in V$. By shrinking W_x and restricting s_x, we may assume $\{W_x\}$ covers V. Thus we have an open cover of V such that for all $s_x \in \mathcal{F}(W_x)$, we have $f_{W_x}(s_x) = \rho_{V,W_x}(t)$.

3. Trivially follows from 1. and 2.
4. (L $\Rightarrow$ R) Use the fact that taking stalks is a functor (Lemma 20.3.0.1).
   (R $\Rightarrow$ L) Let $g_x : \mathcal{G}_x \to \mathcal{F}_x$ be the inverse homomorphism of f_x for each $x \in X$. Using g_x, we can easily construct a sheaf homorphism $g : \mathcal{G} \to \mathcal{F}$ which will be the inverse of f. Indeed, consider the following map for any open $U \subseteq X$

$$\begin{aligned} g_U : \mathcal{G}(U) &\longrightarrow \mathcal{F}(U) \\ t &\longmapsto s \end{aligned}$$

   where $s \in \mathcal{F}(U)$ is formed as the unique glue of the matching family

$$\{s_x \in \mathcal{F}(U_x)\}_{x \in U}$$

   where $(U,t)_x = (U_x, f_{U_x}(s_x))_x$ for each $x \in U$ and $U_x \subseteq U$. In particular, $s_x = g_x((U_x, \rho_{U,U_x}(t))_x)$. This is obtained via the bijectivity of f_x. Consequently, g is a sheaf homomorphism, which is naturally the inverse of f.
5. Follows from 3. and 4.

□

The following theorem further tells us that our intuition about algebra can be globalized, and equality of sheaf morphisms can be checked on each stalk.

Theorem 20.3.0.7. *Let X be a topological space and $\mathcal{F}, \mathcal{G}$ be two sheaves over X. Then, a map $f : \mathcal{F} \to \mathcal{G}$*
1. *is injective if and only if the kernel sheaf* $\mathrm{Ker}\,(f)$ *is the zero sheaf,*
2. *is surjective if and only if the image sheaf* $\mathrm{Im}\,(f)$ *is $\mathcal{G}$,*
3. *is equal to another map $g : \mathcal{F} \to \mathcal{G}$ if and only if $f_x = g_x$ for all $x \in X$.*

Proof. The main idea in most of the proofs below is to either use the definition or the universal property of sheafification.

1. (L $\Rightarrow$ R) Let $f : \mathcal{F} \to \mathcal{G}$ be injective. We wish to show that $\mathrm{Ker}\,(f) = 0$. Since the kernel presheaf $\ker f = 0$, therefore its sheafification $\mathrm{Ker}\,(f) = 0$.
   (R $\Rightarrow$ L) Let $\mathrm{Ker}\,(f) = 0$. We wish to show that f is injective. Suppose to the contrary that f is not injective. We have that $(\mathrm{Ker}\,(f))_x = 0$ for all $x \in X$. Thus there exists an open set $U \subseteq X$ such that $f_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is not injective. Hence, there exists, $0 \neq s \in \mathcal{F}(U)$ such that $f_U(s) = 0$. Thus, we have an element $(U,s)_x \in (\ker f)_x = (\mathrm{Ker}\,(f))_x = 0$ for all $x \in U$. Hence $s = 0$ by Lemma 20.3.0.2, which is a contradiction.
2. (L $\Rightarrow$ R) Let $f : \mathcal{F} \to \mathcal{G}$ be a surjective map. In order to show that $\mathrm{Im}\,(f) = \mathcal{G}$, we will show that $\mathcal{G}$ satisfies the universal property of sheafification (Theorem 20.2.0.1). For this, consider a sheaf $\mathcal{H}$ and a presheaf map $h : \mathrm{im}\,(f) \to \mathcal{H}$. Consider the inclusion map $\iota : \mathrm{im}\,(f) \hookrightarrow \mathcal{G}$.

We will construct a unique sheaf map $\tilde{h} : \mathcal{G} \to \mathcal{H}$ which will be natural such that $\tilde{h} \circ \iota = h$. Pick any open set $U \subseteq X$. We wish to define the map

$$\tilde{h}_U : \mathcal{G}(U) \longrightarrow \mathcal{H}(U).$$

Take $t \in \mathcal{G}(U)$. By surjectivity of f, there exists a covering $\{U_i\}$ of U and matching family $s_i \in \mathcal{F}(U_i)$ for all i such that

$$f_{U_i}(s_i) = \rho_{U,U_i}(t) =: t_i.$$

We shall construct an element $\tilde{h}_U(t) \in \mathcal{H}(U)$. Indeed, we first claim that

$$\{h_{U_i}(t_i) \in \mathcal{H}(U_i)\}_i$$

is a matching family. This can be shown by keeping the following diagram in mind and the fact that $\{s_i\}$ is a matching family:

$$\begin{tikzcd} \mathcal{H}(U_i) \arrow[d, "\rho_{U_i,U_i\cap U_j}"'] & \mathrm{im}\,(f_{U_i}) \arrow[l, "h_{U_i}"'] \arrow[d, "\rho_{U_i,U_i\cap U_j}"] & \mathcal{F}(U_i) \arrow[l, "f_{U_i}"'] \arrow[d, "\rho_{U_i,U_i\cap U_j}"] \\ \mathcal{H}(U_i \cap U_j) & \mathrm{im}\,\left(f_{U_i\cap U_j}\right) \arrow[l, "h_{U_i\cap U_j}"] & \mathcal{F}(U_i \cap U_j) \arrow[l, "f_{U_i\cap U_j}"] \end{tikzcd}.$$

Thus we get a unique glue which we define to be the image of $\tilde{h}_U$ for the section $t \in \mathcal{G}(U)$, denoted $\tilde{h}_U(t) \in \mathcal{H}(U)$. Uniqueness and naturality follows from construction.

(R $\Rightarrow$ L) We have that $(\mathrm{im}\,(f))^{++} = \mathcal{G}$. Pick any open set $U \subseteq X$ and a section $t \in \mathcal{G}$. We wish to find an open cover $\{U_i\}_{i\in I}$ of U and $s_i \in \mathcal{F}(U_i)$ such that $f_{U_i}(s_i) = \rho_{U,U_i}(t)$ for all $i \in I$. Indeed, by Corollary 20.2.0.2, we obtain that $\mathcal{G}_x = \mathrm{im}\,(f)_x$ for all $x \in X$. Hence for the chosen (U,t), we obtain for each $x \in U$, by appropriately shrinking and restricting, an open set $W_x \subseteq U$ containing x and a section $s_x \in \mathcal{F}(W_x)$ satisfying $\rho_{U,W_x}(t) = f_{W_x}(s_x)$.

3. (L $\Rightarrow$ R) Trivial.
(R $\Rightarrow$ L) Suppose for all $x \in X$ we have $f_x = g_x : \mathcal{F}_x \to \mathcal{G}_x$. We wish to show that $f = g$. Pick an open set $U \subseteq X$ and consider $s \in \mathcal{F}(U)$. We wish to show that $f_U(s) = g_U(s)$. For each $x \in U$, we have $(U,s)_x \in \mathcal{F}_x$ and by the fact that $f_x = g_x$, we further have

$$(U, f_U(s))_x = (U, g_U(s))_x.$$

Hence for all $x \in U$, there exists open $x \in W_x \subseteq U$ such that

$$\rho_{U,W_x}(f_U(s)) = \rho_{U,W_x}(g_U(s)).$$

It is then an easy observation that both $\{\rho_{U,W_x}(g_U(s))\}_{x\in U}$ and $\{\rho_{U,W_x}(f_U(s))\}_{x\in U}$ forms the same matching family. Hence we have a unique glue by sheaf axiom of $\mathcal{G}$ to obtain $f_U(s) = g_U(s)$ in $\mathcal{G}(U)$.

$\square$

Lemma 20.3.0.8. *Let X be a topological space. Then, the following are equivalent:*

1. *The following is an exact sequence of sheaves over X*

$$\mathcal{F}' \xrightarrow{f} \mathcal{F} \xrightarrow{g} \mathcal{F}'',$$

that is, $\mathrm{Ker}\,(g) = \mathrm{Im}\,(f)$.

2. *The following is an exact sequence of stalks for each $x \in X$*

$$\mathcal{F}'_x \xrightarrow{f_x} \mathcal{F}_x \xrightarrow{g_x} \mathcal{F}''_x.$$

Proof. (1. $\Rightarrow$ 2.) Pick any $(U,s)_x \in \mathcal{F}_x$ which is in $\ker g_x$. Thus, there exists $V \subseteq U$ open such that $\rho_{U,V}(g_U(s)) = g_V(\rho_{U,V}(s)) = 0$. Thus $\rho_{U,V}(s) \in \mathcal{F}(V)$ is in $\mathrm{Ker}\,(g) = \mathrm{Im}\,(f)$ and thus $(V, \rho_{U,V}(s))_x = (U,s)_x \in \mathcal{F}_x$ is in $\mathrm{im}\,(f_x)$. Conversely, for $(U, f_x(t))_x \in \mathrm{im}\,(f_x)$, we see that since $g \circ f = 0$, then $(U, g_x(f_x(t)))_x = 0$.

(2. $\Rightarrow$ 1.) This is immediate, by looking at a section of $\mathcal{F}$ at any open set (use Remark 20.2.0.4). $\square$

Given an open subset U of X and a sheaf over U, we can extend it to a sheaf over X by zeros. This in particular means extending a sheaf from a subspace in such a way so that stalks outside of the subspace are always zero. This operation would be fundamental in cohomology and other places as it yields a nice exact sequence corresponding to any closed or open subset of X.

Definition 20.3.0.9 (Extending a sheaf by zeros). Let X be a space and $i : Z \hookrightarrow X$ be an inclusion of a closed set and $j : U \hookrightarrow X$ be an inclusion of an open set.

1. If $\mathcal{F}$ is a sheaf over Z, then $i_*\mathcal{F}$ is a sheaf over X called the extension of $\mathcal{F}$ to X by zeros.
2. If $\mathcal{F}$ is a sheaf over U, then the extension of $\mathcal{F}$ to X by zeroes, denoted $j_!\mathcal{F}$ is the sheafification of the presheaf over X given by

$$V \longmapsto \begin{cases} \mathcal{F}(V) & \text{if } V \subseteq U \\ 0 & \text{else.} \end{cases}$$

The main result is as follows.

Proposition 20.3.0.10. [4] *Let X be a space, $i : Z \hookrightarrow X$ be closed and $j : U \hookrightarrow X$ be open. Then,*

1. *If $\mathcal{F}$ is a sheaf over Z, then for any $p \in X$, we have*

$$(i_*\mathcal{F})_p = \begin{cases} \mathcal{F}_p & \text{if } p \in Z \\ 0 & \text{if } p \notin Z. \end{cases}$$

2. *If $\mathcal{F}$ is a sheaf over U, then for any $p \in X$, we have*

$$(j_!\mathcal{F})_p = \begin{cases} \mathcal{F}_p & \text{if } p \in U \\ 0 & \text{if } p \notin U. \end{cases}$$

Moreover, $(j_!\mathcal{F})_{|U} = \mathcal{F}$ and $j_!\mathcal{F}$ is unique w.r.t these two properties.

[4]Exercise II.1.19 of Hartshorne.

Proof. The first item follows immediately from the fact that $Z \subseteq X$ is a closed subset. In particular, if $p \notin Z$, then there is a cofinal collection of open sets containing p on which $i_*\mathcal{F}$ is 0.

For the second item, we proceed as follows. Let G be the presheaf as in Definition 20.3.0.9, 2. Note that

$$G_p = \begin{cases} \mathcal{F}_p & \text{if } p \in V \subseteq U \text{ for some open } V \subset X, \\ 0 & \text{else.} \end{cases}$$

In particular, if $p \in U$, then $G_p = \mathcal{F}_p$ and if $p \notin U$, then $G_p = 0$. Since stalks before and after sheafification are same, therefore we have our result for stalks. Next, $(j_!\mathcal{F})_{|U} = \mathcal{F}$ because over U, the presheaf $G_{|U}$ itself is a sheaf, so sheafification of G will yield a sheaf equal to $\mathcal{F}$ over U. Further $j_!\mathcal{F}$ is unique with the two properties as if for any other sheaf $\mathcal{G}$ which satisfies that $\mathcal{G}_{|U} = \mathcal{F}$, then we get an map of presheaves $G \to \mathcal{G}$ which induces an isomorphism on stalks. By universal property of sheafification (Theorem 20.2.0.1), we deduce that $j_!\mathcal{F} \cong \mathcal{G}$. $\square$

With the above result, we have a useful short exact sequence.

Corollary 20.3.0.11. *Let X be a space and $\mathcal{F}$ be a sheaf over X. Let $i : Z \hookrightarrow X$ be a closed subspace and $j : U = X \setminus Z \hookrightarrow X$ be the corresponding open subspace. Then there is a short exact sequence*

$$0 \longrightarrow j_!\mathcal{F}_{|U} \longrightarrow \mathcal{F} \longrightarrow i_*\mathcal{F}_{|Z} \longrightarrow 0$$

where $\mathcal{F}_{|Z} = i^{-1}\mathcal{F}$. We call this the extension by zero short exact sequence.

Proof. Following the notation of proof of Proposition 20.3.0.10, we see that we have an injective map $G \to \mathcal{F}$, which then by universal property and local nature of injectivity gives an injective map $j_!\mathcal{F}_{|U} \to \mathcal{F}$. The map $\mathcal{F} \to i_*\mathcal{F}_{|Z}$ is obtained by considering the unit map of the adjunction $i_* \vdash i^{-1}$. This is surjective because on the stalks, we obtain $(i_*\mathcal{F}_{|Z})_p = \mathcal{F}_p$ if $p \in Z$ or 0 otherwise by above result. To show exactness at middle, we again go to stalks (Lemma 20.3.0.8) and observe that if $p \in U$, then we get exact sequence $0 \to \mathcal{F}_p \stackrel{\text{id}}{\to} \mathcal{F}_p \to 0 \to 0$ and if $p \in Z$, then we get the exact sequence $0 \to 0 \to \mathcal{F}_p \stackrel{\text{id}}{\to} \mathcal{F}_p \to 0$. $\square$

20.4 Sheaves are étale spaces

Another important and in some sense dual viewpoint of sheaves over X is that they can be equivalently defined as a certain type of bundle over X and all such bundles arises only from a sheaf. This is important because this viewpoint naturally extends the usual concepts of covering spaces, bundles and vector bundles to that of sheaves. In particular, a lot of classical constructs in algebraic topology can be equivalently be seen as specific instantiates of the notion of étale space of the sheaf.

Definition 20.4.0.1. (**Étale space**) Let X be a topological space and let $\pi : E \to X$ be a bundle over X. Then (E, π, X) is said to be étale over X or just étale if for all $e \in E$, there exists an open set $V \ni e$ of E such that $p(V)$ is open and $p|_V : V \to p(V)$ is a homeomorphism, that is, if p is a local homeomorphism. A morphism of étale spaces $(E_1, \pi_1, X), (E_2, \pi_2, X)$ over X is given by a continuous map $f : E_1 \to E_2$ such that $\pi_2 \circ f = \pi_1$. Denote the category of étale spaces over X by $\mathbf{Et}/X$.

Clearly, covering spaces over X are étale spaces over X, but not all étale spaces over X are covering, of-course. We now wish to show that the sheafification functor factors through a functor mapping a presheaf to an étale space. In particular, we want to show the existence of functor F, G so that the following commutes

$$\begin{tikzcd} & \mathbf{Et}/X \arrow[dr, "F"] & \\ \mathbf{PSh}(X) \arrow[ur, "G"] \arrow[rr, "(-)^{++}"', shift left] & & \mathbf{Sh}(X) \arrow[ll, hook', shift left] \end{tikzcd}.$$

Construction 20.4.0.2. (*Étale space of a sheaf*) Let us now show the construction of the above functors:

1. (*The functor G*) Let P be a presheaf over X. The étale space $E := G(P)$ is given by the disjoint union of all stalks:

$$E := \coprod_{x \in X} P_x.$$

The topology on E is given by the initial topology of the map

$$\begin{aligned} \pi : E &\longrightarrow X \\ s_x &\longmapsto x. \end{aligned}$$

In particular, E has a basis given by sets of the form $B_{U,s} \subseteq E$ where $B_{U,s} = \{s_x \in E \mid x \in U\}$ and $s \in P(U)$. Next, we wish to establish that π is a local homeomorphism. So take any $s_x \in E$ and consider the basic open set $B_{U,s} \ni s_x$. The map $\pi|_{B_{U,s}} : B_{U,s} \to \pi(B_{U,s})$ takes $s_x \mapsto x$. This is a homeomorphism because we can construct an inverse given by $x \mapsto s_x$. A simple calculation checks that this is continuous. Hence indeed, (E, π, X) is an étale space over X.

Next consider a map of presheaves $\varphi : F \to G$. We can construct a map of corresponding étale spaces as

$$\begin{aligned} \hat{\varphi} : (E_F, \pi_F, X) &\longrightarrow (E_G, \pi_G, X) \\ s_x &\longmapsto \varphi_x(s_x). \end{aligned}$$

This map is continuous and a valid bundle map over X. This defines the functor G.

2. (*The functor F*) Let $\pi : E \to X$ be an étale space over X. Then, we can construct a sheaf $\mathcal{E}$ over X out of it. This is done in a very natural way by considering the set of sections over U of $\mathcal{E}$ to be quite literally the set of *cross-sections*[5] of map π on U. That is, define:

$$\mathcal{E}(U) := \{s : U \to E \mid \pi \circ s = \mathrm{id}_U\}.$$

The fact that this is indeed a sheaf can be seen by a general phenomenon that for any continuous map $f : X \to Y$, the set of all cross-sections of f over open subsets of Y assembles

[5]In-fact, historically the notion of sheaf was really that of this étale space, and that is why to this day, we still use the terminology of "sections" of a sheaf over an open subset.

itself into a sheaf. Hence, we have constructed a sheaf $\mathcal{E}$ out of an étale space E over X.

Next consider a map of étale spaces $\xi : (E_1, \pi_1, X) \to (E_2, \pi_2, X)$. we can construct a map of corresponding sheaves $\tilde{\xi} : \mathcal{E}_1 \to \mathcal{E}_2$ by defining the following for open $U \subseteq X$:

$$\begin{aligned}\tilde{\xi}_U : \mathcal{E}_1(U) &\longrightarrow \mathcal{E}_2(U)\\ s &\longmapsto \xi \circ s.\end{aligned}$$

One can check that this is indeed a valid sheaf morphism. This defines the functor F.

We then see that the categories $\mathbf{Et}/X$ and $\mathbf{Sh}(X)$ are equivalent.

Theorem 20.4.0.3. [6] *(The étale viewpoint of sheaves) Let X be a topological space. The functors F and G as defined in Construction 20.4.0.2 defines an equivalence of categories*

$$\mathbf{Sh}(X) \equiv \mathbf{Et}/X.$$

We will prove this result in many small lemmas below. We would first like to observe that for any étalé bundle E over X yields a sheaf by $F(E)$ whose stalks are bijective to fibres of E.

Lemma 20.4.0.4. *Let (E, π, X) be an étalé bundle over X and let $\mathcal{E}$ be the sheaf obtained by $F((E, \pi, X))$. Then, for any $x \in X$, the following is a bijection*

$$\begin{aligned}\tau_x : \mathcal{E}_x &\longrightarrow E_x := \pi^{-1}(x)\\ (U, s)_x &\longmapsto s(x).\end{aligned}$$

Proof. We first show that τ_x is injective. Let $(U, s)_x, (V, t)_x$ be two germs such that $p = s(x) = t(x)$. We wish to show that s and t are equal on an open subset in $U \cap V$. As E is étaleé, therefore we have an open $A \subseteq E$ with $p \in A$ such that $\pi|_A : A \to \pi(A)$ is a homeomorphism. Consequently, we see that the open set $W = \pi(A) \cap U \cap V$ would do just fine.

We now show surjectivity. Pick $e \in E_x$. As E is étalé, we thus get an open set $A \ni e$ in E such that $\pi|_A : A \to \pi(A)$ is a homeomorphism. Denote the inverse of this homeomorphism by $g : \pi(A) \to A$. This is therefore a section of E over $\pi(A)$ where $x \in \pi(A)$. Consequently, $(\pi(A), g)_x \in \mathcal{E}_x$ is such that τ_x maps it to e. $\square$

Proof of Theorem 20.4.0.3. We first show that $F \circ G$ is naturally isomorphic to sheafification functor. Let $\mathcal{E}$ be a presheaf, $(E, \pi) = G(\mathcal{E})$ and $F(E, \pi) = \mathcal{E}'$. We wish to show that there is a natural isomorphism $\mathcal{E}^{++} \to \mathcal{E}'$. By Theorem 20.2.0.1 and 20.3.0.6, 3, it suffices to show that there is a map of presheaves $\mathcal{E} \to \mathcal{E}'$ which is isomorphism on stalks.

Consider the map $\varphi : \mathcal{E} \to \mathcal{E}'$ which on an open set $U \subseteq X$ gives the following map

$$\begin{aligned}\varphi_U : \mathcal{E}(U) &\longrightarrow \mathcal{E}'(U)\\ s &\longmapsto U \stackrel{f_{U,s}}{\to} E\end{aligned}$$

[6] Exercise II.1.13 of Hartshorne.

where $f_{U,s} : U \to E$ maps as $x \mapsto (U, s)_x$. Since $f_{U,s}$ is just the stalk map, then as in Construction 20.4.0.2, $f_{U,s}$ is continuous. Now on the stalks, we get the following commutative diagram by Lemma 20.4.0.4:

$$\begin{tikzcd} \mathcal{E}_x \arrow[r, "\varphi_x"] \arrow[d] & \mathcal{E}'_x \arrow[dl, dash, "\cong"] \\ E_x & \end{tikzcd},$$

where the vertical map takes a germ $(U, s)_x$ and maps it to the element represented in $E_x = \mathcal{E}_x$, as $E_x = \pi^{-1}(x) = \{x \in E \mid \pi(e) = x\} = \{(U, s)_y \in E \mid \pi((U, s)_y) = y = x\}$. Consequently, the vertical map is a bijection and thus φ_x is a bijection. The naturality of this isomorphism can be checked trivially.

We now wish to show that $G \circ F$ is naturally isomorphic to the identity functor on $\mathbf{Et}/X$. Pick an étalé bundle (E, π) over X, denote $F(E, \pi) = \mathcal{E}$ and $G(\mathcal{E}) = (E', \pi')$. We wish to find a homeomorphism φ so that the following commutes:

$$\begin{tikzcd} E' \arrow[r, "\varphi"] \arrow[d, "\pi'"'] & E \arrow[dl, "\pi"] \\ X & \end{tikzcd}.$$

Consider the following map

$$\begin{aligned} \varphi : E' &\longrightarrow E \\ (U, s)_x &\longmapsto s(x). \end{aligned}$$

By Lemma 20.4.0.4, φ is a bijective map. We thus reduce to showing that φ is a continuous open map.

To show continuity, consider an open set $A \subseteq E$ and then observe that

$$\begin{aligned} \varphi^{-1}(A) &= \{(U, s)_x \in E' \mid s(x) \in A\} \\ &= \{(U, s)_x \in E' \mid x \in s^{-1}(A)\} \\ &= \bigcup_{U \ni x, s : U \to E} B_{U,s} \end{aligned}$$

and since $B_{U,s} \subseteq E'$ is a basic open, therefore φ is continuous.

Finally, to show that φ is open, one reduces to showing that if $s : U \to E$ is a continuous section of bundle (E, π) and $U \subseteq X$ is an open set, then $s(U)$ is an open set in E (by working with a basic open $B_{U,s} \subseteq E'$). This follows from the fact that since π is a local homeomorphism, therefore for each $e \in s(U)$, there exists an open set $A \ni e$ in E such that $s(U) \cap A \ni e$ and since $\pi : s(U) \cap A \to \pi(s(U) \cap A) = U \cap \pi(A)$ is a homeomorphism, we further get that $s(U) \cap A$ is open (as $U \cap \pi(A)$ is open). Consequently, $s(U)$ is open. $\square$

Remark 20.4.0.5. (*The sheaf associated to a covering space*) By the above equivalence, each covering space space over X, which is an étale map, determines a unique sheaf (upto isomorphism). We analyze this sheaf. Recall that a local system is just a name for locally constant sheaf. We write $\mathrm{LocSys}(X)$ to denote the category of all local systems.

Proposition 20.4.0.6. *Let X be a connected and locally path-connected space. The there is an equivalence of categories*

$$\mathbf{Cov}(X) \equiv \mathrm{LocSys}(X)$$

where $\mathbf{Cov}(X)$ is the category of covering spaces over X and $\mathrm{LocSys}(X)$ is the category of locally constant sheaves of sets over X.

Proof. We will show that this equivalence is induced from the equivalence of Theorem 20.4.0.3. It is sufficient to show that F maps covering spaces to locally constant sheaves and vice-versa for G. Indeed, if (E, p, X) is a covering space and $\mathcal{E}$ is the associated sheaf, then for a connected evenly covered neighborhood $U \subseteq X$ for which $p^{-1}(U) = \amalg_{\alpha \in A_U} V_\alpha$ where $p : V_\alpha \to U$ is a homeomorphism, we get that the set of sections $\mathcal{E}(U)$ is just A_U by connectedness. Moreover, it is clear that $\mathcal{E}(V) = A_U$ again for any connected $V \subseteq U$. This shows that $\mathcal{E}_{|U} = \underline{A_U}$. Hence $\mathcal{E}$ is a local system.

Conversely, if $\mathcal{E}$ is a local system with (E, p, X) its associated étale space, then for $U \subseteq X$ such that $\mathcal{E}_{|U} = \underline{A}$, we get that $p^{-1}(U) = \amalg_{x \in U} \mathcal{E}_x = \amalg_{x \in U} A = \amalg_{\alpha \in A} V_\alpha$ where $V_\alpha = \{\alpha \in A_x \mid x \in U\}$. We first claim that V_α is open. Indeed, it is the basic open set $B_{U,\alpha}$. Next, $V_\alpha \cap V_\beta =$ is clear. Finally, $p : V_\alpha \to U$ being a homeomorphism is also clear as this is a bijection and p is an open map as it is étale. □

20.5 Direct and inverse image

Let $f : X \to Y$ be a continuous map of topological spaces. Then one can derive two functors $f_* : \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ and $f^{-1} : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ which are adjoint of each other, called direct and inverse image functors respectively. While f_* is easy to define, it is usually the inverse image of a sheaf that causes trouble for its obscurity if one works with the definition that inverse image functor is left-adjoint to direct image functor. This is resolved by working with the corresponding étale spaces (Theorem 20.4.0.3). In this section we will show how to construct them.

Let us first define the direct image functor.

Definition 20.5.0.1. (**Direct image**) Let $f : X \to Y$ be a continuous map. Then, for any sheaf $\mathcal{F}$ on X, we can define its direct image under f as $f_*\mathcal{F}$ whose sections on open $V \subseteq Y$ are given by

$$(f_*\mathcal{F})(V) := \mathcal{F}(f^{-1}(V)).$$

This can easily be seen to be a sheaf. For any map of sheaves $\varphi : \mathcal{F} \to \mathcal{G}$ on X, we can define the map of direct image sheaves as

$$\begin{aligned}(f_*\varphi)_V : f_*\mathcal{F}(V) &\longrightarrow f_*\mathcal{G}(V)\\ s &\longmapsto \varphi_{f^{-1}(V)}(s).\end{aligned}$$

This defines a functor

$$f_* : \mathbf{Sh}(X) \longrightarrow \mathbf{Sh}(Y).$$

One defines the inverse image of a sheaf as follows:

Definition 20.5.0.2. (**Inverse image**) Let $f : X \to Y$ be a continuous map and let $\mathcal{G}$ be a sheaf over Y. Consider a presheaf F over X constructed by the data of $\mathcal{G}$ as follows. Let $U \subseteq X$ be open, then define

$$f^+\mathcal{G}(U) := \varinjlim_{\text{open } V \supseteq f(U)} \mathcal{G}(V),$$

where restriction maps of $f^+\mathcal{G}$ is given by the unique map obtained by universality of colimits. Then, $f^+\mathcal{G}$ is a presheaf over X and this construction is functorial again by universal property of colimits:

$$f^+ : \mathbf{PSh}(Y) \longrightarrow \mathbf{PSh}(X).$$

Let $f^{-1}\mathcal{G} = (f^+\mathcal{G})^{++}$ denote the sheafification of $f^+\mathcal{G}$. This sheaf is called the inverse sheaf of $\mathcal{G}$ under f. Now for any map of sheaves $\varphi : \mathcal{G} \to \mathcal{H}$ over Y, we get a corresponding map of inverse image sheaves $f^{-1}\varphi : f^{-1}\mathcal{G} \longrightarrow f^{-1}\mathcal{H}$ by composition of two functors. This yields a functor

$$f^{-1} : \mathbf{Sh}(Y) \longrightarrow \mathbf{Sh}(X).$$

As is visible, this definition is quite obscure if one likes elemental definitions. We thus give some general properties enjoyed by inverse sheaf.

Lemma 20.5.0.3. *Let $f : X \to Y$ be a continuous map and $\mathcal{G}$ be a sheaf over Y.*

1. *If f is open, then $f^{-1}\mathcal{G} = \mathcal{G}(f(-))$.*
2. *If f is constant to $y \in Y$, then $f^{-1}\mathcal{G}$ is the constant sheaf on X with sections $\mathcal{G}_y$.*
3. *If $X = \{x\}$ is a singleton space, then $f^{-1}\mathcal{G}$ is the constant sheaf on X with sections $\mathcal{G}_{f(x)}$.*
4. *If $x \in X$, then*

$$(f^{-1}\mathcal{G})_x \cong \mathcal{G}_{f(x)}.$$

Proof. 1. One notes that $f^+\mathcal{G}(U) := \varinjlim_{V \supseteq f(U)} \mathcal{G}(V) = \mathcal{G}(f(U))$. The mapping $\mathcal{G}(f(-))$ is a sheaf, hence sheafifying it will yield the same sheaf.
2. We see that $f^+\mathcal{G}(U) = \varinjlim_{V \supseteq f(U)} \mathcal{G}(V) = \varinjlim_{V \ni y} \mathcal{G}(V) = \mathcal{G}_y$ and presheaves with constant values are sheaves, as restrictions are identity.
3. We see that $f^+\mathcal{G}(U) = \varinjlim_{V \supseteq f(U)} \mathcal{G}(V) = \varinjlim_{V \ni f(x)} \mathcal{G}(V) = \mathcal{G}_{f(x)}$ and presheaves with constant values are sheaves, as restrictions are identity.
4. By passing to the right adjoint, one observes that for $f : X \to Y$ and $g : Y \to Z$ continuous maps, one can obtain the following natural isomorphism of functors

$$(g \circ f)^{-1} \cong f^{-1} \circ g^{-1}.$$

Consider the composite $f \circ \iota$ where $\iota : \{x\} \hookrightarrow X$ is the inclusion map. Consequently, by 3. above, we obtain the following

$$\begin{aligned}
\mathcal{G}_{f(x)} &\cong (f \circ \iota)^{-1}(\mathcal{G})(\{x\}) \\
&\cong (\iota^{-1} \circ f^{-1})(\mathcal{G})(\{x\}) \\
&\cong \iota^{-1}(f^{-1}\mathcal{G})(\{x\}) \\
&\cong (f^{-1}\mathcal{G})_{f(x)}.
\end{aligned}$$

□

The following is a fundamental duality between inverse and direct image functors.

Theorem 20.5.0.4. [7] *(Direct and inverse image adjunction) Let $f : X \to Y$ be a continuous map. Then the the inverse image functor is the left adjoint of direct image functor* [8]

$$\begin{tikzcd} \mathbf{Sh}(Y) \arrow[r, shift left, "f^{-1}"] & \mathbf{Sh}(X) \arrow[l, shift left, "f_*"] \end{tikzcd} .$$

In particular, we have a natural bijection

$$\mathrm{Hom}_{\mathbf{Sh}(X)}\left(f^{-1}\mathcal{F}, \mathcal{G}\right) \cong \mathrm{Hom}_{\mathbf{Sh}(Y)}\left(\mathcal{F}, f_*\mathcal{G}\right).$$

One situation that we will find ourselves a lot in algebraic geometry is when $f : X \to Y$ will be a closed immersion of topological spaces ($f : X \to f(X)$ is homeomorphism and $f(X) \subseteq Y$ is closed) and for a sheaf $\mathcal{F}$ over X, we would like to find $(f_*\mathcal{F})_{f(x)}$ for each point $x \in X$. This is a situation where the stalk of direct image can be calculated quite easily.

Lemma 20.5.0.5. *Let $f : X \to Y$ will be a closed immersion of topological spaces and $\mathcal{F}$ a sheaf over X. Then, there is a natural isomorphism*

$$(f_*\mathcal{F})_{f(x)} \cong \mathcal{F}_x.$$

Proof. From a straightforward unravelling of definitions of the two stalks, the result follows from the observation that each open set $U \ni x$ in X is in one-to-one correspondence with open set $f(U) \ni f(x)$ in Y. □

Remark 20.5.0.6. We wish to know how the inverse image of sheaves changes the stalk. Let $f : X \to Y$ be a continuous map and let $\mathcal{F}$ be a sheaf on Y. Consider the inverse sheaf $f^{-1}\mathcal{F}$ on X. Let $x \in X$. Then we have that (Lemma 20.5.0.3, 4)

$$(f^{-1}\mathcal{F})_x \cong \mathcal{F}_{f(x)}.$$

The importance of this is that, suppose $f : X \to Y$ is given together with $\mathcal{F}$ and $\mathcal{G}$ are sheaves over X and Y respectively and a map $\varphi^\flat : \mathcal{G} \to f_*\mathcal{F}$ over Y, which is equivalent to $\varphi^\sharp : f^{-1}\mathcal{G} \to \mathcal{F}$ over X. Now, most of the time, our interest in a sheaf is only limited to stalks (functions defined in

[7] Exercise II.1.18 of Hartshorne.

[8] admirers of topoi may see this as a quintessential example of geometric map of topoi.

some open subset around a point), therefore we are mostly interested in considering only the map induced at the level of stalks at a point $f(x) \in Y$:

$$\varphi^\flat_{f(x)} : \mathcal{G}_{f(x)} \longrightarrow (f_*\mathcal{F})_{f(x)}.$$

But the description of the stalk $(f_*\mathcal{F})_{f(x)}$ is usually not simple to derive. But dually, we may ask the map of stalks of the other map at $x \in X$, and we directly land into the stalks

$$\varphi^\sharp_x : \mathcal{G}_{f(x)} \cong (f^{-1}\mathcal{G})_x \longrightarrow \mathcal{F}_x.$$

However, this is a strange map as the stalks are of sheaves which are not on same space. In particular, this map is given as follows. For any open $V \ni f(x)$ in Y, we have the following maps:

$$\mathcal{G}(V) \xrightarrow{\varphi^\flat_V} \mathcal{F}(f^{-1}(V)) \longrightarrow \mathcal{F}_x \ .$$

Passing to colimits ($\varphi^\flat_V$ commutes with restrictions), one can see that we get the map $\varphi^\sharp_x : \mathcal{G}_{f(x)} \to \mathcal{F}_x$ back.

It is a good principle to keep in mind that if we wish to work with explicit local sections, then we should look for the "flat" map and if it is enough to work with germs, then we should look for the "sharp" map, even though the above remark telling us how to construct the map of stalks from the "flat" maps on each open set.

This map $\varphi^\sharp_x$ can be heuristically be defined as the map which on sections which makes sure that a non-invertible section remains non-invertible after going through the map. Hence we mostly work only with maps $f^{-1}\mathcal{G} \to \mathcal{F}$ if we are interested only at the stalk level (which is more than enough for us).

20.6 Category of sheaves

We will discuss some basic properties of the category of sheaves over X, denoted $\mathbf{Sh}(X)$. This is important as we wish to calculate cohomology of its objects, hence we would require the notion of injective and projective resolutions of sheaves. We covered the homological methods necessary for this section in the Homological Methods, Chapter 19. Let us first begin with a more categorical definition of sheaves.

Definition 20.6.0.1. (**Sheaf of sets - categorical defn.**) *Suppose X is a topological space and* **O(X)** *is the posetal category of open sets of X, ordered by inclusion. Then a presheaf*

$$F : \mathbf{O(X)}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$$

is a sheaf if for any open set U and any covering of $U = \bigcup_{i\in I} U_i$, we have that

$$\begin{tikzcd} FU \arrow[dr, dashed, "e"] & & \\ & \prod_{i\in I} FU_i \arrow[r, dashed, shift left, "p"] \arrow[r, dashed, shift right, "q"'] & \prod_{i,j\in I} F(U_i \cap U_j) \end{tikzcd}$$

is an equalizer diagram, where the unique maps e, p & q are given as:

- e *: for a* $f \in FU$*,* e *maps it as*

$$e(f) = \{\underbrace{F(U_i \subset U)}_{FU \to F(U_i)}(f)\} \in \prod_i F(U_i)$$

That is, e *maps each element* f *of the* FU *via the set map under the functor* F *of the inclusion* $U_i \subset U$*.*

- p *: for a sequence* $\{f_i\} \in \prod_{i \in I} FU_i$*,* p *maps it as*

$$p(\{f_i\}) = \{\underbrace{F(U_i \cap U_j \subset U_i)}_{FU_i \to F(U_i \cap U_j)}(f_i)\} \in \prod_{i,j \in I} F(U_i \cap U_j)$$

That is, p *maps each component* y_i *of the sequence* $\{y_i\}$ *via the set map under the functor* F *of the inclusion* $U_i \cap U_j \subset U_i$*.*

- q *: for a sequence* $\{f_i\} \in \prod_{i \in I} FU_i$*,* q *maps it as*

$$q(\{f_i\}) = \{\underbrace{F(U_i \cap U_j \subset U_j)}_{FU_j \to F(U_i \cap U_j)}(f_j)\} \in \prod_{i,j \in I} F(U_i \cap U_j)$$

That is, q *maps each component* y_i *of the sequence* $\{y_i\}$ *via the set map under the functor* F *of the inclusion* $U_i \cap U_j \subset U_j$*.*[9]

20.6.1 Coverings, bases & sheaves

We now quickly discuss some easy properties of sheaves. In the following, a **Subsheaf** of a sheaf F is defined as a subfunctor of F which also satisfies the sheaf property (is a sheaf itself).

Proposition 20.6.1.1. *A subfunctor* S *of a sheaf* F *is a subsheaf if and only if for any open set* U *and it's open covering* $\bigcup_{i \in I} U_i$ *together with an* $f \in FU$*, we have* $f \in SU$ *if and only if* $f|_{U_i} \in SU_i \; \forall \; i \in I$*.*

Proof. (**L** $\implies$ **R**) Suppose S is a subsheaf, then clearly for any $f \in SU \subset FU$, we must have $f|_{U_i} \in SU_i$ for all $i \in I$ and for any such collection of $f|_{U_i}$, by the sheaf property of S, $f \in SU$.
(**R** $\implies$ **L**) Since S is a subfunctor of F, therefore $SV \subset FV$ for any open V. With this, because F is a sheaf, we have the following diagram:

$$\begin{tikzcd} SU \arrow[r] \arrow[d, tail] & \prod_i SU_i \arrow[r, shift left] \arrow[r, shift right] \arrow[d, tail] & \prod_{i,j} S(U_i \cap U_j) \arrow[d, tail] \\ FU \arrow[r] & \prod_i FU_i \arrow[r, shift left] \arrow[r, shift right] & \prod_{i,j} F(U_i \cap U_j) \end{tikzcd}$$

where the bottom row is the equalizer. The condition on the right says that for $f \in FU$, $f \in SU \iff \{f|_{U_i}\} \in \prod_i SU_i$, which means that the left square is a pullback. Now because SU is universal due to it being a pullback, and since the top row infact commutes, therefore SU is universal with top row commuting, hence, it is an equalizer. $\square$

[9]Refraining to write $F(V \subset U) = FU \to FV$ to be equal to the restriction $(-)|_V$ exaggerates the emphasis on the abstract nature of sheaf F, that is, it helps to imagine that FU might not always be a set of specific maps over U, even though in most examples of interest it is the case.

Sheaf itself is local

Define **restriction of a sheaf** F on X restricted to open $U \subset X$ to be the $F|_U(V) = F(V)$ where $V \subset U$, and $F|_U(U) = F(\phi) = \{*\}$ if $V \not\subset U$.

Theorem 20.6.1.2. *Suppose X is a space with a given open covering $X = \bigcup_{k \in I} W_k$. If there are sheaves for each k,*

$$F_k : \mathbf{O(Wk)}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$$

[10]*such that*

$$F_k|_{W_k \cap W_l} = F_l|_{W_k \cap W_l}$$

[11] *then, $\exists$ a sheaf F on X,*

$$F : \mathbf{O(X)}^{\mathrm{op}} \longrightarrow \mathbf{Sets}$$

unique upto isomorphism such that

$$F|_{W_k} \cong F_k.$$

This theorem hence shows that the restriction functor $U \mapsto \mathrm{Sh}(U)$ and $V \subset U \mapsto (\mathrm{Sh}(U) \to \mathrm{Sh}(V), F|_U \mapsto F|_V)$ on **O(X)** is local enough to be *almost* a sheaf. If only for any sheaf F, G on X, we had that $F|_{W_k} = G|_{W_k}\ \forall\ k$ would imply that $F = G$, which is not the case in general however, then we would have said that this restriction functor is also a sheaf.

Sheaf over a basis of X

A **basis** of a space X is a subset of topology $\mathcal{B} \subset \mathcal{O}(X)$ such that for any open $U \in \mathcal{O}(X)$, $\exists \{B_i\} \subseteq \mathcal{B}$ such that $U = \bigcup_i B_i$.
It turns out that the restriction functor $r : \mathrm{Sh}(X) \longrightarrow \mathrm{Sh}(X_{\mathcal{B}})$ which restricts each sheaf over X to that of open sets of basis $\mathcal{B}$ establishes an equivalence of categories!

Theorem 20.6.1.3. *Suppose X is a topological space and $\mathcal{B}$ is a basis for X. Then, the restriction functor*

$$\begin{aligned} r : Sh(X) &\longrightarrow Sh(X_{\mathcal{B}}) \\ F &\longmapsto F|_{\mathcal{B}} \\ \eta : F \Longrightarrow G &\longmapsto \eta|_{\mathcal{B}} : F|_{\mathcal{B}} \Longrightarrow G|_{\mathcal{B}} \end{aligned}$$

establishes an equivalence of categories between $Sh(X)$ and $Sh(X_{\mathcal{B}})$.

Proof. For any sheaves F, G in $\mathrm{Sh}(X)$, we want to show that $\mathrm{Hom}_{\mathrm{Sh}(X)}(F, G) \cong \mathrm{Hom}_{\mathrm{Sh}(X_{\mathcal{B}})}(rF, rG)$, that is, r is fully faithful. One can see that there r is an injection between the above hom-sets as for any $\epsilon, \eta : F \Rightarrow G$, if $rF = F|_{\mathcal{B}} = G|_{\mathcal{B}} = rG$, then due to the commutation of the two squares

[10]where $\mathbf{O(Wk)}^{\mathrm{op}}$ is the opposite category of all open subsets of open set W_k and inclusion.

[11]This condition implies that for any open subsets $V_k \subset W_k$ and $V_l \subset W_l$, $F(V_k \cap W_k \cap W_l) = F(V_l \cap W_k \cap W_l)$ and for arrows $X_1 \subset X_2$ in **O(Wk)** & $Y_1 \subset Y_2$ in **O(Wl)**, $F(X_1 \cap W_k \cap W_l \subset X_2 \cap W_k \cap W_l) = F(Y_1 \cap W_k \cap W_l \subset Y_2 \cap W_k \cap W_l)$.

below because of naturality, (take $U = \bigcup_i B_i$ to be any open set and it's trivial open covering from basic open sets)

$$\begin{tikzcd} FU \arrow[r, "e_F"] \arrow[d, shift right, "\epsilon U"'] \arrow[d, shift left, "\eta U"] & \prod_i FB_i \arrow[d, shift right, "\prod_i \epsilon B_i"'] \arrow[d, shift left, "\prod_i \eta B_i"] \\ GU \arrow[r, "e_G"] & \prod_i GB_i \end{tikzcd}$$

one can infer $\epsilon U = \eta U$ (e_F and e_G are equalizers, so are monic).
With the information $\kappa : rF \Rightarrow rG$, one can construct a natural transformation $\gamma : F \Rightarrow G$ by defining FU and GU, for any open U with it's basic cover $U = \bigcup_i B_i$ where $B_i \in \mathcal{B}$, as the equalizer of the parallel arrows $\prod_i F|_{\mathcal{B}} B_i \rightrightarrows \prod_{i,j} F|_{\mathcal{B}} B_i \cap B_j$ and $\prod_i G|_{\mathcal{B}} B_i \rightrightarrows \prod_{i,j} G|_{\mathcal{B}} B_i \cap B_j$, respectively. Then, one defines $\gamma U : FU \to GU$ by noticing that the former forms a cone over the latter, due to arrows $\prod_i \kappa B_i : \prod_i F|_{\mathcal{B}} B_i \to \prod_i G|_{\mathcal{B}} B_i$ and $\prod_{i,j} \kappa(B_i \cap B_j) : \prod_{i,j} F|_{\mathcal{B}} B_i \cap B_j \Rightarrow \prod_{i,j} G|_{\mathcal{B}} B_i \cap B_j$, so that there exists a unique arrow $FU \to GU$, which we just define as γU.
With this, we see that r is fully faithful. Finally, with the above definitions, $rF \cong F|_{\mathcal{B}}$ where $F \in \mathrm{Sh}(X)$ is the sheaf obtained by the above process from $F|_{\mathcal{B}} \in \mathrm{Sh}(X_{\mathcal{B}})$ because both of them are equalizers of the same diagram for any open set $U = \bigcup_i B_i$ and it's basic covering (note that any covering of U can be decomposed into basic covering). □

20.6.2 Sieves as general covers

This is related to generalization of sheaves to topos theory. As we saw in Definition 13.1.1.1, a subfunctor of $\mathbf{Yon}(C) = \mathrm{Hom}(-, C)$ is a sieve, therefore this notion would allow us to generalize the notion of *covering* of a space, as we will see later. But for now, the *shadow* of that more general notion can still be felt in the usual category **O(X)** of open sets of X.

Definition 20.6.2.1. (**Principal Sieve**) *Suppose X is a topological space and U is open. Then the sieve S,* generated *from U, that is,*

$$S = \{V \;:\; \text{open } V \subset U\}$$

is said to be a principal sieve, denoted $S = \langle U \rangle$, generated by a single open set.

With Definition 20.6.2.1, we can now define a new notion of *covering* of an open set, purely in terms of arrows onto it!

Definition 20.6.2.2. (**Covering Sieve**) *Suppose X is a topological space and U is open in it. A sieve S on U is said to cover U if*

$$U = \bigcup_{W \in S} W.$$

That is, when U is union of all open sets in the sieve S.

Remark 20.6.2.3. It can be seen quite easily that a subfunctor S of $\mathbf{Yon}(U)$ is a principal sieve over U if and only if S is a subsheaf. L $\implies$ R by Proposition 13.1.0.3 and R $\implies$ L by noting that the union of all sets in S would generate it. Remember that you can take covers of only those open sets which are members of S because S is a subsheaf.

The above definition in effect can be replaced with in the definition of sheaves!

Proposition 20.6.2.4. *A presheaf $P : \mathbf{O(X)}^{\text{op}} \longrightarrow \mathbf{Sets}$ on a topological space X is a sheaf if and only if for any open U and a covering sieve S over U, we have that the inclusion nat. trans. $i_S : S \Longrightarrow \boldsymbol{Yon}(U)$ induces an isomorphism:*

$$\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, P) \cong \mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\boldsymbol{Yon}(U), P).$$

Proof. We can re-derive the sheaf condition in terms of the covering sieve as follows. For an open $U = \bigcup_i U_i$, if$\{f_i\} \in \prod_i PU_i$ is such that $f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$, then because S is a covering sieve of U, therefore this condition is equivalent to a sequence $\{f_V\} \in PV$ for all $V \in S$ such that $f_V|_{V'} = f|_{V'}$ whenever $V' \subset V$. It can also be seen that every natural transformation η between S and P can be mapped to an element of $\prod_{V \in S} PV$ by forming the collection $\{\eta_V(*)\}$. Similarly, for any $\{f_V\} \in \prod_{V \in S} PV$ we can construct a nat. trans. $\{f_V : SV = \{*\} \to PV\}$. Now, with this, we can obtain the result by a basic diagram chase around the left square of the following

$$\begin{tikzcd}
\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, P) \arrow[r, "d"] & \prod_i PU_i \arrow[r, shift left] \arrow[r, shift right] & \prod_{i,j} P(U_i \cap U_j) \\
\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\mathbf{Yon}(U), P) \arrow[u, "\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(i_S, P)"] \arrow[r, equal] & PU \arrow[u, "e"'] &
\end{tikzcd}$$

where d is the equalizer of the parallel arrows on the right (the fact that this set is the equalizer is established in the prev. paragraph) $\square$

20.6.3 $\mathrm{Sh}(X)$ has all small limits

We now see that $\mathrm{Sh}(X)$ has all small limits and the inclusion of $\mathrm{Sh}(X)$ to $\widehat{\mathbf{O(X)}}$ preserves these limits.

Proposition 20.6.3.1. *For any topological space X, the category $Sh(X)$ has all small limits and the inclusion functor*

$$i : Sh(X) \rightarrowtail \widehat{\mathbf{O(X)}}$$

preserves all those limits.

Proof. To show that $\mathrm{Sh}(X)$ has all small limits, we can first notice that the singleton functor is a sheaf, which is the terminal object in $\mathrm{Sh}(X)$. Now, to see equalizers, take any parallel arrows in $\mathrm{Sh}(X)$ as $F \rightrightarrows G$. Since $\widehat{\mathbf{O(X)}}$ has all small limits, therefore, we can take the equalizer of this in it, in turn of taking equalizer in $\mathrm{Sh}(X)$. With this, there exists E, the equalizer of $F \rightrightarrows G$ in $\widehat{\mathbf{O(X)}}$. Now because covariant hom-functors preserves limits, therefore for any open U, the $\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\mathbf{Yon}(U), E)$ and $\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, E)$ acts as equalizers in the diagram below:

$$\begin{tikzcd}
\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\mathbf{Yon}(U), E) \arrow[r, "e\circ-"] \arrow[d, "-\circ i_s"', "f_E"] & \mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\mathbf{Yon}(U), F) \arrow[r, shift left, "f\circ-"] \arrow[r, shift right, "g\circ-"'] \arrow[d, "-\circ i_s"', "f_F"] & \mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(\mathbf{Yon}(U), G) \arrow[d, "-\circ i_s"', "f_G"] \\
\mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, E) \arrow[r, "e\circ-"'] & \mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, F) \arrow[r, shift left, "f\circ-"] \arrow[r, shift right, "g\circ-"'] & \mathrm{Hom}_{\widehat{\mathbf{O(X)}}}(S, G)
\end{tikzcd}$$

Using Proposition 20.6.2.4, f_F and f_G are isomorphisms. A simple diagram chase on the left square then shows f_E is also an isomorphism. Binary products exists by the same process. □

The above proposition hence allows us to infer what it means to be a subobject of a sheaf in $\mathrm{Sh}(X)$.

Corollary 20.6.3.2. *For any topological space X, any subobject of a sheaf F in $Sh(X)$ is isomorphic to a subsheaf of F.*

Proof. Suppose $H \Rightarrow F$ is a monic, so a subobject of F. Since $\mathrm{Sh}(X)$ has all limits (Proposition 20.6.3.1), so the kernel pair of this arrow would exist in $\mathrm{Sh}(X)$ and it's inclusion in $\widehat{\mathbf{O(X)}}$ would preserve it. By point-wise construction of presheaves in $\widehat{\mathbf{O(X)}}$, we can see that H would be isomorphic to some some subfunctor of F, which would be a sheaf too because it is isomorphic to H, a sheaf. □

Topology of X ≅ Subobjects of Yon (X) in Sh (X)

Finally, we observe that the topology of X is actually isomorphic to subobjects of $\mathbf{Yon}(X)$[12] in $\mathrm{Sh}(X)$!

Proposition 20.6.3.3. *For any topological space X, there exists an isomorphism of the following posets*

$$\mathcal{O}(X) \cong Sub_{Sh(X)}\left(\boldsymbol{Yon}(X)\right)$$

which is moreover order preserving.[13]

20.6.4 Direct and inverse limits in Sh(X)

Since Grothendieck-abelian categories have all colimits, therefore it also has direct limits. We now show that the direct limits in $\mathbf{Sh}(X)$ are obtained by sheafifying the corresponding direct limit in $\mathbf{PSh}(X)$.

Lemma 20.6.4.1. [14] *Let X be a topological space and $\{\mathcal{F}_i\}$ be a direct system of sheaves over X. Then, the direct limit $\varinjlim_i \mathcal{F}_i$ in $\mathbf{Sh}(X)$ is formed by sheafification of the presheaf $U \mapsto \varinjlim_i \mathcal{F}_i(U)$.*

Proof. Let F denote the presheaf obtained by $U \mapsto \varinjlim_i \mathcal{F}_i(U)$ and further denote $\mathcal{F} = F^{++}$, the sheafification of F. Note that we have $\mathcal{F}_i \xrightarrow{j_i} F \to \mathcal{F}$. We wish to show that $\mathcal{F}$ satisfies the universal property of direct limits in $\mathbf{Sh}(X)$. Indeed, take any other sheaf $\mathcal{G}$ for which there are

[12] Remember that $\mathbf{Yon}(X)$ is the terminal object in $\mathrm{Sh}(X)$.

[13] Remember Proposition 13.1.0.3. Therefore this isomorphism could be extended as:

$$\mathcal{O}(X) \cong \mathrm{Sub}_{\mathrm{Sh}(X)}\left(\mathbf{Yon}(X)\right) \cong \mathrm{Hom}_{\mathrm{Sh}(X)}\left(\mathbf{Yon}(X), \Omega\right)$$

when Ω exists. This is the first sign of how sheaves might be related to topoi.

[14] Exercise II.1.10 of Hartshorne.

maps $f_i : \mathcal{F}_i \to \mathcal{G}$ which further satisfies that for any $j \geq i$ in the direct set indexing the system, we have that the following triangle commutes:

$$\begin{tikzcd} & \mathcal{G} & \\ \mathcal{F}_i \arrow[ur, "f_i"] \arrow[rr, "\varphi_{ij}"'] & & \mathcal{F}_j \arrow[ul, "f_j"'] \end{tikzcd}.$$

We wish to show that there exists a unique map $\tilde{f} : \mathcal{F} \to \mathcal{G}$ such that for all i, the following commutes:

$$\begin{tikzcd} \mathcal{G} & \mathcal{F} \arrow[l, dashed, "\tilde{f}"'] \\ \mathcal{F}_i \arrow[u, "f_i"] \arrow[r, "j_i"'] & F \arrow[u] \end{tikzcd}.$$

But this is straightforward, as by the universal property of direct limits in $\mathbf{PSh}(X)$, we first have a map $f : F \to \mathcal{G}$ which makes the bottom left triangle in the above commute. Then, by the universal property of sheafification (Theorem 20.2.0.1), we get a corresponding $\tilde{f} : \mathcal{F} \to \mathcal{G}$ which makes the top right triangle in the above commute. Consequently, we have obtained $\tilde{f}$ which makes the square commute. $\square$

20.7 Classical Čech cohomology

Sheaf cohomology becomes an important tool to any attempt at understanding any sophisticated geometric situation in topology. It is a tool which measures the obstructions faced in extending a local construction (which are usually not too difficult to make) to a global one (which are most of the time very difficult to make). To get a feel of why such questions and tools developed to solve them are important, one may look no further than basic analysis; say in case of $\mathbb{R}^n$, we wish to extend a local isometry from an open set of $\mathbb{R}^n$ to $\mathbb{R}^m$, into a global one between $\mathbb{R}^n$ and $\mathbb{R}^m$. Clearly the former is much, much easier than the latter. In the same vein, we wish to understand obstructions faced in making local-to-global leaps in the context of schemes, which covers almost all range of algebro-geometric situations.

Construction 20.7.0.1 (*Čech cochain complex and Čech cohomology of an abelian presheaf.*)**.** Let X be a topological space and F be an abelian presheaf over X. We will construct and discuss the Čech cohomology groups $\check{H}^q(X;F)$. After giving the basic constructions, we will specialize to the case of schemes in Chapter 1, to prove the Serre's theorem on invariance of affine refinements of cohomology of coherent sheaves.

We first construct the Čech cochain complex of F w.r.t. to an open cover $\mathcal{U}$. Let $\mathcal{U} = \{U_\alpha\}_{\alpha\in I}$ be a fixed open cover of X. We can then define for each $i = 0, 1, 2, \dots$, a group called the group of *i-cochains of F w.r.t. $\mathcal{U}$*:

$$C^i(\mathcal{U}, F) := \prod_{(\alpha_0,\dots,\alpha_i)\in I^{i+1}} F(U_{\alpha_0} \cap U_{\alpha_1} \cap \dots \cap U_{\alpha_i}).$$

where the product runs over all increasing $i+1$-tuples with entries in I[15]. A typical element $s \in C^i(\mathcal{U}, F)$ is called an i-cochain, whose part corresponding to $(\beta_0, \dots, \beta_i) \in I^{i+1}$ is denoted by $s(\beta_0, \dots, \beta_i) \in F(U_{\beta_0} \cap \dots \cap U_{\beta_i})$. For example, the set of all 0-cochains is $\prod_{\alpha_0\in I} F(U_{\alpha_0})$, which is equivalent to choosing a section for each element of the cover. Similarly, choosing an element from $C^1(\mathcal{U}, F)$ can be thought of as choosing a section for each intersection of two open sets from $\mathcal{U}$. Similarly one can interpret the higher cochains.

Next, we give the sequence of groups $\{C^i(\mathcal{U}, F)\}_{i\in\mathbb{N}\cup 0}$ the structure of a cochain complex. Indeed, one defines the required *differential* in quite an obvious manner, if one knows the construction of singular homology. Define a map

$$\begin{aligned} d : C^i(\mathcal{U}, F) &\longrightarrow C^{i+1}(\mathcal{U}, F) \\ s = (s(\alpha_0, \dots, \alpha_i)) &\longmapsto ds \end{aligned}$$

where the components of ds are given as follows for $\beta_0, \dots, \beta_{i+1} \in I$:

$$\begin{aligned} (ds)(\beta_0, \dots, \beta_{i+1}) &:= \sum_{j=0}^{i+1} (-1)^j \rho_j(s(\beta_0, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_{i+1})) \\ &= \sum_{j=0}^{i+1} (-1)^j \rho_j(s(\widehat{\beta_j})) \end{aligned}$$

[15]we choose increasing tuples only to make sure we don't repeat an open set in the product.

where ρ_j is the following restriction map of the presheaf F:

$$\rho_j : F(U_{\beta_0} \cap \cdots \cap U_{\beta_{j-1}} \cap U_{\beta_{j+1}} \cap U_{\beta_{i+1}}) \longrightarrow F(U_{\beta_0} \cap \cdots \cap U_{\beta_{j-1}} \cap U_{\beta_j} \cap U_{\beta_{j+1}} \cap U_{\beta_{i+1}}),$$

that is, the one where the open set U_{β_j} is dropped from intersection.

This differential can be understood in the simple case of $i = 0$ as follows. Take $s = (s(\alpha_0)) \in C^0(\mathcal{U}, F)$. Then $ds \in C^1(\mathcal{U}, F)$ and it corresponds to a choice of a section in the intersection on each pair of open sets in $\mathcal{U}$. For $\beta_0, \beta_1 \in I$, this choice is given by

$$(ds)(\beta_0, \beta_1) = \rho_0(s(\beta_1)) - \rho_1(s(\beta_0)).$$

This is interpreted as "how much far away $s(\beta_1) \in F(U_{\beta_1})$ and $s(\beta_0) \in F(U_{\beta_0})$ are in the intersection $U_{\beta_1} \cap U_{\beta_0}$". If $d(s) = 0$, then $s \in C^0(\mathcal{U}, F)$ corresponds to a matching family.

Similarly, for a $s \in C^1(\mathcal{U}, F)$, we can think of it as a choice of a section on each intersecting pair of open sets of $\mathcal{U}$. Then, the differential $ds \in C^2(\mathcal{U}, F)$ for any $(\beta_0, \beta_1, \beta_2) \in I^3$ has the component

$$(ds)(\beta_0, \beta_1, \beta_2) = \rho_0(s(\beta_1, \beta_2)) - \rho_1(s(\beta_0, \beta_2)) + \rho_2(s(\beta_0, \beta_1)).$$

If this quantity is non zero, then it measures "how much the three elements $s(\beta_1, \beta_2) \in F(U_{\beta_1} \cap U_{\beta_2})$, $s(\beta_0, \beta_2) \in F(U_{\beta_0} \cap U_{\beta_2})$ and $s(\beta_0, \beta_1) \in F(U_{\beta_0} \cap U_{\beta_1})$ differs in the combined intersection $U_{\beta_0} \cap U_{\beta_1} \cap U_{\beta_2}$". Indeed, suppose the three agree on $F(U_{\beta_0} \cap U_{\beta_1} \cap U_{\beta_2})$. Then, we have $\rho_0(s(\beta_1, \beta_2)) = \rho_1(s(\beta_0, \beta_2)) = \rho_2(s(\beta_0, \beta_1))$. Consequently, $ds(\beta_0, \beta_1, \beta_2) = \rho_2(s(\beta_0, \beta_1))$.

Now it is quite obvious that in order to measure the failure of an element of $C^i(\mathcal{U}, F)$ to "match up in one level above" will be measured by the homology of the cochain complex. Indeed that is what we do now.

For any $s \in C^i(\mathcal{U}, F)$, it is observed by doing the summation and using the fact that the restriction maps ρ are group homomorphisms that

$$d^2 = 0.$$

Hence, we have a cochain complex, called the **Čech cochain complex w.r.t.** $\mathcal{U}$:

$$C^0(\mathcal{U}, F) \xrightarrow{\ d\ } C^1(\mathcal{U}, F) \xrightarrow{\ d\ } C^2(\mathcal{U}, F) \xrightarrow{\ d\ } \cdots$$

The cohomology of this complex is denoted by

$$H^q(\mathcal{U}; F) := \frac{\operatorname{Ker}(d)}{\operatorname{Im}(d)} =: \frac{Z^q\mathcal{U}, \mathcal{F}}{B^q(\mathcal{U}, \mathcal{F})}$$

for $C^{q+1}(\mathcal{U}, F) \leftarrow C^q(\mathcal{U}, F) \leftarrow C^{q-1}(\mathcal{U}, F)$. The subgroup $B^q(\mathcal{U}, \mathcal{F}) = \operatorname{Im}(d) =\subseteq C^q(\mathcal{U}, F)$ is called the *group of q-coboundaries*, whereas the group $Z^q(\mathcal{U}, \mathcal{F}) = \operatorname{Ker}(d) \subseteq C^q(\mathcal{U}, F)$ is called the *group of q-cocycles*.

To define the general Čech cohomology groups, we need to take limit of cohomology groups with respect to finer and finer open covers. To this end, we first define the following. Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in I}$ and $\mathcal{V} = \{V_\beta\}_{\beta \in J}$ be two open covers. Then, $\mathcal{V}$ is said to be *finer* than $\mathcal{U}$ if for all $j \in J$, there is an $i \in I$ such that $V_j \subseteq U_i$. We therefore obtain a function $\sigma : J \to I$ such that $V_j \subseteq U_{\sigma(j)}$.

For two open covers $\mathcal{U}, \mathcal{V}$ where $\mathcal{V}$ is finer than $\mathcal{U}$ as above, we first get a map of cochain complexes given by

$$\begin{aligned} r_{\mathcal{U},\mathcal{V}} : C^q(\mathcal{U}, F) &\longrightarrow C^q(\mathcal{V}, F) \\ s &\longmapsto r_{\mathcal{U},\mathcal{V}}(s) \end{aligned}$$

where for any $(\beta_0, \dots, \beta_q) \in J^{q+1}$, we define

$$r_{\mathcal{U},\mathcal{V}}(s)(\beta_0, \dots, \beta_q) = \rho\left(s(\sigma\beta_0, \dots, \sigma\beta_q)\right)$$

for $\rho : F(U_{\sigma\beta_0} \cap \dots \cap U_{\sigma\beta_q}) \longrightarrow F(V_{\beta_0} \cap \dots \cap V_{\beta_q})$ is the restriction map of F. As restriction homomorphisms commute with themselves, therefore we have that the following square commutes

$$\begin{tikzcd} C^q(\mathcal{U}, F) \arrow[r, "d"] \arrow[d, "r_{\mathcal{U},\mathcal{V}}"'] & C^{q+1}(\mathcal{U}, F) \arrow[d, "r_{\mathcal{U},\mathcal{V}}"] \\ C^q(\mathcal{V}, F) \arrow[r, "d"'] & C^{q+1}(\mathcal{V}, F) \end{tikzcd},$$

showing that $r_{\mathcal{U},\mathcal{V}} : C^\bullet(\mathcal{U}, F) \to C^\bullet(\mathcal{V}, F)$ is a map of cochain complexes. Consequently, we get a map at the level of cohomology also denoted by

$$r_{\mathcal{U},\mathcal{V}} : H^q(\mathcal{U}, F) \longrightarrow H^q(\mathcal{V}, F).$$

We call the above the *refinement homomorphism*.

We now wish to show that if $\mathcal{V}$ is a refinement of $\mathcal{U}$ via $\sigma : J \to I$, then the refinement homomorphism $r_{\mathcal{U},\mathcal{V}}$ on cohomology doesn't depend on σ; there might be many such σ making $\mathcal{V}$ finer than $\mathcal{U}$, but all give same refinement homomorphism on cohomology.

Lemma 20.7.0.2. *The refinement homomorphism $r_{\mathcal{U},\mathcal{V}}$ is independent of σ.*

Proof. Let $r, r' : C^q(\mathcal{U}, F) \to C^q(\mathcal{V}, F)$ be the refinement homomorphisms on cochain level for $\sigma, \tau : J \to I$ respectively. Pick any q-cocycle $s \in C^q(\mathcal{U}, F)$. We wish to show that $r(s) - r'(s)$ is a q-coboundary w.r.t. $\mathcal{V}$. The following $t \in C^{q-1}(\mathcal{V}, F)$

$$t(\alpha_0, \dots, \alpha_{q-1}) := \sum_{j=0}^{q-1} (-1)^j \rho\left(s\left(\sigma\alpha_0, \dots, \sigma\alpha_j, \tau\alpha_j, \tau\alpha_{j+1}, \dots, \tau\alpha_{i-1}\right)\right)$$

where $\rho : F(U_{\sigma\alpha_0} \cap \dots \cap U_{\sigma\alpha_j} \cap U_{\tau\alpha_j} \cap \dots \cap U_{\tau\alpha_{i-1}}) \longrightarrow F(V_{\alpha_0} \cap \dots \cap V_{\alpha_j} \cap \dots \cap V_{\alpha_{i-1}})$ is such that

$$r(s) - r'(s) = dt$$

in $C^q(\mathcal{V}, F)$. This can be checked by expanding dt and using the fact that $ds = 0$. This calculation is omitted for being too cumbersome to write. $\square$

This finally allows us to define Čech cohomology of a presheaf over a topological space as follows. Let $\mathcal{O}$ be the poset of all open covers of X ordered by refinement. The **Čech cohomology groups of presheaf** F are then defined to be

$$\check{H}^q(X, F) := \varinjlim_{\mathcal{U} \in \mathcal{O}} H^q(\mathcal{U}, F).$$

Diagrammatically, we have for any two open covers $\mathcal{U}$ and $\mathcal{V}$ where $\mathcal{V}$ is a refinement of $\mathcal{U}$ the following

$$\begin{tikzcd} & \check{H}^q(X,F) \\ H^q(\mathcal{U},F) \arrow[ur] \arrow[r, "r_{\mathcal{U},\mathcal{V}}"'] & H^q(\mathcal{V},F) \arrow[u] \end{tikzcd}.$$

This completes the construction of Čech cohomology groups.

Let us first see something that we hinted during the construction.

Lemma 20.7.0.3. *Let X be a space and $\mathcal{F}$ be a sheaf over X. Then, for any open cover $\mathcal{U}$ of X, we have*

$$H^0(\mathcal{U},\mathcal{F}) \cong \Gamma(X,\mathcal{F}).$$

Consequently, we have $\check{H}^0(X,\mathcal{F}) \cong \Gamma(X,\mathcal{F})$.

Proof. We first have $H^0(X,F) = \mathrm{Ker}\,(d)$ where $d : C^0(\mathcal{U},F) \to C^1(\mathcal{U},F)$. But any $s \in \mathrm{Ker}\,(d)$ is equivalent to the data of a matching family over $\mathcal{U}$. As $\mathcal{F}$ is a sheaf, therefore this gives rise to a unique element in $\Gamma(X,\mathcal{F})$. Conversely, by restriction, we get an element of $\mathrm{Ker}\,(d)$ via a global section. □

Let us first see an example computation of $\check{H}^1(X,F)$.

Example 20.7.0.4. Let $X = S^1$ and $F = \mathcal{K}$ be the constant sheaf of a field K. Further, let $\mathcal{U}$ be the open cover obtained by dividing S^1 into n-open intervals $U_1,\dots,U_n$ where $U_i \cap U_{i+1}$ and $U_i \cap U_{i-1}$ are non-empty and $U_i \cap U_j$ is empty for all $j \neq i, i+1, i-1$. We wish to calculate $H^1(\mathcal{U},\mathcal{K})$. To this end, we first see that

$$C^0(\mathcal{U},\mathcal{K}) = \prod_{i=1}^n \mathcal{K}(U_i) = K^{\oplus n}$$

and

$$C^1(\mathcal{U},\mathcal{K}) = \prod_{i=1}^n \mathcal{K}(U_i \cap U_{i+1}) = K^{\oplus n}.$$

For $q \geq 2$, we clearly have $C^q(\mathcal{U},\mathcal{K}) = 0$ as there are no higher intersections. The differential $d : C^0(\mathcal{U},\mathcal{K}) \to C^1(\mathcal{U},\mathcal{K})$ maps as

$$d(x_1,\dots x_n) = (x_2 - x_1, x_3 - x_2, \dots, x_1 - x_n).$$

Consequently,

$$H^0(\mathcal{U},\mathcal{K}) = \mathrm{Ker}\,(d) = \{(x_1,\dots,x_n) \in C^0(\mathcal{U},\mathcal{K}) \mid x_1 = x_2 = \dots = x_n\} \cong K$$

and

$$H^1(\mathcal{U},\mathcal{K}) = \frac{C^1(\mathcal{U},\mathcal{K})}{\mathrm{Im}\,(d)} \cong K$$

as $C^1(\mathcal{U},\mathcal{K})$ is an n-dimensional K-vector space and $\mathrm{Im}\,(d)$ is of dimension $n-1$ because its defined by one equation deeming the sum of all entries to be 0.

Construction 20.7.0.5 (Map in cohomology)**.** Any map of abelian sheaves over X yields a map in the cohomology as well. Indeed, let $\varphi : \mathcal{F} \to \mathcal{G}$ be a map of sheaves. Then we get a map

$$\begin{aligned} \varphi^q : C^q(\mathcal{U}, \mathcal{F}) &\longrightarrow C^q(\mathcal{U}, \mathcal{G}) \\ s = (s(\alpha_0, \dots, \alpha_q)) &\longmapsto \varphi^q(s) = \left(\varphi_{\alpha_0 \dots \alpha_q}(s(\alpha_0, \dots, \alpha_q))\right) \end{aligned}$$

where $\varphi_{\alpha_0 \dots \alpha_q} = \varphi_{U_{\alpha_0} \cap \dots \cap U_{\alpha_q}}$.

It then follows quite immediately from the fact that each $\varphi_{\alpha_0 \dots \alpha_q}$ is a group homomorphism that $d\varphi^q = \varphi^{q+1} d$. It follows that we get a map of chain complexes

$$\varphi^\bullet : C^\bullet(\mathcal{U}, \mathcal{F}) \longrightarrow C^\bullet(\mathcal{U}, \mathcal{G}).$$

Hence, we get a map in cohomology

$$\varphi^q : H^q(\mathcal{U}, \mathcal{F}) \longrightarrow H^q(\mathcal{U}, \mathcal{G}).$$

Finally, this gives by universal property of direct limits a unique map

$$\varphi^q : \check{H}^q(X, \mathcal{F}) \longrightarrow \check{H}^q(X, \mathcal{G})$$

such that for every open cover $\mathcal{U}$, the following diagram commutes:

$$\begin{tikzcd} \check{H}^q(X, \mathcal{F}) \arrow[r, dashed, "\varphi^q"] & \check{H}^q(X, \mathcal{G}) \\ H^q(\mathcal{U}, \mathcal{F}) \arrow[u] \arrow[r, "\varphi^q"'] & H^q(\mathcal{U}, \mathcal{G}) \arrow[u] \end{tikzcd}$$

where vertical maps are the maps into direct limits.

The main tool for calculations with cohomology theories is the cohomology long exact sequence. We put below, without proof, the main theorem of Čech cohomology which gives a condition for an exact sequence of sheaves to induce this long exact sequence in cohomology. Recall X is paracompact if it is Hausdorff and every open cover has a locally finite refinement. Such spaces are always normal. We first give an explicit description of the first connecting homomorphism.

Construction 20.7.0.6 (Connecting homomorphism)**.** Let X be a topological space and

$$0 \longrightarrow \mathcal{F} \xrightarrow{\varphi} \mathcal{G} \xrightarrow{\psi} \mathcal{H} \longrightarrow 0$$

be an exact sequence of sheaves on X. We define the connecting homomorphism

$$\check{H}^0(X, \mathcal{H}) \xrightarrow{\delta} \check{H}^1(X, \mathcal{F})$$

as follows. First, pick any $h \in H^0(X, \mathcal{H}) = \Gamma(X, \mathcal{H})$. As ψ is surjective therefore there exists an open covering $\mathcal{U} = \{U_i\}_{i \in I}$ of X and $g_i \in \mathcal{G}(U_i)$ such that $\psi_{U_i}(g_i) = h|_{U_i}$. Using (g_i) and (U_i) we construct a 1-cocycle for $\mathcal{F}$ as follows. Observe that for each $i, j \in I$, we have $\psi_{U_i \cap U_j}(g_i - g_j) = 0$ in $\mathcal{H}(U_i \cap U_j)$. Thus, $g_i - g_j \in \operatorname{Ker}\left(\psi_{U_i \cap U_j}\right)$. By exactness guaranteed by Lemma 20.3.0.8, it follows that there exists $f_{\alpha_0\alpha_1} \in \mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1})$ such that $\varphi_{U_{\alpha_0}\cap U_{\alpha_0}}(f_{\alpha_0\alpha_1}) = g_{\alpha_0} - g_{\alpha_1}$, for each $\alpha_0, \alpha_1 \in I$. We claim that the element

$$f := (f_{\alpha_0\alpha_1})_{\alpha_0,\alpha_1} \in \prod_{(\alpha_0,\alpha_1)\in I^2} \mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1}) = C^1(\mathcal{U}, \mathcal{F})$$

is a 1-cocycle. Indeed, we need only check that $df = 0$ in $C^2(\mathcal{U}, \mathcal{F})$. Pick any $(\alpha_0, \alpha_1\alpha_2) \in I^3$. We wish to show that $df(\alpha_0, \alpha_1\alpha_2) = 0$. Indeed,

$$\begin{aligned} df(\alpha_0, \alpha_1\alpha_2) &= \sum_{j=0}^{2} (-1)^j \rho_j \left(f_{\alpha_0\hat{\alpha}_j\alpha_2}\right) \\ &= f_{\alpha_1\alpha_2} - f_{\alpha_0\alpha_2} + f_{\alpha_0\alpha_1} \end{aligned}$$

in $\mathcal{F}(U_{\alpha_0} \cap U_{\alpha_1} \cap U_{\alpha_2})$. We claim the above is zero. Indeed, By Lemma 20.3.0.8 on $V := U_{\alpha_0} \cap U_{\alpha_1} \cap U_{\alpha_2}$ we get that φ_V is injective. But since

$$\begin{aligned} \varphi_V(f_{\alpha_1\alpha_2} - f_{\alpha_0\alpha_2} + f_{\alpha_0\alpha_1}) &= \varphi_V(f_{\alpha_1\alpha_2}) - \varphi_V(f_{\alpha_0\alpha_2}) + \varphi_V(f_{\alpha_0\alpha_1}) \\ &= g_{\alpha_1} - g_{\alpha_2} - (g_{\alpha_0} - g_{\alpha_2}) + g_{\alpha_0} - g_{\alpha_1} \\ &= 0, \end{aligned}$$

hence it follows that $df(\alpha_0, \alpha_1\alpha_2) = 0$, as required. Hence $f \in C^1(\mathcal{U}, \mathcal{F})$ is a 1-cocycle. Thus we get an element $[f] \in H^1(\mathcal{U}, \mathcal{F})$. This defines a group homomorphism $\check{H}^0(X, \mathcal{H}) \to H^1(\mathcal{U}, \mathcal{F})$. Further by passing to direct limit, we get an element $[f] \in \check{H}^1(X, \mathcal{F})$. We thus define

$$\delta(f) := [f] \in \check{H}^1(X, \mathcal{F}).$$

This defines the required group homomorphism δ.

Theorem 20.7.0.7. *Let X be a paracompact space and the following be an exact sequence of sheaves over X*

$$0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to 0.$$

Then, there is a long exact sequence in cohomology

$$\begin{tikzcd} 0 \arrow[r] & \check{H}^0(X, \mathcal{F}_1) \arrow[r] & \check{H}^0(X, \mathcal{F}_2) \arrow[r] & \check{H}^0(X, \mathcal{F}_3) \arrow[dll] \\ & \check{H}^1(X, \mathcal{F}_1) \arrow[r] & \check{H}^1(X, \mathcal{F}_2) \arrow[r] & \check{H}^1(X, \mathcal{F}_3) \arrow[dl, dashed, no head] \\ & & {} & \end{tikzcd}.$$

20.8 Derived functor cohomology

We will here define the cohomology of abelian sheaves over a topological space as right derived functors of the left exact global-sections functor (see Section 19.2 for preliminaries on derived functors).

Let X be a topological space. In Section 20.6, we showed that the category of abelian sheaves $\mathbf{Sh}(X)$ has enough injectives. We now use it to define cohomology of $\mathcal{F} \in \mathbf{Sh}(X)$.

Definition 20.8.0.1. (**Sheaf cohomology functors**) Let X be a topological space and $\mathbf{Sh}(X)$ be the category of abelian sheaves over X. The i^{th}-cohomology functor $H^i(X,-) : \mathbf{Sh}(X) \to \mathbf{AbGrp}$ is defined to be the i^{th}-right derived functor of the global sections functor $\Gamma(-,X) : \mathbf{Sh}(X) \to \mathbf{AbGrp}$. In other words, $H^i(X,\mathcal{F})$ for $\mathcal{F} \in \mathbf{Sh}(X)$ is defined by choosing an injective resolution $0 \to \mathcal{F} \xrightarrow{\epsilon} \mathcal{I}^\bullet$ in $\mathbf{Sh}(X)$ and then

$$H^i(X,\mathcal{F}) := h^i(\Gamma(X,\mathcal{I}^\bullet)).$$

As sheaf cohomology functors are in particular derived functors, so they satisfy results from Section 19.2.3. The main point in particular being that sheaf cohomology induces a long exact sequence in cohomology from a short exact sequence of sheaves. This will be our primary source of computations.

20.8.1 Flasque sheaves & cohomology of $\mathcal{O}_X$-modules

We would like to see the following theorem.

Theorem 20.8.1.1. *Let $(X,\mathcal{O}_X)$ be a ringed space. Then the right derived functors of $\Gamma(-,X) : \mathbf{Mod}(\mathcal{O}_X) \to \mathbf{AbGrp}$ is equal to the restriction of the cohomology functors $H^i(X,-) : \mathbf{Sh}(X) \to \mathbf{AbGrp}$.*

Remember that $\mathbf{Mod}(\mathcal{O}_X)$ has enough injectives (Theorem 3.5.2.2) but, apriori, the above two functors might be different because an injective object in $\mathbf{Mod}(\mathcal{O}_X)$ may not be injective in $\mathbf{Sh}(X)$. Consequently, the above result is important because its relevance in rectifying the cohomology of $\mathcal{O}_X$-modules (which are of the only utmost interest in algebraic geometry) to that of the usual sheaf cohomology functors. Hence, we may completely work inside the module category $\mathbf{Mod}(\mathcal{O}_X)$. Clearly to prove such a result, we need a bridge between injective modules in $\mathbf{Mod}(\mathcal{O}_X)$ and either injective or acyclic objects in $\mathbf{Sh}(X)$. Indeed, we will see that this bridge is provided by the realization that injective modules in $\mathbf{Mod}(\mathcal{O}_X)$ are acyclic because they are *flasque*.

Definition 20.8.1.2 (**Flasque sheaves**)**.** A sheaf $\mathcal{F}$ on X is said to be flasque if all restriction maps of $\mathcal{F}$ are surjective.

The following is a simple, yet important class of examples of flasque sheaves.

Example 20.8.1.3. Let X be an irreducible topological space and $\mathcal{A}$ be the constant sheaf over X for an abelian group A. We claim that $\mathcal{A}$ is flasque. Indeed, first recall that any open subspace $U \subseteq X$ is irreducible, therefore connected. Conseuqently, all restrictions are $\rho : \mathcal{A}(V) \to \mathcal{A}(U)$ are identity maps $\mathrm{id} : A \to A$ (see Remark 20.1.0.3). In-fact this shows that on an irreducible space, any constant sheaf $\mathcal{A}$ of abelian group A has section over any open set U as $\mathcal{A}(U) = A$ and all restrictions are identities.

An important property of flasque sheaves is that they have no obstruction to lifting of sections, a hint to their triviality in cohomology. However, the proof of this is quite non-constructive and thus a bit enlightening.

Theorem 20.8.1.4. *Let X be a space. If $0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to 0$ is an exact sequence of sheaves and $\mathcal{F}_1$ is flasque, then we have an exact sequence of sections over any open $U \subseteq X$*

$$0 \to \mathcal{F}_1(U) \to \mathcal{F}_2(U) \to \mathcal{F}_3(U) \to 0.$$

Proof. By left-exactness of global sections functor, we need only show the surjectivity of $\Gamma(\mathcal{F}_2, X) \to \Gamma(\mathcal{F}_3, X)$. To this end, pick any $s \in \Gamma(\mathcal{F}_3, X)$. We wish to lift this to an element of $\Gamma(\mathcal{F}_2, X)$. Consider the following poset

$$\mathcal{P} = \{(U, t) \mid U \subseteq X \text{ open } \& \ t \in \mathcal{F}_2(U) \text{ is a lift of } s|_U\}$$

where $(U, t) \leq (U', t')$ iff $U' \supseteq U$ and $t'|_U = t$. We reduce to showing that $\mathcal{P}$ has a maximal element and it is of the form (X, t). This will conclude the proof.

To show the existence of a maximal element, we will use Zorn's lemma. Pick any toset of $\mathcal{P}$ denoted $\mathcal{T}$. We wish to show that it is upper bounded. Indeed, let $V = \bigcup_{(U,t)\in\mathcal{T}} U$ and $\tilde{t} \in \mathcal{F}_2(V)$ be the section obtained by gluing $t \in \mathcal{F}_2(U)$ for each $(U, t) \in \mathcal{T}$ (they form a matching family because $\mathcal{T}$ is totally ordered). We thus have $(V, \tilde{t})$ which we wish to show is in $\mathcal{P}$. Indeed, as $\tilde{t}$ is obtained by lifts of restrictions of s, therefore $\tilde{t}$ is a lift of $s|_V$ by locality of sheaf $\mathcal{F}_3$. This shows that $\mathcal{P}$ has a maximal element, denote it by $(V, \tilde{t})$.

We finally wish to show that $V = X$. Indeed, if not, then $V \subsetneq X$. Pick any point $x \in X \setminus V$. Since we have a surjective map on stalks $\mathcal{F}_{2,x} \to \mathcal{F}_{3,x} \to 0$, hence the germ $(X, s)_x \in \mathcal{F}_{3,x}$ can be lifted to $(U, a)_x$ for some open $U \ni x$ and $a \in \mathcal{F}_2(U)$. We now have two cases. If $U \cap V = \emptyset$, then $(V \cup U, \tilde{t} \amalg a)$ is a lift of $s|_{V\cup U}$, contradicting the maximality of $(V, \tilde{t})$. On the other hand, suppose we have $U \cap V \neq \emptyset$. Let $W = U \cap V$. Since $W \subseteq V$, therefore we have $t_W \in \mathcal{F}_2(W)$ a lift of $s|_W$. Moreover, by restriction, we have $a \in \mathcal{F}_2(W)$ also a lift of $s|_W$. It follows that $a - t_W \in \mathcal{F}_1(W)$. As $\mathcal{F}_1$ is flasque, therefore there exists $b \in \Gamma(\mathcal{F}_1, X)$ which extends $a - t_W$. Consequently, we have $a - b = t_W \in \mathcal{F}_2(W)$. Observe that $a - b \in \mathcal{F}_2(U)$ is also a lift of $s|_U$ because $b = 0$ in $\Gamma(\mathcal{F}_3, X)$ by the left-exactness of global sections functor. It follows that $(U, a - b)$ and $(V, \tilde{t})$ is a matching family, which glues to $(U \cup V, c)$ where c is a lift of $s|_{U\cup V}$ as well, contradicting the maximality of $(V, \tilde{t})$. ☐

Corollary 20.8.1.5. *Let X be a space. If $0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to 0$ is an exact sequence of sheaves where $\mathcal{F}_2$ is flasque, then $\mathcal{F}_3$ is flasque.*

Proof. This is immediate from Theorem 20.8.1.4 and the following diagram where $U \supseteq V$ an inclusion of open subsets of X:

$$\begin{tikzcd}
\mathcal{F}_2(U) \arrow[r] \arrow[d, two heads, "\rho"'] & \mathcal{F}_3(U) \arrow[r] \arrow[d, "\rho"] & 0 \\
\mathcal{F}_2(V) \arrow[r] & \mathcal{F}_3(V) \arrow[r] & 0
\end{tikzcd}.$$

☐

Lemma 20.8.1.6. *Let $(X, \mathcal{O}_X)$ be a ringed space and $\mathcal{F}$ be an $\mathcal{O}_X$-module. Denote $\mathcal{O}_U = i_!\mathcal{O}_{X|U}$ to be the extension by zeros of $\mathcal{O}_{X|U}$ for any open set $i : U \hookrightarrow X$. Then,*

$$\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_U, \mathcal{F}) \cong \mathcal{F}(U).$$

Proof. Indeed, we have the following isomorphisms

$$\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_U, \mathcal{F}) \cong \mathrm{Hom}_{\mathcal{O}_{X|U}}(\mathcal{O}_{U|U}, \mathcal{F}_{|U}) \cong \mathrm{Hom}_{\mathcal{O}_{X|U}}(\mathcal{O}_{X|U}, \mathcal{F}_{|U}) \cong \mathcal{F}(U).$$

The first isomorphism follows from the universal property of sheafification. The second isomorphism follows from the observation that $\mathcal{O}_{U|U} = \mathcal{O}_{X|U}$ as is clear from Definition 20.3.0.9 and the fact that sheafification of a sheaf is that sheaf back. The last isomorphism follows from Lemma 3.5.1.20, 2. $\square$

Proposition 20.8.1.7. *Let $(X, \mathcal{O}_X)$ be a ringed space. If $\mathcal{I}$ is an injective $\mathcal{O}_X$-module, then $\mathcal{I}$ is flasque.*

Proof. Let $i : U \hookrightarrow X$ be an open set. Denote $\mathcal{O}_U = i_! \mathcal{O}_{X|U}$ (see Definition 20.3.0.9). We know from Lemma 20.8.1.6 that $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_U, \mathcal{I}) \cong \mathcal{I}(U)$ for any open $U \subseteq X$. Now, let $U \subseteq V$ be an inclusion of open sets. To this, we get $\rho : \mathcal{I}(V) \to \mathcal{I}(U)$ the restriction map. Restricting to open set V, we get the following injective map by Corollary 20.3.0.11

$$0 \to \mathcal{O}_U \to \mathcal{O}_V.$$

Using injectivity of $\mathcal{I}$, we obtain a surjection

$$\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_V, \mathcal{I}) \to \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_U, \mathcal{I}) \to 0.$$

Consequently, we have

$$\mathcal{I}(V) \to \mathcal{I}(U) \to 0$$

where the map is the restriction map of sheaf $\mathcal{I}$. Indeed, this follows from the explicit isomorphism $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_X, \mathcal{I}) \cong \mathcal{I}(X)$ constructed in the proof of Lemma 3.5.1.20, 2. $\square$

Finally, we see that flasque sheaves have trivial cohomology.

Proposition 20.8.1.8. *Let X be a space and $\mathcal{F}$ be a flasque sheaf over X. Then*

$$H^i(X, \mathcal{F}) = 0$$

for all $i \geq 1$. That is, flasque sheaves are acyclic for the global sections functor.

Proof. Let $0 \to \mathcal{F} \to \mathcal{I}$ be an injective map where $\mathcal{I}$ is an injective sheaf. Consequently, we have an exact sequence of sheaves

$$0 \to \mathcal{F} \to \mathcal{I} \to \mathcal{G} \to 0$$

where $\mathcal{G} = \mathcal{I}/\mathcal{F}$. It follows from Proposition 20.8.1.7 that $\mathcal{I}$ is flasque. By Corollary 20.8.1.5 it follows that $\mathcal{G}$ is flasque. By Theorem 19.2.3.5, we have a long exact sequence in cohomology

$$\begin{tikzcd}
H^i(X,\mathcal{F}) \arrow[r] & H^i(X,\mathcal{I}) \arrow[r] & H^i(X,\mathcal{G}) \arrow[dll, "\delta_i"'] \\
H^{i+1}(X,\mathcal{F}) \arrow[r] & H^{i+1}(X,\mathcal{I}) \arrow[r] & H^{i+1}(X,\mathcal{G}) \arrow[dll, dotted, "\delta_{i+1}"'] \\
{} & &
\end{tikzcd}.$$

Since $\mathcal{I}$ is injective, therefore by Remark 19.2.3.4, we have $H^i(X, \mathcal{I}) = 0$ for all $i \geq 1$. It follows from exactness of the above diagram that δ_i are isomorphisms for each $i \geq 1$, that is,

$$H^i(X, \mathcal{G}) \cong H^{i+1}(X, \mathcal{F}).$$

But since $\mathcal{G}$ is also flasque, therefore by repeating the above process, we deduce that $H^{i+1}(X, \mathcal{F}) \cong H^1(X, \mathcal{H})$ where $\mathcal{H}$ is some other flasque sheaf. It thus suffices to show that $H^1(X, \mathcal{F}) = 0$. This follows immediately as we have an exact sequence

$$0 \to \Gamma(\mathcal{F}, X) \to \Gamma(\mathcal{I}, X) \to \Gamma(\mathcal{G}, X) \to H^1(X, \mathcal{F}) \to 0$$

where by Theorem 20.8.1.4, the map $\Gamma(\mathcal{I}, X) \to \Gamma(\mathcal{G}, X)$ is surjective and since $\Gamma(\mathcal{G}, X) \to H^1(X, \mathcal{F})$ is surjective by exactness, it follows that the map $\Gamma(\mathcal{G}, X) \to H^1(X, \mathcal{F})$ is the zero map and $H^1(X, \mathcal{F}) = 0$, as required. □

An immediate corollary is the proof of Theorem 20.8.1.1.

Proof of Theorem 20.8.1.1. Pick any $\mathcal{F} \in \mathbf{Mod}(\mathcal{O}_X)$ and pick an injective resolution of $\mathcal{F}$ in $\mathbf{Mod}(\mathcal{O}_X)$

$$0 \to \mathcal{F} \xrightarrow{\epsilon} \mathcal{I}^\bullet.$$

By Proposition 20.8.1.7, it follows that each $\mathcal{I}^i$ is flasque. By Proposition 20.8.1.8, it follows that the above is an acyclic resolution for the sheaf $\mathcal{F}$ in $\mathbf{Sh}(X)$. Denote by $\bar{\Gamma} : \mathbf{Mod}(\mathcal{O}_X) \to \mathbf{AbGrp}$ the restriction of the global sections functor. We wish to show that $R^i\bar{\Gamma}(\mathcal{F}) \cong H^i(X, \mathcal{F})$. By Proposition 19.2.3.9, we have the following isomorphism

$$R^i\bar{\Gamma}(\mathcal{F}) \cong h^i(\bar{\Gamma}(\mathcal{I}^\bullet)) = h^i(\Gamma(\mathcal{I}^\bullet)) \cong H^i(X, \mathcal{F}),$$

as needed. □

An important property of flasque sheaves over noetherian spaces is that it is closed under direct limits.

Proposition 20.8.1.9. *Let X be a noetherian space and $\{\mathcal{F}_\alpha\}$ be a directed system of flasque sheaves. Then $\varinjlim \mathcal{F}_\alpha$ is a flasque sheaf as well.*

Proof. **TODO.** □

Examples

We now present some computations.

Example 20.8.1.10. [16] Let $X = \mathbb{A}^1_k$ be the affine line over an infinite field k and $\mathbb{Z}$ be the constant sheaf over X. Let $P, Q \in X$ be two distinct closed points and let $U = X \setminus C$ where $C = \{P, Q\}$ be an open set. Denote $\mathbb{Z}_U$ to be the extension by zero sheaf of $\mathbb{Z}_{|U}$ over X. We claim that

$$H^1(X, \mathbb{Z}_U) \neq 0.$$

[16] Exercise III.2.1, a) of Hartshorne.

We will use the extension by zero short exact sequence of Corollary 20.3.0.11. Denote $i : C \hookrightarrow X$ to be the inclusion. Then, we have

$$0 \to \mathbb{Z}_U \to \mathbb{Z} \to i_*\mathbb{Z}_{|C} \to 0.$$

By Theorem 19.2.3.5 and Example 20.8.1.3, it follows that the following sequence is exact

$$0 \to \Gamma(\mathbb{Z}_U, X) \to \Gamma(\mathbb{Z}, X) \to \Gamma(i_*\mathbb{Z}_{|C}, X) \to H^1(X, \mathbb{Z}_U) \to 0.$$

Now suppose that $H^1(X, \mathbb{Z}_U) = 0$. It follows that the map $\Gamma(\mathbb{Z}, X) \to \Gamma(i_*\mathbb{Z}_{|C}, X)$ is surjective. Since X is irreducible and hence connected, we yield $\Gamma(\mathbb{Z}, X) = \mathbb{Z}$. Consequently, we have a surjective map $\mathbb{Z} \to \Gamma(i_*\mathbb{Z}_{|C}, X)$. It follows that $\Gamma(i_*\mathbb{Z}_{|C}, X) = \mathbb{Z}$ or $\mathbb{Z}/n\mathbb{Z}$. We claim that this is not possible by showing that $\Gamma(i_*\mathbb{Z}_{|C}, X)$ is isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$, which will yield a contradiction.

We first observe that $\Gamma(i_*\mathbb{Z}_{|C}, X) = \Gamma(\mathbb{Z}_{|C}, C)$. Recall that $\mathbb{Z}_{|C} = i^{-1}\mathbb{Z}$. Note that $(\mathbb{Z}_{|C})_P = \mathbb{Z}_p = \mathbb{Z} = (\mathbb{Z}_{|C})_Q$ by Lemma 20.5.0.3. Hence, by Definition 20.5.0.2 and Remark 20.2.0.4, we deduce that $(i^+\mathbb{Z})_P = (i^+\mathbb{Z})(\{P\}) = \mathbb{Z}_P = \mathbb{Z} = (i^+\mathbb{Z})_Q$ and

$$\Gamma(\mathbb{Z}_{|C}, C) = \begin{Bmatrix} (s,t) \in \mathbb{Z} \oplus \mathbb{Z} \mid \exists \text{ opens } U_P \ni P, U_Q \ni \\ Q \text{ in } C \ \& \ s' \in i^+\mathbb{Z}(U_P) \ \& \ t' \in \\ i^+\mathbb{Z}(U_Q) \text{ s.t. } s = s'_P,\ t = t'_Q,\ s = t'_P \text{ if } P \in \\ U_Q \ \& \ t = s'_Q \text{ if } Q \in U_P. \end{Bmatrix}$$

With this, we observe that for each $(s,t) \in \mathbb{Z} \oplus \mathbb{Z}$, if we keep $U_P = \{P\}$ and $U_Q = \{Q\}$ (which is possible since $P \neq Q$ are closed points in X), we obtain $i^+\mathbb{Z}(U_P) = \mathbb{Z} = i^+\mathbb{Z}(U_Q)$. Then, we may take $s' = s$ and $t' = t$ to obtain that $\Gamma(\mathbb{Z}_{|C}, C) \cong \mathbb{Z} \oplus \mathbb{Z}$. This completes the proof.

Moreover, one can see that the only properties of $\mathbb{A}^1_k$ that we needed was that it is irreducible and $P, Q \in \mathbb{A}^1_k$ are distinct closed points. Consequently, the above result holds true for X an arbitrary irreducible space and $U = X \setminus \{P, Q\}$ where P, Q are two distinct closed points.

Example 20.8.1.11. Consider the notations of Example 20.8.1.10. We give another simple calculation of

$$\Gamma(i_*\mathbb{Z}_{|C}, X) = \mathbb{Z} \oplus \mathbb{Z}.$$

TODO.

Example 20.8.1.12. Consider the notations of Example 20.8.1.10. As an exercise in working with sheaves and sheafification, we also show that

$$\Gamma(\mathbb{Z}_U, X) = 0.$$

TODO.

20.8.2 Čech-to-derived functor spectral sequence

We wish to now observe how the Čech cohomology and derived functor cohomology are related. This is done by a spectral sequence.